FACULTY OF MEDICINE
AND HEALTH SCIENCES

MASTER'S DISSERTATION PART 2

IMPROVED COCHLEAR SYNAPTOPATHY DIAGNOSTICS USING EEG MEASUREMENTS WITH FLUCTUATING SOUNDS

Word count: 10 259 words

Pauline Devolder
Student number: 01700175

Supervisor(s): Prof. Dr. Sarah Verhulst, Prof. Dr. Ingeborg Dhooge, Dr. Emmanuel Ponsot

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Master of Science in Speech Language and Hearing Sciences

Academic year: 2021 - 2022

GHENT
UNIVERSITY

Contents

Word of thanks

First of all I would like to thank my promoter Prof. Verhulst and co-promoters Prof. Dhooge and Dr. Ponsot for the guidance and constructive feedback during the last two years and Dr. Ponsot in particular for the elaboration of the protocol, the stimuli and MATLAB scripts for testing and analysis.

Also, a special thanks to Drs. Heleen Van der Biest for the practical guidance during the investigations, all administrative matters and the explanation of data analysis. In addition, thanks to Dr. Sarineh Keshishzadeh and Drs. Nele De Poortere for writing the analysis scripts and for helping interpret the data respectively.

I would also like to thank all test subjects for participating and audiologist Stephen Vlaeminck for his help in recruiting hearing impaired subjects.

Finally, thanks to my parents, sisters, brothers and close friends for the positive and motivating support during my studies and master's thesis.

Abbreviations

ABG: Air bone gap
ABR: Auditory brainstem response
AM: Amplitude modulation
ANF: Auditory nerve fiber
BB: Broad band
DP: Distortion product
ECV: Ear canal volume
EEG: Electroencephalogram
EFR: Envelope following response
ENV: Envelope
FFR: Frequency following response
FFT: Fast fourrier transformation
HFA: High frequency audiometry
HP: High pass
IHC: Inner hair cell
LP: Low pass
MEP: Middle ear pressure
MPS: Modulation power spectrum
OAE: Otoacoustic emissions
oHI: Older hearing impaired
oNH: Older normal hearing
PTA: Pure tone threshold
SAA: Statical acoustical admittance
SM: Spectral modulation
SNR: Signal to noise ratio
SPIN: Speech in noise
SPIQ: Speech in quiet
SPL: Sound pressure level
SRT: Speech reception threshold
STM: Spectrotemporal modulation
TFS: Temporal fine structure
TLA: Tonal liminary audiometry
yNH: Younger normal hearing

Abstract

Abstract (Ne)

Achtergrond: In de audiologische praktijk is er veel variabiliteit te zien in spraakverstaan tussen verschillende patiënten, ondanks hun gelijkaardige (normale of verhoogde) gehoordrempels. Het spraakverstaan is complex en kan beïnvloed worden door verschillende individuele factoren (gehoorverlies, cognitie en suprathreshold codering). Vooral personen van middelbare leeftijd hebben vaak klachten van spraakverstaan, voornamelijk in rumoerige omgeving. Cochleaire synaptopathie (schade aan de synapsen tussen inwendige haarcellen (IHC) in de cochlea en afferente zenuwen) zou een belangrijke rol spelen bij deze klacht. Recent toonden verschillende studies aan dat de envelope following response (EFR) een goede maatstaaf zou zijn voor cochleaire synaptopathie. Hiermee wordt nagegaan in hoeverre men de temporale enveloppe (ENV) van een spraaksignaal kan detecteren. Echter bestaan realistische spraaksignalen in het dagelijks leven niet enkel uit temporele modulaties, maar ook spectrale, waarbij de temporele fijnstructuur (TFS) een belangrijke rol speelt. Onderzoek met gedragsmatige psycho-akoestische testen toont aan dat spectrotemporeel gemoduleerde (STM) stimuli goede maatstaven zouden zijn voor spraakverstaanbaarheid. Ondanks de belovende resultaten, kunnen deze gedragsmatige testen beïnvloed worden door cognitieve factoren. Een manier om spectrale verwerking objectief op te meten, kan de frequency following response (FFR) zijn, meestal gemeten aan de hand van syllabes of klinkers in de huidige literatuur.

