Aus der

Universitätsklinik für Hals-, Nasen- und

Ohrenheilkunde mit Poliklinik Tübingen

# Differentielle Analyse altersabhängiger cochleärer Synaptopathien in Probanden mit und ohne Sprachverständlichkeits-Einschränkungen

**Inaugural-Dissertation**
**zur Erlangung des Doktorgrades**
**der Medizin**

**der Medizinischen Fakultät**
**der Eberhard Karls Universität**
**zu Tübingen**


**vorgelegt von**

**Schirmer, Jakob Leonard**


**2026**

Dekanin: Professorin Dr. S. Y. Brucker

1. Berichterstatter: Professorin Dr. M. Knipper-Breer
2. Berichterstatter: Professor Dr. C.Plewnia

Tag der Disputation: 20.03.2026

I dedicate this dissertation to my parents, for enabling me in all my decisions.

Without you, I could not have reached where I am today.

# Contents

# List of figures



# List of tables

# List of Appendices



# List of Abbreviations

| | |
|---|---|
| ABR | Auditory brainstem response |
| AMPA receptor | α-amino-3-hydroxy-5-methyl-isoxazolepropionic acid receptor |
| ANF | Auditory nerve fibre |
| ARHL | Age-related hearing loss |
| ASSR | Auditory steady-state response |
| ATPase | AdenosineTriPhosphatase |
| BDI-II | Beck's depression index II |
| Ca2+ | Calcium |
| dB | Decibel |
| DPOAE | Distortion product otoacoustic emissions |
| ENT | Ear-nose-throat; Otorhinolaryngology |
| GDS | Geriatric depression scale |
| HHL | Hidden hearing loss |
| HL | Hearing level |
| Hz | Hertz |
| IHC | Inner hair cell |
| IPL | Inter-peak latency |
| K+ | Potassium |
| $L_{EDPT}$ | Level of the estimated distortion product threshold |
| MEMR | Middle-ear muscle reflex |
| MMSE | Mini-mental status examination |
| ms | Millisecond |
| Na+ | Sodium |
| NIHL | Noise-induced hearing loss |

| | |
|---|---|
| OAE | Otoacoustic emissions |
| OHC | Outer hair cell |
| OLSA | Oldenburger Satztest |
| PC | Principal components |
| pDPOAE | Pulsed DPOAE |
| PLL | Phase-locking limit |
| PNOT | Pure-tone normalized threshold |
| PTA | Pure-tone audiometry |
| PTT | Pure-tone threshold |
| SD | Standard deviation |
| SGC | Spiral ganglion cell |
| SNR | Signal-to-noise ratio |
| SR | Spontaneous rate |
| SOAE | Spontaneous otoacoustic emissions |
| SPL | Sound-pressure level |
| $SRT_{50}$ | Speech Reception Threshold of 50% comprehension |
| TENV | Temporal envelope |
| TEOAE | Transient evoked otoacoustic emissions |
| TFS | Temporal fine-structure |
| WHO | World Health Organisation |

# 1 Introduction

## 1.1 Global relevance of hearing loss

Hearing-loss is one of the health challenges with increasing relevance across all aging populations at the moment. According to World Health Organisation (WHO) projections from 2018, the estimated number of 466 million people worldwide with hearing loss can be expected to keep rising substantially within the next decades, as a result of demographic developments (World Health Organization, 2018). In the United States alone, numbers are estimated to rise from 38 million in 2016 to 73 million people in 2060 (Goman and Lin, 2016, Goman et al., 2017). These numbers only consider "disabling hearing-loss" as part of the statistic, while a larger, but often overlooked number of people may already suffer from the consequences of developing hearing loss that does not fit into practical diagnostic criteria. While one may argue that hearing loss does not compare to other more acute illnesses, its long-term impacts on society cannot be ignored. The WHO estimates annual health costs from hearing loss at around 750 billion dollars worldwide, but it is clear that these health costs in the form of diagnostic and treatment are spread unevenly across countries. Not only does hearing-loss lead to decreasing quality of life of those affected but it is also thought to impact the development of other severe (and expensive) illnesses, such as depression or dementia. Livingston et al. identified hearing-loss as the third most relevant factor for the development of dementia, while being the largest potentially modifiable factor (Livingston et al., 2017). Even subclinical hearing loss has been shown to correlate with dementia prevalence (Golub et al., 2020, Hoppe et al., 2022). While the exact mechanisms leading from hearing-loss to cognitive decline are yet to be proven (Griffiths et al., 2020), common factors like aging, microvascular disease and education (Lin and Albert, 2014) could explain the association of hearing loss and dementia. Mechanistic pathways leading from hearing loss to dementia are also being evaluated, such as impacts on cognitive load (Tun et al., 2009), brain structure (Lin et al., 2014) and reduced social engagement (Fratiglioni et al., 2000). Improved treatment of hearing-loss could contribute to decelerating the increasing prevalence of dementia (Montero-Odasso et al., 2020).

One aspect of averting hearing loss is the reduction or prevention of modifiable risk factors. These have been shown to include smoking (Ferrite and Santana, 2005), cardiovascular disease (Tan et al., 2018), fitness (Chen et al., 2015, Han et al., 2016), diet (Kopke et al., 2005), and ototoxic medication. Exposure to loud music or noise is another relevant, modifiable factor. Whether of recreational or professional nature, exposure to loud sounds over prolonged time can cause irreversible damage to inner-ear hair cells. Especially teenagers and young adults can significantly decrease their risk of developing noise-induced hearing loss by using hearing protection in disco/club and concert settings (Daniel, 2007). To fully understand the different aetiologies of hearing loss, and deriving how to treat them, one must first understand the physiological mechanisms involved in transmitting an acoustic sound wave from the outer ear to the auditory cortex.

## 1.2 Physiology of hearing

Sound waves reaching the auricle are channelled into the outer ear canal. This approximately 2-2.5 cm canal leads into the petrous bone of the skull and ends with the tympanic membrane. Here, the energy of the sound wave is transmitted as oscillations through the tympanic membrane and then onto the chain of ossicles – malleus, incus and stapes (Brandes et al., 2019). The ossicles mechanically couple the tympanic membrane with the oval window at the base of the cochlea, while improving the efficiency of transmission through amplification of the sound waves by lever action of the ossicles and the area reduction between both membranes. By this mechanism, the impedance is patched from air to liquid, and the sound waves are amplified at least 18-fold. The amplification can also be regulated, by contraction of two muscles within the eardrum, the *musculus tensor tympani* (tightens the tympanic membrane) and the *musculus stapedius* (reduces transmission to the oval window) (Brandes et al., 2019). The oval window leads to a fluid-filled system of three spiral ducts within the cochlea (Fig. 1). These are named the scala media, scala vestibuli and scala tympani, with the latter two being connected at the apex of the cochlea – the helicotrema. The scala vestibuli and scala tympani are both filled with perilymph, a fluid that resembles extracellular fluids and is rich in sodium ions ($Na^+$). The scala media, situated in between scala vestibuli and scala tympani, holds the specialized potassium ($K^+$) rich endolymph. Endolymph is produced and maintained in its ion balance by the stria vascularis, a capillary loop in the outer wall of the cochlear duct (Brandes et al., 2019). Sound waves spread along the cochlea, from base to apex as a travelling wave. Sitting on top of the basilar membrane in the scala media is the Organ of Corti, the primary sensory organ within the cochlea (Fig. 1). The

Organ of Corti contains mechanical receptor cells, 3 rows of outer hair cells (OHCs) and a single row of inner hair cells (IHCs). Vibrations of the travelling wave along the tectorial- and basilar membranes of the cochlea deflect the tips of the stereocilia of hair cells, activating the opening of $K^+$-channels. The flow of $K^+$ into the hair cell depolarizes the hair cell. The cochlea is arranged tonotopically, meaning that each frequency of sound wave activates a different but specific group of cells along the cochlea. A gradient of increasing flexibility and width of the basilar membrane from base to apex is responsible for high frequencies to activate the hair cells at the base, and low frequencies to yield the highest deflection of hair cells at the apex. This tonotopic distribution of frequencies along the basilar membrane is projected on within the fibres of the auditory nerve into the auditory cortex (Brandes et al., 2019).

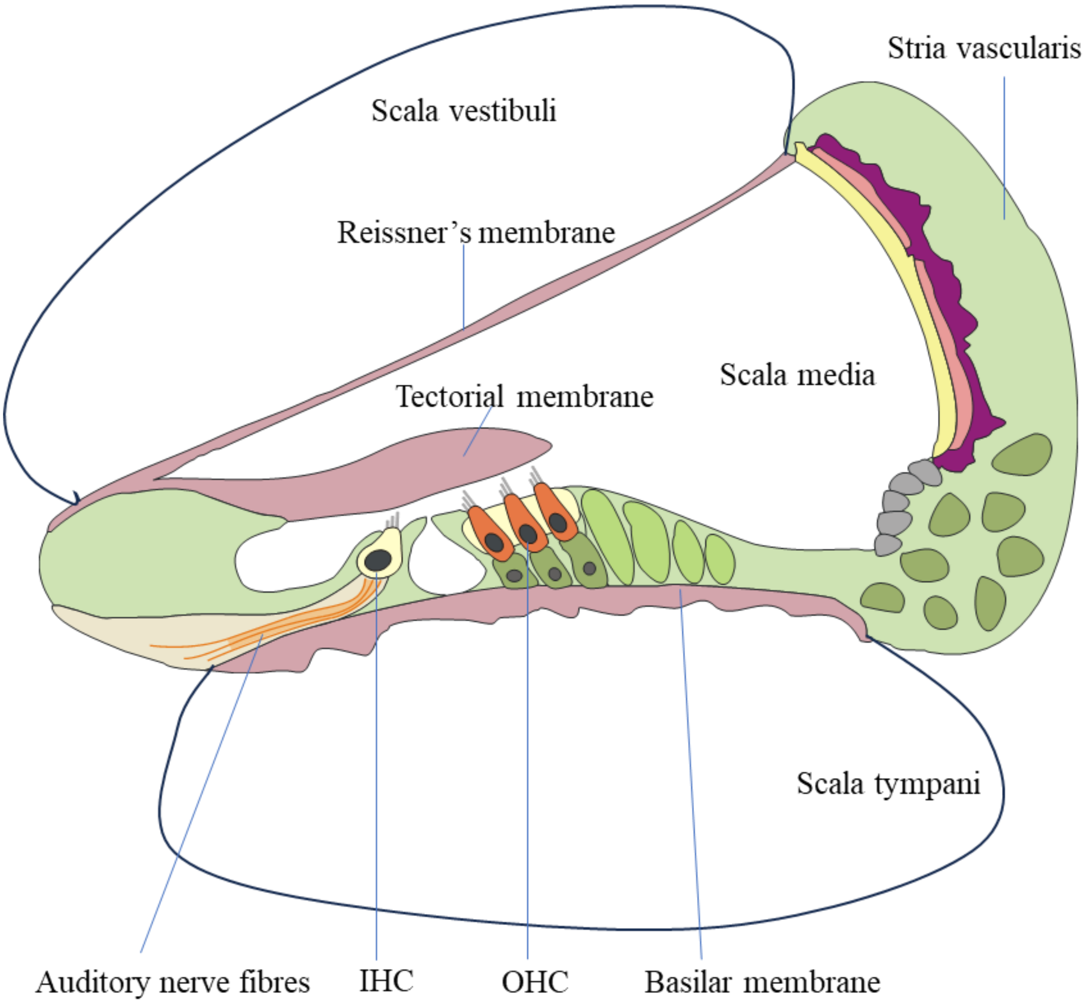



**Fig. 1** Schematic cross-section of the cochlear duct.
The K+ rich scala media lies in between the Na+ rich scala vestibuli and scala tympani. Separated by the Reissner's membrane and the basilar membrane. The electrolyte gradient is upheld by the stria vascularis, which secretes the endolymph for the scala media. The organ of Corti sits on the basilar membrane and holds 3 rows of outer hair cells (OHC) and one row of inner hair cells (IHC). Auditory signals to the cochlea are transmitted as vibrations of the tectorial membrane, deflecting the stereocilia of the OHC's. The OHC's act as mechanical receptors to amplify the auditory signal by fast conformation changes of the prestin molecule. Deflections of IHC stereocilia then leads to depolarization and exocytosis of glutamate into the synapse to activate the bipolar sensory neurons of the auditory nerve (Brandes et al., 2019).
.

OHCs are responsible for the nonlinear pre-amplification of the travelling wave at the frequency-specific place along the cochlea, before the wave reaches the sensory receptors. Deflection of stereocilia on the three rows of OHCs causes an influx of $K^+$-ions into the hair cell, depolarizing it. Depolarization of the OHC leads to fast conformation changes of prestin-molecules, which contracts the hair cell and amplifies the vibrations of the tectorial membrane (Zheng et al., 2000). The loss of this cochlear amplifier leads to hearing losses of 40-60 dB in mice (Liberman et al., 2002).

Deflection of stereocilia of the IHCs leads to depolarization of the hair cell, which activates voltage-controlled calcium ($Ca^{2+}$) channels. The resulting $Ca^{2+}$ inflow activates synaptic vesicles through exocytosis, releasing glutamate into the synapse. At the postsynapse, glutamate binds to α-amino-3-hydroxy-5-methyl-isoxazolepropionic acid receptors (AMPA receptors) of the bipolar sensory neurons, where the action potential is generated (Klinke, 1986, Eybalin, 1993, Puel, 1995). The approximately 35000 bipolar sensory neurons make up the human auditory nerve, their nuclei lie within the spiral ganglion, which is situated in the modiolus, the conical central axis of the cochlea. (Spoendlin and Schrott, 1989). Counts of innervation densities showed that 5% of these neurons are unmyelinated and connect the nuclei in the spiral ganglion with the OHCs in the organ of Corti (Spoendlin, 1972). Through these efferent fibres, neurons of the superior olivary complex exhibit inhibitory innervation of the OHCs. In response to loud sounds, the cochlear amplification at specific frequencies on the cochlea can be inhibited by regulating OHC function (Siegel and Kim, 1982). The remaining 95% of fibres are responsible for afferent innervation of the IHCs (Spoendlin, 1972). Each IHC is innervated by approximately 16 ANFs (Wu et al., 2019). The axons of the spiral ganglion neurons transmit their electrical signal towards the neurons of the cochlear nucleus (CN). The signal is then transmitted across the superior nucleus olivaris (SOC), the lateral lemniscus, the inferior colliculus (IC), and the medial geniculate nucleus (MGN), before reaching the auditory cortex, situated in the temporal lobe. Starting at the CN, signals are crossed to the contralateral side. On the level of the superior olive, information from both ears is compared for the first time. This is especially relevant for directional hearing, where tiny delays between both ears and differences in sound pressure level give indications on the direction of the source of sound (Brandes et al., 2019).

## 1.3 Distinction and aetiology of acquired hearing loss

Hearing loss can be categorized in different ways, mainly by its onset – congenital, pre- or postlingual, adult or presbycusis – or by the affected components of the auditory pathway. Here we can distinguish conductive hearing loss, which takes place in the outer ear or the ossicles of the middle ear, from sensorineural hearing loss, which occurs as a result of damaged components of the inner ear. Mixed hearing loss combines losses at levels of both conductive and sensorineural hearing, while central auditory dysfunction describes malfunctions at the level of auditory nerve, auditory brainstem or cerebral cortex. (Shearer et al., 1993).

Hereditary factors also play a large role in the occurrence of hearing loss. There are currently more than 150 genes known to be related to hereditary hearing loss. These make up over 80% of prelingual hearing loss and can be associated with combination of other disabilities as syndromes (Shearer et al., 1993). Whole Exome Sequencing with targeted analysis of 120 gene panels can help identify the responsible loci (Zazo Seco et al., 2017).

Unlike conductive hearing loss, where conservative and surgical treatments have long been used to improve or at least maintain a status quo of hearing levels, sensorineural and auditory pathway malfunctions remain difficult to treat. The reason being that diagnostics often lack the possibility of identifying the specific locus of the malfunction. Furthermore, there are no reliable ways to access or target the tiny anatomical areas of the inner ear, for drug application or surgical intervention.

While there are certainly studies attempting to restore hair-cell function by injection of drugs into the cochlea through the oval or round window, there are no approved treatments in Germany as of yet. For this study, the focus will lie on hearing loss that originates from the cochlea and the auditory pathway, with onset in adult life.

Noise-induced hearing loss (NIHL) is a significant, potentially avoidable cause for rising numbers of hearing loss across the globe. While occupational noise-exposure is controllable through occupational health and safety acts, recreational exposure through portable audio devices and attendance of noisy venues (concerts, festivals, clubs) frequently exposes listeners to harmful levels of sound for too long. The WHO estimates that this puts up to 50% of 12-35 year-olds at risk of hearing loss (World Health Organization, 2021). It is generally accepted, that when exposed to high noise for a limited time, the ear may experience temporary threshold shifts, which can be recovered by regeneration processes of the ear within hours to days and the initial loss of threshold sensitivity recovered back to the original level. Histologically, in these cases studies have made different observations of contributors to this temporary threshold shift. These include effects on the stereocilia bundles, observed as loss of tip-links (Zhao et al., 1996), or disarray of stereocilia (Mulroy and Whaley, 1984), as well as collapse of supporting cells of the organ of Corti (Flock et al., 1999, Nordmann et al., 2000) and temporary swelling of afferent dendrites of the IHCs. The swelling is most likely caused as a consequence of overstressed activation by noise exposure, when mitochondrial activity is exhausted (Spoendlin, 1971, Liberman and Mulroy, 1982, Robertson, 1983).

When repeated or prolonged exposure to such sounds eventually leads to structural damage of the cochlea, a permanent threshold loss is possible. Examinations of animal cochlea have shown that within minutes of the exposure, damage to hair-cells can be observed, followed by hair-cell death within several days (Wang et al., 2002). Meanwhile numbers of SGCs have also shown to deteriorate after noise-exposure, but this process happens much slower, even progressing for years (Spoendlin, 1971, Kujawa and Liberman, 2006).

Permanent threshold loss most commonly occurs at the higher frequencies, with a characteristic dip at 4 kHz (Passchier-Vermeer, 1974, Imam and Hannan, 2017). Molecular and biochemical mechanisms of noise-induced cell-death of the cochlea are still being researched. One accepted approach is the mechanical destruction of hair-cell membranes and supporting structures within the Organ of Corti by the sound wave (Spoendlin, 1971). Metabolic mechanisms in response to noise exposure include the increased production of reactive oxygen species (ROS). These oxygen radicals are required for numerous cell signalling and homeostasis processes, their excess production in response to noise exposure (Ohlemiller et al., 1999, Yamane et al., 1995a, Yamane et al., 1995b) can however induce apoptosis and necrosis of cells in the organ of Corti when lipids of cell membranes undergo peroxidation by ROS (Henderson et al., 2006). ROS generation is also linked to reduced blood flow in the cochlea. The resulting ischemia provides further stress to cochlear cells, in turn promoting more ROS generation (Ohinata et al., 2000).

Studies with application of glutamate or glutamate agonists have shown that another mechanism of auditory cell-death is that of glutamate excitotoxicity, especially at the SGC cells. Acoustic overstimulation releases toxic amounts of glutamate into the synapse, which activates large sodium and potassium influx at the post-synapse.

Eventually, the osmotic imbalance can cause the SGC neuron to swell and rupture (Janssen et al., 1991, Puel et al., 1994, Le Prell et al., 2001).

In contrast to NIHL, presbycusis, or age-related hearing loss (ARHL) is diagnosed when hearing deteriorates in aging patients, without any responsible otologic diseases, genetic factors or ototoxins. It is difficult to differentiate the impact of lifelong noise exposures in relation to normal age-related hearing declines in hearing capabilities, especially since noise-overexposure has also been suggested to accelerate effects of ARHL (Kujawa and Liberman, 2006). In standard Otorhinolaryngology (ear-nose-throat; ENT) diagnostics, the classic pattern of ARHL is an increase and threshold loss with higher hearing thresholds at higher frequencies. Meanwhile low frequency hearing is mostly maintained. As ARHL advances, the threshold loss at high frequencies begins to extend to lower frequencies as well. As Gordon-Salant evaluates from both the Framingham Heart Cohort study (Mościcki et al., 1985) and the Beaver Dam Hearing Loss Study (Cruickshanks et al., 1998), consistent declines in hearing thresholds can be observed across the whole study (Gordon-Salant, 2005). However, in unscreened cohorts, self-reported history of noise exposure, otologic disease and ototoxicity also played a role in developing hearing loss. ARHL generally includes a combination of conductive hearing loss from the middle ear, as well as sensorineural hearing loss. The decline in inner ear function has frequently been shown in humans and animals to be of metabolic character instead of hair-cell loss, with atrophy of the stria vascularis diminishing its capability of maintaining the electric potential of the cochlea through ion transport (Ramadan and Schuknecht, 1989, Schmiedt et al., 2002, Dubno et al., 2013). Wu et al. have contradicted these findings, stating that ARHL in humans only correlates with hair-cell loss (OHCs and IHCs)(Wu et al., 2020). According to their study, strial atrophy only occurs after OHC loss, when the effect of a dysfunctional cochlear amplifier greatly exceeds the hearing loss caused by reduced endolymphatic potentials. Hair-cell death generally begins at the basal, high-frequency end of the cochlea (Schuknecht, 1993).

## 1.4 Cochlear synaptopathy as cause of hidden hearing loss

Standard ENT-diagnostics mainly uses audiometric thresholds to evaluate hearing. While this method may be sufficient for many otologic diseases and hearing impairments, this is not the case for all patients. Gordon-Salant recognized that, while word recognition as a measure of comprehension problems generally correlated with threshold losses, this does not apply for all participants (Gordon-Salant, 2005). The so-called “hidden hearing-loss” (HHL) (Schaette and McAlpine, 2011) has undergone more focused research since Kujawa and Liberman showed that noise-induced damage could be more prevalent within the population than expected (Kujawa and Liberman, 2009). Instead, hidden hearing-loss is masked by physiological thresholds, while affected patients experience significant comprehension difficulties during their everyday lives.

This concept of cochlear synaptopathy without hair-cell loss has since then been shown both in animals (Kujawa and Liberman, 2009, Sergeyenko et al., 2013, Möhrle et al., 2016) and humans (Bharadwaj et al., 2014, Plack et al., 2014, Liberman and Kujawa, 2017). It is currently believed, that while HHL is associated with intact auditory thresholds, well synchronized ABRs and intact OHCs, the processing disorder originates instead from a subgroup of auditory fibres, called low spontaneous-rate (SR) fibres. ANFs can be divided into two subpopulations, those with a low SR and high threshold, and those with high SR and low thresholds (Liberman, 1978). In consequence, high-SR fibres are mostly required at threshold level stimuli and thus important for pure-tone audiometry (PTA) measurements, but are quickly saturated at louder stimulus levels, whereas low-SR fibres are only sensitive at louder stimulus levels, such as when background noise is present (Winter et al., 1990). Animal studies have shown that low-SR fibres are more vulnerable to noise (Furman et al., 2013), aging (Schmiedt et al., 1996) and ouabain-exposure, a cardiac glycoside used to study the effects $Na^+$-$K^+$-adenosinetriphosphatase (ATPase) inhibition on neurons (Bourien et al., 2014). While the exact reasons are still being discussed, this indicates that low SR fibres degenerate much earlier than high SR fibres, both due to noise exposure and aging (Liberman and Kujawa, 2017). As a result, PTA threshold sensitivity can stay normal for much longer, while speech-in-noise understanding deteriorates.

The loss of ANFs was shown to be part of the ARHL-process, even though PTA-thresholds are rarely affected (Wu et al., 2020), in fact, PTA-thresholds have been shown to only deteriorate when more than 80% of ANFs are destroyed (Schuknecht and Woellner, 1955, Lobarinas et al., 2013). The current understanding is, that different degrees of ANF-loss could therefore explain the varying performances in speech comprehension tests in noisy environments between people with the same PTA-thresholds (Plack et al., 2014). Counts of SGCs also do not accurately represent the extent of cochlear synapse loss, as spiral ganglion cell-bodies degenerate far later than their axons (Spoendlin and Schrott, 1989, White et al., 2000, Sergeyenko et al., 2013). Postmortem analysis of human temporal bones without hair-cell loss has shown that SGC degeneration rate seems to vary drastically between individuals, with environmental factors and noise exposure suggested as the reason (Makary et al., 2011), but overall age seems to play a big role in SGC counts (Otte et al., 1978).

## 1.5 Audiometric methods to evaluate hearing loss

Standard clinical practice for hearing diagnostics, but also studies claiming to test with "normal audiograms", do not necessarily guarantee that no threshold elevation is present. In most clinical settings, thresholds are measured only at defined frequency steps, usually between 0.25 and 6-8 kHz. Furthermore, the definition of "normal hearing" includes audiograms up to threshold losses of 25 dB (World Health Organization, 1991), or with the newer classification suggested by the WHO, which includes hearing losses of up to 20 dB as "normal" (Stevens et al., 2013). It should also be noted that different studies focus on different audiogram frequencies when matching groups by threshold sensitivity. Finally, audiometric thresholds can only provide some differentiation of the origins of the hearing loss. While conductive hearing loss and some other specific illnesses may be recognized from the combined diagnostics with bone-conduction or the shape of the PTA-curve, it does not provide information if the loss originates from the OHCs, the IHCs, the ANF, auditory synapses or anywhere else along the auditory pathway.

To further distinguish between types of hearing loss, different audiometric measurements can be used additionally. Tympanometry is used as a fast and sensitive method to diagnose middle-ear pathologies by measuring the impedance of the eardrum (Terkildsein and Thomsen, 1959)(see 2.3.2), Pathologic results are usually complemented by PTA-thresholds, in the form of conductive hearing loss. Changes in pressure in the middle ear, or any pathologies altering the function of the eardrum or the ossicles would be recognized as discrepancy between the bone and air-conduction. Measuring PTA-thresholds is the standard clinical practice for any hearing-related question. Despite not being an objective measurement by nature, experienced medical technicians can provide accurate measurements. Furthermore, most clinical diagnostic and therapeutic decisions are based on threshold losses in PTA.

To test for speech comprehension, the German "Oldenburger Satztest" (OLSA) (Wagener et al., 1999b, Wagener et al., 1999c, Wagener et al., 1999a) was chosen for quantification of comprehension in quiet and noise conditions. The use of entire sentences compared to single words like in the "Freiburger Wörtertest" can be advantageous in its representation of day-to-day hearing situations (Wardenga et al., 2015). The Speech Reception Threshold of 50% comprehension ($SRT_{50}$ ) was measured in quiet, ipsilateral and contralateral noise, each with an unfiltered "broadband" speech condition (OLSA-BB), a low pass filtered speech condition (OLSA-LP) and a high-pass filtered condition (OLSA-HP), where frequencies above 1.5 kHz (OLSA-LP) or below (OLSA-HP) are removed from the frequency spectrum of the stimulus (Schirmer et al., 2024). The filtered stimulus speech aims to separate temporal fine-structure (TFS) cues (Lorenzi et al., 2006) from temporal envelope (TENV) cues (Shaheen et al., 2015, Parthasarathy et al., 2019) to evaluate their individual impacts on speech comprehension.
Otoacoustic emissions (OAEs) are used in clinical and scientific settings as an objective, non-invasive method to selectively measure OHC function. Intact cochlear amplifier function of the OHCs is required for normal hearing thresholds (Dallos, 2008). OAEs are created by an active source in the ear as a response to external acoustic stimulation and can be measured as mechanical sound energy in the outer ear canal. Upon discovery in 1978, OAEs were attributed to be a by-product of the active cochlear amplifier, located at the OHCs (Kemp, 1978). Otoacoustic emissions can be grouped into spontaneous OAEs (SOAE) and evoked OAEs. SOAEs can be recorded in the outer ear canal without external stimulation. These narrow-band signals are stable in frequency in repeated measurements over months. However, their amplitude can vary substantially between each measurement (Probst et al., 1991). Recent studies suggest SOAEs to be prevalent in 30-40% of ears, with decreased prevalence in participants with increased hearing loss. (Kuroda, 2007, Jedrzejczak et al., 2016).

Distortion product otoacoustic emissions (DPOAEs) are detectable in humans up to a hearing loss of 50 dB and present additional diagnostic options for higher frequency ranges when comparing to transient evoked otoacoustic emissions (TEOAEs) (Hoth and Neumann, 2006). The DPOAE signal is created by presenting two primary tones with frequencies $f_1$ and $f_2$ simultaneously. When stimulus parameters are presented with $f_2/f_1=1.2$ and $L_1 \geq L_2$, the travelling waves overlap on the cochlea near the f2 frequency due to the nonlinear output mechanics of the cochlea(Zelle et al., 2017a). Upon deflection of OHC stereocilia, the mechano-electrical transducer propagates the receptor voltages into cochlear fluids. While its purpose is to modulate and amplify the signal for the IHCs, some of the vibration travels retrogradely through the tympanum and is then detectible in the outer ear canal as the OAE (Avan et al., 2013). For humans, the largest DPOAE response can be recorded at the cubic difference frequency $f_{DP} = 2f_1\text{-}f_2$ (Goldstein, 1967). The measurement of this DPOAE response can then be used to analyse the function of the cochlear amplifier.

Auditory steady-state responses (ASSRs) are recorded in order to obtain a metric for the early stages of sound transmission around the stimulus onset which is not reflected in pure-tone thresholds (PTTs). Repeated stimuli evoke a phase locked neural activity that is mainly coded by TENV (Rance, 2008)(see 2.4.4 for details).

ABRs originate from the summed neuronal activity of the ascending auditory pathway (Buran et al., 2010) as a response to short auditory stimulations and are used to assess the functional integrity of the auditory pathway in the brainstem. In the clinical setting, ABRs can be used in newborn hearing screening for auditory pathway deficits, to detect retrocochlear pathologies such as acoustic neuromas or for intraoperative and intensive care monitoring (Habib and Habib, 2021). These potentials of short latency can be recorded within 10 milliseconds (ms) after stimulus onset. When filtering and averaging the recorded ABRs over many repeated stimuli clicks, distinct waves I-VI can be identified in humans (Møller et al., 1994). Wave I represents potentials in the distal parts of the auditory nerve (AN) (Portmann et al., 1980), wave II is thought to be generated by globular cells of the cochlear nucleus (CN) (Melcher and Kiang, 1996), wave III originates from the superior olivary complex (SOC) and the lateral lemniscus (Melcher and Kiang, 1996). Wave V and VI originate from the inferior colliculus (IC) (Møller et al., 1994) and the medial geniculate body (MGB) (Hashimoto et al., 1981, Melcher and Kiang, 1996).

In contrast to the utilized click stimuli, which usually rely on either TENV or TFS coding separately depending on the stimulus frequencies, language comprehension can be broken down into the correct discrimination of vowels (Won et al., 2016) and consonants (Hornickel et al., 2009). These require both TFS coding and TENV coding to correctly understand the phonemes above and below the PLL (phase-locking limit) (Verschooten et al., 2019). By creating a phoneme discrimination task, we aimed to link different comprehension performances to varying TFS or TENV coding (Huet et al., 2018). The measurement used different decomposed narrowband signals representing phoneme pairs with formant frequencies either above or below the expected PLL of 1.5 kHz in humans (Schirmer et al., 2024).

### 1.6 Aim of the study

The aim of this study is to identify factors that contribute to decreased speech comprehension, apart from the commonly known factors of age and threshold shifts. Establishing a cohort of participants of different ages and hearing levels for measurements of the standard hearing diagnostics, as well as more experimental measurements could help to assess the impact of these other factors that can reduce speech comprehension. This may in the future help to improve hearing diagnostics of developing speech comprehension impairments.

## 2 Methods

### 2.1 Ethics application

The ethics application with project number 392/2021BO2 was approved by the Ethics Committee of the University of Tübingen in accordance with the Declaration of Helsinki for human experiments. All study participants were informed before examinations about risks, privacy protection and study processes. Written consent and informed consent of every participant was required. All digital data was anonymized, consent forms and other non-anonymized data was stored in a locked cabinet in the lab. The participants received compensation for their participation effort. Methods and inclusion/exclusion criteria were predefined as listed below.

### 2.2 Recruitment of the cohort

Participants were recruited locally, by contacting target groups through mailing lists of the university and the university hospital in Tübingen, Germany (Appendix A) and a poster at the shared entrance of the ENT and eye clinic Tübingen. Participants were awarded 100€ as compensation for the participation in all four sessions of the study (total measurement time of 8 hours). Participants were recruited between 18 (lowest age permitted by ethics application) and 76 years so that age dependent changes in hearing can be discerned. Study plans aimed for a cohort size of 90 participants, with 30 participants per age group to allow adequate statistical analysis of results. Volunteers were recruited with subjective comprehension deficits, as well as participants with subjective good hearing.

### 2.2.1 Inclusion and exclusion criteria

Exclusion criteria, defined in the ethics application (392/2021BO2) were verified for all participants at different points of the study. During recruitment, potential participants were asked about tinnitus, previous hearing-related treatments (e.g. ear-surgery), other medical history and active medication through phone or email. As a further criterion, participants were verified to be of German mother tongue, due to the usage of a German matrix comprehension test. In cases where the participant had moved to Germany after birth, we required them to have joined the German education system in the first years of their school career.

A list of elimination-criteria on the Q1 questionnaire at the beginning of session 1 was used to cover all exclusion criteria (Appendix B). Relevant hearing-related diseases as well as other comorbidities and medications (Appendix C) were recorded with the Q1 questionnaire as exclusion criteria.

Recruitment was targeted at persons with subjectively normal hearing, speech comprehension problems or known mild hearing impairment that had not yet led to treatment with hearing aids. The allocation of good and impaired hearing was done after pure-tone audiometry and OLSA comprehension, according to criteria described in results.

In total, 112 participants between 18 and 76 years of age were initially recruited. After subsequent exclusion of a number of participants due to lack of compliance (failure to attend measuring sessions several times), hearing-related conditions (e.g. tinnitus, ear surgery) or systemic diseases known to affect hearing, 89 participants remained in the study, divided into 3 age-groups: young (18-29 years, n=29), middle-aged (30-55 years, n=32) and older (56-76 years, n=28) (Appendix C),.

## 2.3 Standard audiometric diagnostics for grouping of the cohort

### 2.3.1 Otoscopy of the ear canal

To begin with, each participant underwent a standard otoscopic ear examination by an ENT-specialist of the Department of Otolaryngology, Head and Neck Surgery Tübingen. Using a Zeiss ear-microscope, the outer and inner ear canal and the tympanic membrane were examined for intactness. Cerumen was removed when necessary to allow sufficient vision of the tympanic membrane and to prevent blockage of the earphone tubes by cerumen in the following measurements. Confirmation of intactness of the tympanic membrane was signed by the ENT-specialist before the next measurements were undertaken.

### 2.3.2 Tympanometry for exclusion of ear ventilation disorders

Evaluation of the condition of the tympanic membrane and the ossicles was conducted using the AT235 (Interacoustics, Middelfart, Denmark) tympanometry system. The probe tip, inserted into the outer ear canal, adjusts the pressure in the ear canal in relation to the middle ear and measures the reflection of a 226Hz pure-tone on the ear-drum at different pressures (Shanks, 1984). In a healthy ear, the best sound transmission (and least reflection) occurs when ambient air pressure and middle ear pressure are equal, this is named a type A tympanogram. Abnormal tympanograms may reveal fluid in the middle ear, perforations of the tympanic membrane or patent pressure equalization tube. In these cases, the peak in the tympanogram either disappears (type B), or shifts to either side, where the pressure on both sides of the tympanic membrane is not equal (Shanks and Shohet, 2009). Given that participants should have intact tympanic membranes and normal pressures in the middle ear, participants can be expected to show Type A tympanograms, otherwise rejection from the study must be considered. Since problematic ear canals can also lead to abnormal tympanograms when the probes do not fit tightly, an ENT-specialist was consulted, taking into account the audiogram and the otoscopy before making the final decision.

### 2.3.3 Anamnesis for subjective hearing and exclusion criteria

The custom Q1 questionnaire focuses on a subjective self-evaluation of individual hearing capabilities in different conversation situations, as well as some general anamnestic information such as education and handedness (Appendix B). A checklist to inquire for any of the comorbidities mentioned in the ethics application as reasons for elimination from the study is included at the beginning of the questionnaire. This includes other hearing-related illnesses including tinnitus or previous ear surgery, as well as systemic diseases that could affect hearing or processing in the brain. (Appendix B) (Vande Maele et al., 2021).

2.3.4 Questionnaires screening for depression and cognition for exclusion of depression- and dementia-associated hearing loss

To eliminate hearing loss associated with depression or dementia (Livingston et al., 2017), participants were screened for depression and dementia in the first measuring session. The Beck's Depressionsinventar II (BDI-II) (Appendix D) and the Geriatric depression scale (GDS) (Appendix E) were sent to participants on the day before their first measurement. The BDI-II has shown to effectively screen for depression with sufficient sensitivity in a clinical setting (Beck et al., 1996). It consists of 21 items, with each item awarding 0-3 points. Mild depression can be expected at scores starting at 14 points out of 63, medium depression at 20 and heavy depression above 29 points. Additionally, we used the Geriatric depression scale (GDS) (Yesavage et al., 1982). Each of the 15 items can award one point, at scores of 5-10 points, mild to moderate depression should be suspected, severe depression can be expected starting at 11 points. The GDS is deemed more suitable for older patients, because it does not include somatic criteria such as sexual interest, sleep, appetite, energy level but focuses on affective and cognitive domains like sadness, apathy, crying and hopelessness. Nonetheless, both questionnaires offer good internal consistency and high correlation with each other (Jefferson et al., 2001).

The Mini Mental Status Examination (MMSE, or MMST in German) was used as a screening test for cognitive impairment and developing dementia (Appendix F). 11 items test for memory, orientation in time, reading, writing, ability to communicate (Folstein et al., 1975). Patients with less than 27 of the possible 30 points should undergo further professional diagnostics as this could be an indication for beginning cognitive decline and were consequently excluded from the study and offered to contact a professional psychiatrist for advanced psychiatric diagnostics.

### 2.3.5 Pure-tone audiometry up to high frequencies to assess conductive and sensorineural hearing loss

As gold standard audiometric measurement, pure-tone audiometry was conducted for all participants. The thresholds were measured at standard the frequencies of 0.25, 0.5, 1, 1.5, 2, 4, 6, 8, 10 kHz, as well as additional 4 frequencies in the extended frequency region: 11.2, 12.5, 14, 16 kHz. The measurement was done by experienced Audiometrists of the ENT-clinic Tübingen using the AT 1000 Audiometer (Auritec, medizindiagnostische Geräte Gmbh, Hamburg, Germany). As headphones, Beyerdynamic AT1350A (Beyerdynamic, Heilbronn, Germany) were used for frequencies up to 10 kHz and the Sennheiser HDA300 (Sennheiser, Wedemark-Wennebostel, Germany) for the extended frequency region. The participants were seated in a sound-proof chamber (Industrial Acoustics Company GmBH, Niederkrüchten, Germany). Participants activated a button as soon as the detection threshold was reached. To confirm the absence of conductive hearing loss, bone conduction was also measured using the B71Wbone transducer (Radioear, Middelfart, Denmark) at frequencies 0.25, 0.5, 1, 1.5, 2, 4, 6 kHz. In both measurements, the tone volume was increased continuously until the button was activated. The uncomfortable loudness level (ULL) was measured at 0.25, 0.5, 1, 2, 4, 6 kHz, participants were asked to activate the button when they experienced the sound to be uncomfortable.

Pure-tone averages were calculated from the right ear for low frequencies (PTA-LF) (0.125, 0.25, 0,5 and 1 kHz), high frequencies (PTA-HF) (6, 8, and 10 kHz), extended high frequencies (PTA-EHF) (11.2, 12.5, 14, and 16 kHz) and (PTA4) (0.5, 1, 2, and 4 kHz). These specific PTA groups were chosen to simplify and analyse thresholds for specific frequency ranges: (i) PTA-LF, for hearing below the phase-locking limit (PLL) i. e. < 1500 Hz in humans, this is also the limit where we set the frequency filter on the syllable task (see 2.4.2) (ii) PTA4, because this is the worldwide standard for defining hearing loss severity, (iii) PTA-HF, because these frequencies are regularly measured in the clinical setting, but are partly above the frequency range of the PTA4, as well as above the PLL, and (iv) PTA-EHF, because those frequencies display a disputed role for speech perception (Hunter et al., 2020) and are above the standard frequency ranges measured in a clinical setting.

### 2.3.6 Speech reception thresholds with Oldenburger Satztest with high- and low-pass filter

The Oldenburger Satztest (OLSA) (Wagener et al., 1999d), a German speech comprehension matrix test, was carried out with additional frequency filters and noise conditions for the study. The participant was seated in a sound-proof audiometry chamber (Industrial Acoustics company GmBH, Niederkrüchten, Germany). Stimuli were generated by a laptop with a sound card (Scarlett 8i6 3rd Gen, Focusrite, United Kingdom) using ER2 transducers and ER1-14A earpieces (Etymotic Research, Elk Grove Village, Illinois, USA). Transducers were shielded with µ-metal and calibrated to 70 dB (RMS). The entire setup was grounded to the measuring chamber. Due to the time-limitations of the measurement sessions, only the right ear was presented with the sentence stimuli. The standard matrix word list of the clinical OLSA was used (Fig. 2 D). It consists of 5 categories of 10 words to make a total of $10^5$ possible, random 5-word sentences. After presentation of each sentence, the participant repeats understood words and if possible, the entire sentence. While the grammatical structure of the 5-word sentences remains the same - name, verb, number, adjective, noun - the random choice of words from the list guarantees that the rest of the sentence cannot be anticipated by the participant. Furthermore, there should be no additional learning effect after an initial training phase (Brand and Kollmeier, 2002). The adaptive volume of the presented sentences is automatically adjusted, depending on the number of correct words of the previous sentence, aiming to converge to a specific volume threshold at which 50% comprehension is achieved (Fig. 2 B). The participant speaks into a microphone, while the investigator selects all reproduced words from the matrix list. In order to reduce experimenter effects and increase objectivity, the investigator is blinded to the presented sentence. The speech recognition threshold ($SRT_{50}$) is calculated from sentence blocks of 20. The SRT represents the volume threshold, at which comprehension of 50% of the sentences is achieved. The adaptive volume typically converges towards a fixed volume, the $SRT_{50}$ (Fig. 2 B), within several sentences and should be reached in all cases within 20 sentences, thus each condition was presented for a fixed number of 20 sentences (same as in a clinical setting).

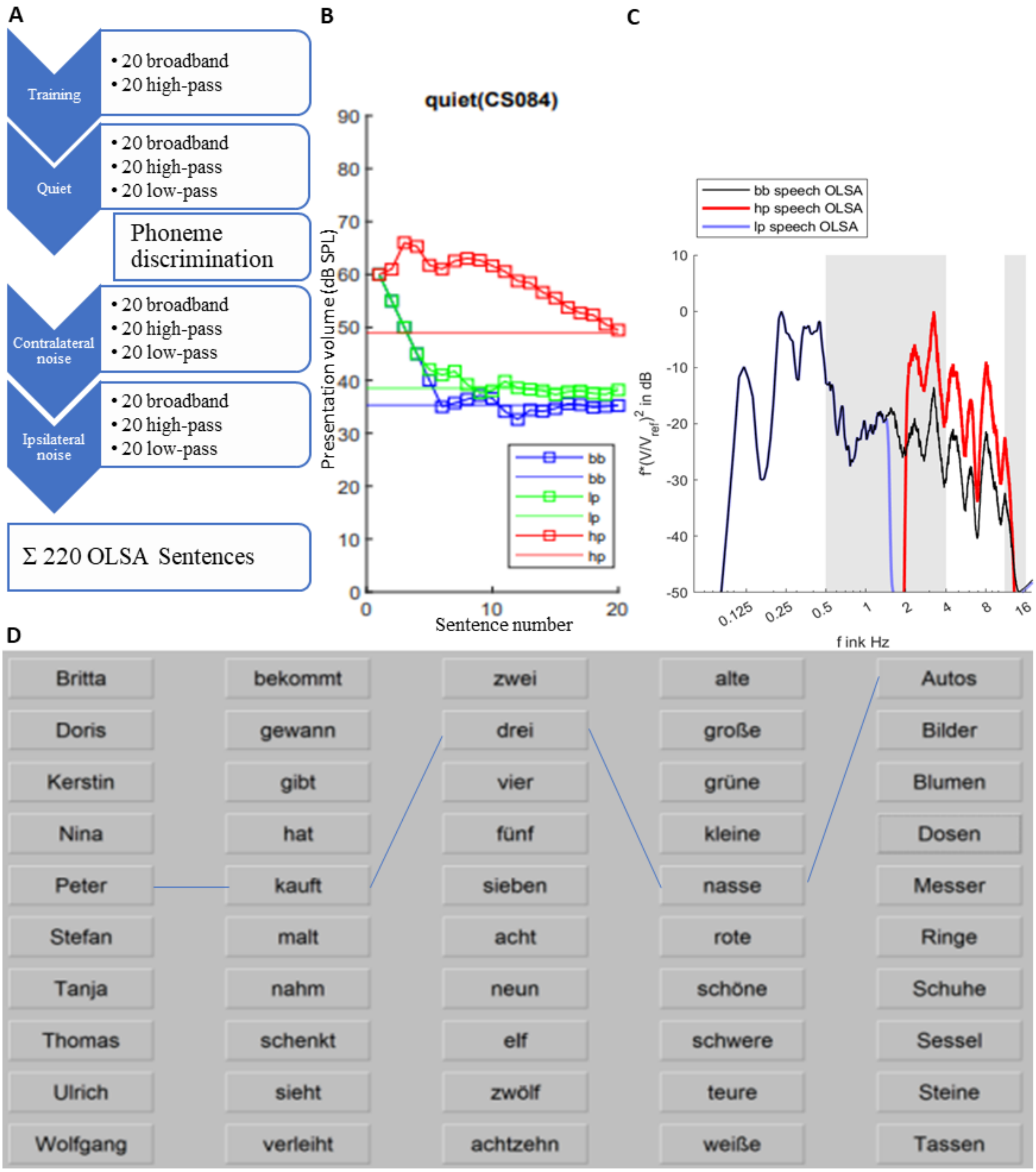



**Fig. 2** **A** OLSA conditions overview. Training block is followed by quiet condition sentences. After the psychoacoustic task, OLSA blocks contralateral and ipsilateral noise follow. **B** Exemplary OLSA $SRT_{50}$ of an individual participant for the quiet condition. The adaptive volume of the presented sentences approaches the estimated threshold level ($SRT_{50}$), represented by the horizontal lines in equivalent colours for broadband (blue), low-pass (green) and high-pass (red) filtered stimuli. **C** Frequency spectra of broadband (black), low-pass (blue) and high-pass (red) stimulus conditions of the OLSA. **D** List of words used for sentences, one word from each column makes up a sentence in following order from left to right: name, verb, number, adjective, noun. Blue lines represent an example sentence: “Peter kauft drei nasse Autos” (“Peter buys three wet cars”).

The test was divided into 4 condition blocks, namely the training block, quiet, contralateral noise and the ipsilateral noise blocks. The psychoacoustic phoneme-discrimination task was moved between OLSA blocks 2 and 3 to keep up the participant’s attention in this rather exhausting task (Fig. 2 A). Each block 2-4 consists of 3x 20 sentences: a broadband speech stimulus, a high-pass (low cut-off) filtered and a low-pass (high cut-off) filtered part, with the filter division set at 1500 Hz (Fig. 2 C). The filter was picked in accordance to the syllables presented in 2.4.2 and is placed at the estimated human phase-locking-limit at approximately 1500 Hz (Brughera et al., 2013) allowing the differentiation of temporal processing through both temporal fine structure processing and temporal envelope coding (Oxenham, 2018, Verschooten et al., 2019, Dapper et al., 2025)

To eliminate learning effects that typically occur during the first few sentences (getting to know the stimuli words, structure of the task etc.) (Hörzentrum Oldenburg, 2019), participants were presented with 20 sentences of both the broadband and high-pass filtered speech in quiet to get used to the measurement.

2.3.7 Pure-tone normalized threshold (PNOT)

To evaluate the impact of speech comprehension factors other than age and PTT, we implemented the pure-tone normalized threshold, from here on referred to as PNOT (Schirmer et al., 2024). Three independent groupings were calculated, one for the quiet (n=89), ipsilateral noise (n=63) and contralateral noise (n=63) conditions each. OLSA thresholds were normalized for all 15 available frequencies between 0.125 and 16 kHz with a multivariate regression between the three OLSA thresholds (broadband, low-pass, high-pass) and the first five principal components of the pure-tone thresholds (Schirmer et al., 2024). A principal component analysis was performed using a singular-value decomposition algorithm (MatLab Version 2021b), which captured approximately 93% of the variations in threshold when using the first five components. OLSA predictions were then evaluated for each participant individually based on the linear regression model. The three OLSA threshold predictions were then subtracted from measured OLSA SRTs and averaged. The so-called “PT-normalized OLSA threshold” (PNOT) in quiet, ipsilateral noise and contralateral noise were each used to divide the cohort into 3 groups of equal size, aiming to maximize the spread of speech performance across them as “good”, “standard” and “poor” performance (Schirmer et al., 2024).

## 2.4 Audiometric measurements to analyse cochlear amplifier efficiency and synchronicity changes

### 2.4.1 Distortion product otoacoustic emissions to measure cochlear amplification

Distortion product otoacoustic emissions (DPOAE) allow the objective evaluation of the transduction mechanism of the cochlea (Kemp, 1978). The DPOAE were measured on the setup described by Vetesnik et al. and Zelle et al. (Vetešník et al., 2009, Zelle et al., 2015, Zelle et al., 2017b).

Participants were seated in a double-walled soundproof chamber (Industrial Acoustics Company, Niederkrüchten, Germany) and stayed awake during the 50-minute measurement. Input/Output-function were performed on one ear at a time. A DPOAE probe system ER-10C (Etymotic Research, Elk Grove Village, Illinois, USA) was used as earphone and microphone. Flexible tubes and Etymotic foam ear tips ER10C-14A connected the probe system with the ear-canal. The compressible foam allows for sufficient sealing of the ear canal when positioned at around 1.5-2 cm from the eardrum. Calibration of probe sound sources was carried out before measurement of the right ear as well as when switching to the left ear by correcting for probe-to-tympanic membrane transfer function, based on an artificial ear simulator (B&K type 4157, Bruel & Kjær, Nærum, Denmark). A custom-designed software (NI LabViewTM Full Development System, Version 8.01; LabViewTM, National Instruments, Austin, Texas, USA) was used for stimulus generation and data acquisition. Signal post-processing and data analysis were done in MATLAB (Version 2021b, MathWorks, Natick, MA).

Utilizing short-pulse DPOAE (pDPOAE) for the second primary tone (**f2**), the artefactual interference of primary and secondary DPOAE source can be reduced, allowing relatively easy acquisition of the primary-source DPOAE (as described by Zelle et al., 2017). Furthermore, onset decomposition was applied, a technique to capture the primary source component at a single pre-defined time instant, before the secondary source component begins to interfere (Vetešník et al., 2009). 8 frequencies ($f_2$=0.8, 1.2, 1.5, 2, 3, 4, 6, 8 kHz) were measured for input-output functions (I/O), with $f_1$ frequency consequently chosen to satisfy the constant frequency ratio of $f_2/f_1$=1.2. An adaptive algorithm aimed to measure at least four pDPOAE values for a frequency. From these, estimated distortion-product thresholds ($L_{EDPT}$) were calculated based on a linear regression of the semi-logarithmic I/O functions (Zelle et al., 2017b, Boege and Janssen, 2002). The $L_{EDPT}$ (level of the estimated distortion product threshold) derived from said I/O-function pDPOAEs have been shown in previous studies to correlate in an almost 1:1 relationship with auditory thresholds with a consequently exceptionally small standard deviation of the residual of approximately 6 dB (Zelle et al., 2017b, Zelle et al., 2020) and to be highly reproducible (Bader et al., 2021). The previously mentioned 8 frequencies were presented as two sets of four two-tone pulse pairs. The first set consisted of $f_2$=0.8, 1.5, 3, and 6 kHz in a block of 180 ms length, while the second set had $f_2$=1.2, 2, 4, and 8 kHz in a block of 120 ms length. The sets were averaged over 100 ensembles, where one ensemble consists of 4 blocks with primary-tone phase variation technique, where phase-shifts deliver a time-domain DPOAE response, while contribution of the two stimulus tones vanishes by averaging (Whitehead et al., 1996). The adaptive procedure started with $L_2$=45 dB SPL (sound pressure level) for each $f_2$-pulse, followed by lower or higher stimulus levels depending on SNR (signal-to-noise ratio), a minimum $L_2$ step size of 3 dB, and estimations based on population data of I/O slopes (Krokenberger, 2019). As acceptance criteria for the subsequent DPOAE threshold estimation, each single DPOAE value was required to have SNR ≥ 10 dB. $L_1$ values are chosen according to a frequency-specific scissor paradigm (Zelle et al., 2017b) aimed to equalize both travelling-wave amplitudes of the stimulus tones $f_1$ and $f_2$ at the $f_2$-place in the cochlea (Zelle et al., 2020). The $f_2$-pulses were followed by presentation of the corresponding $f_1$-pulse 2.5ms later, each pulse with a width of 40 to 10 ms, adjusted so that the intracochlear travelling wave may reach a steady state before the $f_2$-pulse is presented. The $f_2$-pulse half-maximum width was decreased between 11.9 and 3.0 ms for frequencies from 0.8 to 8 kHz to permit settling of the DPOAE nonlinear-distortion component before the second coherent-reflection DPOAE component builds, in order to allow visual distinction of both component in the time domain (Zelle et al., 2013).

To avoid large extrapolations errors for the estimated thresholds, acceptance criteria were applied to the linear regressions. The squared correlation coefficient was limited to $r^2_{I/O} \geq 0.8$, and the standard deviation of the $L_{EDPT}$, $\sigma_{EDPT} \leq 10$ dB (for further details, see Zelle et al., 2017). Based on these criteria, we calculated an acceptance rate in Results (3.2.5). This is the ratio between accepted I/O functions within groups of all members at all eight frequencies and both ears and the total number of measurements. Acceptance rate depends on DPOAE amplitudes, background noise, and the linearity of the measured I/O functions, to mention some important contributors. Especially participants that struggled to keep still during the measurement often created a lot of background noise, reducing the quality of the measurement.

2.4.2 Psychoacoustic phoneme-discrimination

The phoneme discrimination task was established using analysis-re-synthesis as implemented in the WORLD vocoder (Morise et al., 2016), to generate the stimuli syllables from recordings of a male speaker. The two pairs of steady-state vowels were set to an average fundamental frequency ($F_0$) of 116 Hz, matching the average F0 of the OLSA speaker. From the vowels /o/, /u/, /i/ and /y/ (Table 1), a total of 8 phonemes were used, two pairs of steady state vowels /ie/-/y/ and /o/-/u/ and two pairs of consonant-vowel syllables /di/-/bi/ and /du/-/bu/.

The vowel pairs /o/ (like in *oder*, "or" in German) and /u/ (like in *Du*, "you" in German), that differed in their first formant (F1, see Table 1) and are located well below the supposed PLL in humans at 1.5 kHz, were synthesized with a 30 ms raised cosine ramp at the onset and offset, and had a total duration of approximately 414 ms (corresponding to 48 $F_0$ cycles). Similarly, the differences in the /du/-/bu/ syllable pair only appeared at frequencies below the PLL, and only within the first 100 ms. No differences between the syllables were recognizable for the following 371 ms. We also made sure that the vowel segment of this syllable pair, /u/, was identical to the isolated steady-state /u/ used in the /o/-/u/ vowel pair, except that the duration was trimmed to an overall syllable duration of 471 ms. The spectral differences of vowel pairs /i/ (like in *sie*, "she" in German) and /y/ (like in *üben*, "practice" in German) only appeared in their second and third formants (F2 and F3, Table 1), which were above the human PLL of 1.5kHz. We can therefore expect that the encoding of the /i/-/y/ contrast could not rely on TFS, but on TENV coding instead. In accordance with /du/-/bu/, the /di/-/bi/ pair was built to only differ in frequencies above the PLL, with differences in spectral power only appearing in the first 100 ms. Again, the /i/ from these syllables was identical in spectral shape to the /i/ used in the vowel pair /i/-/y/. Finally, all stimuli were spectrally tilted to ensure similar signal-to-noise ratios above and below the PLL when presented in the speech-shaped noise used in the OLSA task (Schirmer et al., 2024).

**Table: 1 Phoneme formant frequencies.** Frequencies of the first four formants (F1-F4) of each individual vowel sound, in Hz. *Table adapted from Schirmer et al., 2024*

| | **/o/** | **/u/** | **/i/** | **/y/** |
|---|---|---|---|---|
| F1 | 380 | 200 | 350 | 350 |
| F2 | 750 | 750 | 1950 | 1780 |
| F3 | 3500 | 3500 | 2800 | 2100 |
| F4 | 10000 | 10000 | 4000 | 3000 |

During piloting stages, a 9-step continuum was generated for each of the four stimulus pairs, where formant frequencies were gradually modified on a log-scale. For each stimulus pair, a large and small contrast (indicated as "easy" and "difficult" respectively) was selected, aiming to best reduce floor and ceiling effects. Furthermore, the average stimulus levels were adjusted to 60 dB SPL. Stimuli were presented to the participant using the same Setup as for the OLSA task, using ER2 earphones, the noise was the same speech-shaped noise used during the OLSA measurement and presented at 0 dB SNR.

Each syllable pair was tested at both difficulty levels and for the three different noise masking conditions (quiet, ipsilateral and contralateral noise). The resulting 6 conditions per syllable pair are each repeated 9 times, producing 54 total trials per syllable pair. The /ie/-/y/ and /o/-/u/ pairs were tested during Session 2 of the study, /di/-/bi/ and /du/-/bu/ took place as part of session 3. The syllables were measured using a three alternative forced choice paradigm, where two options were identical and one option presented the modified frequency contrast. The participant was asked to select the syllable that was identified as different from the other two syllables.

To account for learning effects between both syllable pairs it is randomly decided whether the task begins with \i\-\y\ or \o\-\u\ in session 2 and /di/-/bi/ or /du/-/bu/ in session 3. Before each subsection of the test, a short training task is presented with feedback, so that the participant can get familiar with the task. These results were not used for analysis. During the main test, there is no feedback on correctness of answers.

Recorded results include the number of correct answers in each noise condition and difficulty. Response time was also recorded, but finally not used for analysis, because participants were not motivated to answer quickly. Therefore, response time could be able to indicate perceived difficulty of the task but showed large inter-individual differences.

### 2.4.3 Evoked responses of the brainstem

Brainstem auditory evoked potentials originate from the summed neuronal activity of the ascending auditory pathway (Buran et al., 2010) as a response to short auditory stimulations. These potentials of short latency can be recorded within 10 ms after stimulus onset. When filtering and averaging the recorded ABRs over many repeated stimuli clicks, distinct waves I-VI can be identified in humans (Møller and Jannetta, 1982). Wave I represents potentials in the distal parts of the auditory nerve (AN) (Portmann et al., 1980), wave II is thought to be generated by globular cells of the cochlear nucleus (CN) (Melcher and Kiang, 1996), wave III originates from the superior olivary complex (SOC) and the lateral lemniscus (Melcher and Kiang, 1996). Wave V and VI originate from the inferior colliculus (IC) (Møller et al., 1994) and the medial geniculate body (MGB) (Melcher and Kiang, 1996, Hashimoto et al., 1981).

The soundcard was used to generate both the trigger signal and auditory stimulation. In order to eliminate electrical interference by the trigger signal, the triggers were offset by -5ms. Recording of ABRs was carried out using the actiCHamp Plus64 (Brainproducts GmbH, Gilching, Germany). The sample rate was set to 48 kHz and fifty times preamplification (EP50Amp) was used. This two-channel system was used with 3 Neuroline 720 electrodes (Ambu, Bad Nauheim, Germany), The active electrode was placed behind the right ear on the mastoid, one as ground electrode between the eyebrows and a reference electrode on the middle sagittal line near the hairline. Before electrode placement, the skin was cleaned, superficial skin rubbed off with fine sandpaper and defatted with disinfectant spray to reach a low impedance of < 2 kOhm on all electrodes. If the impedance was too high after electrode placement, the above-described procedure was repeated until satisfactory impedance levels were obtained. The participants were measured lying on an examination table and were asked to stay as relaxed as possible to reduce muscular artefacts. Electrodes were plugged into a pre-amplifier. The supra-threshold ABR measurement consisted of two blocks of 3000 epochs (1500 positive peaks and 1500 negative peaks) of acoustic supra-threshold broadband click stimuli at 70dB and 80 dB SPL. The duration of a single click was set to 83 ms.