Doelstelling: Met dit onderzoek wilden we nagaan of een FFR met spectraal gemoduleerde stimuli (harmonischen) objectieve informatie kan bieden over de verwerking van TFS cues en of deze gerelateerd zijn aan spraakverstaan bij personen met normale en verhoogde gehoordrempels.

Methode: Twee groepen participanten van middelbare leeftijd (40 tot 60 jaar) werden geïncludeerd in de studie; een normaalhorende (noNH=16) en een slechthorende groep (noHI=14), gebaseerd op tonale audiometrie. Om na te gaan of de FFR een goed meetinstrument is voor spraakverstaan, werden resultaten vergeleken met de Vlaamse Matrix zinstest. Ook gedragsmatige S(T)M detectiedrempels werden opgemeten en vergeleken om na te gaan of deze dezelfde onderliggende mechanismes hebben. Daarnaast wilden we nagaan of de FFR beïnvloed wordt door neurosensorieel gehoorverlies gebaseerd op tonale audiometrie, oto-akoestische emissies (OAE), de EFR en de auditieve hersenstamrespons (ABR).

Resultaten: Zowel de S(T)M detectiedrempels als de FFR correleerden niet met spraakverstaan in ruis. De S(T)M detectiedrempels toonden ook geen significante verschillen tussen de normaalhorende en slechthorende proefgroep, waren niet gecorreleerd aan de auditieve gehoordrempels, maar wel aan verlies van uitwendige haarcellen (gebaseerd op OAEs). De FFR resultaten werden opgesplitst in TFS en ENV cues. Ondanks de laagfrequente stimulus, toonde de FFR_Env correlaties met hoogfrequent gehoorverlies en cochleaire synaptopathie, terwijl FFR_Tfs gecorreleerd was aan laagfrequent gehoorverlies.

Conclusie: De FFR_Env zou eerder beïnvloed zijn door cochleaire synaptopathie en waarschijnlijk ook door leeftijd, terwijl de FFR_Tfs aangetast zou zijn door IHC-verlies in de overeenkomende frequentierange van de stimulus. We vermoeden dat jongere participanten meer zouden vertrouwen op ENV cues dan onze oudere proefgroep. Toekomstige studies zouden een jong normaalhorende proefgroep moeten includeren. Ook zou de spraakverstaanbaarheidstest in ruis gecompenseerd moeten worden voor gehoorverlies om na te gaan in welke mate de ENV en TFS bijkomende informatie kan bieden.

Abstract (En)

Background: In the audiological practice, much variability in speech intelligibility is observed between patients, despite similar (normal or increased) hearing thresholds. Speech perception is complex and could be influenced by tree different individual factors (audibility, cognition and suprathreshold coding). Especially middle-aged people experience complaints about speech understanding, particularly in noisy environments. Cochlear synaptopathy (damage to the synapses between the inner hair cells (IHC) of the cochlea and the afferent nerve fibers) would play an important role in this complaint. Recently, the envelope following response (EFR), an EEG-response to amplitude modulated sounds, has shown to be a good measurement for cochlear synaptopathy. These would measure the ability to process the temporal envelope (ENV) cues in speech signals. Nevertheless, realistic speech signals in daily life are not only temporally modulated, but also spectrally, where the temporal fine structure (TFS) cues play an important role. Research with behavioral psycho-acoustic measurements has shown that spectrotemporal modulated (STM) stimuli appear to be good representatives of realistic speech signals. Despite the promising results, these behavioral tests can be influenced by cognitive factors. A way to measure spectral processing objectively could be the frequency following response (FFR), mostly measured with syllables or vowels in the recent literature.

Purpose: The study aims to analyze whether an FFR measurement with spectral modulated stimuli (harmonics) could offer objective information about the processing of TFS cues and whether this is related to speech perception in people with normal and impaired hearing thresholds.

Method: Two groups of middle-aged participants (40 to 60 years old) were included; a normal hearing (noNH=16) and a hearing-impaired group (noHI=14), based on tonal audiometry. To consider whether the FFR is a good measurement for speech understanding, results were compared with the Flemish Matrix sentence-test. Also behavioral S(T)M detection thresholds were measured and compared in order to find out if these have the same underlying mechanisms. Besides, we wanted to investigate whether the FFR is influenced by other sensorineural hearing loss based on tonal audiometry, otoacoustic emissions (OAE), the EFR and the auditory brainstem response (ABR).

Results: Neither the S(T)M detection thresholds, nor the FFR correlated with speech perception in noise. The S(T)M detection thresholds showed no significant differences between normal hearing and hearing impaired participants, were not correlated to auditory thresholds, but did significantly to loss of outer hair cells (based on OAE). The FFR results were split up into TFS and ENV cues. Despite its low frequent stimulus, the FFR_Env showed correlations with high-frequent hearing thresholds and cochlear synaptopathy, while the FFR_Tfs correlated with low frequency hearing thresholds.

Conclusion: The FFR_Env seems to be more affected by cochlear synaptopathy and possibly by age, while the FFR_Tfs detectability would be affected by IHC-loss in the frequency range of the presented stimulus. We suspect that younger participants would rely more on ENV cues than our middle-aged test group. Future studies should include a young normal hearing test group. Also, speech in noise tests should be compensated for hearing loss to consider in which amount that TFS could add additional information.

Introduction

Speech understanding, especially in noise, is one of the most reported complaints in the audiologists' daily practice. Nevertheless, adults with normal hearing thresholds also often complain of difficulties with speech understanding in noisy environments. This phenomenon is referred to 'hidden hearing loss' and is believed to be common in older adults, especially those with a history of extensive noise exposure. Indeed, an individual's speech intelligibility cannot only be understood from his hearing thresholds, at least three components of hearing must be considered (Bernstein et al., 2016; Füllgrabe et al., 2015): (i) the audibility of the speech signal (auditory thresholds), (ii) the suprathreshold coding (the quality of the auditory signal above threshold) and (ii) cognitive capacities (involving working memory, attention…) of the listener. In each individual, these components may contribute in a different proportion to the degree of speech intelligibility, but the current view in the field is that suprathreshold coding deficits significantly contribute to 'hidden hearing loss'.

Among the potential sources of deterioration in suprathreshold coding, cochlear synaptopathy is thought to have a substantial contribution. Cochlear synaptopathy refers to the loss of synapses between the inner hair cells (IHCs) and their auditory nerve fibers (ANFs) in the cochlea and could be caused by overexposure to extensive noise, aging and ototoxicity, even if the (outer) hair cells are unaffected (yet) (Bharadwaj et al., 2014; Kujawa & Liberman, 2009; Sergeyenko et al., 2013). The major subjective complaint of listeners with suspected cochlear synaptopathy is poor speech perception in noisy listening environments (Ruggles et al., 2012). Also, problems with sound localization, musical pitch perception and tinnitus or hyperacusis are other suggested symptoms of cochlear synaptopathy (Plack et al., 2014).

In the human ear, ANFs can be divided into three groups, based on their spontaneous spike rate (SR): high-SR (>18 spikes/s), medium-SR (0.5 to 18 spikes/s) and low-SR fibers (≤ 0.5 spikes/s). High-SR fibers have, due to their increased spike rates, the lowest firing thresholds, but already saturate at 20 to 30 dB above that threshold. These fibers would be important for sound detection (tested with tonal audiometry), while low-SR fibers have higher firing thresholds, but a larger dynamic range and would be necessary for suprathreshold coding (Bharadwaj et al., 2014). The temporal characteristics of a speech signal can be separated into two parts: temporal fine structure (TFS) and the temporal envelope (ENV). Low-frequency ANFs (< 1-2 kHz in humans) would convey both TFS and ENV information, while high-frequency ANFs (> 1-2 kHz) would not phase-lock to TFS and only convey temporal information by the ENV (Bharadwaj et al., 2014; Joris et al., 2004). The fluctuating characteristics of the ENV cannot be detected by high-SR fibers, due to early saturation of the spike rate (Bharadwaj et al., 2014). Therefore, low-SR fibers are useful in situations where TFS is less detectable (noisy environments). Thus, as the high-SR fibers better detect sound in quiet (because of their low thresholds), low-SR fibers are needed for speech understanding in noise (Bharadwaj et al., 2014; Furman et al., 2013).