During the measurement, real-time averaging of all clicks was used for monitoring. If the measurement looked noisy or did not show typical appearance of ABR curves, the measurement was repeated, provided that this was possible within the permitted session time of two hours. In these cases, waves were examined after filtering, and the better measurement was chosen according to noise and especially distinctness of waves I, III and V. After the measurement, ABR signals were band-pass filtered between 30-2000 Hz (First order FIR filter, Hamming windowed) and averaged for the 3000 clicks of the same stimulus volume.

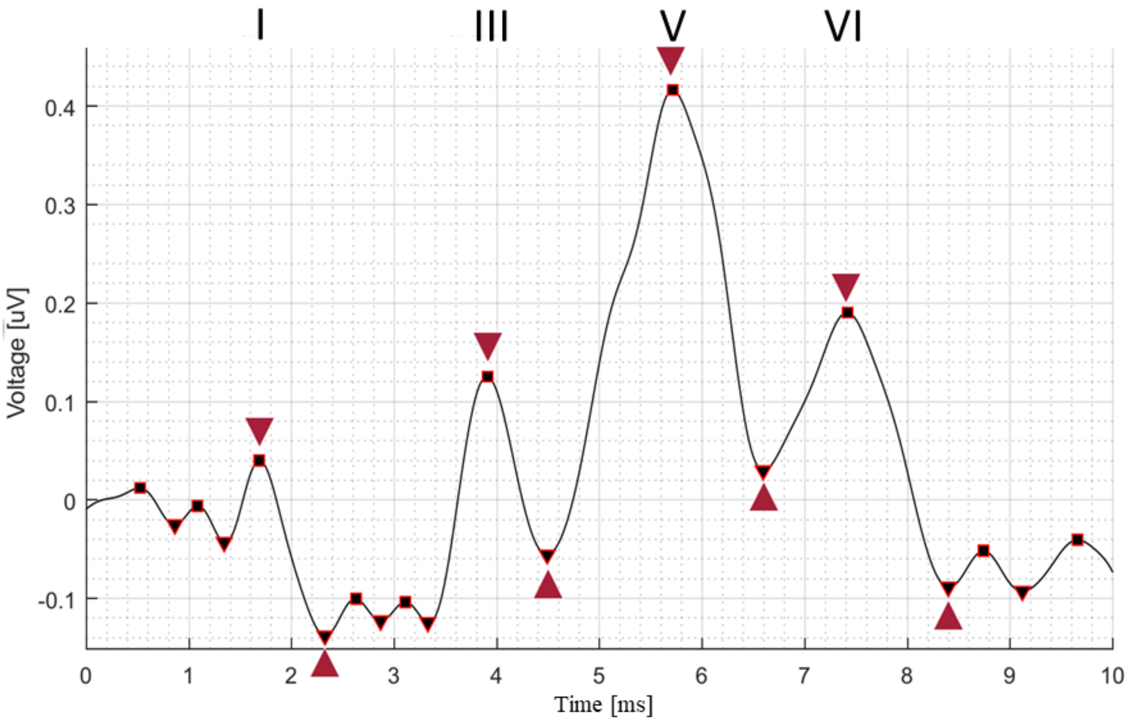



**Fig. 3** Exemplary filtered and averaged human ABR wave. Horizontal axis shows time in ms after click onset, the vertical axis presents the volatage measured in µV. All positive (black squares) and negative (black triangles) deflections are marked. Once the reflevant deflections are selected, latency is calculated as the time in ms after the click onset at the positive deflection peak. Amplitude is calculated as the difference in volatage between positive and following negative deflection of the wave. Waves I-VI are marked with dark-red traingles, wave II and IV are not clearly distinguishable in this example. Wave IV was rarely clearly seen, in most cases the inflection (black arrow) only indicated its presence.

For analysis, the ABR of a single ear and stimulus volume was averaged. Latency in ms was defined as time after stimulus onset (0 ms) and the window of 10 ms after stimulus onset was evaluated. Wave amplitudes were calculated in µV as the difference between leading positive and trailing negative deflections/peaks (Fig. 3). Their latency was determined at the leading positive peak. When available, Waves I, II, III, V and VI were recorded at approximately 1-2 ms, 2-3 ms, 3-4 ms, 5-6 ms and 6-7 ms respectively. In cases where the negative peak of the wave was not properly distinguishable, only the positive peak was determined as to calculate the latency but not the amplitude of said wave. Inter-peak latencies I-III, III-V and I-V as well as V/I and III/I ratio were calculated (Hofmeier et al., 2018, Hofmeier et al., 2021). It was attempted to record wave IV as well, however only few measurements showed a distinct positive and negative peak, while most other measurements only presented two inflections suggesting the presence of another wave between III and V, but not allowing any exact values to be determined.

### 2.4.4 Auditory steady-state response

The 20-minute auditory steady-state response (ASSR) was measured immediately after the ABR measurement using the same setup, the participants remained lying down and was measured with a modulation frequency of 116 Hz. Two blocks of 800 clicks each were measured with a carrier frequency of 4 and 6 kHz at 70 dB SPL (rectangular modulation, i.e., RAM-EFR; described by Vasilkov et al., 2021). Stimulus duration was 400ms and epoch duration was set to 500 ± 10 ms. All epochs were averaged to calculate the spectral power by FFT (Matlab 2021b). The peak amplitudes of the first three harmonics (Vasilkov et al., 2021) were averaged. Measurements with inadequate signal-to-noise ratios (SNR below two) or ASSR peak amplitudes higher than 0.15µV, were excluded from the statistical evaluation.

## 2.5 Procedure of the study

Participants were invited for a total of 4 measuring sessions for this project. Sessions 3 and 4 were conducted and analysed by a colleague, from these two sessions only the phoneme discrimination task is included in this thesis. The overall structure of the four measuring sessions with the individual measurements includes in each session is shown in Fig. 4. Per ethics application, each of the four sessions was limited to 120 minutes.

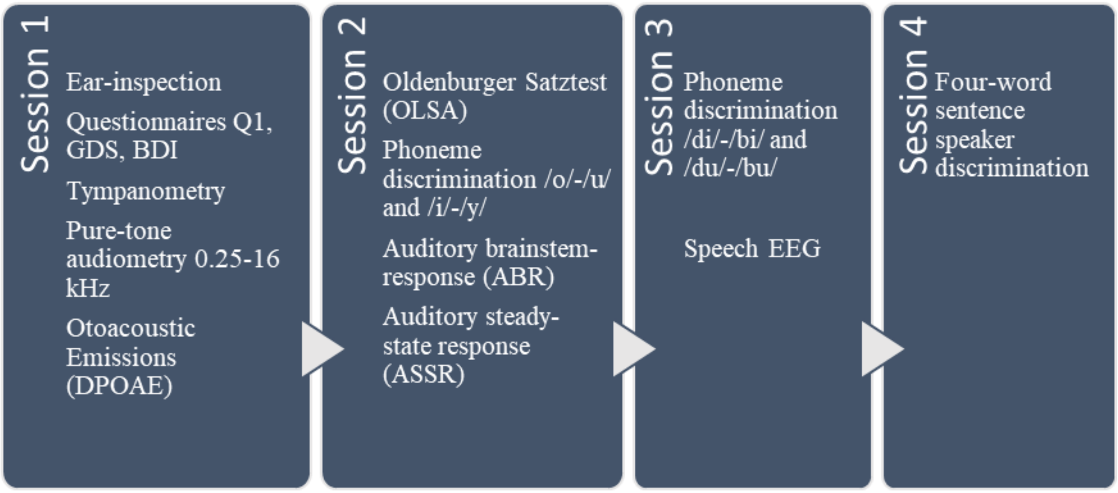



**Fig 4.** Study overview. Participants underwent a total of four measuring sessions on separate days, lasting 2 hours each. Only sessions 1, 2 and the phoneneme discrimination task in session 3 are part of this dissertation. Pre-screening before Session 1 was undertaken to filter out subjects that did not meet criteria. If questionnaires or audiometry during Session 1 showed reason to reject participant, the measuring session was cancelled.

## 2.6 Calculation and statistical analysis

Statistical analysis of group differences was calculated using tests for non-normally distributed data were applied. ABR wave amplitudes and latencies were compared for group differences using a one-way ANOVA (analysis of variance).

Calculated p-values were compared against the criterion of $\alpha = 0.05$, $p \leq 0.05$ was considered as statistically significant, while $p \leq 0.1$ was noted with an asterisk in brackets to inform for a trend in the distribution, though not reaching statistical significance. Correlations of two measurement parameters were verified with Pearson Correlation Coefficients (r).

ASSR response amplitude differences between PNOT groups were compared using Mann-Whitney-U tests. A 1-sided hypothesis was applied between good and poor, good and standard, and standard and poor performers. Phoneme discrimination performance was also compared between the groups categorized for good, standard and poor discrimination performance based on PNOT, the percentage-correct score of answers was compared by Mann-Whitney-U tests with a 1-sided hypothesis: The group of poor/standard performers contains more participants with low percentage correct scores compared to the standard/good group. (Tab. 5). Statistical comparisons were done for the percentage of correct scores obtained for the "difficult" discrimination task (small spectral and temporal syllable contrast) and for the scores obtained for the "easy" discrimination task (larger spectral and temporal syllable contrast).

Analysis of variance for speech perception thresholds beyond pure tone thresholds was performed by least square multivariate linear fitting of the 5 added PCs (principal components) derived from pure tone thresholds and one additional observable which 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 5 PCs and the tested observable. This can be understood as removing the influence of PTT on ABR wave amplitude or latency or other parameters like $L_{EDPT}$ thresholds, ASSR amplitudes or phoneme discrimination. An inherent risk of the increase of dimensions of the regression model is the possibility for overfitting. In order to eliminate this effect, we compared the observed increase in explainable variance in the observed 6-dimensional model to the variance of 10.000 pseudo-models in which we randomly shuffled the additional observable before fitting the model. This gives us a reliable estimate of what gain in explained variance was achieved based on chance (Schirmer et al., 2024).

# 3 Results

## 3.1 Pre-screening participants and final recruited cohort

After pre-screenings, 112 participants were invited to participate in the first session of the study as described under methods (see 2.2). Using questionnaires and pure-tone audiometry as described under methods, a final 89 participants were included in the study, as seen in Appendix C.

### 3.1.1 Otoscopy of the ear canal and tympanic membrane to exclude pathologies

Otoscopy examinations of both ears, as described under methods (see 2.3.1) for all 89 participants showed no relevant pathologies, with only few participants showing sclerosis of the tympanic membrane after childhood paracentesis or otitis media (number not recorded, as this was not considered rejection criteria). Cerumen was removed by an ENT-specialist if complete vision of the tympanic membrane was not possible. When no relevant pathologies were evident, the consent form was signed by the ENT-specialist (which was the case for all participants, including those that were later excluded) and advancing to the audiometric measurements.

3.1.2 Questionnaires screening for depression and cognition for exclusion of depression- and dementia-associated hearing loss show increased depression scores for young participants.

To gain insight into any depressive-like behaviour of the study participants, BDI-II and GDS were used as questionnaires to evaluate its potential impact on hearing, as described in 2.3.4. BDI-II and GDS scores placing participants at "moderate depression" or higher could provide insight in contributions to hearing (see 2.3.4). Both Depression questionnaires BDI-II (Fig. 5 A) and GDS (Fig. 5 B) showed that the young group had an overall higher mean depression score than the middle-aged or older group, while the median BDI and GDS scores between young and older groups are equal and the standard error of the mean (SEM) shows a higher variance for the young group (Table 2). This makes it unlikely that our cohort results are skewed by interactions between depression level and hearing capabilities, elevated mean scores occur as a result of few individual elevated scores. Looking at individual depression scores, four participants from the cohort scored above 19 points in the BDI, indicating moderate depression: CS008, CS023, CS053 and CS074 (middle-aged, older, young and young respectively, Fig. 5 A). All four participants also showed elevated scores in the GDS (> 5). They were identified and marked in red in Fig. 5 and will be observed individually for the following measurements.


**Fig. 5** Questionnaires for depression and cognition. Individual test-scores plotted over age for **A** BDI-II (Beck et al., 1996), **B** GDS (Yesavage et al., 1982) and **C** MMSE (Folstein et al., 1975). 4 participants particularly stand out with a "moderate depression" in BDI (score >19). Elevated BDI scores are spread-out between young, middle-aged and elderly participants. **C** No participant scored lower than the 27-point limit, which indicates suspected cognitive decline.

The MMSE did not detect any relevant cognitive decline, with no participant of the otherwise included cohort testing below the 27-point score, which would otherwise indicate cognitive declines (Fig. 5 C). A single participant, who was excluded after session 1 due to insufficient proficiency in the German language (participant provided incorrect information during recruitment) scored 27 points. This could have been due to not being fluent in German, he was offered further diagnostics nonetheless.

In conclusion, 4 participants stand out as exhibiting moderate depression according to BDI-II, while on average the young group scores higher in both depression scores than the older cohorts. Screening for cognitive decline using MMSE indicated intact cognitive functions of the entire study cohort.

**Table 2: Questionnaire results grouped by ages.** Mean, median and SEM values of BDI, GDS and MMSE questionnaires, divided by age group.

| | BDI | | | GDS | | | MMSE | | |
|---|---|---|---|---|---|---|---|---|---|
| | mean | median | SEM | mean | median | SEM | mean | median | SEM |
| young | 6.06 | 4 | 1.18 | 1.96 | 1 | 0.47 | 29.46 | 30 | 0.11 |
| middle | 3.96 | 2 | 1.01 | 0.93 | 0 | 0.27 | 29.56 | 30 | 0.12 |
| older | 4.75 | 4 | 0.84 | 1.41 | 1 | 0.34 | 29.62 | 30 | 0.10 |

### 3.1.3 Tympanometry for exclusion of Eustachian tube dysfunction

Out of the 89 participants that went through both measuring sessions1 and 2, 73 showed regular type A tympanograms (Fig. 6 A, B, C peak shows between -100 and 100 daPa, amplitude within dotted box), while 7 right ears showed a type AS tympanogram (Fig. 6 D lower compliance, but peak appears around 0 daPa), 6 ears showed type AD tympanograms (elevated compliance centre peak), 2 ears with type B (no visible peak) and 1 ear with a type C tympanogram (peak appears at negative pressure <100 daPa) (Lehnhardt and Laszig, 2009). These results were discussed individually with a clinically experienced doctor to evaluate study inclusion. In combination with inconspicuous otoscopic examination and pure tone audiometry measurements, these findings were deemed to rather be caused by badly fitting earphones and no action to exclude said patients was taken. Tympanograms did however support the decisions to exclude individual participants when PTA or checklist criteria found reasons suggesting such exclusion from the study. In conclusion, 73 out of 89 participants showed normal type A tympanogram middle-ear function.

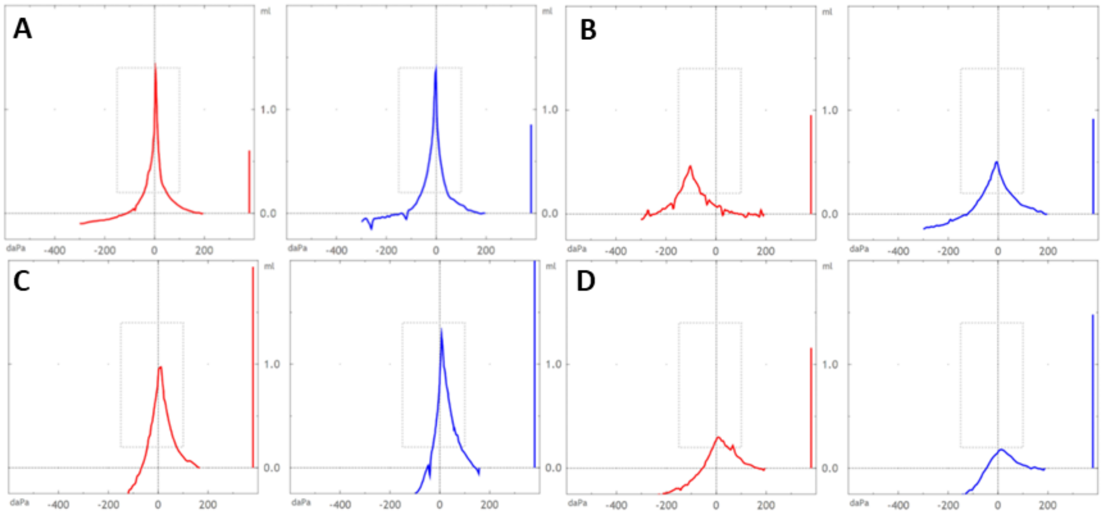



**Fig. 6** Tympanograms measured at 226 Hz of four subjects. Right ears represented in red, left ear in blue. **A,B, C** physiological ear-drum compliance –Type A. **D** Reduced ear-drum compliance – Type As.

3.1.4 Pure-tone audiometry thresholds to assess conductive and sensorineural hearing loss

Pre-screening during recruitment of participants (through e-mail and phone), aimed to filter out participants with any relevant previous medical history, especially related to hearing. Thus, only few exclusions from the study took place after or during measurements. Only one participant was excluded for hearing related problems, namely a conductive hearing loss of approximately 20 dB on both ears (CS055, Appendix C).

3.1.5 Final cohort of participants included in the study

The final cohort of participants included in the study consists of 89 participants, with an overall gender ratio of 29:60 (m:f) and a median age of 47 years. Grouped by age, 29 young participants (18-29 years) with a median of 24 years and gender ratio 11:18 (m:f), 32 middle-aged (30-55 years) participants at median 48 years and gender ratio 8:24 (m:f) and 28 older (56-76 years) participants at median 62 years old and gender ratio 10:18 (m:f) were included in the study. A more equal gender-ratio was not possible within the given study-timeframe, without reducing the total number of participants. Appendix C shows the complete list of 89 included participants with their age, gender, handedness, questionnaire and tympanometry results (Schirmer et al., 2024).

## 3.2 Audiometric measurements

### 3.2.1 Pure-tone audiometry shows age-dependant increases in thresholds, most pronounced in the extended high frequency region

Pure-tone audiometric thresholds were measured as the first-line hearing exam, as performed in a clinical setting (see 2.3.5). The recruited cohort shows the range of hearing loss that could be expected for a healthy, but aging group of participants, with the largest differences at higher frequencies but only small threshold differences at the lower frequencies. In the young group (Fig. 7 A left), variation between participants is minimal at lower frequencies. Only at high frequencies, above approximately 8-10 kHz, some participants do diverge towards higher thresholds. Middle-aged participants (Fig. 7 A middle) behave similarly to the young group, but more participants show threshold increases towards the high frequencies at higher ages. Furthermore, there are some participants with visible threshold losses at the lower frequencies as well. As expected for an older population group where ARHL can be suspected, the older group shows more pronounced high-frequency hearing losses (Fig. 7 A right). While thresholds at low frequencies do not differ greatly from younger participants, threshold elevations begin at lower frequencies when going up to frequency scale and are more pronounced (Schirmer et al., 2024).

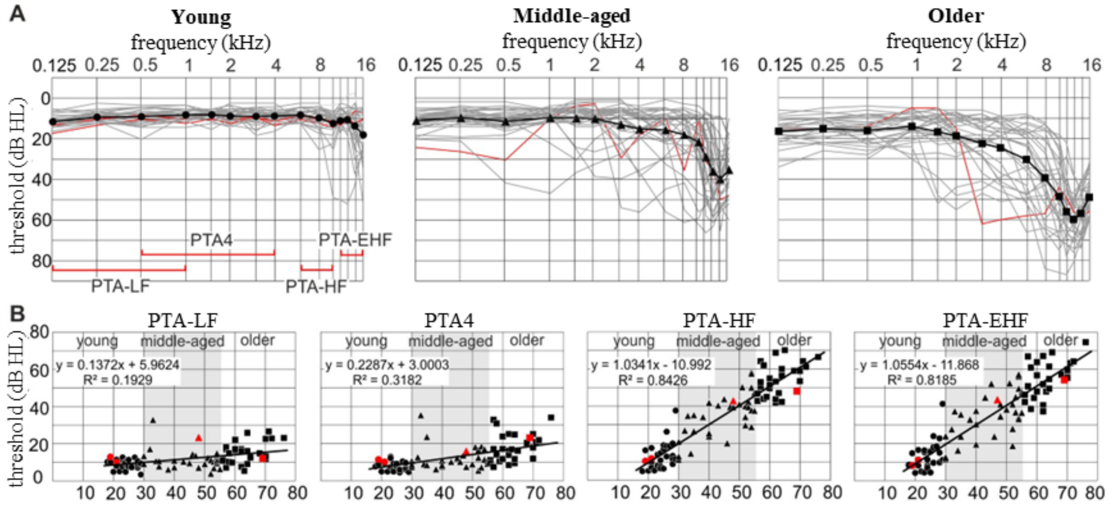



**Fig. 7** Pure-tone thresholds all correlate with age, particularly at higher frequencies. **A** Individual pure tone thresholds for the three age groups young (left), middle-aged (center) and older (right) measured of the right ear for tone frequencies between 125 Hz and 16 kHz, assigned to four different frequency ranges "PTA-LF" [0.125 - 1 kHz], "PTA4" [0.5 - 4 kHz], "PTA-HF" [6 – 10 kHz] and "PTA-EHF" [11.2-16 kHz], illustrated on the abscissa of the leftmost audiogram. The group mean thresholds are plotted in black. **B** Scatterplots for individual hearing thresholds as a function of age, split for the four frequency ranges. The shaded area delineates the age range of the middle-aged group. Furthermore, age groups are represented by circles, triangles and squares for young, middle-aged and older participants respectively. The four subjects with elevated depression scores are represented with a red symbol. Threshold loss over age correlates stronger across frequencies from PTA-LF to PTA-EHF (Pearsons r, n=89: PTA-LF: r=.4392, p=0.000017; PTA4: r=.5641, p<0.00001; PTA-HF: r=.9179, p<0.00001; PTA-EHF: r=0.9047, p<0.00001). *Figure adapted from Schirmer et al., 2024*

To expand the standardized PTA4 average (International Organization for Standardization, 2010), we similarly calculated three more frequency range averages to also represent low and high frequencies. PTA-LF includes low frequencies 0.125, 0.25, 0.5 and 1 kHz, PTA4 range includes 0.5, 1, 2 and 4kHz, PTA-HF includes high frequencies 6, 8 and 10 kHz and PTA-EHF includes the extended high frequency region of 11.2, 12.5, 14 and 16 kHz. All mentioned averages are calculated as the mean threshold of the right ear across said frequencies. Since the remaining tests were only conducted on the right ears, thresholds for the left ear were not considered for pure-tone averages.

While all averages demonstrate a correlation with age, PTA-HF and PTA-EHF show the strongest correlations with $R^2$(PTA-HF) = 0.84, $p < 0.00001$ and $R^2$(PTA-EHF) = 0.82, $p < 0.00001$. Regression lines depicting the correlation between age and the average thresholds PTA-LF, PTA4, PTA-HF and PTA-EHF (Fig. 7 B) were then used to group participants into "low threshold" and "high threshold" groups of each frequency range (above and below the regression line). The aim is to be able to quantify hearing/comprehension losses while eliminating the effects of age. The groups contain between 37 and 52 participants, with median age difference being within 3 years difference for PTA4, PTA-HF and PTA-LF, while for PTA-EHF, the “low threshold” group is 9.5 years older than the “high threshold” group (Table 3) (Schirmer et al., 2024). These groupings will be referred to for some of the following analysis.

**Table 3: PTA groups.** Overview of PTA groups, with n, median age, mean PTA’s and Pearson correlations. Cells comparing mean threshold at a certain frequency range with low or high threshold groups at the same frequency range are highlighted.

| | | n | median age | mean PTA-LF | mean PTA4 | mean PTA-HF | mean PTA-EHF | Pearson's r | p |
|---|---|---|---|---|---|---|---|---|---|
| PTA-LF | PTA-LF low Thr | 52 | 49 | 9.1 | 11.2 | 21.8 | 34.0 | | |
| (0.125, 0.25, 0.5, 1 kHz) | PTA-LF high Thr | 37 | 46 | 16.0 | 15.7 | 23.5 | 35.1 | 0.4392 | 0.000017 |
| PTA4 | PTA4 low Thr | 50 | 47 | 9.7 | 9.6 | 19.0 | 33.4 | | |
| (0.5, 1, 2, 4 kHz) | PTA4 high Thr | 39 | 47 | 14.9 | 17.4 | 27.0 | 35.7 | 0.5641 | < .00001 |
| PTA-HF | PTA-HF low Thr | 50 | 47.5 | 11.4 | 11.4 | 15.1 | 32.8 | | |
| (6, 8, 10 kHz) | PTA-HF high Thr | 39 | 45 | 12.7 | 15.2 | 32.1 | 36.5 | 0.9179 | < .00001 |
| PTA-EHF | PTA-EHF low Thr | 45 | 52 | 11.0 | 11.7 | 19.9 | 29.5 | | |
| (11.2, 12.5, 14, 16 kHz) | PTA-EHF high Thr | 44 | 41.5 | 13.0 | 14.4 | 25.2 | 39.6 | 0.9047 | < .00001 |

To sum up, thresholds correlate with age across all frequencies measured, increasing towards the higher frequencies, supporting the notion that ARHL is most pronounced as a high-frequency hearing-loss.

### 3.2.2 German word matrix test Oldenburger Satztest

Implementation of the OLSA allows to expand the hearing measurements to analyse additional factors that affect speech comprehension other than pure-tone thresholds. Preliminary analysis showing the adaptive volume progress along the 20 sentences for each measurement showed that the algorithm used in the clinical OLSA was mostly accurate for broadband and low-pass filters, but the high-pass filter curves sometimes calculated the $SRT_{50}$ too far off from the last presented sentences when the adaptive volume curve still showed a gradient after 20 sentences. Alternatively, the mean volume of the last 3 sentences showed a more promising approach to calculating the $SRT_{50}$ and was thus used for all calculations.

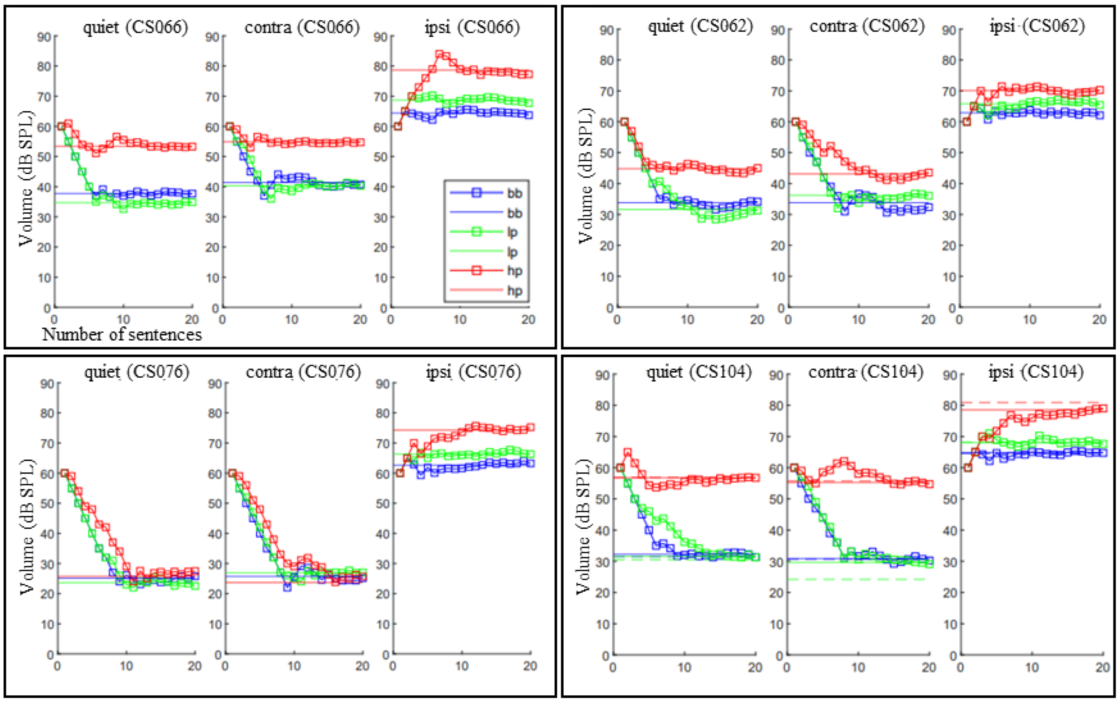



**Fig. 8** OLSA results of four individuals (top and bottom row) in all three noise conditions: quiet, contralateral and ipsilateral noise. Blue represents broadband stimuli, green low-pass and red high-pass stimuli. Each data-point of a curve shows the volume adjustments (y-axis) across the 20 sentences (along the x-axis) of said stimulus condition. Dotted lines are present when standard OLSA algorithm SRTs did not properly represent the curve. Solid lines show final SRT used from mean of last 3 volume values.

Qualitative analysis of results of individual performances almost always shows very similar performance in quiet and with contralateral noise between corresponding frequency filters, as is the case for all four example participants in Fig. 8. For the entire cohort, the mean $SRT_{50}$ for broadband stimulus with contralateral noise masking is only 0.950 ±0.255 dB (SEM) worse than in quiet. Furthermore, broadband and low-pass performance is very close to each other, the mean comprehension in quiet is only worse by 0.163 dB ± 0.340 dB (SEM) for low-pass $SRT_{50}$ compared to quiet. For most participants, high-pass filtered stimuli are the hardest to understand, resulting in a large gap in SRT between broadband/low-pass and high-pass. However, some participants, such as CS076 also performed equally well (or even better) in high-pass filtered speech than with the other filters (Fig. 8., CS076).

When plotted against participant's age (Fig. 9), OLSA performance decreased the most and therefore correlations are most pronounced with the high-pass filter. Here, the highest correlation is observed for high-pass contralateral noise stimulus, with Pearsons $r = 0.6562$, $p < 0.0001$. This goes in agreement with the pure-tone audiometric correlations over age, which are also more pronounced at the higher frequencies. Ipsilateral noise affected comprehension performance in a way that there is almost no difference observed in OLSA performance at all ages (Fig. 9 D, E, F), yet there remains a significant correlation for each.