Today, the envelope following response (EFR) generated by amplitude-modulated (AM) signals, is considered a promising objective, electro-encephalographic based (EEG), diagnostic test for cochlear synaptopathy (Bramhall et al., 2019; Shaheen et al., 2015; Verhulst et al., 2018). These provide a signature of temporal coding in the ascending auditory pathway. The responses, recorded through electrodes on the scalp, reflect the phase-locking capacities of the auditory nerve to the ENV of these signals. According to recent studies, this temporal coding (particularly the ENV) of supra-threshold sound would be impaired in patients with synaptopathy (Bharadwaj et al., 2014; Plack et al., 2014; Ruggles et al., 2012). In general, the EFR is considered as a better diagnostic test for synaptopathy than the auditory brainstem response (ABR) (Shaheen et al., 2015; Vasilkov et al., 2021). In contrast to the onset responses of the ABR, the steady state responses of the EFR appear to be of the same magnitude in high-SR rates as in low- and medium-SR rates, which results in a better representation of cochlear synaptopathy (Bharadwaj et al., 2014). Furthermore, EFR would be less affected by outer hair cell (OHC)-loss, provided that the correct parameters are used (Verhulst et al., 2018). Following this idea, Vasilkov et al. (2021) optimized the parameters for the EFR to maximally reduce any influence of increased auditory thresholds on the magnitude of the response of the EFR, making this diagnostic tool specifically sensitive to cochlear synaptopathy.

However, the spectrogram of natural speech reveals that speech does not only consist of temporal modulation: it also shows spectral modulations, which reflect the frequency characteristics of the sound at different scales, namely its pitch and timbre. Elliott and Theunissen (2009) visualized these modulations of speech in the ‘modulation power spectrum’ (MPS), using a two-dimensional Fourier transformation of the spectrogram. The sounds with little spectral structure, but fast temporal changes would correspond to non-vocalic sounds, whereas sounds with rich spectral structure and slow temporal modulations would correspond to vocalic sounds, such as vowels (Elliott & Theunissen, 2009). The intermediate spectro-temporal modulations have less power, but would be similar to formant-transitions.

The coding of spectral modulations is also studied in the current literature. One way to measure the ability to detect spectral modulations is the spectral ripple detection test. A spectral ripple is a broadband noise, modulated in energy across frequencies (Aronoff & Landsberger, 2013), modeling a formant-like structure of energy as it is present in vowel sounds. These signals can be varied in density (ripples/octave or cycles/octave) and modulation depth (d; with d=1 corresponding to full modulation). The spectral ripple detection test is a psychoacoustic test that measures the detection threshold i.e., the smallest modulation depth that the listener can distinguish from non-modulated sounds (i.e. white noises). For older adults with symmetrical hearing loss, spectral ripple thresholds have been shown to predict speech intelligibility in conditions where subjects are wearing their fitted hearing aids (Miller et al., 2018). This decreased ripple sensitivity in participants with decreased speech understanding, is currently interpreted as being caused by a reduction of frequency selectivity and/or ability to encode the TFS (Bernstein et al., 2013).

Those spectral ripples can also be measured objectively by using EEG techniques, namely the frequency following response (FFR). Several studies measured the FFR on vowels and developed a technique to distinguish between the ENV and TFS based on alternating polarities (Aiken & Picton, 2008; Ananthakrishnan et al., 2016). The energy on the formant peaks of the used vowels was detected and analyzed. Results showed that the FFR_Env would be more affected by high-frequency, age-related hearing loss and intensity changes of the stimulus, than the FFR_Tfs (Ananthakrishnan et al., 2016; Märcher-Rørsted et al., 2022; Ruggles et al., 2012). Ruggles et al. (2012) also showed correlations with age for both FFR_Env and FFR_Tfs. Nevertheless, no correlations were found between FFR_Env and FFR_Tfs, which would mean that both parameters describe other mechanisms (Ruggles et al., 2012). One study of Mai et al. (2018) found correlations between FFR_Env and SPIN, making it a possible objective diagnostic tool.

As mentioned above, the sensitivity to temporal modulations assessed through EFR measurements to AM-sounds, is currently proposed as an interesting candidate for the diagnosis of cochlear synaptopathy (Parthasarathy & Kujawa, 2018; Shaheen et al., 2015). In the present work, we asked whether sensitivity to modulations in the spectral dimension could give additional information to temporal modulation sensitivity regarding speech processing capacities. In order to approach natural speech, spectro-temporal modulated (STM) stimuli can be used. STM stimuli are broadband noises whose envelopes are modulated in both the spectral and temporal domains, which can be defined by three important parameters: temporal rate, spectral density and direction (Bernstein et al., 2013). Several studies investigated the correlation between detection thresholds of STM-stimuli and speech understanding. As expected based on previous results of other spectral-ripple studies, these studies showed that STM thresholds also correlate with speech perception thresholds (Bernstein et al., 2016; Bernstein et al., 2013; Miller et al., 2018). Interestingly, correlations between SM and STM detection thresholds suggest that their perception is driven by the same underlying processing (Miller et al., 2018; Sabin et al., 2012), while the absence of a relationship between AM (pure temporal modulations) and STM thresholds suggest that they reflect different underlying processing (Sabin et al., 2012).