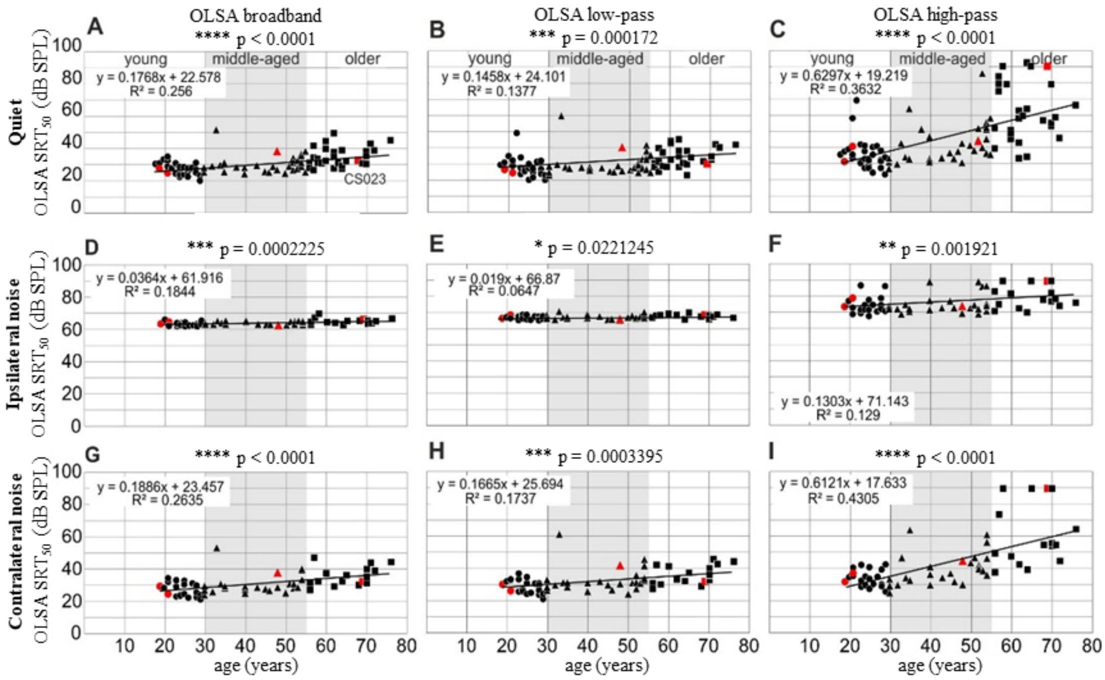



**Fig. 9** Individual OLSA Speech recognition thresholds ($SRT_{50}$) for all 9 conditions plotted against age of the participant. $SRT_{50}$ represents the volume threshold, at which 50% of the presented sentences are correctly understood. Left column shows broadband stimulus sentences, middle column low-pass filter (frequencies <1500Hz) and right column shows high-pass filter (frequencies >1500Hz). Top row shows blocks carried out in quiet (without noise), middle row with 70 dB noise presented on the ipsilateral (right) ear and bottom row with noise presented at the contralateral (left) ear. All 9 conditions show significant correlations, with **E** presenting the weakest and **I** the strongest correlation (Pearson's r for **A**-**I** respectively: r= 0.5062, 0.371, 0.6026, 0.4313, 0.2586, 0.3595, 0.0513, 0.4162, 0.6562). Broadband and low-pass stimulus show barely any age-dependent changes in OLSA performance. Especially Ipsilateral noise has no age-dependent OLSA performance. High-pass filter shows best age-dependency, likely due to an ability to hear higher frequencies. *Figure adapted from Schirmer et al., 2024*

It is important to note that while the quiet condition includes the 89 participants mentioned earlier, it was only possible to evaluate 63 of them for ipsi- and contralateral noise. This was due to an error in the measurement protocol, where noise volume was incorrectly presented for the first 26 participants that underwent measurement session 2, before this problem was fixed. All following measurements had the correct noise level and were included in the cohort of 63 participants for ipsilateral and contralateral noise.

Comparing the mean OLSA differences between different conditions (Table 4), it can be concluded that the contralateral noise masking had little impact on speech comprehension, with the average $SRT_{50}$ being 1.67 ± 0.28, 1.01 ± 0.52 and 2.75 ± 0.29 dB (± SEM) higher for contralateral noise compared to quiet for broadband, high-pass and low-pass filtering respectively. Furthermore, low-pass filtering similarly did not have a large impact on speech comprehension overall, with exception of the condition with ipsilateral noise masking, where low-pass filtered speech had a 4.21 ± 0.15 dB higher $SRT_{50}$ (±SEM). Finally speech comprehension shifted the most when applying the high-pass filter, compared to broadband stimulus, as seen qualitatively in Fig. 9. The largest threshold increase occurred in quiet, with a 16.52 ± 1.72 dB-shift, followed by a 13.25 ± 0.68 dB in ipsilateral noise and a 12.30 ± 1.74 dB (± SEM) shift in contralateral noise shift (Table 4).

**Table 4: OLSA $SRT_{50}$ mean differences.** Comparison of different OLSA conditions shown as absolute mean differences in OLSA SRT50, comparing either noise condition (top row), or filter conditions (middle and bottom row)

| | | Mean (dB) | n | SD (dB) | SEM (dB) |
|---|---|---|---|---|---|
| Difference between **contralateral noise** and **quiet** condition | broadband | 1.67 | 63 | 2.20 | 0.28 |
| | high-pass filtered | 1.01 | 63 | 4.10 | 0.52 |
| | low-pass filtered | 2.75 | 63 | 2.30 | 0.29 |
| Difference between **low-pass** and **broadband** condition | in quiet | 0.16 | 89 | 16.27 | 0.34 |
| | in ipsilateral noise | 4.21 | 63 | 1.21 | 0.15 |
| | in contralateral noise | 1.29 | 63 | 2.22 | 0.28 |
| Difference between **high-pass** and **broadband** condition | in quiet | 16.52 | 89 | 16.27 | 1.72 |
| | in ipsilateral noise | 13.25 | 63 | 5.40 | 0.68 |
| | in contralateral noise | 12.30 | 63 | 13.82 | 1.74 |

When plotting the same OLSA SRTs against the corresponding pure-tone average (Fig. 10 – A, D, G: broadband against PTA4, B, E, H: low-pass against PTA-LF and C, F, I: high-pass against PTA-HF or PTA-EHF), the strongest correlation appears for OLSA low-pass in contralateral noise (Fig. 10 H, $r = 0.7908$, $p < 0.00001$, $n = 63$). All three ipsilateral conditions, while still significant, show very weak correlations with their corresponding pure-tone threshold.

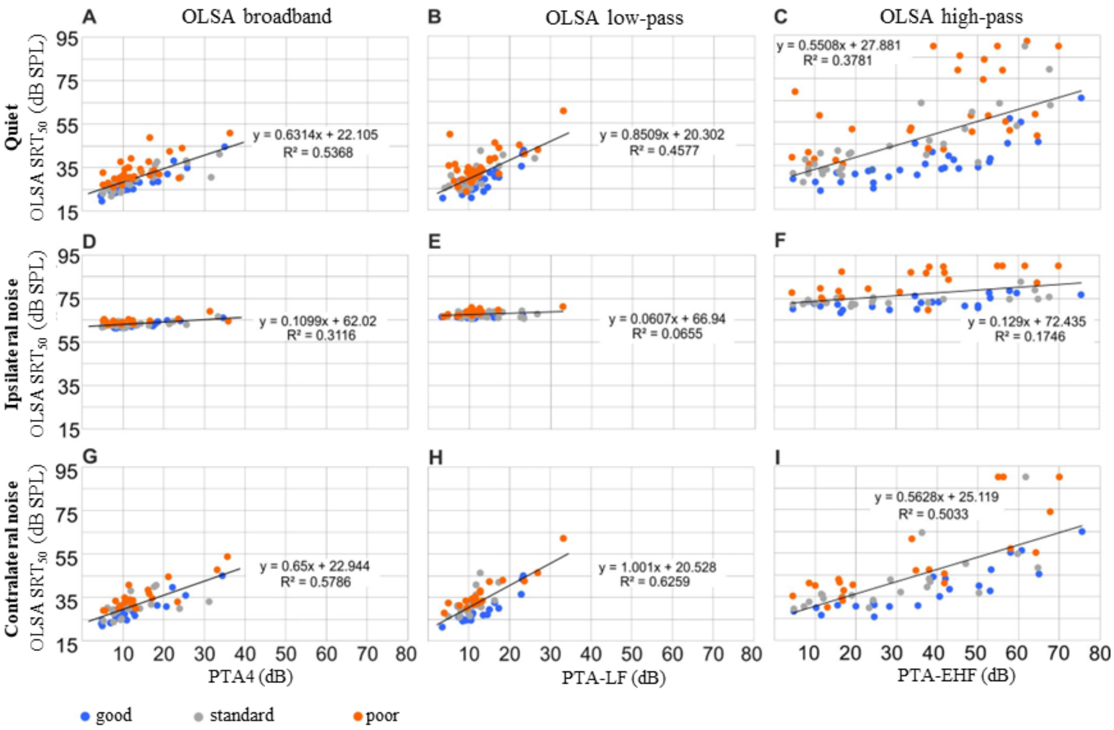



**Fig. 10** OLSA results in quiet condition plotted against PTA4, PTA-MF and PTA-HF. Top Row shows OLSA with broadband stimulus, middle row with high pass (>1500Hz) filter, bottom row with low pass (<1500Hz) filter. PNOT groupings for quiet (**A**-**C**), ipsilateral noise (**D**-**F**) and contralateral noise (**G**-**I**) are marked in colour as good (blue), standard (grey) and poor (orange). Correlation between pure-tone threshold and OLSA $SRT_{50}$ is weakest (but significant) for **E**, and strongest for **H** (Pearson‘s r for **A**-**I** respectively: r= 0.7323, 0.6763, 0.6148, 0.5578, 0.2529, 0.4181, 0.7615, 0.7908, 0.7094). *Figure adapted from Schirmer et al., 2024*

At this point it was deemed reasonable to establish another grouping, this time focusing on the OLSA performance, not just pure-tone thresholds to consequently allow the comparison of performance in the subjective comprehension measurement (OLSA) with objective measurements such as ABR and ASSR (see methods 2.4.3 and 2.4.4).

**Table 5: PNOT group characteristics.** Pure tone normalized speech comprehension (PNOT) groups for quiet, contralateral and ipsilateral noise OLSA conditions shown for age and averaged hearing thresholds. Sub-group sizes, mean age, and PTA4 and PTA-EHF thresholds (dB HL) for the participants with good, standard and poor speech comprehension, Standard error of the mean (SEM) added behind thresholds. *Table adapted from Schirmer et al., 2024*

| | | n | mean age (SEM) | mean PTA4 (SEM) | mean PTA-EHF (SEM) |
|---|---|---|---|---|---|
| speech-deficit (PNOT) quiet | good | 30 | 45.40 +- 2.87 | 12.33 +- 1.21 | 36.22 +- 3.24 |
| | standard | 29 | 38.48 +- 3.10 | 13.41 +- 1.44 | 28.67 +- 3.73 |
| | poor | 30 | 47.60 +- 3.25 | 13.40 +- 1.18 | 38.28 +- 3.73 |
| speech-deficit (PNOT) contra | good | 21 | 45.52 +-3.67 | 12.69 +- 1.65 | 36.85 +- 4.30 |
| | standard | 21 | 40.24 +- 3.44 | 12.36 +- 1.38 | 29.85 +- 4.09 |
| | poor | 21 | 42.62 +- 4.04 | 14.57 +- 1.75 | 33.01 +- 4.70 |
| speech-deficit (PNOT) ipsi | good | 21 | 45.24 +- 3.64 | 14.02 +- 1.49 | 36.00 +- 4.19 |
| | standard | 21 | 41.05 +- 3.87 | 13.21 +- 1.53 | 29.42 +- 4.67 |
| | poor | 21 | 42.10 +- 3.70 | 12.38 +- 1.78 | 34.30 +- 4.25 |

Table 5 shows characteristics of PNOT groups. Average PPTs of the groups were well matched, as well as participant age. While poor and good performers showed almost the same mean age, the standard PNOT-group was on average slightly younger (most pronounced for PNOT quiet) and exhibited better PTA-EHF-thresholds, with differences that were not significant however: PNOT-quiet: p(age) = 0.100; p(PTA-EHF) = 0.150; PNOT-contra: p(age) = 0.606; p(PTA-EHF) = 0.529; PNOT-ipsi: p(age) = 0.713; p(PTA-EHF) = 0.547 (one-way ANOVA) (Schirmer et al., 2024).

Comparing the groupings of individual participants with themselves for the other noise conditions, 77.8% with good, standard or poor speech comprehension in quiet were also grouped into the same tier in contralateral noise. However, only 39.9% with good, standard or poor speech comprehension in quiet were also grouped into the same tier for their comprehension with ipsilateral noise, implying that participants that performed good or poorly in quiet (independent of PTT and age), are not the same participants that performed equally good or poor when presented with ipsilateral masking noise (Schirmer et al., 2024). Together with the previously mentioned observation that OLSA in contralateral noise generally does not differ much from OLSA in quiet comprehension, results with OLSA contra or PNOT contra were from here on not shown for simplification purposes. Interestingly, none of the 4 participants with elevated BDI/GDS scores (Fig. 5 A), were grouped in the good PNOT in quiet group or the good PNOT in ipsilateral noise group (not shown).

To sum up, OLSA $SRT_{50}$ showed significant age-dependant correlations under all filtering conditions, with the strongest correlations occurring for the high-pass filtered OLSA. Similarly, OLSA $SRT_{50}$ also correlated with the PTT's themselves.

### 3.2.3 Principle-component analysis to calculate variance of speech-comprehension contributors

Speech-in-quiet comprehension, which was essentially used here as a control, depended strongly on PTT with a slope of more than 0.6 dB/dB. When removing the effect of PTT on PNOT (Appendix G), the multivariate regression of PTT on OLSA thresholds returned $R^2 = 0.49$. The remaining variance of speech intelligibility in quiet, not explained by PTT was attributed by 38.7% of the variance to OLSA-BB, which corresponds to 3.7 dB SD in SRT (Schirmer et al., 2024). In contrast, broadband speech-in-noise comprehension depended on PTT only with a slope of approximately 0.1 dB/dB, leaving 51.4% of variance unexplained when removing the effect of PTT, corresponding to 1.0 dB SD in SRT.

To explain the remaining variance of OLSA-SRT after removal of age and PTT dependance, diagnostic approaches with identification of changes in signal transmission at stimulus onset were used. Participants with good or poor speech performance based on PNOT grouping were measured with respect to **(i)** subjective speech understanding, **(ii)** ASSR, **(iii)** cochlear amplifier function using short-pulsed DPOAEs, **(iv)** click-evoked central suprathreshold auditory brainstem responses (ABR) and **(v)** phoneme discrimination ability (Schirmer et al., 2024). Their respective shares of the contribution to variance in speech comprehension of each of these measurements are depicted in Appendix G and calculated as the percentage of significant contributions to the variance of OLSA SRT (Schirmer et al., 2024).

### 3.2.4 Subjective hearing assessment with custom questionnaire

Using the custom questionnaire described in methods, subjective self-evaluation of participants was acquired. Question 6 rates subjective hearing into 5 categories of excellent, very good, good, moderate and bad hearing (Fig. 11 A, B). Overall, no participant rated themself as having “bad” hearing. On the other extreme, only 4 out of 89 participants rate their hearing as “excellent” (Schirmer et al., 2024). Comparison of self-assessment when grouped by age (Fig. 11 A), where we could expect young participants to self-rate better than older participants, coheres much worse with actual hearing performance, than by PNOT quiet (Fig. 11 B), where a clear difference in self-rating between good, standard and poor performers can be seen (good group with inverted pyramid, poor comprehension performance with normal pyramid spread). This is, however, not statistically significant (Fig. 11 B, $p = 0.10$, one-sided Fisher Probability Test for “very good” and “good” assessments). When grouped for PNOT in ipsilateral, no significance was found ($P = 0.50$) but again more congruent for contralateral noise ($P = 0.02$, one-sided Fisher Exact Probability Test for “very good” and “good” assessment, not shown in Fig. 11). In an attempt to reproduce results from Motlagh Zadeh et al. (Fig. 11 C), we were not able to show a relation of EHF hearing loss with self-reported hearing difficulty in noisy environments as previously suggested, a feature which is possibly best explained by the different content quality of the questionnaire used in the present study (Motlagh Zadeh et al., 2019).

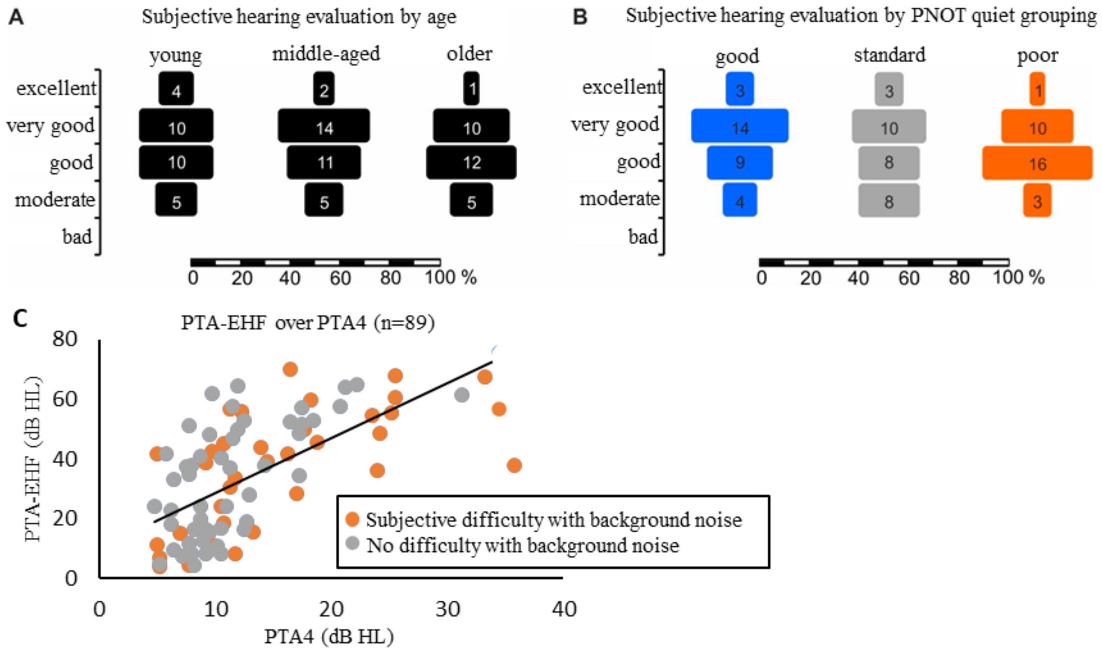


**Figure 11** Subjective hearing evaluation from custom questionnaire, grouped by **A** age groups and **B** PNOT quiet grouping. Participants were asked to assess their own hearing as bad, moderate, good, very good or excellent. y-axis: subjective answer given. x-axis: percentage of all responses given by all participant in age group (A) and in PNOT-quiet-group (B). **C** Thresholds of PTA-EHF in relation to PTA4 of all subjects, as presented in Hunter et al., 2020. Orange data points show subjects that report "difficulty following a conversation with background noise" (question #8), in contrast to subjects that do not, in grey. Regression line of entire cohort. Unlike Hunter et al. (2020), no obvious correlation between difficulty in background noise and position above/below regression line was found. *Figure adapted from Schirmer et al., 2024.*

Zadeh et al. were able to show that increased threshold in the extended frequency region in relation to PTA4 thresholds made it more likely for participants to report problems understanding with background noise (Motlagh Zadeh et al., 2019). Our data does not support this result (Fig. 11 C), neither in high frequencies below 10 kHz, nor in extended high frequencies. The participants reporting problems understanding in background noise (Appendix F) are equally spread above and below the population regression line for both PTA-MF and PTA-HF.

Overall, subjective hearing evaluation showed few correlations with comprehension measurements. Nonetheless, when grouping by PNOT comprehension in quiet, participants showed the most accurate self-rating compared to their peers.

### 3.2.5 Distortion product otoacoustic emissions correlate with speech comprehension

The next step was to look at more objective variables of the auditory pathway. To establish the possible contribution of cochlear amplifier performance on speech comprehension regardless of PTT and age, we analysed the pDPOAE responses. For speech recognition in quiet, when comparing only good (Fig. 12 A-D, blue) and poor performers (Fig. 12 A-D, orange) four of the ten factors (slope of I/O functions not shown here – no significances with one exception for ipsilateral noise) show significant differences, plus one tendency (Fig. 12 D right). **(i)** Comparing the acceptance rates of extrapolated $L_{EDPT}$'s, participants with good speech recognition had a significantly higher rate than participants with poor speech recognition, for both left (n = 60 ears; p = 0.039) and right (n = 60 ears; p = 0.039) ears (Fig. 12 A) (Schirmer et al., 2024)**.** This was tested using a chi-squared test, with the two speech performance groups as the first dimension, and above or below-average $L_{EDPT}$ acceptance rate as the second dimension. **(ii)** Overall, the PTT was not different between groups for left (n = 60 ears; p = 0.86) and right (n = 60 ears; p = 0.61), despite some visible differences towards the higher frequencies being observable (Fig. 12 B). **(iii)** Only the left ear showed significantly higher $L_{EDPT}$ (n = 60; p = 0.012) for the poor group compared to the good group (Fig. 12 C), keeping in mind however, that the left ear was measured only in DPOAE, whereas the grouping relies solely on right ear hearing. **(iv)** When normalizing the $L_{EDPT}$ for PTT (i.e. the PTT was subtracted from the $L_{EDPT}$), again participants of the poor comprehension group show a significantly higher threshold for their left ears (n = 60; p = 0.017), whereas on the right ear, a tendency (n = 60; p = 0.084) can be observed (Fig. 12 D) (Schirmer et al., 2024)

The difference in normalized thresholds (Fig. 12 D: $L_{EDPT}$-PTT) between both groups was calculated as 2.8 and 3.1 dB for the right and left ear, respectively. When excluding the 8 kHz frequency - where the largest differences between both groups, but also the lowest acceptance rates occurred - higher $L_{EDPT}$-PTT differences are observed for participants with poor speech performance, again showing significant differences in the left ear and a tendency in the right ear.

So, semi-logarithmic DPOAE I/O functions, which correlate with cochlear amplification at sound pressure levels near the threshold ($L_{EDPT}$) and the acceptance rate, which is indicative of cochlear amplification at stimulus levels up to 55 dB SPL, demonstrate a more robust cochlear amplifier (lower $L_{EDPT}$ values, higher acceptance rates) for participants with strong speech-in-quiet recognition, or a weaker amplifier for those with poor speech-in-quiet recognition. It is noteworthy however, that with respect to $L_{EDPT}$ results and the normalized $L_{EDPT}$ ($L_{EDPT}$ – PTT), the left ear (contralateral ear to which

OLSA speech recognition was measured) showed the more significant results. Disregarding the absence of absolute consistency, it may be concluded that a more potent pre-neural input signal is advantageous for performance in speech-in-quiet recognition for participants with comparable PTTs and could therefore unveil the “hidden” impact of cochlear amplification of speech recognition in quiet (Schirmer et al., 2024).

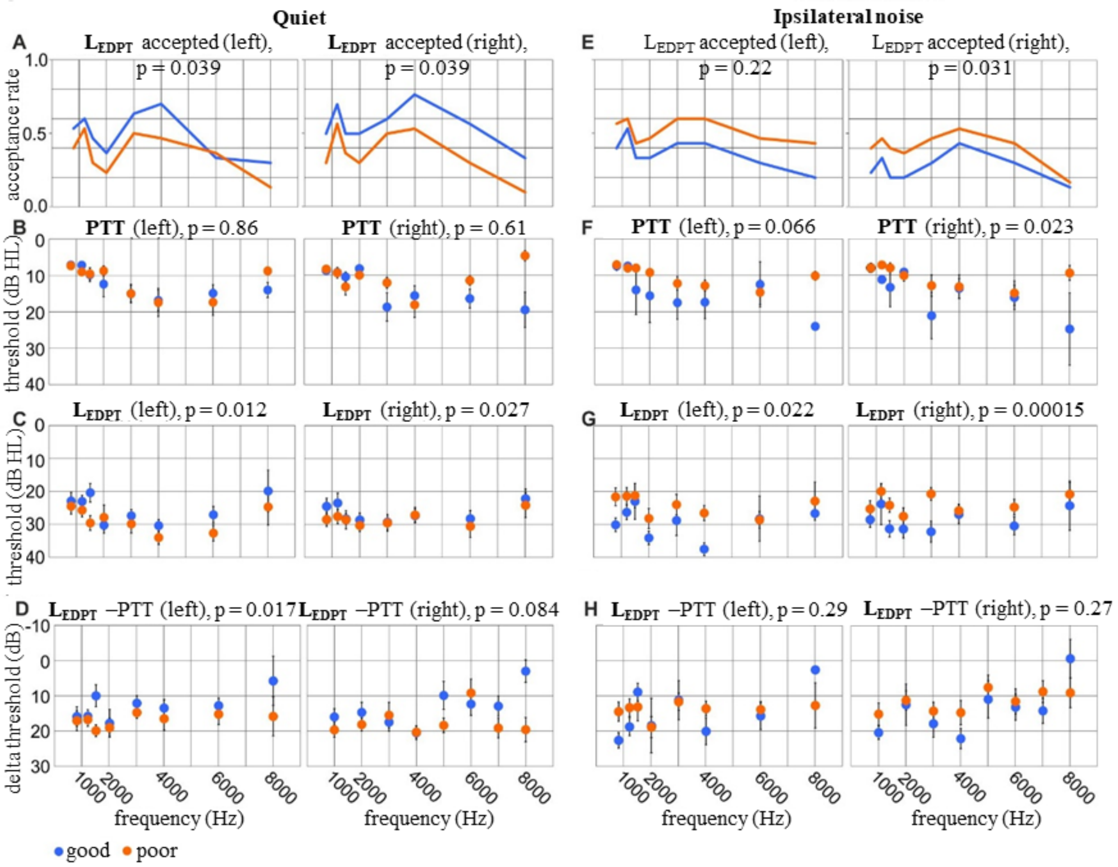



**Fig. 12 A, E** $L_{EDPT}$ acceptance rates; **B, F** PTT; **C, G** $L_{EDPT;}$ **D, H** difference between PTT and $L_{EDPT}$, presented for left and right ears and grouped by PNOT good/poor in quiet **(A-D)** and ipsilateral noise **(E-H).** Compared to participants with poor comprehension-in-quiet performance, participants of the good group (blue) demonstrated higher acceptance rates **(A)**, equal PTT **(B)**, inconclusive $L_{EDPT}$ **(C)**, but a consistent 3 dB lower threshold for $L_{EDPT}$ –PTT, although not significant for the right ear with p = 0.084 **(D)**. Participants with good speech-in-ipsilateral-noise performance (blue), compared to the poor group, demonstrate lower acceptance rates **(E)**, reduced PTT **(F)**and $L_{EDPT}$ **(G)**, but no difference for $L_{EDPT}$ –PTT **(H)**. *Figure adapted from Schirmer et al., 2024*

When looking at DPOAE differences between participants that performed good (Fig. 12 E-H, blue) or poor (Fig. 12 E-H, orange) in speech-in-ipsilateral-noise recognition, five out of ten factors became significant, plus one tendency (slope of I/O functions not shown here – non-significant with one exception) (Schirmer et al., 2024). Unlike the quiet conditions, poor performers showed a significantly higher acceptance rate for the right ear (Fig. 12 E; $p = 0.031$). The PTT (Fig. 12 F) was significantly lower for poor performers on the right ear ($p = 0.023$) but only showing a tendency for the left ear ($p = 0.066$). Poor performers showed significantly lower $L_{EDPT}$ (Fig. 12 G) for both left ($p = 0.0022$) and right ears ($p = 0.00015$). Though not shown as a figure, the slope of the I/O function of the right ear was significantly steeper for the poor performers ($p = 0.041$). When normalized for PTT, the $L_{EDPT}$ (Fig. 8H) did not show any significant differences between good and poor performers. Thus, we might conclude that both the acceptance rate and the $L_{EDPT}$ represent stronger cochlear amplification (lower $L_{EDPT}$, higher acceptance rates) for participants with poor speech-in-ipsilateral-noise (Schirmer et al., 2024).

When classified by speech recognition performance in contralateral noise (not shown in figures), only the slope of the I/O function for the left ear was significantly steeper for good performers ($p = 0.039$). Thus, we could conclude that speech recognition in the contralateral noise conditions shares no correlation with the performance of the cochlear amplifier (Schirmer et al., 2024).

Finally, for each of the three OLSA filter versions, we calculated the DPOAE contributions to total speech comprehension performances. Of the DPOAE contributors, the I/O function acceptance rates (Fig. 12 A) and PTT corrected $L_{EDPT}$ ($L_{EDPT}$ – PTT) (Fig. 12 D) survived the most restrictive post hoc linear mixed model after permutation. This explained 2.0 - 8.3% of the variance, or 0.8 dB, and 3.2 - 4.8 dB of SRT variation in the broadband and high-pass condition, respectively (Appendix G). Acceptance rates of $L_{EDPT}$ measurements for the ipsilateral noise setting were significant for explaining SRT variance in broadband and high-pass conditions, but not quite significant for the low pass condition ($p = 0.051$). Acceptance rates accounted for 3.1-5.5% of the variance, or 0.3 – 1.0 dB in SRT. The $L_{EDPT}$ – PTT thresholds explained up to 7% of variance, or 0.3 dB in SRT in the low pass condition (Schirmer et al., 2024).

To summarize, we find that poor speech comprehension in quiet conditions, independent of PTT and age, is linked to elevated pDPOAE thresholds. This could indicate a weaker pre-neural input signal at stimulus onset. On the contrary, a stronger pre-neural input signals at stimulus onset could be suggested by poor speech comprehension in ipsilateral noise which can be associated with lower pDPOAE thresholds.

### 3.2.6 Psychoacoustic Phoneme discrimination

With the use of phoneme pairs that exhibit formant contrasts above and below the PLL, we attempted to examine the connection between TFS or TENV coding and good or poor speech comprehension independent of PTT and age. The formant pairs with contrasts below the PLL /o/-/u/ and /du/-/bu/ require TFS coding, while the formant pairs /i/-/y/ and /di/-/bi/, due to their frequency spectrum (Table 1), require TENV coding. All pairs were tested with a "difficult" and an "easy" contrast, with each of the three masking noises (stimuli explained in 2.4.2).

Plotting discrimination ability in percentage as a function of age for discrimination in quiet and ipsilateral noise only shows a weak correlation for /di/-/bi/ in the quiet condition and for /du/-/bu/ with ipsilateral masking noise (data not shown). Instead, syllable discrimination was plotted against the corresponding PNOT groups (Fig. 13) established in 3.2.3. Here, discrimination ability from easy and difficult conditions were averaged. Table 6 shows statistical differences for discrimination ability between PNOT groups in quiet and ipsilateral noise.

**Table 6. Phoneme discrimination statistical analysis.** P-values of phoneme discrimination score differences between poor, standard, and good speech performers in quiet and with ipsilateral noise. Easy and difficult task items are shown separately. Significance calculation by 1-sided Mann-Whitney U2 test for syllable discrimination scores between good, standard, and poor speech performers, classified by PNOT. Values in bold highlight $p \leq 0.05$. *Table adapted from Schirmer et al., 2024*

| Listening condition | | quiet | | | Ipsilateral noise masking | | |
|---|---|---|---|---|---|---|---|
| Group comparison | | poor vs. good | poor vs. standard | good vs. poor | poor vs. good | poor vs. standard | good vs. poor |
| /o/-/u/ | (difficult) | 0.059 | **0.011** | 0.159 | **0.032** | 0.337 | **0.035** |
| /i/-/y/ | (difficult) | 0.496 | 0.444 | **0.041** | 0.081 | **0.031** | 0.433 |
| /du/-/bu/ | (difficult) | **0.0314** | 0.0584 | 0.334 | 0.248 | 0.448 | 0.233 |
| /di/-/bi/ | (difficult) | 0.291 | 0.102 | 0.264 | **0.042** | 0.187 | 0.189 |
| /o/-/u/ | (easy) | **0.007** | **0.001** | 0.264 | 0.212 | 0.337 | 0.125 |
| /i/-/y/ | (easy) | **0.014** | 0.084 | 0.236 | 0.129 | 0.248 | 0.356 |
| /du/-/bu/ | (easy) | 0.076 | 0.284 | 0.271 | 0.179 | 0.061 | 0.245 |
| /di/-/bi/ | (easy) | 0.264 | 0.176 | 0.378 | 0.284 | 0.326 | 0.492 |

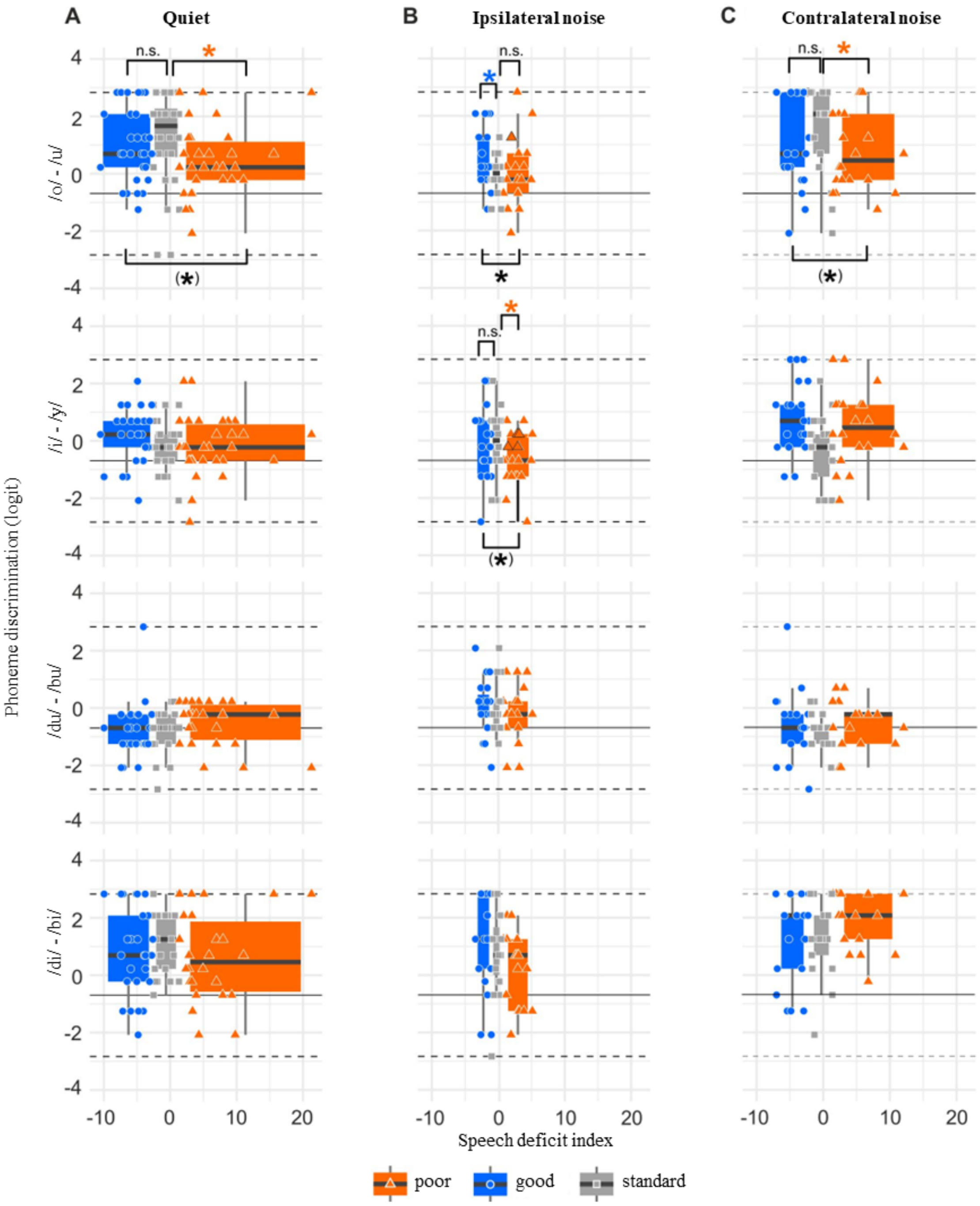



**Fig. 13** Phoneme discrimination scores for all four syllable pairs (/o/-/u/, /i/-/y/, /du/-/bu/, /di/-/bi/), grouped by PNOT speech comprehension in **A** quiet, **B** ipsilateral noise and **C** contralateral noise. Comprehension performance by PNOT groups is color-coded for good(blue), standard (grey) and poor(orange). Each plot consists of a boxplot with task performance (%correct) as a function of the PNOT score (x-axis). Chance level performance of 33% correct is shown by solid line, dotted lines represent 0% and 100% correct on the logit scale. Finally, there is a graphical representation of the significance assessed by Mann–Whitney U tests, significant differences are shown as asterisks color-coded to reflect the three groups. *Figure adapted from Schirmer et al., 2024.*

In general, all participants performed better for the discrimination of /di/-/bi/ than of /du/-/bu/, seen by higher average performances (Fig. 13) (Schirmer et al., 2024).

Performance for the /du/-/bu/ pair (Figure 13 A, B, /du/-/bu/) showed the smallest variation in behavioral results across the cohort, with only a single participant of the group (CS083) performing above 66%. On average and regardless of age, 29.4% to 58.9% of the participants responded at or below chance level (33.3% correct - answers are merely guessed out of 3 options), depending on noise condition. The on average highest percentage correct results was achieved for syllable pairs /o/-/u/ and /di/-/bi/, followed by /i/-/y/ (Figure 13 A, B). Looking at differences between participants with good/standard/poor PNOT comprehension, the most notable and significant differences are shown for the /o/-/u/ syllable, where formant contrasts appear below the PLL, for both quiet and ipsilateral noise (Figure 13 A, B; Table 6). More specifically, even for the easy condition, participants with poor PNOT comprehension in quiet performed significantly worse for /o/-/u/ than both good and standard PNOT groups. Under ipsilateral masking noise, this difference was only observable for the difficult conditions (Table 6). For /i/-/y/ syllables, participants with good PNOT comprehension in quiet performed significantly better than their peers with standard comprehension (Schirmer et al., 2024). For ipsilateral noise, this significant gap in performance can be seen between standard and poor PNOT performers. The discrimination of /du/-/bu/ and /di/-/bi/ syllables showed the least significant differences between different PNOT groups. Performance for /du/-/bu/ was especially poor, with almost no participant exceeding 66% correct answers, and many participants scoring below chance level. Performance for /di/-/bi/ syllables almost always exceeded 90%, even for the difficult condition (Schirmer et al., 2024). Difficulty levels were set too high or low to achieve a sufficient dynamic range, thus no statistical differences could be found.

Overall, data analysis shows that good and poor speech comprehension (PNOT) in quiet differs from comprehension in ipsilateral masking noise in its discriminatory ability of formant contrasts below the PLL of 1.5 kHz, thus requiring TFS coding, and above the PLL, thus requiring TENV coding. More specifically, under quiet conditions, poor speech comprehension was linked to worse phoneme discrimination of syllable pairs with formant contrasts below the PLL, as seen with the /o/-/u/ pair. Meanwhile good speech comprehension in quiet was rather beneficial for phoneme discrimination with contrasts above the PLL, seen for the /i/-/y/ pair. For ipsilateral noise masking, poor speech comprehension had a stronger impact on comprehension of the /i/-/y/ syllables, hence formants above the PLL, but good comprehension was linked to better discrimination of the /o/-/u/ pair instead, hence below the PLL (Schirmer et al., 2024).

### 3.2.7 Auditory brainstem response

Individual peak-to-peak analysis was sometimes limited by too high levels of background noise on the measurement. When visible during monitoring of the measurement session, attempts to improve the measurement were undertaken, such as new electrodes for lower impedance, or restarting the measurement to give the participant the opportunity to stay still. Different methods of filtering were attempted, but this did not make a notable difference for the peak-to-peak analysis. If possible, all identifiable waves were recorded. If only the positive deflection was clearly identifiable, only the wave latency, but not the amplitude was recorded. Wave VI amplitude was difficult to accurately identify, because of interference between negative deflection and noise around 9-10ms after click onset.

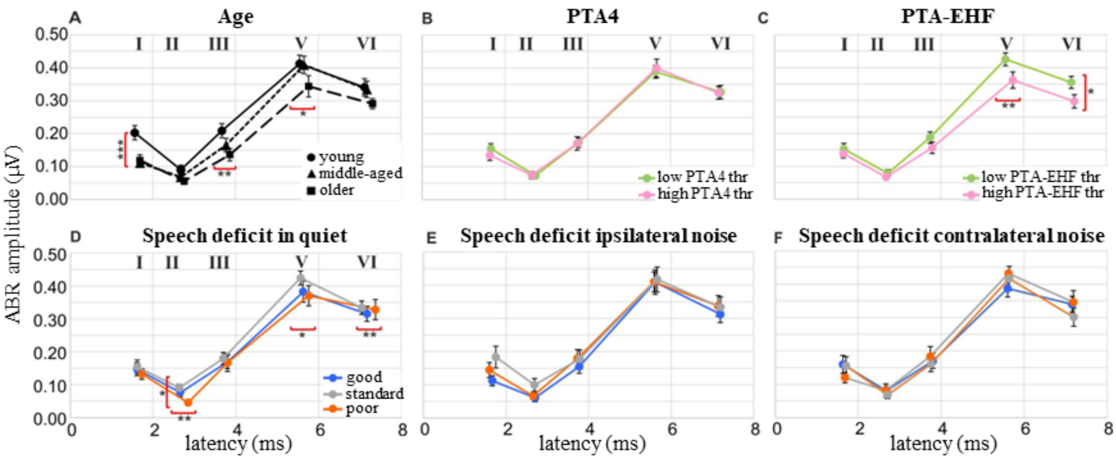



**Fig. 14** Amplitudes and latencies of ABR waves I-VI at 80 dB stimulus clicks. **A** ABR waves averages separated by age groups. Circles represent young, triangles middle-aged and squares older participants. **B, C** display ABR waves grouped by low (pink) versus high (green) threshold of PTA4 (B) and PTA-EHF (C) in relation to their individual age. Wave averages for individuals with matched PTA-thresholds (PNOT groups) for good (blue), standard (grey) or poor (orange) speech comprehension in **D** quiet, **E** ipsi- and **F** contralateral noise are shown. For speech comprehension in quiet (D), significant shifts in latency in low comprehension in comparison to the group with high speech comprehension threshold was observed (ABR wave I: n=29, 27, 24, p=0.218242; ABR wave II, n=24, 22, 16, p= 0.007707, ABR wave III: n=30, 28, 26, p= 0.182784; ABR wave V: n=30,28, 28, p= 0.026617 and VI: n=27, 27, 24, p=0.001055. *Figure adapted from Schirmer et al., 2024*

ABR waves, which represent the sum of all activity of neurons in the ascending auditory pathway (Buran et al., 2010) were then analysed according to the established groupings. The ability for precise synchronous firing of ANFs is critical for the amplitudes of the

ABR wave. In order to maintain synchronous firing rates, physiological function of high SR/low threshold ANFs is particularly important.

To begin with, 70 dB and 80 dB stimuli were compared with respect to the expected amplitude and latency shifts. As anticipated, the evoked responses are weaker in amplitude (with exception of wave II) and have a latency shift at the 70 dB stimulus across all waves (data not shown), compared to the 80 dB stimulus. The comparison is however lacking with regard to the limited number of averaged waves at 70 dB stimulus that were accurately identifiable. For most participants, only wave V was clearly identifiable, while the remaining waves could not be properly distinguished from artefact noise. Consequently, data from the 70 dB stimulus was not used for further analysis from this point onwards. Examining the cohort when divided by age-group confirms the expected results (Fig. 14 A). Older participants are expected to have experienced more ARHL and lower amplitudes, especially in waves I and V (Schirmer et al., 2024). Comparing the young, middle-aged and older participant groups (circles, triangles and squares respectively), we observe that at the 80dB stimulus, the young group shows a significantly higher wave I amplitude compared to middle-aged and older (Fig. 14 A, wave I: young = 0.203 µV, n = 27; middle-age = 0.111 µV; n = 30; older= 0.119 µV; n = 19; p = 0.00029). This indicates an age-dependent cochlear synaptopathy, resulting in lower wave amplitude already at the level of the ANFs. At wave III, there is no significant difference between the three groups, yet the post-hoc Tukey test reveals a slight significance between young and older groups at wave III (Fig. 14 A, wave III amplitude post-hoc Tukey HSD young = 0.209; olde r= 0.136; p = 0.042). At wave V, the reduced amplitudes seem to have completely recovered statistically, even though the older group is still separated by a gap from the numerically almost identical young and middle-aged groups (Fig. 14 A; wave V amplitude; young = 0.413, n = 29; middle-aged = 0.409, n = 31; older = 0.344, n = 21; p = 0.202). This suggests a certain level of amplification through central gain, where the middle-aged and older group are able to fully, or at least partly compensate the reduced signal from the auditory nerve by amplification processes. Latency-wise, a significant difference between age-groups occurs only for wave III and V, where the older group has the largest delay (Fig. 14 A; latency wave III: p = 0.009497, latency wave V: p = 0.018063) (Schirmer et al., 2024).

The next step was to inquire whether these changes are also observable when eliminating the age factor. Applying the PTA groups from 3.2.1 and comparing good and impaired groups in each category, the most distinct observations can be made between the PTA4 and PTA-EHF categories. PTA4 (Fig. 14 B) seems to have no impact on the ABR waves, with no significant differences between PTA4 good and impaired (groups normalized for age to distinguish "good for their age" compared to "bad for their age"). When dividing by PTA-EHF performance (Fig. 14 C), wave I does not differ between good and impaired, but the two waves diverge after wave II, resulting in a visible gap between the subgroups at wave V (Fig. 14 C wave V amplitude: n = 40, 41; p = 0.050773; latency: n = 41, 40; p = 0.003) and VI (Fig. 14 C wave VI amplitude: n = 38, 40; p = 0.040637; latency: n = 38, 40; p = 0.328052). At the same time, the PTA-EHF impaired group has a significant latency delay at wave V compared to the good group (Fig. 14 C wave V latency: n = 38, 40, p = 0.003). This is interesting, considering that the click-stimulus does not reach up to the EHF-region at all, but the group with better hearing at those high frequencies still shows the mentioned gap. One might argue that the PTA-EHF good group also has better hearing at lower frequencies (Table 3), but if this was the case, similar shifts would be expected in the other groupings (Schirmer et al., 2024)

When analysing ABR results for PNOT groups (Fig. 14 D-F), we observed significant wave differences for the PNOT in quiet grouping including an amplitude difference for wave II (Fig. 14 D, good: 0.0767 µV, n = 24, standard: 0.0905 µV, n = 22, poor: 0.0458 µV, n = 16 , p = 0.0458), as well as for latencies of wave II (Fig. 14 D, good: 2.66 ms, n = 24, standard: 2.63 ms, n = 22, poor: 2.85 ms, n = 16, p= 0.00771), for wave V (Fig. 14 D, good: 5.63 ms, n =30, standard: 5.56 ms, n = 28, poor: 5.76 ms, n =28, p = 0.027) or wave VI (Fig. 14 D, good: 7.17 ms, n = 27, standard: 7.04 ms, n = 27, poor: 7.37 ms, n = 24, p= 0.0011) (Schirmer et al., 2024). Interestingly, in all cases of significance, it was not the PNOT quiet – good group, but the indifferent/standard group with the highest average amplitude or lowest latency compared to the poor group. This could be due to the slightly lower average age of the standard group compared to good and poor groups (Table 4). In comparison, ABR wave averages showed no significant differences when grouped by speech comprehension in ipsilateral (Fig. 14 E PNOT ipsi) or contralateral noise (Fig. 14 F PNOT contra). It is however important to recognize, that the size of the cohort for ipsilateral and contralateral noise is smaller than for the quiet condition.

To summarize the ABR analysis, the latency shifts of waves I and V showed the highest significances after restrictive post hoc linear mixed model analysis after permutation (p = 0.004 - 0.030), explaining 2.2% and 2.2% of OLSA-BB variance and 4.6% and 3.2% of OLSA-LP variation in quiet, which corresponds to 0.8 - 1.3 dB of $SRT_{50}$ variance (Appendix G) (Schirmer et al., 2024).

### 3.2.8 Auditory steady state response

A Fast-Fourrier transformation (FFT) analysis was done to show the spectral components of the recorded ASSR signal. As expected, peaks appear at the modulation frequency of 116 Hz (the fundamental) and multiples of 116 Hz (the harmonics). The modulation frequency above 100 Hz allows the expectation that temporal processing of amplitude modulated stimuli in subcortical areas of the brain, rather than neocortical regions can be made (Engelien et al., 2000, Kuwada et al., 2002, Purcell et al., 2004, Lu et al., 2022).

The absolute amplitudes of the fundamental as well as the first and second harmonic were averaged (Vasilkov et al., 2021). Results where the corresponding signal-to-noise ratio (SNR) was below 2 were not considered for our analysis. Furthermore, absolute values greater than 0.15µV were also excluded. A slight negative correlation of lower amplitude with increasing age was noted for ASSR amplitudes at both 4 and 6 kHz carrier frequencies. At 6 kHz there was a slight positive correlation for participants by ipsi/contra PNOT grouping (shown for the average of 4 kHz and 6 kHz ASSR responses in Fig. 15 A). However, these correlations are all minimal across age and most likely occur from individual outliers. The four participants who were outside the normal levels in the BDI and GDS tests (Fig. 5) were all distributed within normal levels for all ASSR conditions (data not shown). The grouped ASSR amplitudes were not found to show significant differences between group averages in speech comprehension between good, standard and poor groups (highest significance occurs for 4kHz, by PNOT in quiet: p = 0.205) Interestingly, in the post hoc linear mixed model analysis after permutation, ASSR was the only electrophysiological measure that significantly explained a considerable amount (7%) of speech-in-noise comprehension (p = 0.012); this corresponded to 0.4 dB in the broadband condition (Appendix G). Therefore, correcting SRTs and ASSR amplitudes for PTT-related variance using the normalization by PCA revealed a significant association between ASSR amplitude and speech recognition threshold (Fig. 15 B, p = 0.0037).

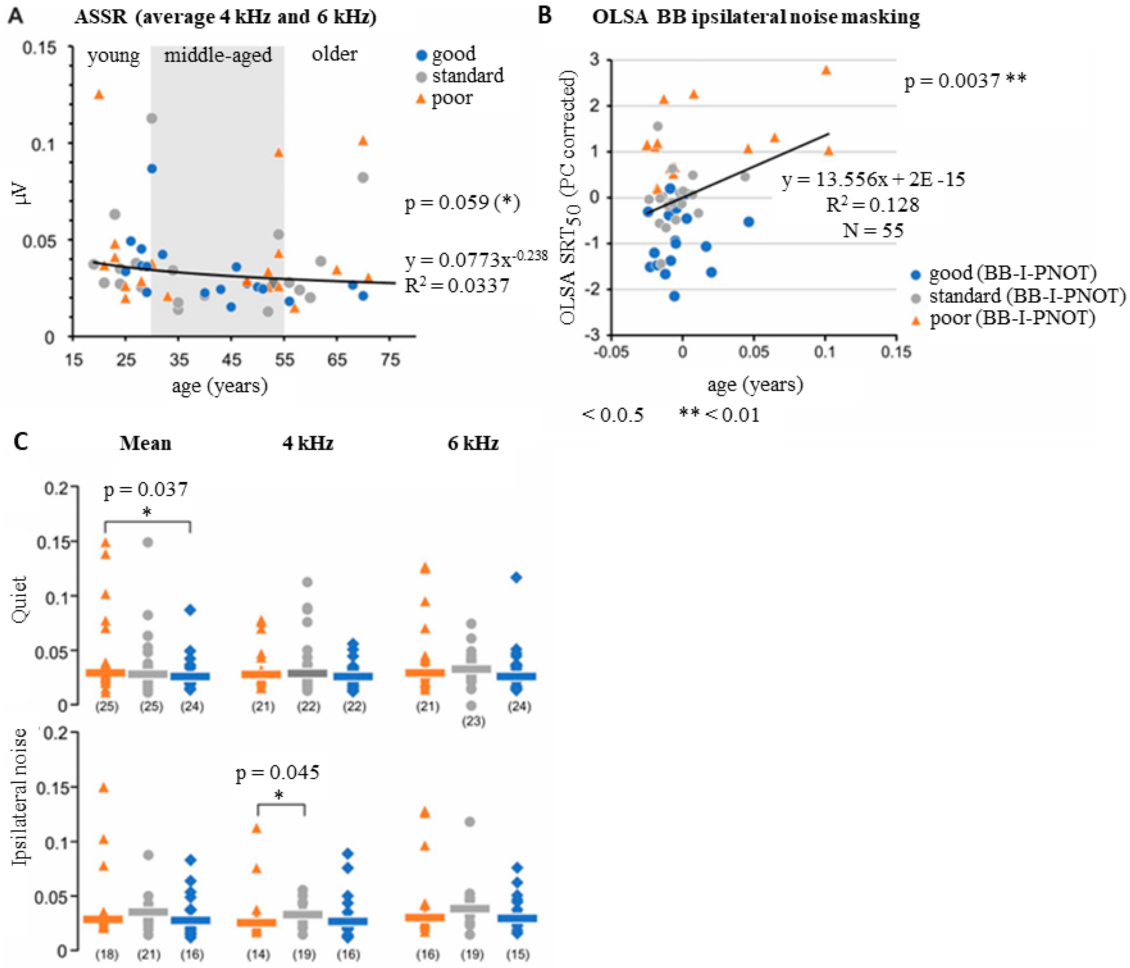



**Fig. 15 A** averaged ASSR response amplitudes in µV for 4 and 6 kHz carriers as a function of participant age in years. Each data point represents an individual participant, color-coded in orange, grey and blue to represent poor, standard and good PNOT comprehension, respectively. **B** Regression line (black) of the dependence of OLSA $SRT_{50}$ in ipsilateral noise on ASSR amplitudes (averaged for 4 and 6 kHz carriers) normalized for PTT. The y-intersection, R2 value, and p-value of regression are given close to the trend line. **C** Median (horizontal bar) and individual participants (symbols) ASSR amplitude averaged for 4 and 6 kHz carriers (Mean, left), 4 kHz carrier (middle), and 6 kHz carrier (right) for the quiet listening condition (upper row), or in ipsilateral noise (lower row). Numbers in brackets indicate the number of participants included in the analyses. *Figure adapted from Schirmer et al., 2024.*

# 4 Discussion

The aim of the study was to examine the factors that contribute to speech comprehension in quiet and in noise (here represented by the OLSA in noise measurements) on a cohort of participants between 18 and 76 years and with normal or mild hearing losses. Expectations that PTTs represent the largest contributing factor were met, but particular focus was put on the remaining variance that was not explained by PTT or age, which we calculated with the variance analysis explained in 3.2.3. In quiet, 38.7 % and in ipsilateral noise, 51.4% of the $SRT_{50}$ variance were not accounted for by PTT and age (Schirmer et al., 2024). In order to assess the influence of cochlear amplifier efficiency and synchronicity at stimulus onset, the cohort was split into 3 groups according to individual OLSA speech comprehension performance in relation to their PTT using the described PNOT method. These relative groupings of poor, standard and good speech comprehension were established for quiet and noise conditions separately (Schirmer et al., 2024).

Pure-tone thresholds and OLSA comprehension deteriorate in aging participants. Effects of the aging ear were confirmed both with pure-tone audiometry, as well as by speech comprehension thresholds, as previously known (Füllgrabe et al., 2014, Monson et al., 2014, Wu and Liberman, 2022). This effect was small at low frequencies (PTA-LF, PTA4), but increases at PTA-HF and even more so at PTA-EHF frequencies (Fig. 7) (Schirmer et al., 2024). Correspondingly, SRT's in the OLSA also correlate significantly with age, in all tested filter conditions (Vasilkov et al., 2021, Garrett et al., 2024)(Fig. 9). Here as well, high-pass filtered speech (frequencies below 1.5 kHz removed) showed the strongest correlation with age (Fig. 9 C, F, I). As the consensus remains that frequencies > 8 kHz have a limited effect on overall hearing perception, standard clinical audiometry rarely includes the extended high frequency spectrum tested here. Previous observations on the relevance of EHFs should however be reconsidered, as they have been shown to provide highly directional cues for speech localization (Heffner and Heffner, 2008) and talker's head rotation (Monson et al., 2014, Hunter et al., 2020). Furthermore, tests comparing comprehension showed a significantly better speech perception with broad filtered masking noise compared to a broadband masking noise (Schirmer et al., 2024), indicating that EHF frequencies do impact speech comprehension (Motlagh Zadeh et al., 2019). Though we can show the expected correlations at extended high frequencies, the study cannot sufficiently discern the isolated impact of EHF threshold loss on speech comprehension (Schirmer et al., 2024).

ABR wave amplitudes and latency shift with age and PTA-EHF thresholds. As shown in animal models, ABR wave I amplitudes are correlated with cochlear synapse numbers (Kujawa and Liberman, 2009, Sergeyenko et al., 2013). However this correlation is based on postmortem counts of temporal bone data, which is not realistically obtainable for humans (Bramhall, 2021). At a supra-threshold level, peak amplitudes are defined by discharge rates of the IHCs (Buran et al., 2010), as well as the synchronicity of the ANF discharges at the stimulus onset (Johnson and Kiang, 1976). As stimulus intensity level is increased, higher frequency ANFs are recruited and the synchronicity increases, leading to reduced latency and increased amplitude of the peak (Harris et al., 2018). Whereas low-SR ANFs are mainly required at hearing threshold but provide little for ANF synchronicity, the synchronous firing rate relies on the sensitivity of the high-SR ANFs instead, as these are recruited at supra-threshold level when stimulus intensity level is increased (Liberman, 1978, Huet et al., 2019, Rhode and Smith, 1986, Meddis, 2006). Middle-aged and older participants are shown to have reduced ABR wave I amplitudes, when compared to the young group (Fig. 14 A), which suggests the presence of age-dependent synaptopathy (Bramhall, 2021, Vasilkov et al., 2021, Schirmer et al., 2024). The observed amplitude reduction and latency delay of wave III-VI in older participants, compared to young and middle-aged participants (Fig. 14A) indicates that the older group is not able to compensate as effectively for the occurring synaptopathy (Schirmer et al., 2024). Considering that the in-ear headphones used in this study were able to present a frequency bandwidth up to 8 kHz, but the most pronounced PTT differences between ages occurred at these high frequencies, a possible explanation could relate to previously observed influences of PTA-EHF losses on lower frequences (Märcher-Rørsted et al., 2022). Märcher-Rørsted et al. used frequency-following responses (FFR), which are phase-locked evoked responses generated by the auditory brainstem in response to periodic stimuli. These FFRs, which are most robust at frequencies below 1.5 kHz and thus below the PLL (Kuwada et al., 2002, Batra et al., 1986), were nonetheless affected in their ANF synchronicity at lower frequencies when elevated thresholds were shown in the EHF region (Märcher-Rørsted et al., 2022). At the same time, OHC dysfunction did not affect the temporal coding of the phase-locked ANF response in the FFR measurements (Märcher-Rørsted et al., 2022) and since it is mainly the high-SR/low threshold fibers (compared to low-SR/high threshold fibers) that contribute to ANF synchronicity (Huet et al., 2019), we may infer that the high-SR fibers are responsible for the reduced central compensation when comparing the older with the middle-aged group. This raises the open question, if elevated PTA-EHF thresholds can have an impact on the possible transformation of high-SR low threshold auditory fibers to low-SR high threshold ANFs, as predicted by Liberman et al. in 1984 for the acoustic overexposure damage to stereocilia when thresholds elevate (Liberman and Dodds, 1984). The current data could support this theory (Fig. 14 C), where the group with better PTA-EHF thresholds (and normalized for age) shows a significant increase in ABR wave V compared to the other half of the cohort. This question might in the future receive more attention in studies analysing the effect of the extended high frequencies on speech comprehension (Motlagh Zadeh et al., 2019, Hunter et al., 2020).

By recording ABR waves at 2 supra-threshold fixed volumes of 70 and 80 dB, the study allowed comparison of brainstem responses between participants. During the study it was recognized that the 70 dB measurement was often not sufficiently loud to achieve reliable ABR wave results, as a result this measurement could not be used for analysis altogether. In hindsight, it was not sufficiently taken into account that young participants without hearing loss used for testing during the pilot phase would show better ABR results at 70 dB than the later cohort composed of much older participants with developing hearing loss. As suggested by Bramhall et al., stimuli at 90-100 dB SPL could have yielded better results and allowed comparisons in individual wave latency shifts as the stimulus volume is increased (Bramhall, 2021). Alternatively, an individual threshold-dependent approach could also be considered.

Self-rating of participants is not a good indicator or actual hearing ability. By recording the subjective hearing abilities in everyday settings, it was possible to compare perceived hearing disability with (more) objective audiometric measurements. Here it was shown that when grouped by age (which was shown to correlate well with hearing thresholds and SRTs), no relevant correlation was observed. Instead, the grouping by PNOT, especially in quiet, showed a statistical tendency for the self-assessment (Fig. 11) (Schirmer et al., 2024). While this result is in line with previous findings about self-assessed hearing (Motlagh Zadeh et al., 2019, Shehabi et al., 2022), it does not fully support the hypothesis that the here identified factors affecting speech comprehension regardless of age and PTT are relevant for the self-rating of individual comprehension ability.

ASSR amplitudes have a low impact on speech comprehension in quiet and in noise. Current study results did not quite meet previous studies showing that ASSR amplitudes decline with age (Boettcher et al., 2001, Rumschlag and Razak, 2021), even though temporal coding war measured at relatively high envelope frequency (Grose et al., 2009). Instead, a slightly declining amplitude could only be shown with a statistical tendency (Fig. 15 A). Furthermore, there appeared to be no trend between PNOT grouping and ASSR amplitude (Fig. 15 A) (Schirmer et al., 2024) Surprisingly, participants with poor comprehension (PNOT quiet and ipsilateral noise, Fig. 15 C), provided higher ASSR amplitudes and correlated strongly with OLSA SRT in ipsilateral noise (Fig. 15 B) (Schirmer et al., 2024). Searching to explain this finding, it can be observed that the data correlation may mainly be created by a small subgroup within the poor performer group, which demonstrate comparatively much higher ASSR amplitudes (Fig. 15 C) and also significantly increased UCL (Appendix G), indicating a maladaptive loudness sensation. This could occur because of weakened compressive nonlinearity of the aging (Abdala et al., 2021), which is then perceived as steeper scaling loudness. This weakened compressive nonlinearity could be linked to changes in cochlear amplifier function, as observed in participants of poor PNOT comprehension in ipsilateral noise (as discussed next) (Schirmer et al., 2024).

DPOAE responses show effects of cochlear amplification on speech comprehension. Using pDPOAE measurements, the current study was able to associate PNOT comprehension differences with varying levels of cochlear amplifier functionality. As seen in Fig. 14 D, participants with poor and good speech comprehension in quiet showed a significant difference for the $L_{EDPT}$ when normalized for PTT ($L_{EDPT}$-PTT), which explains 2% and 8.3% of OLSA-broadband and OLSA-hp variance, respectively (Appendix G) (Schirmer et al., 2024). As a possible explanation, $L_{EDPT,}$ known to correlate well with PTTs (Kimberley et al., 1994, Boege and Janssen, 2002), do not depend on auditory fiber firing rate adaptation (Kiang et al., 1966). Since the pDPOAE stimulus pulse length used here for I/O-functions is much smaller than the time constant of the medial olivocochlear reflex (Kim et al., 2001), $L_{EDPT}$ should not be affected by its adaptation. DPOAE instead provide results of the pre-neural input signal that reaches the IHC, before adaptation (Zelle et al., 2017a). In contrast, PTTs as measured in a clinical setting are essentially integrated in an adapted state of nerve firing over approximately 500 ms (Goutman, 2017)). Weakened nerve adaptation at stimulus onset is linked to vesicle depletion at the IHC synapse in mammals, which can cause synaptic fatigue or desensitization kinetics of postsynaptic receptors (Goutman, 2017, Moser and Beutner, 2000, Peterson et al., 2018, Willmore and King, 2023). Increased levels of synaptic fatigue at the IHC synapse would therefore present themselves as a reduced firing rate adaptation, which would be reflected as a smaller $L_{EDPT}$-PTT difference (and poorer speech comprehension in Fig. 12 A, D) and vice versa (Schirmer et al., 2024), and should in future be considered as possible causes for detection of stimulus onset at speech presentation. The calculated variances of this study allocated 7.0% and 5.5% of the $SRT_{50}$ variance in OLSA-LP condition to the LEDPT-PTT difference and the acceptance rate of DPOAE growth functions under ipsilateral noise conditions (Fig. 12, Appendix G), which corresponds to 0.3 dB in $SRT_{50}$ for both measurements. In the OLSA-HP condition, LEDPT-PTT and I/O-function acceptance rate make up 0.7% and 3.1% of the variance respectively, which corresponds to 0.5 dB and 1.1 dB (Schirmer et al., 2024). The acceptance rate variance sign switch between the quiet and ipsilateral condition (Appendix G) indicates that in ipsilateral noise condition, a larger pDPOAE growth function, higher acceptance rate and larger $L_{EDPT}$-PTT correlates to poorer speech comprehension (Fig. 12 E-H). This could be caused by compression of the cochlear input signal to the auditory pathway. Considering that for the OLSA in noise task temporal information can only be utilized within a narrow dynamic range, a comparatively lower DPOAE I/O function acceptance rate would occur at a level where basilar membrane compression ends (Schirmer et al., 2024). At this level, the growth behaviour approaches to becoming linear. This effect could be advantageous for analysis, since the mostly linear dynamic range would give better modulation contrast during the measurement (Marrufo-Pérez and Lopez-Poveda, 2022). This recruitment describes the phenomenon when subjective loudness perception increases abnormally with sound intensity (Fowler, 1936, Kubota et al., 2019, Denes and Naunton, 1950). Nonetheless, recruitment does not only affect DPOAE growth functions (Rasetshwane et al., 2013), but also loudness scaling (Abdala et al., 2021) and limits TENV coding (Schirmer et al., 2024).

PNOT comprehension grouping is affected by ABR peak amplitudes and latencies. As mentioned earlier, amplitude reduction in ABR wave I and latency shifts of wave II-VI of participants of the poor PNOT in quiet group (Fig. 12) revealed a functional synaptopathy, keeping in mind that the poor PNOT in quiet group was of similar age as the corresponding good group (Table 4). ABR wave I explained 2.8% (1.0 dB) in amplitude and 4.6% (1.3 dB) in latency of OLSA-LP $SRT_{50}$ (Appendix G). Previous studies describe the impact of cochlear synaptopathy, independent of OHC loss, on OLSA $SRT_{50}$ to lie at approximately 1 dB. A model to simulate loss of more than 50% of synapses showed an intensity level difference on basic perceptual tasks of 1.4 dB (Suthakar and Liberman, 2022, Oxenham, 2016). As evident by reduced amplitudes delayed latencies of early and late ABR waves in poor PNOT in quiet (Fig. 14 D), we can therefore show in this study that cochlear synaptopathy seems to be present and affect hearing, independent of age and pure-tone thresholds, This contradicts studies arguing that speech comprehension deficits without background noise occur mainly in combination with elevated PTTs (Johannesen et al., 2016), and results, which assert that cochlear synaptopathy does not have a direct impact on coding of moderate and high intensity speech (Carney, 2018). Amplitudes of ABR wave peaks are known to represent the firing rate of individual ANFs (Buran et al., 2010) as well as the precision of simultaneous and synchronous discharge of ANFs at the onset (Johnson and Kiang, 1976). The ability of synchronous discharge and onset relies heavily on the function of high-SR/low threshold fibers, which is in turn reflected by wave latencies and thresholds (Rhode and Smith, 1986, Meddis, 2006, Zohar et al., 2011, Heil et al., 2008) and its contribution to the increase in spike rate at stimulation onset (Chimento and Schreiner, 1991, Young and Sachs, 1973). This leads to the interpretation that the observed amplitude changes of waves I/II and the latency shift of late ABR waves between poor and good PNOT in quiet (Fig. 14 D) most likely occur due to differences in mentioned high-SR/low threshold auditory fibers, which changes the ability to fire in synchrony at stimulus onset. In connection with the described reduced DPOAE acceptance rate in quiet, pre- and postsynaptic changes in occurrence with synaptic fatigue or desensitization should be investigated. These could lead to reduced firing rates at peak onsets (Goutman, 2017, Willmore and King, 2023, Moser and Beutner, 2000, Peterson et al., 2018) and should be evaluated as an additional contributor affecting speech reception without background noise, independent of age and thresholds.

Discrimination of different phonemes by PNOT groups depends on formant frequencies. The most noteworthy observation of the phoneme discrimination task, when analysed by PNOT groups, was that the poor PNOT in quiet group showed poorer speech coding for the /o/-/u/ syllable (Fig. 13A), thus below the PLL (Schirmer et al., 2024). This result can be traced back to the functionality of high-SR low threshold fibers, which have previously been shown to crucially contribute to the perceptional threshold for frequencies below the PLL (Huet et al., 2018, Bourien et al., 2014)). For the syllables with formant contrasts above the phase-locking range (Fig. 13 B, /i/-/y/), where participants with poor PNOT performed worse than standard PNOT, such situation can be explained in context with the pDPOE results of this group. The higher acceptance rate and higher $L_{EDPT}$-PTT difference (Fig. 15 E,H) (Schirmer et al., 2024) could be related to reduced basilar membrane compression, which limits the dynamic range width of low-SR and medium-SR fibers. As these fibers are necessary for TENV coding, phoneme discrimination above the PLL is reduced (Huet et al., 2019).

In conclusion, we showed that speech comprehension depends strongest on thresholds, making up approximately 50% of the SRT variance (Schirmer et al., 2024). Taking a closer look at factors that affect speech discrimination independent of PTT, we discovered that these factors differ in their importance, depending on whether the speech stimulus is close to threshold, tested as speech-in-quiet, or supra-threshold, which we tested as speech-in-noise. We showed that diagnostic procedures to record changes in auditory processing at the start of the stimulus, such as pDPOAE, ABR peak amplitudes and latencies, phoneme discrimination ability, and ASSR can contribute to speech discrimination, independent of the otherwise dominating factors of age and hearing threshold. Factors contributing to comprehension in quiet were identified as the state of the cochlear amplifier and the cochlear synaptopathy of high-SR fibres, which affects the synchronicity of ANF signals at stimulus onset. The effect of reduced cochlear amplifier function in noise appeared to be counteracted partially by the recruitment phenomenon, reducing discrimination deficits. This suggests that nerve adaptation rate at stimulus onset and recruitment phenomena should also be included as part of the remaining 50% as factors that affect speech comprehension in addition to PTTs. These mechanisms seem to depend on the PLL as well, with different coding principles taking over below or above the PLL. Overall, the understanding of individual coding mechanisms of speech comprehension is vital to improving clinical routine measurements that are necessary to diagnose early hearing impairment or improve its treatment in the future.

# 5 Summary

## 5.1 English Summary

As part of demographic population changes, rising prevalence of hearing-loss is becoming a more prominent challenge of our healthcare in the current year. On the other hand, hearing-loss makes up the largest modifiable risk factor for developing dementia, which further underlines the need for early diagnosis and treatment of hearing loss. Developing hearing-loss, usually age-related or acquired, is often preceded by speech comprehension deficits, especially in background noise, which are usually not noticed or acted upon until much later. The aim of this study was to analyse the impact of different objective measures at the auditory pathway on the developing comprehension deficit. A cohort of 89 participants ranging from 18 to 76 years of age and grouped into young, middle-aged and older individuals with and without perceived hearing loss (but without diagnosed hearing loss) underwent a series of clinical and experimental auditory measurements comprised of pure-tone (PT) audiometry, short pulsed distortion-product otoacoustic emissions (pDPOAE), auditory brainstem responses (ABR), auditory steady-state responses (ASSR), speech comprehension testing (OLSA), and syllable discrimination in quiet and noise. Furthermore, questionnaires to collect data depression, cognitive function and their subjective hearing assessment were used. Hearing thresholds decreased most significantly in the extended high-frequency region (11.2-16 kHz), which also influenced the ABR-wave amplitudes at lower frequencies. To account for hearing-threshold differences when comparing comprehension levels, the OLSA thresholds were normalized for PT thresholds. Here we observed differences in speech comprehension in both noise and quiet conditions, and at all ages tested. More specifically, participants grouped for their relatively poor speech comprehension in the quiet condition showed comparatively lower pDPOAEs, indicating a reduced cochlear amplifier performance, as well as relatively lower amplitude and delayed ABR waves and also a worse performance in vowel-phoneme discrimination below the phase-locking limit (/o/-/u/). Inversely, Participants grouped as poor speech comprehension in the noise condition showed larger pDPOAEs, indicating better cochlear amplifier performance, as well as larger ASSR amplitudes, higher uncomfortable loudness levels and a lower performance of the vowel-phoneme discrimination of syllables above the phase-locking limit (/i/-/y/). This led to

the overall conclusion that impaired basilar membrane compression is associated with reduced speech comprehension in noise due to the effects on envelope coding. In contradiction to previous assumptions, these results lead to the conclusion that both good and poor speech comprehension can exist independently of differences in PTTs and age. This observation should be further studied with improved techniques and evaluated for the implementation in clinical diagnostic routine to differentiate hearing disabilities.

## 5.2 German summary

Die zunehmede Prävalenz von Hörverlust stellt eine wachsende Herausforderung des demografischen Wandels dar. Hierbei fallen häufig Sprachverstehensdefizite, insbesondere im Störgeräusch, auf, bevor ein manifester Hörschwellenverlust nachweisbar wird. Meist sind altersbedingte oder erworbene Hörstörungen Ursache der Hörminderung. Ziel dieser Studie war es, den Einfluss verschiedener Abschnitte der auditorischen Bahn auf die Entwicklung von Sprachverstehensdefiziten zu analysieren.

Die Kohorte von 89 Teilnehmern im Alter von 18 bis 76 Jahren wurde in junge, mittelalte und ältere Personen, mit und ohne subjektiv wahrgenommenen (aber nicht diagnostizierten) Hörverlust eingeteilt. Die Teilnehmer durchliefen eine Reihe audiologischer Untersuchungen, bestehend aus Tonaudiometrie, gepulsten Distorsionsprodukten otoakustischer Emissionen (pDPOAEs), akustisch evozierten Hirnstammpotenzialen (ABR), auditiven steady-state-responses (ASSR), Sprachverstehen (OLSA) sowie Silbenunterscheidung in Ruhe und im Störgeräusch.

Die Hörschwellen nahmen am deutlichsten im erweiterten Hochfrequenzbereich (11,2–16 kHz) ab, was sich auch auf die Amplituden der ABR-Wellen bei niedrigeren Frequenzen auswirkte. Um Unterschiede der Sprachverstehensleistung unabhängig von Hörschwellenunterschieden und Alterseffekten zu untersuchen, wurde die OLSA-Schwelle normalisiert berechnet, jeweils in Ruhe und im Störgeräusch. Dabei zeigten sich Unterschiede im Sprachverstehen sowohl in Ruhe als auch im Störgeräusch – und zwar in allen untersuchten Altersgruppen.

Konkret zeigten Teilnehmende mit schlechtem Sprachverstehen in Ruhe vergleichsweise geringere pDPOAE-Werte, was auf eine reduzierte Funktion des cochleären Verstärkers hinweist, sowie mittels ABR niedrigere Amplituden und verzögerte Wellen. Zudem schnitten sie schlechter bei der Unterscheidung von Vokalphonemen unterhalb des Phase-Locking-Limits (/o/-/u/) ab.

Umgekehrt zeigten Teilnehmende mit schlechtem Sprachverstehen im Störgeräusch höhere pDPOAE-Werte – ein Hinweis auf eine bessere Funktion des cochleären Verstärkers – sowie größere ASSR-Amplituden, höhere unangenehme Lautheitspegel und schlechtere Leistungen bei der Unterscheidung von Vokalphonemen oberhalb der Phasen-Locking-Grenze (/i/-/y/).

Dies führte zur Gesamtinterpretation, dass eine gestörte basiliäre Membrankompression mit reduziertem Sprachverstehen im Störgeräusch assoziiert ist – verursacht durch beeinträchtigte Hüllkurvencodierung. Deutlich wurde dabei – entgegen bisheriger Annahmen –, dass sowohl gutes als auch schlechtes Sprachverstehen unabhängig von Hörschwellen (PTTs) und Alter auftreten kann. Dieses Phänomen macht deutlich, dass dringend verbesserte diagnostische Verfahren zur Erfassung der Schallverarbeitung beim Reizbeginn in der klinischen Routine erforderlich sind.

# 6 References


ABDALA, C., ORTMANN, A. J. & GUARDIA, Y. C. 2021. Weakened Cochlear Nonlinearity During Human Aging and Perceptual Correlates. *Ear Hear,* 42**,** 832-845.

AVAN, P., BÜKI, B. & PETIT, C. 2013. Auditory distortions: origins and functions. *Physiol Rev,* 93**,** 1563-619.

BADER, K., DIERKES, L., BRAUN, L. H., GUMMER, A. W., DALHOFF, E. & ZELLE, D. 2021. Test-retest reliability of distortion-product thresholds compared to behavioral auditory thresholds. *Hear Res,* 406**,** 108232.

BATRA, R., KUWADA, S. & MAHER, V. L. 1986. The frequency-following response to continuous tones in humans. *Hearing Research,* 21**,** 167-177.

BECK, A. T., STEER, R. A. & BROWN, G. K. 1996. *BDI-II, Beck depression inventory : manual,* San Antonio, Tex., Boston, Psychological Corp. ; Harcourt Brace.

BHARADWAJ, H. M., VERHULST, S., SHAHEEN, L., LIBERMAN, M. C. & SHINN-CUNNINGHAM, B. G. 2014. Cochlear neuropathy and the coding of supra-threshold sound. *Frontiers in Systems Neuroscience,* Volume 8 - 2014.

BOEGE, P. & JANSSEN, T. 2002. Pure-tone threshold estimation from extrapolated distortion product otoacoustic emission I/O-functions in normal and cochlear hearing loss ears. *J Acoust Soc Am,* 111**,** 1810-8.

BOETTCHER, F. A., POTH, E. A., MILLS, J. H. & DUBNO, J. R. 2001. The amplitude-modulation following response in young and aged human subjects. *Hear Res,* 153**,** 32-42.

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**,** 1025-39.

BRAMHALL, N. F. 2021. Use of the auditory brainstem response for assessment of cochlear synaptopathy in humans. *J Acoust Soc Am,* 150**,** 4440.

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**,** 2801-10.

BRANDES, R., LANG, F. & SCHMIDT, R. F. 2019. *Physiologie des Menschen,* Berlin, Heidelberg, Springer.

BRUGHERA, A., DUNAI, L. & HARTMANN, W. M. 2013. Human interaural time difference thresholds for sine tones: the high-frequency limit. *J Acoust Soc Am,* 133**,** 2839-55.

BURAN, B. N., STRENZKE, N., NEEF, A., GUNDELFINGER, E. D., MOSER, T. & LIBERMAN, M. C. 2010. Onset coding is degraded in auditory nerve fibers from mutant mice lacking synaptic ribbons. *J Neurosci,* 30**,** 7587-97.

CARNEY, L. H. 2018. Supra-Threshold Hearing and Fluctuation Profiles: Implications for Sensorineural and Hidden Hearing Loss. *J Assoc Res Otolaryngol,* 19**,** 331-352.

CHEN, D. S., BETZ, J., YAFFE, K., AYONAYON, H. N., KRITCHEVSKY, S., MARTIN, K. R., HARRIS, T. B., PURCHASE-HELZNER, E.,

SATTERFIELD, S., XUE, Q. L., PRATT, S., SIMONSICK, E. M. & LIN, F. R. 2015. Association of hearing impairment with declines in physical functioning and the risk of disability in older adults. *J Gerontol A Biol Sci Med Sci,* 70**,** 654-61.

CHIMENTO, T. C. & SCHREINER, C. E. 1991. Adaptation and recovery from adaptation in single fiber responses of the cat auditory nerve. *J Acoust Soc Am,* 90**,** 263-73.

CRUICKSHANKS, K. J., WILEY, T. L., TWEED, T. S., KLEIN, B. E., KLEIN, R., MARES-PERLMAN, J. A. & NONDAHL, D. M. 1998. Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin. The Epidemiology of Hearing Loss Study. *Am J Epidemiol,* 148**,** 879-86.

DALLOS, P. 2008. Cochlear amplification, outer hair cells and prestin. *Curr Opin Neurobiol,* 18**,** 370-6.

DANIEL, E. 2007. Noise and hearing loss: a review. *J Sch Health,* 77**,** 225-31.

DAPPER, K., WOLPERT, S. M., SCHIRMER, J., FINK, S., GAUDRAIN, E., BAŞKENT, D., SINGER, W., VERHULST, S., BRAUN, C., DALHOFF, E., RÜTTIGER, L., MUNK, M. H. J. & KNIPPER, M. 2025. Age dependent deficits in speech recognition in quiet and noise are reflected in MGB activity and cochlear onset coding. *Neuroimage,* 305**,** 120958.

DENES, P. & NAUNTON, R. F. 1950. The clinical detection of auditory recruitment. *J Laryngol Otol,* 64**,** 375-98.

DUBNO, J. R., ECKERT, M. A., LEE, F. S., MATTHEWS, L. J. & SCHMIEDT, R. A. 2013. Classifying human audiometric phenotypes of age-related hearing loss from animal models. *J Assoc Res Otolaryngol,* 14**,** 687-701.

ENGELIEN, A., SCHULZ, M., ROSS, B., AROLT, V. & PANTEV, C. 2000. A combined functional in vivo measure for primary and secondary auditory cortices. *Hear Res,* 148**,** 153-60.

EYBALIN, M. 1993. Neurotransmitters and neuromodulators of the mammalian cochlea. *Physiol Rev,* 73**,** 309-73.

FERRITE, S. & SANTANA, V. 2005. Joint effects of smoking, noise exposure and age on hearing loss. *Occup Med (Lond),* 55**,** 48-53.

FLOCK, A., FLOCK, B., FRIDBERGER, A., SCARFONE, E. & ULFENDAHL, M. 1999. Supporting cells contribute to control of hearing sensitivity. *J Neurosci,* 19**,** 4498-507.

FOLSTEIN, M. F., FOLSTEIN, S. E. & MCHUGH, P. R. 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res,* 12**,** 189-98.

FOWLER, E. P. 1936. A METHOD FOR THE EARLY DETECTION OF OTOSCLEROSIS: A STUDY OF SOUNDS WELL ABOVE THRESHOLD. *Archives of Otolaryngology,* 24**,** 731-741.

FRATIGLIONI, L., WANG, H. X., ERICSSON, K., MAYTAN, M. & WINBLAD, B. 2000. Influence of social network on occurrence of dementia: a community-based longitudinal study. *Lancet,* 355**,** 1315-9.

FÜLLGRABE, 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.

FURMAN, A. C., KUJAWA, S. G. & LIBERMAN, M. C. 2013. Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. *J Neurophysiol,* 110**,** 577-86.

GARRETT, M., VASILKOV, V., MAUERMANN, M., DEVOLDER, P., WILSON, J. L., GONZALES, L., HENRY, K. S. & VERHULST, S. 2024. Deciphering compromised speech-in-noise intelligibility in older listeners: the role of cochlear synaptopathy. *bioRxiv***,** 2020.06.09.142950.

GOLDSTEIN, J. L. 1967. Auditory nonlinearity. *J Acoust Soc Am,* 41**,** 676-89.

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.

GOMAN, A. M. & LIN, F. R. 2016. Prevalence of Hearing Loss by Severity in the United States. *Am J Public Health,* 106**,** 1820-2.

GOMAN, A. M., REED, N. S. & LIN, F. R. 2017. Addressing Estimated Hearing Loss in Adults in 2060. *JAMA Otolaryngol Head Neck Surg,* 143**,** 733-734.

GORDON-SALANT, S. 2005. Hearing loss and aging: new research findings and clinical implications. *J Rehabil Res Dev,* 42**,** 9-24.

GOUTMAN, J. D. 2017. Mechanisms of synaptic depression at the hair cell ribbon synapse that support auditory nerve function. *Proc Natl Acad Sci U S A,* 114**,** 9719-9724.

GRIFFITHS, T. D., LAD, M., KUMAR, S., HOLMES, E., MCMURRAY, B., MAGUIRE, E. A., BILLIG, A. J. & SEDLEY, W. 2020. How Can Hearing Loss Cause Dementia? *Neuron,* 108**,** 401-412.

GROSE, J. H., MAMO, S. K. & HALL, J. W., 3RD 2009. Age effects in temporal envelope processing: speech unmasking and auditory steady state responses. *Ear Hear,* 30**,** 568-75.

HABIB, S. H. & HABIB, S. S. 2021. Auditory brainstem response: An overview of neurophysiological implications and clinical applications -A Narrative Review. *J Pak Med Assoc,* 71**,** 2230-2236.

HAN, C., DING, D., LOPEZ, M. C., MANOHAR, S., ZHANG, Y., KIM, M. J., PARK, H. J., WHITE, K., KIM, Y. H., LINSER, P., TANOKURA, M., LEEUWENBURGH, C., BAKER, H. V., SALVI, R. J. & SOMEYA, S. 2016. Effects of Long-Term Exercise on Age-Related Hearing Loss in Mice. *J Neurosci,* 36**,** 11308-11319.

HARRIS, K. C., VADEN, K. I., JR., MCCLASKEY, C. M., DIAS, J. W. & DUBNO, J. R. 2018. Complementary metrics of human auditory nerve function derived from compound action potentials. *J Neurophysiol,* 119**,** 1019-1028.

HASHIMOTO, I., ISHIYAMA, Y., YOSHIMOTO, T. & NEMOTO, S. 1981. Brain-stem auditory-evoked potentials recorded directly from human brain-stem and thalamus. *Brain,* 104**,** 841-59.

HEFFNER, H. & HEFFNER, R. 2008. High-Frequency Hearing. *Handbook of the senses: Audition,* 3**,** 55-60.

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.

HENDERSON, D., BIELEFELD, E. C., HARRIS, K. C. & HU, B. H. 2006. The role of oxidative stress in noise-induced hearing loss. *Ear Hear,* 27**,** 1-19.

HOFMEIER, B., WERTZ, J., REFAT, F., HINRICHS, P., SAEMISCH, J., SINGER, W., RÜTTIGER, L., KLOSE, U., KNIPPER, M. & WOLPERT, S. 2021. Functional biomarkers that distinguish between tinnitus with and without hyperacusis. *Clin Transl Med,* 11**,** e378.

HOFMEIER, B., WOLPERT, S., ALDAMER, E. S., WALTER, M., THIERICKE, J., BRAUN, C., ZELLE, D., RÜTTIGER, L., KLOSE, U. & KNIPPER, M. 2018. Reduced sound-evoked and resting-state BOLD fMRI connectivity in tinnitus. *Neuroimage Clin,* 20**,** 637-649.

HOPPE, U., HOCKE, T. & IRO, H. 2022. Age-Related Decline of Speech Perception. *Front Aging Neurosci,* 14**,** 891202.

HORNICKEL, J., SKOE, E., NICOL, T., ZECKER, S. & KRAUS, N. 2009. Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proc Natl Acad Sci U S A,* 106**,** 13022-7.

HOTH, S. & NEUMANN, K. 2006. Die diagnostische Aussagekraft der otoakustischen Emissionen. *Praktische Arbeitsmedizin, im Druck,* 6.

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.

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.

IMAM, L. & HANNAN, S. A. 2017. Noise-induced hearing loss: a modern epidemic? *Br J Hosp Med (Lond),* 78**,** 286-290.

JANSSEN, R., SCHWEITZER, L. & JENSEN, K. F. 1991. Glutamate neurotoxicity in the developing rat cochlea: physiological and morphological approaches. *Brain Res,* 552**,** 255-64.

JEDRZEJCZAK, W. W., KOCHANEK, K., PILKA, E. & SKARZYNSKI, H. 2016. Spontaneous otoacoustic emissions in schoolchildren. *International Journal of Pediatric Otorhinolaryngology,* 89**,** 67-71.

JEFFERSON, A., POWERS, D. & POPE, M. 2001. Beck Depression Inventory-II and the Geriatric Depression Scale (GDS) in older women. *Clinical Gerontologist,* 22**,** 3-12.

JOHANNESEN, P. T., PÉREZ-GONZÁLEZ, 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.

JOHNSON, D. H. & KIANG, N. Y. 1976. Analysis of discharges recorded simultaneously from pairs of auditory nerve fibers. *Biophysical Journal,* 16**,** 719-734.

KEMP, D. T. 1978. Stimulated acoustic emissions from within the human auditory system. *J Acoust Soc Am,* 64**,** 1386-91.

KIANG, N. Y., WATANABE, T., THOMAS, E. C. & CLARK, L. F. 1966. *Discharge patterns of single fibers in the cat's auditory nerve,* Oxford, England, M.I.T. Press.
KIM, D. O., DORN, P. A., NEELY, S. T. & GORGA, M. P. 2001. Adaptation of Distortion Product Otoacoustic Emission in Humans. *Journal of the Association for Research in Otolaryngology,* 2**,** 31-40.
KIMBERLEY, B. P., HERNADI, I., LEE, A. M. & BROWN, D. K. 1994. Predicting Pure Tone Thresholds in Normal and Hearing-Impaired Ears with Distortion Product Emission and Age. *Ear and Hearing,* 15.
KLINKE, R. 1986. Neurotransmission in the inner ear. *Hearing Research,* 22**,** 235-243.
KOPKE, R., BIELEFELD, E., LIU, J., ZHENG, J., JACKSON, R., HENDERSON, D. & COLEMAN, J. K. 2005. Prevention of impulse noise-induced hearing loss with antioxidants. *Acta Otolaryngol,* 125**,** 235-43.
KROKENBERGER, M. 2019. *Adaptive DPOAE-Wachstumsfunktionen zur objektiven Hörschwellenschätzung bei normalhörenden und hörgeschädigten Ohren.* Universität Tübingen.
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. 2006. Acceleration of age-related hearing loss by early noise exposure: evidence of a misspent youth. *J Neurosci,* 26**,** 2115-23.
KUJAWA, S. G. & LIBERMAN, M. C. 2009. Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss. *J Neurosci,* 29**,** 14077-85.
KURODA, T. 2007. Clinical investigation on spontaneous otoacoustic emission (SOAE) in 447 ears. *Auris Nasus Larynx,* 34**,** 29-38.
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.
LE PRELL, C., BLEDSOE, S. C., BOBBIN, R. P. & PUEL, J.-L. 2001. Neurotransmission in the inner ear: Functional and molecular analyses.
LEHNHARDT, E. & LASZIG, R. 2009. *Praxis der Audiometrie,* Stuttgart, Thieme.
LIBERMAN, M. C. 1978. Auditory-nerve response from cats raised in a low-noise chamber. *J Acoust Soc Am,* 63**,** 442-55.
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., GAO, J., HE, D. Z., WU, X., JIA, S. & ZUO, J. 2002. Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier. *Nature,* 419**,** 300-4.
LIBERMAN, M. C. & KUJAWA, S. G. 2017. Cochlear synaptopathy in acquired sensorineural hearing loss: Manifestations and mechanisms. *Hear Res,* 349**,** 138-147.
LIBERMAN, M. C. & MULROY, M. J. 1982. Acute and chronic effects of acoustic trauma: cochlear pathology and auditory nerve pathophysiology. *In:* HAMERNIK, R. P., HENDERSON, D. & SALVI, R. (eds.) *New Perspectives on Noise-Induced Hearing Loss.* New York: Raven Press.

LIN, F. R. & ALBERT, M. 2014. Hearing loss and dementia - who is listening? *Aging Ment Health,* 18**,** 671-3.
LIN, F. R., FERRUCCI, L., AN, Y., GOH, J. O., DOSHI, J., METTER, E. J., DAVATZIKOS, C., KRAUT, M. A. & RESNICK, S. M. 2014. Association of hearing impairment with brain volume changes in older adults. *Neuroimage,* 90**,** 84-92.
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., SELBÆK, G., TERI, L. & MUKADAM, N. 2017. Dementia prevention, intervention, and care. *Lancet,* 390**,** 2673-2734.
LOBARINAS, E., SALVI, R. & DING, D. 2013. Insensitivity of the audiogram to carboplatin induced inner hair cell loss in chinchillas. *Hear Res,* 302**,** 113-20.
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**,** 18866-9.
LU, H., MEHTA, A. H. & OXENHAM, A. J. 2022. Methodological considerations when measuring and analyzing auditory steady-state responses with multi-channel EEG. *Current Research in Neurobiology,* 3**,** 100061.
MAKARY, C. A., SHIN, J., KUJAWA, S. G., LIBERMAN, M. C. & MERCHANT, S. N. 2011. Age-related primary cochlear neuronal degeneration in human temporal bones. *J Assoc Res Otolaryngol,* 12**,** 711-7.
MÄRCHER-RØRSTED, J., ENCINA-LLAMAS, G., DAU, T., LIBERMAN, M. C., WU, P. Z. & HJORTKJÆR, J. 2022. Age-related reduction in frequency-following responses as a potential marker of cochlear neural degeneration. *Hear Res,* 414**,** 108411.
MARRUFO-PÉREZ, M. I. & LOPEZ-POVEDA, E. A. 2022. Adaptation to noise in normal and impaired hearing. *J Acoust Soc Am,* 151**,** 1741.
MEDDIS, R. 2006. Auditory-nerve first-spike latency and auditory absolute threshold: a computer model. *J Acoust Soc Am,* 119**,** 406-17.
MELCHER, J. R. & KIANG, N. Y. S. 1996. Generators of the brainstem auditory evoked potential in cat III: identified cell populations. *Hearing Research,* 93**,** 52-71.
MÖHRLE, D., NI, K., VARAKINA, K., BING, D., LEE, S. C., ZIMMERMANN, U., KNIPPER, M. & RÜTTIGER, L. 2016. Loss of auditory sensitivity from inner hair cell synaptopathy can be centrally compensated in the young but not old brain. *Neurobiol Aging,* 44**,** 173-184.
MØLLER, A. R. & JANNETTA, P. J. 1982. Comparison between intracranially recorded potentials from the human auditory nerve and scalp recorded auditory brainstem responses (ABR). *Scand Audiol,* 11**,** 33-40.
MØLLER, A. R., JANNETTA, P. J. & JHO, H. D. 1994. Click-evoked responses from the cochlear nucleus: a study in human. *Electroencephalogr Clin Neurophysiol,* 92**,** 215-24.
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.

MONTERO-ODASSO, M., ISMAIL, Z. & LIVINGSTON, G. 2020. One third of dementia cases can be prevented within the next 25 years by tackling risk factors. The case "for" and "against". *Alzheimers Res Ther,* 12**,** 81.
MORISE, M., YOKOMORI, F. & OZAWA, K. 2016. WORLD: A Vocoder-Based High-Quality Speech Synthesis System for Real-Time Applications. *IEICE Transactions on Information and Systems,* E99**,** 1877-1884.
MOŚCICKI, E. K., ELKINS, E. F., BAUM, H. M. & MCNAMARA, P. M. 1985. Hearing loss in the elderly: an epidemiologic study of the Framingham Heart Study Cohort. *Ear Hear,* 6**,** 184-90.
MOSER, T. & BEUTNER, D. 2000. Kinetics of exocytosis and endocytosis at the cochlear inner hair cell afferent synapse of the mouse. *Proc Natl Acad Sci U S A,* 97**,** 883-8.
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.
MULROY, M. J. & WHALEY, E. A. 1984. Structural changes in auditory hairs during temporary deafness. *Scan Electron Microsc***,** 831-40.
NORDMANN, A. S., BOHNE, B. A. & HARDING, G. W. 2000. Histopathological differences between temporary and permanent threshold shift. *Hear Res,* 139**,** 13-30.
OHINATA, Y., YAMASOBA, T., SCHACHT, J. & MILLER, J. M. 2000. Glutathione limits noise-induced hearing loss. *Hear Res,* 146**,** 28-34.
OHLEMILLER, K. K., MCFADDEN, S. L., DING, D. L., FLOOD, D. G., REAUME, A. G., HOFFMAN, E. K., SCOTT, R. W., WRIGHT, J. S., PUTCHA, G. V. & SALVI, R. J. 1999. Targeted deletion of the cytosolic Cu/Zn-superoxide dismutase gene (Sod1) increases susceptibility to noise-induced hearing loss. *Audiol Neurootol,* 4**,** 237-46.
OTTE, J., SCHUNKNECHT, H. F. & KERR, A. G. 1978. Ganglion cell populations in normal and pathological human cochleae. Implications for cochlear implantation. *Laryngoscope,* 88**,** 1231-46.
OXENHAM, A. J. 2016. Predicting the Perceptual Consequences of Hidden Hearing Loss. *Trends Hear,* 20**,** 2331216516686768.
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.
PASSCHIER-VERMEER, W. 1974. Hearing loss due to continuous exposure to steady-state broad-band noise. *J Acoust Soc Am,* 56**,** 1585-93.
PETERSON, A. J., HUET, A., BOURIEN, J., PUEL, J. L. & HEIL, P. 2018. Recovery of auditory-nerve-fiber spike amplitude under natural excitation conditions. *Hear Res,* 370**,** 248-263.
PLACK, C. J., BARKER, D. & PRENDERGAST, G. 2014. Perceptual consequences of "hidden" hearing loss. *Trends Hear,* 18.
PORTMANN, M., CAZALS, Y., NEGREVERGNE, M. & ARAN, J. M. 1980. Transtympanic and surface recordings in the diagnosis of retrocochlear disorders. *Acta Otolaryngol,* 89**,** 362-9.

PROBST, R., LONSBURY-MARTIN, B. L. & MARTIN, G. K. 1991. A review of otoacoustic emissions. *J Acoust Soc Am,* 89**,** 2027-67.
PUEL, J.-L. 1995. Chemical synaptic transmission in the cochlea. *Progress in Neurobiology,* 47**,** 449-476.
PUEL, J. L., PUJOL, R., TRIBILLAC, F., LADRECH, S. & EYBALIN, M. 1994. Excitatory amino acid antagonists protect cochlear auditory neurons from excitotoxicity. *J Comp Neurol,* 341**,** 241-56.
PURCELL, D. W., JOHN, S. M., SCHNEIDER, B. A. & PICTON, T. W. 2004. Human temporal auditory acuity as assessed by envelope following responses. *J Acoust Soc Am,* 116**,** 3581-93.
RAMADAN, H. H. & SCHUKNECHT, H. F. 1989. Is there a conductive type of presbycusis? *Otolaryngol Head Neck Surg,* 100**,** 30-4.
RANCE, G. 2008. *The Auditory Steady-State Response: Generation, Recording, and Clinical Application*, Plural Publishing, Incorporated.
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**,** 2803-17.
RHODE, W. S. & SMITH, P. H. 1986. Encoding timing and intensity in the ventral cochlear nucleus of the cat. *J Neurophysiol,* 56**,** 261-86.
ROBERTSON, D. 1983. Functional significance of dendritic swelling after loud sounds in the guinea pig cochlea. *Hear Res,* 9**,** 263-78.
RUMSCHLAG, J. A. & RAZAK, K. A. 2021. Age-related changes in event related potentials, steady state responses and temporal processing in the auditory cortex of mice with severe or mild hearing loss. *Hear Res,* 412**,** 108380.
SCHAETTE, R. & MCALPINE, D. 2011. Tinnitus with a normal audiogram: physiological evidence for hidden hearing loss and computational model. *J Neurosci,* 31**,** 13452-7.
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. *Journal of Clinical Medicine* [Online], 13.
SCHMIEDT, R. A., LANG, H., OKAMURA, H. O. & SCHULTE, B. A. 2002. Effects of furosemide applied chronically to the round window: a model of metabolic presbyacusis. *J Neurosci,* 22**,** 9643-50.
SCHMIEDT, R. A., MILLS, J. H. & BOETTCHER, F. A. 1996. Age-related loss of activity of auditory-nerve fibers. *J Neurophysiol,* 76**,** 2799-803.
SCHUKNECHT, H. F. 1993. *Pathology of the Ear,* Baltimore, Lea & Febiger.
SCHUKNECHT, H. F. & WOELLNER, R. C. 1955. An experimental and clinical study of deafness from lesions of the cochlear nerve. *J Laryngol Otol,* 69**,** 75-97.
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**,** 13686-94.

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**,** 727-45.
SHANKS, J. & SHOHET, J. 2009. Tympanometry in clinical practice. *Handbook of Clinical Audiology***,** 157-188.
SHANKS, J. E. 1984. Tympanometry. *Ear and Hearing,* 5.
SHEARER, A. E., HILDEBRAND, M. S., ODELL, A. M. & SMITH, R. J. H. 1993. Genetic Hearing Loss Overview. *In:* ADAM, M. P., FELDMAN, J., MIRZAA, G. M., PAGON, R. A., WALLACE, S. E. & AMEMIYA, A. (eds.) *GeneReviews(®).* Seattle (WA): University of Washington, Seattle
Copyright © 1993-2025, University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle. All rights reserved.
SHEHABI, A. M., PRENDERGAST, G. & PLACK, C. J. 2022. The Relative and Combined Effects of Noise Exposure and Aging on Auditory Peripheral Neural Deafferentation: A Narrative Review. *Front Aging Neurosci,* 14**,** 877588.
SIEGEL, J. H. & KIM, D. O. 1982. Efferent neural control of cochlear mechanics? Olivocochlear bundle stimulation affects cochlear biomechanical nonlinearity. *Hear Res,* 6**,** 171-82.
SPOENDLIN, H. 1971. Primary structural changes in the organ of Corti after acoustic overstimulation. *Acta Otolaryngol,* 71**,** 166-76.
SPOENDLIN, H. 1972. Innervation densities of the cochlea. *Acta Otolaryngol,* 73**,** 235-48.
SPOENDLIN, H. & SCHROTT, A. 1989. Analysis of the human auditory nerve. *Hear Res,* 43**,** 25-38.
STEVENS, G., FLAXMAN, S., BRUNSKILL, E., MASCARENHAS, M., MATHERS, C. D. & FINUCANE, M. 2013. Global and regional hearing impairment prevalence: an analysis of 42 studies in 29 countries. *Eur J Public Health,* 23**,** 146-52.
SUTHAKAR, K. & LIBERMAN, M. C. 2022. Noise Masking in Cochlear Synaptopathy: Auditory Brainstem Response vs. Auditory Nerve Response in Mouse. *J Neurophysiol,* 127**,** 1574-85.
TAN, H. E., LAN, N. S. R., KNUIMAN, M. W., DIVITINI, M. L., SWANEPOEL, D. W., HUNTER, M., BRENNAN-JONES, C. G., HUNG, J., EIKELBOOM, R. H. & SANTA MARIA, P. L. 2018. Associations between cardiovascular disease and its risk factors with hearing loss-A cross-sectional analysis. *Clin Otolaryngol,* 43**,** 172-181.
TERKILDSEIN, K. & THOMSEN, K. A. 1959. The influence of pressure variations on the impedance of the human ear drum. A method for objective determination of the middle-ear pressure. *J Laryngol Otol,* 73**,** 409-18.
TUN, P. A., MCCOY, S. & WINGFIELD, A. 2009. Aging, hearing acuity, and the attentional costs of effortful listening. *Psychol Aging,* 24**,** 761-6.
VANDE MAELE, T., KESHISHZADEH, S., DE POORTERE, N., DHOOGE, I., KEPPLER & VERHULST, S. 2021. *The variability in potential biomarkers for cochlear synaptopathy after recreational noise exposure*.
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. *Hearing Research,* 377, 109-121.
VETEŠNÍK, A., TURCANU, D., DALHOFF, E. & GUMMER, A. W. 2009. Extraction of sources of distortion product otoacoustic emissions by onset-decomposition. *Hearing Research,* 256, 21-38.
WAGENER, K., BRAND, T. & KOLLMEIER, B. 1999a. Development and evaluation of a German sentence test I: Design of the Oldenburg sentence test. 38, 4-15.
WAGENER, K., BRAND, T. & KOLLMEIER, B. 1999b. Development and evaluation of a German sentence test III: Evaluation of the Oldenburg sentence test (in German). *Z Audiol,* 38, 86-95.
WAGENER, K., BRAND, T. & KOLLMEIER, B. 1999c. Development and evaluation of a German sentence test. Part II: Optimization of the Oldenburg sentence test. *Z Audiol/Audiological Acoustics,* 38, 44-56.
WAGENER, K., KUEHNEL, V. & KOLLMEIER, B. 1999d. Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests. *Zeitschrift für Audiologie,* 38, 4–15.
WANG, Y., HIROSE, K. & LIBERMAN, M. C. 2002. Dynamics of noise-induced cellular injury and repair in the mouse cochlea. *J Assoc Res Otolaryngol,* 3, 248-68.
WARDENGA, N., BATSOULIS, C., WAGENER, K. C., BRAND, T., LENARZ, T. & MAIER, H. 2015. Do you hear the noise? The German matrix sentence test with a fixed noise level in subjects with normal hearing and hearing impairment. *Int J Audiol,* 54 Suppl 2, 71-9.
WHITE, J. A., BURGESS, B. J., HALL, R. D. & NADOL, J. B. 2000. Pattern of degeneration of the spiral ganglion cell and its processes in the C57BL/6J mouse. *Hear Res,* 141, 12-8.
WHITEHEAD, M. L., STAGNER, B. B., MARTIN, G. K. & LONSBURY-MARTIN, B. L. 1996. Visualization of the onset of distortion-product otoacoustic emissions, and measurement of their latency. *J Acoust Soc Am,* 100, 1663-79.
WILLMORE, B. D. B. & KING, A. J. 2023. Adaptation in auditory processing. *Physiol Rev,* 103, 1025-1058.
WINTER, I. M., ROBERTSON, D. & YATES, G. K. 1990. Diversity of characteristic frequency rate-intensity functions in guinea pig auditory nerve fibres. *Hear Res,* 45, 191-202.
WON, J. H., TREMBLAY, K., CLINARD, C. G., WRIGHT, R. A., SAGI, E. & SVIRSKY, M. 2016. The neural encoding of formant frequencies contributing to vowel identification in normal-hearing listeners. *J Acoust Soc Am,* 139, 1-11.
WORLD HEALTH ORGANIZATION 1991. Report of the Informal Working Group on Prevention of Deafness and Hearing Impairment: Programme Planning. Geneva, Switzerland: World Health Organization.
WORLD HEALTH ORGANIZATION 2018. Addressing the rising prevalence of hearing loss.
WORLD HEALTH ORGANIZATION 2021. World report on hearing. World Health Organization.

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.
WU, P. Z. & LIBERMAN, M. C. 2022. Age-related stereocilia pathology in the human cochlea. *Hear Res,* 422**,** 108551.
WU, P. Z., O'MALLEY, J. T., DE GRUTTOLA, V. & LIBERMAN, M. C. 2020. Age-Related Hearing Loss Is Dominated by Damage to Inner Ear Sensory Cells, Not the Cellular Battery That Powers Them. *J Neurosci,* 40**,** 6357-6366.
YAMANE, H., NAKAI, Y., TAKAYAMA, M., IGUCHI, H., NAKAGAWA, T. & KOJIMA, A. 1995a. Appearance of free radicals in the guinea pig inner ear after noise-induced acoustic trauma. *Eur Arch Otorhinolaryngol,* 252**,** 504-8.
YAMANE, H., NAKAI, Y., TAKAYAMA, M., KONISHI, K., IGUCHI, H., NAKAGAWA, T., SHIBATA, S., KATO, A., SUNAMI, K. & KAWAKATSU, C. 1995b. The emergence of free radicals after acoustic trauma and strial blood flow. *Acta Otolaryngol Suppl,* 519**,** 87-92.
YESAVAGE, J. A., BRINK, T. L., ROSE, T. L., LUM, O., HUANG, V., ADEY, M. & LEIRER, V. O. 1982. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res,* 17**,** 37-49.
YOUNG, E. & SACHS, M. B. 1973. Recovery from sound exposure in auditory-nerve fibers. *J Acoust Soc Am,* 54**,** 1535-43.
ZAZO SECO, C., WESDORP, M., FEENSTRA, I., PFUNDT, R., HEHIR-KWA, J. Y., LELIEVELD, S. H., CASTELEIN, S., GILISSEN, C., DE WIJS, I. J., ADMIRAAL, R. J., PENNINGS, R. J., KUNST, H. P., VAN DE KAMP, J. M., TAMMINGA, S., HOUWELING, A. C., PLOMP, A. S., MAAS, S. M., DE KONING GANS, P. A., KANT, S. G., DE GEUS, C. M., FRINTS, S. G., VANHOUTTE, E. K., VAN DOOREN, M. F., VAN DEN BOOGAARD, M. H., SCHEFFER, H., NELEN, M., KREMER, H., HOEFSLOOT, L., SCHRADERS, M. & YNTEMA, H. G. 2017. The diagnostic yield of whole-exome sequencing targeting a gene panel for hearing impairment in The Netherlands. *Eur J Hum Genet,* 25**,** 308-314.
ZELLE, D., BADER, K., DIERKES, L., GUMMER, A. W. & DALHOFF, E. 2020. Derivation of input-output functions from distortion-product otoacoustic emission level maps. *The Journal of the Acoustical Society of America,* 147**,** 3169-3187.
ZELLE, D., DALHOFF, E. & GUMMER, A. W. 2017a. Objective audiometry with DPOAEs : New findings for generation mechanisms and clinical applications. *Hno,* 65**,** 122-129.
ZELLE, D., GUMMER, A. W. & DALHOFF, E. 2013. Extraction of otoacoustic distortion product sources using pulse basis functions. *J Acoust Soc Am,* 134**,** El64-9.
ZELLE, D., LORENZ, L., THIERICKE, J. P., GUMMER, A. W. & DALHOFF, E. 2017b. Input-output functions of the nonlinear-distortion component of distortion-product otoacoustic emissions in normal and hearing-impaired human ears. *J Acoust Soc Am,* 141**,** 3203.
ZELLE, D., THIERICKE, J. P., DALHOFF, E. & GUMMER, A. W. 2015. Level dependence of the nonlinear-distortion component of distortion-product otoacoustic emissions in humans. *J Acoust Soc Am,* 138**,** 3475-90.

ZHAO, Y., YAMOAH, E. N. & GILLESPIE, P. G. 1996. Regeneration of broken tip links and restoration of mechanical transduction in hair cells. *Proc Natl Acad Sci U S A,* 93**,** 15469-74.
ZHENG, J., SHEN, W., HE, D. Z., LONG, K. B., MADISON, L. D. & DALLOS, P. 2000. Prestin is the motor protein of cochlear outer hair cells. *Nature,* 405**,** 149-55.
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. *The Journal of Neuroscience,* 31**,** 9192.

# 7 Declaration of contribution (in German)

Die Arbeit wurde in der HNO-Klinik Tübingen unter Betreuung von Prof. Dr. Marlies Knipper durchgeführt.

Die Konzeption der Studie erfolgte durch das CoSy-speech Konsortium.

Die Messungen der Seesion 1 und Session 2 wurden nach Einarbeitung durch Labormitglieder Konrad Dapper und Moritz Rühle von mir eigenständig und vollständig durchgeführt. Lediglich die Messung der Session 1 wurde von Proband CS001 bis CS038 von Moritz Rühle durchgeführt.

Auswertung: Die Auswertung erfolgte durch mich, mit Unterstützung durch Prof. Dr. Marlies Knipper, Prof. Dr. Lukas Rüttiger, Dr. med. Stephan Wolpert, PD Dr. med. Matthias Munk, Dr. Ernst Dalhoff, Konrad Dapper und Jakob Wertz.

Die statistische Auswertung erfolgte selbständig in Betreuung durch Prof. Lukas Rüttiger.

Ich versichere, das Manuskript selbständig verfasst zu haben und keine weiteren als die von mir angegebenen Quellen verwendet zu haben.

# 8 Publications

Schirmer, J.; Wolpert, S.; Dapper, K.; Rühle, M.; Wertz, J.; Wouters, M.; Eldh, T.; Bader, K.; Singer, W.; Gaudrain, E.; et al. 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. 2024, 13, 2725. https://doi.org/10.3390/jcm13092725

Dapper K, Wolpert SM, Schirmer J, Fink S, Gaudrain E, Başkent D, Singer W, Verhulst S, Braun C, Dalhoff E, Rüttiger L, Munk MHJ, Knipper M. Age dependent deficits in speech recognition in quiet and noise are reflected in MGB activity and cochlear onset coding. Neuroimage. 2025 Jan; 305:120958. doi: 10.1016/j.neuroimage.2024.120958. Epub 2024 Nov 30. PMID: 39622462.

# 9 Appendix

## **Appendix A**: Recruitment email for university staff & students

**Date:** 03/31/2022 [02:31:26 PM CEST]
**From:** [redacted]
**To:** [redacted]
**Subject:** Die HNO sucht Probanden mit leichter Hörstörung (4 Sessions – 100€)

```
************************************************************************
* Die Universitätsleitung hat dem Versand dieser Rundmail zugestimmt. *
************************************************************************
Liebe Studien-interessierte,

wir suchen gesunde Versuchspersonen mit leichter Hörstörung oder
Sprachverständnisschwierigkeiten im Alter von 18 bis 30 Jahren für eine
Hörstudie zu cochleären Synaptopathien im Tübinger Hörforschungszentrum.

Wir erforschen den Zusammenhang des Sprachverständnisses mit Besonderheiten in
der Empfindlichkeit des Gehörs. Unser Ziel ist es in Zukunft
Sprachverständnisstörungen mit Hilfe individueller Anpassung der Versorgung
besser therapieren zu können.
Teilnehmende erhalten eine Aufwandentschädigung von 100€.
Die Studie umfasst 4 Sitzungen mit jeweils ca. 2 Stunden:
1) Ohruntersuchung und Hörtest
2) Sprachverständnistest
3) Messung der elektrischen Gehirnaktivität beim Hören (ABR)
4) Messung der magnetischen Gehirnaktivität (MEG)
Alle personenbezogenen Angaben werden anonymisiert, nur zu diesem
Forschungszweck verwendet und nicht an Dritte weitergegeben.
Auf die Freiwilligkeit der Teilnahme wird ausdrücklich hingewiesen.
Vielen Dank für Ihre Unterstützung !

Bei Interesse an der Teilnahme und Fragen zum Studienablauf
kontaktieren sie uns bitte unter:
Jakob Schirmer (cand. med.), Konrad Dapper (cand. Dr. rer. nat.)
Email: [redacted]
Tel: [redacted]

Wir freuen uns über ihre Teilnahme!
Jakob Schirmer und Konrad Dapper

-----------------------------------
Studienleiterin:
Prof. Dr. Marlies Knipper
Hearing Research Center Tübingen
Universitätsklinik für Hals-, Nasen- und Ohrenheilkunde
Elfriede Aulhornstr. 5
72076 Tübingen
Tel: [redacted]
E-mail: [redacted]
```

**Appendix B**: Checklist and Custom Questionnaire

**Cosy Studie**

**Einschlusskriterien**

| | |
|---|---|
| Kontrollgruppe (18-25J ohne Hörschaden) | |
| Fallgruppe (18-25J mit Hörschaden) | |
| Kontrollgruppe (45-52J ohne Hörschaden) | |
| Fallgruppe (45-52J mit Hörschaden) | |
| Kontrollgruppe (53-75J ohne Hörschaden) | |
| Fallgruppe (53-75J mit Hörschaden) | |

Betroffenes Ohr (zutreffendes bitte ankreuzen): rechts // links // beide

**Ausschlusskriterien**

Hörsystem

| | ja | nein |
|---|---|---|
| Tinnitus oder Hyperakusis | | |
| Retrocochleäre Hörstörung (Akustikusneurinom) | | |
| Vertigo/Schwindel | | |
| Z.n Knalltrauma | | |
| Morbus Menier/ endolymphatischer Hydrops | | |
| Ein oder beidseitige Ertaubung | | |
| Zustand nach jahrelanger Lärmexposition | | |
| Schallleitungsschwerhörigkeit mit einer Schallleitungskomponente von mehr als 15 dB in einer oder mehr Frequenzen | | |
| Chronische Gehörgangs /Mittelohrentzündung | | |

Allgemeine Krankengeschichte

| | ja | nein |
|---|---|---|
| Schädelhirntrauma (Grad II oder III) | | |
| Herz-Kreislauferkrankungen | | |
| Diabetes | | |
| Nierenerkrankungen | | |
| Behandlung von Krebsleiden (Leukämie) | | |
| ototox. Medikamente (Schleifendiuretika, Aminoglykoside, Chemotherapie) | | |
| Behandlung von neurologischen oder psychiatrischen Erkrankungen (auch | | |

| | | |
|---|---|---|
| medikamentös: Neuroleptika, Haloperidol, L-Dopa) | | |
| Alkohol, Drogen | | |
| Immunsuppressiva Einnahme | | |
| Anamnestisch Epilepsieleiden, Parkinson | | |
| Dementielle Erkrankungen | | |
| Anamnestisch eingeschränktes Temperaturempfinden | | |
| Nicht Einwilligungsfähigkeit des Probanden | | |

| | ja | nein |
|---|---|---|
| Schwangerschaft | | |

Therapie Hörsystem

| | ja | nein |
|---|---|---|
| Hörgeräteversorgung | | |
| Ohroperationen | | |

# Fragebogen Hörvermögen

**1. Wie lautet Ihr Geburtsjahr?**

**2. Was ist Ihr Geschlecht?**
**a) Mann**
**b) Frau**
**c) x**

**3. Was ist der höchste Abschluss, den Sie erworben haben?**
**Grundschulbildung (Primarschule oder kein Schulabschluss)**
**Realschulabschluss**
**Gymnasialschulabschluss/ allgemeine Fachhochschulreife**
**Bachelor-Studiengang**
**Masterstudiengang**
**Staatsexamen**

**4. Haben Sie derzeit das Gefühl auf den Ohren Druck zu verspüren**
**Nein**
**Ja**

**5. Rauchen Sie oder haben Sie geraucht?**
**Früher habe ich geraucht, jetzt nicht mehr**
**Ich rauche derzeit und bin regelmäßiger Raucher (täglich)**
**Ich rauche derzeit und bin Gelegenheitsraucher (weniger als täglich)**
**Ich habe nie geraucht**

**6. Wie gut empfinden Sie Ihr Gehör?**
**Ich höre ausgezeichnet (Übergehen zu Frage 8)**
**Ich höre sehr gut (Übergehen zu Frage 8)**
**Ich höre gut (Übergehen zu Frage 8)**
**Ich höre vernünftige**
**Ich höre schlecht**

**7. Auf welchem Ohr hören Sie weniger gut?**
**Auf meinem rechten Ohr**
**Auf meinem linken Ohr**
**Auf beiden Ohren**
**Ich weiß es nicht**

**8. Fällt es Ihnen schwer, einem Gespräch zu folgen, wenn es Hintergrundgeräusche gibt, z. B. Fernsehen**
**Ja**
**Nein**

**9. Können Sie in einem Gespräch mit mehreren Personen klar und deutlich hören, was gesagt wird?**
**Ja**
**Nein**

**10. Können Sie deutlich hören, was in einem Gespräch von einer anderen Person gesagt wird?**
**Ja**
**Nein**

**11. Kreuzen Sie für jede Aussage nur 1 Kästchen an. Hatten Sie jemals Schwierigkeiten alles gut verstehen?**

**A) In einer ruhigen Umgebung mit höchstens 2 anderen Personen?**

**Immer Oft Manchmal Selten Nie**

**B) In einer ruhigen Umgebung mit vielen anderen Menschen?**

**Immer Oft Manchmal Selten Nie**

**C) In lauten Situationen**

**Immer Oft Manchmal Selten Nie**

**12. Sind Sie jetzt oder in der Vergangenheit regelmäßig Lärm ausgesetzt?**

**Ja**

**Nein**

Notiz:

Datum, Ort | Unterschrift

**Appendix C** Cohort information of each participant included in the study with anonymized ID, age, gender (f/m), handedness and results from depression evaluation according to Beck Depression Inventory II (BDI-II), Geriatric Depression Scale (GDS) as well as dementia test Mini Mental Status Test (MMST) and Tympanometry results.

| **ID** | **Age (y)** | **Gender** | **Handed-ness** | **BDI*** | **GDS*** | **MMST** | **Tympano-metry** |
|---|---|---|---|---|---|---|---|
| CS001 | 26 | f | right | minimal | minimal | 30 | Type A |
| CS002 | 63 | f | right | minimal | minimal | 30 | Type A |
| CS003 | 59 | m | right | minimal | minimal | 30 | Type A |
| CS004 | 62 | m | right | minimal | minimal | 30 | Type A |
| CS005 | 55 | f | right | minimal | minimal | 30 | Type A |
| CS006 | 60 | f | right | mild | minimal | 30 | Type A |
| CS007 | 38 | f | right | minimal | minimal | 30 | Type A |
| CS008 | 48 | f | right | moderate | mild/moderate | 30 | Type B |
| CS009 | 21 | f | right | minimal | minimal | 30 | Type A |
| CS010 | 47 | f | right | minimal | minimal | 30 | Type A |
| CS011 | 19 | f | right | minimal | minimal | 30 | Type C |
| CS012 | 22 | m | right | minimal | minimal | 29 | Type A |
| CS013 | 26 | f | right | minimal | minimal | 30 | Type A |
| CS014 | 25 | m | right | minimal | minimal | 30 | Type A |
| CS015 | 25 | m | right | mild | minimal | 30 | Type A |
| CS018 | 24 | m | right | minimal | minimal | 29 | Type A |
| CS023 | 69 | m | right | moderate | mild/moderate | 29 | Type A |
| CS025 | 64 | f | right | minimal | minimal | 28 | Type A |
| CS027 | 60 | f | right | minimal | minimal | 30 | Type A |
| CS029 | 64 | m | right | minimal | minimal | 29 | Type A |
| CS030 | 23 | f | right | minimal | minimal | 30 | Type A |
| CS031 | 53 | m | right | minimal | minimal | 29 | Type A |
| CS033 | 26 | f | right | minimal | minimal | 29 | Type A |
| CS034 | 24 | f | right | minimal | minimal | 29 | Type A |
| CS035 | 45 | m | right | minimal | minimal | 28 | Type A |
| CS037 | 45 | m | right | minimal | minimal | 29 | Type A |
| CS038 | 57 | f | right | minimal | minimal | 30 | Type A |
| CS039 | 18 | f | right | minimal | minimal | 29 | Type A |
| CS040 | 57 | f | right | minimal | minimal | 30 | Type A |
| CS041 | 58 | f | right | minimal | minimal | 30 | Type A |
| CS042 | 54 | f | right | mild | minimal | 29 | Type AS |
| CS044 | 62 | f | right | minimal | minimal | 30 | Type AS |
| CS047 | 55 | m | right | minimal | minimal | 29 | Type A |
| CS050 | 62 | f | right | minimal | minimal | 29 | Type A |
| CS051 | 29 | m | left | mild | minimal | 30 | Type AS |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| CS053 | 19 | m | right | moderate | mild/moderate | 30 | Type A |
| CS056 | 28 | f | right | minimal | minimal | 30 | Type A |
| CS057 | 28 | f | right | minimal | minimal | 29 | Type AD |
| CS058 | 21 | f | right | minimal | minimal | 30 | Type A |
| CS059 | 29 | m | left | minimal | minimal | 29 | Type A |
| CS060 | 24 | f | right | minimal | minimal | 29 | Type AD |
| CS061 | 58 | f | right | minimal | minimal | 30 | Type A |
| CS062 | 52 | f | right | minimal | minimal | 30 | Type A |
| CS063 | 28 | m | right | minimal | minimal | 30 | Type A |
| CS065 | 23 | f | right | minimal | minimal | 29 | Type A |
| CS066 | 70 | m | right | minimal | minimal | 30 | Type AD |
| CS067 | 33 | m | right | mild | minimal | 30 | Type AD |
| CS068 | 52 | f | right | minimal | minimal | 30 | Type A |
| CS069 | 24 | f | right | minimal | minimal | 30 | Type A |
| CS070 | 20 | f | left | mild | mild/moderate | 28 | Type A |
| CS071 | 60 | f | right | minimal | minimal | 30 | Type A |
| CS073 | 43 | f | right | minimal | minimal | 29 | Type A |
| CS074 | 21 | m | right | moderate | mild/moderate | 29 | Type A |
| CS075 | 54 | f | right | minimal | minimal | 30 | Type A |
| CS076 | 30 | m | right | minimal | minimal | 30 | Type A |
| CS077 | 52 | f | right | minimal | minimal | 30 | Type A |
| CS078 | 35 | m | right | minimal | minimal | 30 | Type A |
| CS079 | 34 | f | right | minimal | minimal | 28 | Type A |
| CS080 | 35 | m | left | minimal | minimal | 30 | Type A |
| CS081 | 54 | f | right | minimal | minimal | 30 | Type A |
| CS082 | 48 | f | right | minimal | minimal | 29 | Type AD |
| CS083 | 76 | m | right | minimal | minimal | 29 | Type A |
| CS084 | 54 | f | right | minimal | minimal | 30 | Type A |
| CS085 | 40 | f | right | minimal | minimal | 28 | Type A |
| CS086 | 56 | f | right | minimal | minimal | 30 | Type AD |
| CS087 | 32 | f | right | minimal | minimal | 29 | Type A |
| CS088 | 30 | f | right | minimal | minimal | 29 | Type A |
| CS089 | 53 | f | right | minimal | minimal | 30 | Type A |
| CS090 | 56 | f | right | minimal | minimal | 30 | Type A |
| CS091 | 30 | f | right | minimal | minimal | 30 | Type A |
| CS093 | 40 | m | right | minimal | minimal | 30 | Type A |
| CS095 | 50 | m | right | minimal | minimal | 29 | Type A |
| CS096 | 51 | f | right | minimal | minimal | 30 | Type AS |
| CS097 | 40 | f | right | minimal | minimal | 30 | Type A |
| CS098 | 54 | f | right | minimal | minimal | 30 | Type A |
| CS099 | 27 | f | right | minimal | minimal | 29 | Type A |
| CS100 | 25 | m | right | minimal | minimal | 30 | Type A |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| CS101 | 46 | f | right | minimal | minimal | 30 | Type A |
| CS102 | 21 | m | right | mild | minimal | 30 | Type A |
| CS103 | 28 | f | right | minimal | minimal | 28 | Type A |
| CS104 | 68 | f | right | minimal | minimal | 30 | Type AS |
| CS105 | 65 | m | left | minimal | minimal | 30 | Type A |
| CS106 | 23 | f | left | minimal | minimal | 30 | Type A |
| CS107 | 64 | m | right | minimal | minimal | 30 | Type B |
| CS108 | 70 | f | right | minimal | minimal | 29 | Type A |
| CS109 | 70 | w | right | minimal | minimal | 29 | Type A |
| CS110 | 72 | m | right | minimal | minimal | 29 | Type AS |
| CS111 | 62 | m | right | minimal | minimal | 30 | Type A |
| CS112 | 71 | f | right | minimal | minimal | 29 | Type As |

## **Appendix D**: BDI-II

Patientenname:

Datum:

UNIVERSITÄTS KLINIKUM TÜBINGEN

Hertie-Institut für klinische Hirnforschung

DZNE Deutsches Zentrum für Neurodegenerative Erkrankungen in der Helmholtz-Gemeinschaft

### BECK DEPRESSIONSFRAGEBOGEN - II

**Anleitung:** Dieser Fragebogen besteht aus 21 Gruppen von Aussagen. Bitte lesen Sie jede dieser Gruppen von Aussagen sorgfaltig durch und suchen Sie sich dann in jeder Gruppe **eine Aussage** aus, die am besten beschreibt, wie Sie sich **in den letzten zwei Wochen, einschließlich heute,** gefühlt haben. Kreuzen Sie die Zahl neben der Aussage an, die Sie sich herausgesucht haben. Wenn in einer Gruppe mehrere Aussagen gleichermaßen auf Sie zutreffen, kreuzen Sie die Aussage mit der höheren Zahl an. Bitte achten Sie darauf, dass Sie in jeder Gruppe nicht mehr als eine Aussage ankreuzen. Das gilt auch für Gruppe 16 (Veränderungen der Schlafgewohnheiten) oder Gruppe 18 (Veränderungen des Appetits).

**1. Traurigkeit**
- 0 Ich bin nicht traurig.
- 1 Ich bin oft traurig.
- 2 Ich bin ständig traurig.
- 3 Ich bin so traurig oder unglücklich, dass ich es nicht aushalten kann.

**2. Pessimismus**
- 0 Ich bin nicht mutlos, was meine Zukunft angeht.
- 1 Ich bin mutloser als früher, was meine Zukunft angeht.
- 2 Ich glaube nicht, dass sich meine Lage verbessert.
- 3 Ich habe das Gefühl, dass es keine Hoffnung gibt für meine Zukunft und es nur noch schlimmer wird.

**3. Frühere Misserfolge**
- 0 Ich fühle mich nicht als Versager.
- 1 Ich habe öfter versagt als ich sollte.
- 2 Wenn ich zurückblicke, sehe ich eine Menge Misserfolge.
- 3 Ich fühle mich persönlich als totaler Versager.

**4. Verlust von Freude**
- 0 Ich habe so viel Freude wie immer an den Dingen, die mir Spaß machen.
- 1 Ich habe nicht mehr so viel Spaß an den Dingen wie früher.
- 2 Ich habe sehr wenig Freude an den Dingen, die mir früher Spaß gemacht haben.
- 3 Ich habe keine Freude an den Dingen, die mir früher Spaß gemacht haben.

**5. Schuldgefühle**
- 0 Ich habe keine besonderen Schuldgefühle.
- 1 Ich habe bei vielen Dingen, die ich getan habe oder hätte tun sollen, Schuldgefühle.
- 2 Ich habe die meiste Zeit Schuldgefühle.
- 3 Ich habe ständig Schuldgefühle.

**6. Gefühle, bestraft zu werden**
- 0 Ich habe nicht das Gefühl, für etwas bestraft zu werden.
- 1 Ich habe das Gefühl, dass ich vielleicht für etwas bestraft werde.
- 2 Ich glaube, dass ich für etwas bestraft werde.
- 3 Ich habe das Gefühl, für etwas bestraft zu werden.

**7. Abneigung gegen sich selbst**
- 0 Meine Gefühle mir gegenüber sind die gleichen geblieben.
- 1 Ich habe das Vertrauen in mich verloren.
- 2 Ich bin von mir selbst enttäuscht.
- 3 Ich mag mich nicht.

**8. Selbstvorwürfe**
- 0 Ich bin mir selbst gegenüber nicht kritischer als sonst und mache mir nicht mehr Vorwürfe als sonst.
- 1 Ich bin mir selbst gegenüber kritischer als früher.
- 2 Ich mache mir Vorwürfe für alle meine Fehler.
- 3 Ich gebe mir die Schuld für alles Schlimme, was passiert.

**9. Selbstmordgedanken oder- wünsche**
- 0 Ich denke nie daran, mich umzubringen.
- 1 Ich habe Selbstmordgedanken, aber ich würde sie nicht ausführen.
- 2 Ich möchte mich umbringen.
- 3 Ich würde mich umbringen, wenn ich die Möglichkeit hätte.

**10. Weinen**
- 0 Ich weine nicht mehr als früher.
- 1 Ich weine mehr als früher.
- 2 Ich weine wegen jeder Kleinigkeit.
- 3 Mir ist nach Weinen zumute, aber ich kann nicht.

## BECK DEPRESSIONSINVENTAR- II

**11. Unruhe**
0 Ich bin nicht unruhiger oder erregter als sonst.
1 Ich bin unruhiger oder erregter als sonst.
2 Ich bin so unruhig oder erregt, dass es schwer ist, mich nicht zu bewegen.
3 Ich bin so unruhig oder erregt, dass ich ständig in Bewegung bleiben oder etwas tun muss.

**12. Interesselosigkeit**
0 Ich habe das Interesse an anderen Menschen oder an Tätigkeiten nicht verloren.
1 Ich bin weniger an anderen Menschen oder Dingen interessiert als vorher.
2 Ich habe mein Interesse an anderen Menschen oder Dingen zum größten Teil verloren.
3 Es ist schwer, für irgendetwas Interesse aufzubringen.

**13. Entschlussunfähigkeit**
0 Ich treffe Entscheidungen etwa so leicht wie immer.
1 Es fällt mir schwerer als sonst, Entscheidungen zu treffen.
2 Ich habe viel größere Schwierigkeiten, Entscheidungen zu treffen, als früher.
3 Ich habe Mühe, überhaupt Entscheidungen zu treffen.

**14. Wertlosigkeit**
0 Ich fühle mich nicht wertlos.
1 Ich halte mich nicht für so wertvoll und nützlich wie früher.
2 Ich habe das Gefühl, weniger wert zu sein als andere Menschen.
3 Ich habe das Gefühl, völlig wertlos zu sein.

**15. Verlust an Energie**
0 Ich habe so viel Energie wie immer.
1 Ich habe weniger Energie als früher.
2 Ich habe nicht genügend Energie, sehr viel zu tun.
3 Ich habe nicht genügend Energie, irgendetwas zu tun.

**16. Veränderungen der Schlafgewohnheiten**
0 Meine Schlafgewohnheiten haben sich nicht geändert.
1a Ich schlafe etwas mehr als sonst.
1b Ich schlafe etwas weniger als sonst.
2a Ich schlafe viel mehr als sonst.
2b Ich schlafe viel weniger als sonst.
3a Ich schlafe die meiste Zeit des Tages.
3b Ich wache 1-2 Stunden zu früh auf und kann dann nicht mehr einschlafen.

**17. Reizbarkeit**
0 Ich bin nicht reizbarer als sonst.
1 Ich bin reizbarer als sonst.
2 Ich bin viel reizbarer als sonst.
3 Ich bin ständig reizbar.

**18. Veränderungen des Appetits**
0 Mein Appetit hat sich nicht verändert.
1a Mein Appetit ist etwas kleiner als sonst.
1b Mein Appetit ist etwas größer als sonst.
2a Mein Appetit ist viel kleiner als vorher.
2b Mein Appetit ist viel größer als vorher.
3a Ich habe überhaupt keinen Appetit.
3b Ich habe ständig großen Hunger.

**19. Konzentrationsschwierigkeiten**
0 Ich kann mich so gut konzentrieren wie immer.
1 Ich kann mich nicht so gut konzentrieren wie sonst.
2 Es fällt mir schwer, mich sehr lange auf etwas zu konzentrieren.
3 Ich kann mich auf gar nichts konzentrieren.

**20. Müdigkeit**
0 Ich bin nicht müder als sonst.
1 Ich werde schneller müde als sonst.
2 Ich bin für viele Dinge, die ich früher getan habe, zu müde.
3 Ich bin für die meisten Dinge, die ich früher getan habe, zu müde.

**21. Verlust des Interesses am Sex**
0 Ich habe in letzter Zeit keine Veränderung meines Interesses am Sex bemerkt.
1 Ich habe weniger Interesse am Sex als früher.
2 Ich habe jetzt viel weniger Interesse am Sex.
3 Ich habe das Interesse am Sex völlig verloren.

**Appendix E**: GDS

Universitätsklinikum Tübingen

Geriatrische Depressionsskala nach Yesavage et al. (1983)

| Nr. | Frage | JA | NEIN |
|---|---|---|---|
| 1. | Sind Sie grundsätzlich mit Ihrem Leben zufrieden? | | |
| 2. | Haben Sie viele Ihrer Aktivitäten und Interessen aufgegeben? | | |
| 3. | Haben Sie das Gefühl, Ihr Leben sei unausgefüllt? | | |
| 4. | Ist Ihnen oft langweilig? | | |
| 5. | Sind Sie die meiste Zeit guter Laune? | | |
| 6. | Haben Sie Angst, dass Ihnen etwas Schlimmes zustoßen wird? | | |
| 7. | Fühlen Sie sich die meiste Zeit glücklich? | | |
| 8. | Fühlen Sie sich oft hilflos? | | |
| 9. | Bleiben Sie lieber zuhause, anstatt auszugehen und Neues zu unternehmen? | | |
| 10. | Glauben Sie, mehr Probleme mit dem Gedächtnis zu haben als die meisten anderen? | | |
| 11. | Finden Sie, es sei schön, jetzt zu leben? | | |
| 12. | Kommen Sie sich in Ihrem jetzigen Zustand ziemlich wertlos vor? | | |
| 13. | Fühlen Sie sich voller Energie? | | |
| 14. | Finden Sie, dass Ihre Situation hoffnungslos ist? | | |
| 15. | Glauben Sie, dass es den meisten Leuten besser geht als Ihnen? | | |

Auswertung: Für Antwort NEIN auf die Fragen 1, 5, 7, 11, 13 sowie für Antwort JA auf die übrigen Fragen gibt es je einen Punkt

| | | |
|---|---|---|
| | Summe max. 15 Pkte. | |

Mehr als 5 Pkte.: Es besteht möglicherweise eine Depression
➔ Weitergehende Diagnostik erforderlich

**Appendix F**: MMSE

## Mini-Mental-Status-Test (MMST)

(modifiziert nach Folstein, Folstein & McHugh)

**Zeitliche Orientierung**

*(Frage: „Welchen Tag haben wir heute?")*

- Tag ........................................ ☐
- Monat ........................................ ☐
- Jahr ........................................ ☐
- Wochentag ........................................ ☐
- Jahreszeit ........................................ ☐

**Örtliche Orientierung**

*(Frage: „Wo sind wir jetzt?")*

- Stadt ........................................ ☐
- Stadtteil ........................................ ☐
- Bundesland ........................................ ☐
- Klinik/Pflegeheim/Praxis ........................................ ☐
- Station/Stockwerk ........................................ ☐

**Merkfähigkeit**

*(Folgende 3 Gegenstände nennen, dann zur Wiederholung auffordern)*

- Apfel ........................................ ☐
- Schlüssel ........................................ ☐
- Ball ........................................ ☐

**Aufmerksamkeit und Rechnen**

*(Jeweils 7 von 100 abziehen oder „STUHL" rückwärts buchstabieren)*

- 93 oder „L" ........................................ ☐
- 86 oder „H" ........................................ ☐
- 79 oder „U" ........................................ ☐
- 72 oder „T" ........................................ ☐
- 65 oder „S" ........................................ ☐

**Erinnern**

*(Frage: „Was waren die Dinge, die Sie sich vorhin gemerkt haben?")*

- Apfel ........................................ ☐
- Schlüssel ........................................ ☐
- Ball ........................................ ☐

**Benennen**

*(Die Testperson soll die folgenden zwei Gegenstände benennen)*

- Armbanduhr ........................................ ☐
- Bleistift/Kugelschreiber ........................................ ☐

**Wiederholen**

*(Die Testperson soll den folgenden Satz nachsprechen; nur ein Versuch ist erlaubt; die Redewendung „Kein wenn und aber" ist nicht erlaubt.)*

- „Kein wenn und oder aber" ........................................ ☐

**Dreiteiliger Befehl**

*(„Nehmen Sie das Blatt Papier, falten es in der Mitte und lassen es auf den Boden fallen")*

- „Nehmen Sie das Blatt Papier," ........................................ ☐
- „falten es in der Mitte" ........................................ ☐
- „und lassen es auf den Boden fallen" ........................................ ☐

**Reagieren**

*(Die Testperson soll den Satz: „Schließen Sie die Augen" lesen und befolgen)*

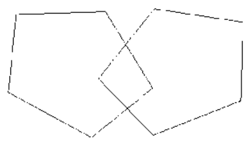
- Testperson schließt die Augen ........................................ ☐

Mini-Mental-Status-Test (Teil 1)

**Schreiben**
*(Die Testperson soll einen beliebigen vollständigen Satz aufzuschreiben)*

- Sinnhafter Satz mit Subjekt und Verb............................................... ☐

**Abzeichnen**
*(Testperson soll die folgende Zeichnung abzeichnen)*



- Zwei sich an einer Ecke überschneidende 5-Ecke................................. ☐

**Gesamtpunktzahl.......................................................................... ___ (max. 30)**

**Interpretation des Testergebnisses**

| **Punkte** | **Beurteilung** |
|---|---|
| 30-27 | Keine Demenz |
| 26-20 | Leichte Demenz |
| 19-10 | Mittelschwere Demenz |
| ≤9 | Schwere Demenz |

**Appendix G** Post-hoc variance analysis of OLSA performance based on a linear-mixed model. *Adapted from Schirmer et al., 2024*

| | | | sign of correlation | | | Explained total Variance in % | | | Explain std dB | | | p value | | | Total std dB | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | N | bb | lp | hp | bb | lp | hp | bb | lp | hp | bb | lp | hp | bb | lp | hp |
| LEDPT-PTT | Quiet | 81 | 1 | 1 | 1 | 2.0 | 1.2 | 8.3 | 0.8 | 0.7 | 4.8 | 0.033 | 0.161 | 0.001 | 5.5 | 6.6 | 16.6 |
| | Ipsi | 56 | 1 | 1 | 1 | 0.1 | 7.0 | 0.7 | 0.0 | 0.3 | 0.5 | 0.783 | 0.037 | 0.519 | 1.1 | 1.3 | 6.0 |
| Number of DPOAE | Quiet | 88 | -1 | -1 | -1 | 2.0 | 1.8 | 3.3 | 0.8 | 0.9 | 3.2 | 0.037 | 0.105 | 0.019 | 5.9 | 6.7 | 17.8 |
| | Ipsi | 63 | 1 | 1 | 1 | 4.8 | 5.5 | 3.1 | 0.3 | 0.3 | 1.1 | 0.023 | 0.051 | 0.079 | 1.4 | 1.3 | 6.1 |
| ABR amp I | Quiet | 76 | -1 | -1 | 1 | 2.5 | 2.8 | 0.2 | 0.9 | 1.0 | 0.8 | 0.022 | 0.032 | 0.561 | 5.4 | 6.2 | 16.7 |
| | Ipsi | 56 | -1 | -1 | 1 | 0.3 | 0.3 | 0.5 | 0.1 | 0.1 | 0.4 | 0.634 | 0.671 | 0.525 | 1.5 | 1.3 | 6.0 |
| ABR lat I | Quiet | 80 | 1 | 1 | -1 | 2.2 | 4.6 | 0.1 | 0.8 | 1.3 | 0.5 | 0.027 | 0.004 | 0.727 | 5.6 | 6.3 | 16.5 |
| | Ipsi | 60 | -1 | 1 | -1 | 0.5 | 0.1 | 3.7 | 0.1 | 0.0 | 1.2 | 0.491 | 0.790 | 0.063 | 1.5 | 1.2 | 6.0 |
| ABR lat II | Quiet | 62 | 1 | 1 | 1 | 4.0 | 4.3 | 2.1 | 1.0 | 1.2 | 2.3 | 0.018 | 0.053 | 0.128 | 5.2 | 5.8 | 16.1 |
| | Ipsi | 47 | -1 | 1 | -1 | 0.3 | 0.2 | 3.6 | 0.1 | 0.1 | 1.1 | 0.675 | 0.745 | 0.110 | 1.6 | 1.2 | 5.9 |
| ABR lat III | Quiet | 84 | 1 | 1 | -1 | 0.1 | 1.1 | 0.0 | 0.2 | 0.7 | 0.3 | 0.612 | 0.211 | 0.811 | 5.7 | 6.7 | 17.1 |
| | Ipsi | 61 | 1 | -1 | -1 | 0.7 | 0.0 | 0.5 | 0.1 | 0.0 | 0.4 | 0.427 | 0.908 | 0.487 | 1.4 | 1.2 | 6.0 |
| ABR lat V | Quiet | 86 | 1 | 1 | 1 | 2.2 | 3.2 | 1.6 | 0.8 | 1.2 | 2.2 | 0.024 | 0.030 | 0.112 | 5.6 | 6.6 | 17.6 |
| | Ipsi | 62 | 1 | -1 | -1 | 0.0 | 0.3 | 0.1 | 0.0 | 0.1 | 0.2 | 0.858 | 0.667 | 0.720 | 1.4 | 1.3 | 6.2 |
| ABR lat VI | Quiet | 78 | 1 | 1 | 1 | 1.0 | 2.0 | 4.2 | 0.5 | 0.9 | 3.3 | 0.153 | 0.105 | 0.011 | 5.4 | 6.7 | 16.2 |
| | Ipsi | 58 | -1 | -1 | -1 | 0.1 | 0.4 | 0.3 | 0.1 | 0.1 | 0.3 | 0.726 | 0.613 | 0.650 | 1.4 | 1.3 | 5.9 |
| ASSR | Quiet | 74 | 1 | 1 | 1 | 0.4 | 0.1 | 0.9 | 0.4 | 0.2 | 1.6 | 0.387 | 0.749 | 0.273 | 5.8 | 6.8 | 16.3 |
| | Ipsi | 55 | 1 | 1 | 1 | 7.0 | 4.3 | 2.0 | 0.4 | 0.3 | 0.9 | 0.012 | 0.109 | 0.219 | 1.4 | 1.3 | 6.2 |
| ASSR 4k | Quiet | 65 | 1 | -1 | 1 | 0.0 | 0.7 | 0.5 | 0.1 | 0.6 | 1.1 | 0.875 | 0.336 | 0.447 | 5.9 | 6.9 | 15.1 |
| | Ipsi | 49 | 1 | 1 | -1 | 0.3 | 0.9 | 0.1 | 0.1 | 0.1 | 0.2 | 0.612 | 0.487 | 0.805 | 1.2 | 1.2 | 5.9 |
| ASSR 6k | Quiet | 68 | 1 | 1 | 1 | 1.2 | 0.5 | 0.2 | 0.6 | 0.5 | 0.7 | 0.150 | 0.407 | 0.632 | 5.2 | 6.7 | 17.7 |
| | Ipsi | 50 | 1 | 1 | 1 | 3.8 | 0.4 | 3.3 | 0.3 | 0.1 | 1.1 | 0.085 | 0.661 | 0.147 | 1.4 | 1.3 | 6.3 |
| di bi easy | Quiet | 73 | -1 | -1 | 1 | 2.3 | 1.7 | 0.2 | 0.9 | 0.9 | 0.8 | 0.038 | 0.146 | 0.603 | 6.1 | 7.0 | 18.2 |
| | Ipsi | 51 | -1 | 1 | -1 | 0.1 | 0.1 | 0.2 | 0.0 | 0.0 | 0.3 | 0.761 | 0.842 | 0.667 | 1.4 | 1.3 | 6.2 |
| ou difficult | Quiet | 88 | -1 | -1 | -1 | 2.0 | 0.0 | 0.2 | 0.9 | 0.1 | 0.7 | 0.035 | 0.863 | 0.603 | 6.0 | 6.7 | 17.8 |
| | Ipsi | 62 | -1 | -1 | -1 | 0.4 | 1.7 | 1.6 | 0.1 | 0.2 | 0.8 | 0.533 | 0.289 | 0.216 | 1.4 | 1.3 | 6.2 |
| ou easy | Quiet | 88 | -1 | -1 | -1 | 2.0 | 0.7 | 1.2 | 0.8 | 0.5 | 1.9 | 0.039 | 0.319 | 0.169 | 6.0 | 6.7 | 17.8 |
| | Ipsi | 62 | 1 | 1 | -1 | 0.1 | 0.0 | 0.7 | 0.0 | 0.0 | 0.5 | 0.765 | 0.899 | 0.406 | 1.4 | 1.3 | 6.2 |
| iy difficult | Quiet | 88 | -1 | -1 | -1 | 0.6 | 0.5 | 0.5 | 0.5 | 0.5 | 1.2 | 0.256 | 0.384 | 0.398 | 6.0 | 6.7 | 17.8 |
| | Ipsi | 62 | -1 | 1 | -1 | 4.1 | 0.1 | 2.6 | 0.3 | 0.0 | 1.0 | 0.041 | 0.748 | 0.106 | 1.4 | 1.3 | 6.2 |
| iy easy | Quiet | 88 | 1 | 1 | -1 | 0.0 | 0.1 | 0.5 | 0.1 | 0.2 | 1.3 | 0.892 | 0.664 | 0.359 | 6.0 | 6.7 | 17.8 |
| | Ipsi | 62 | -1 | 1 | -1 | 1.3 | 0.3 | 1.0 | 0.2 | 0.1 | 0.6 | 0.248 | 0.680 | 0.334 | 1.4 | 1.3 | 6.2 |
| UCL 250 | Quiet | 55 | -1 | 1 | 1 | 0.0 | 0.0 | 0.1 | 0.1 | 0.1 | 0.6 | 0.853 | 0.839 | 0.729 | 6.2 | 7.2 | 15.9 |
| UCL 250 | Ipsi | 42 | 1 | 1 | -1 | 0.0 | 10.1 | 0.7 | 0.0 | 0.4 | 0.5 | 0.969 | 0.024 | 0.497 | 1.5 | 1.2 | 6.2 |
| UCL 500 | Quiet | 69 | -1 | 1 | 1 | 0.0 | 0.0 | 1.5 | 0.1 | 0.1 | 2.1 | 0.863 | 0.839 | 0.182 | 6.2 | 7.2 | 17.4 |

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| UCL 500 | Ipsi | 52 | 1 | 1 | -1 | 0.3 | 7.6 | 0.0 | 0.1 | 0.4 | 0.1 | 0.619 | 0.036 | 0.906 | 1.5 | 1.3 | 6.5 |
| UCL 1000 | Quiet | 72 | -1 | -1 | 1 | 0.2 | 0.1 | 1.9 | 0.3 | 0.2 | 2.4 | 0.544 | 0.726 | 0.117 | 6.2 | 7.1 | 17.8 |
| UCL 1000 | Ipsi | 53 | 1 | 1 | -1 | 0.4 | 8.6 | 0.0 | 0.1 | 0.4 | 0.1 | 0.533 | 0.022 | 0.895 | 1.5 | 1.3 | 6.4 |
| UCL2000 | Quiet | 80 | -1 | -1 | 1 | 0.0 | 0.1 | 2.1 | 0.0 | 0.2 | 2.6 | 0.915 | 0.776 | 0.078 | 6.0 | 6.8 | 17.8 |
| UCL2000 | Ipsi | 56 | 1 | 1 | -1 | 0.1 | 4.4 | 0.2 | 0.1 | 0.3 | 0.2 | 0.718 | 0.099 | 0.696 | 1.4 | 1.3 | 6.3 |
| UCL4000 | Quiet | 77 | 1 | 1 | 1 | 0.1 | 0.2 | 1.2 | 0.2 | 0.3 | 1.9 | 0.692 | 0.644 | 0.209 | 5.9 | 6.9 | 17.4 |
| UCL4000 | Ipsi | 56 | 1 | 1 | -1 | 0.3 | 4.2 | 0.1 | 0.1 | 0.3 | 0.2 | 0.572 | 0.098 | 0.796 | 1.4 | 1.3 | 6.3 |
| UCL 6000 | Quiet | 42 | 1 | -1 | 1 | 0.1 | 0.5 | 1.0 | 0.2 | 0.6 | 1.9 | 0.782 | 0.616 | 0.417 | 7.1 | 8.4 | 19.0 |
| UCL 6000 | Ipsi | 25 | -1 | -1 | -1 | 1.2 | 1.4 | 0.0 | 0.2 | 0.2 | 0.0 | 0.455 | 0.515 | 0.947 | 1.4 | 1.4 | 5.9 |

# 10. Acknowledgements

I would like to thank Marlies Knipper, Lukas Rüttiger, Stephan Wolpert, Matthias Munk, Ernst Dalhoff, Jakob Werz, Konrad Dapper, Wibke Singer, Moritz Rühle, Therese Eldh, Kerstin Just, Antje Knock and everyone else from the AG Knipper lab who relentlessly supported the development of this dissertation and without whom this work would not have been possible.

I thank the entire CoSy-Speech consortium for enabling me to take part in this research.

Finally, I thank the IZKF-Promotionskolleg for providing valuable advice and means to support the writing of this dissertation and making the participation at IEB and DGHNO possible.