Thus, sensitivity to STM stimuli appears to account for a significant part of the variance in speech performance (in noise) (Bernstein et al., 2016; Bernstein et al., 2013). In combination with the use of the speech intelligibility index, these results can account for up to 80% of the variance of variability across listeners with sensorineural hearing loss (Bernstein et al., 2013). The correlation between SM and STM thresholds is currently interpreted as reflecting the ability of the listeners to use TFS information (Bernstein et al., 2016; Bernstein et al., 2013; Miller et al., 2018). With this reasoning in mind, EEG recordings with S(T)M-stimuli could be a potentially interesting test to assess both cochlear synaptopathy and speech understanding capacities. The main advantage of using EEG measurements is that these are objective and thus could also be applied to children or adults with mental retardation. Moreover, the results are not influenced by cognitive capacities of the participants.

In the present work, we aimed to investigate and compare the information that can be derived from measurements with SM stimuli with the information derived from measurements with STM stimuli. The objectives are thus to find out:

1. Whether the parameters of the FFR of the SM-EEG and behavioral S(T)M-detection thresholds could constitute individual speech intelligibility metrics, i.e. that would account for performance in the Matrix sentence test.
2. Whether the FFR of the SM-EEG is as selective for cochlear synaptopathy as the EFR of the RAM-EEG in middle-aged people while remaining insensitive to co-existing OHC damage.

We will include two groups of participants: one with normal hearing (oNH) thresholds and one with clinical hearing loss (oHI) on tonal audiometry. In this way, we can investigate whether there is an influence of other cochlear damage. Hypotheses are summed up below:

1. We will measure otoacoustic emissions (OAEs), representing OHC-loss. In our study, we expect the oHI-subjects to have OHC-dysfunction, congruent to the auditory threshold-configuration (high-frequency sloping). In the oNH-group, OAEs are expected to be present. nevertheless, OHC-loss is not always visible with tonal audiometry (also a kind of ‘hidden hearing loss’), whereby OAEs can still be absent due to age or noise exposure. We also expect an age-effect on both audiometric thresholds and OAEs for the normal hearing participants. Also, ABR measurements will be used in order to describe the hearing loss of the test groups. Liberman et al. (2016) measured ABRs in three test groups (young normal hearing (yNH), oNH and oHI). The amplitudes were only decreased in both older groups compared to the yNH-group. Therefore, we do not expect a significant difference in the ABR-amplitudes in our test-groups (oNH and oHI). The latency depended on the stimulus intensity (Liberman et al., 2016); there was a significant difference in the latencies of wave V between oNH and oHI participants at 70 dB pe SPL, which was reduced at 100 dB peSPL. Wave I was not always present in the oHI-group. We expect similar findings in our study.
2. The SPIN test will be measured in order to investigate if the S(T)M detection threshold test and FFR could represent aspects of speech intelligibility in daily life. The speech stimuli will be filtered in low-pass (LP), high-pass (HP) and broad band (BB) conditions. In previous studies of our group using a

similar Matrix test and different filtering conditions (Garrett et al., 2020; Verhulst & Warzybok, 2018), the results for speech-in-quiet (SPIQ) differed significantly between groups with and without clinical hearing loss (ages between 60 and 70 yrs.) and were correlated with the DPOAEs and hearing thresholds. The thresholds for speech-in-noise (SPIN) (HP-condition) were increased in older compared to younger adults, but not between oHI and oNH test groups. These thresholds were also correlated with the EFR, which would reflect the degree of cochlear synaptopathy (Garrett et al., 2020). In the present study, we expect similar results (difference in SPIQ, but similar SPIN) with middle aged subjects.

3. Our S(T)M detection threshold test will be similar to the one of Miller et al. (2018) with both static (SM with harmonics or with noise in our study) and moving ripples (STM with harmonics in our study). They found significantly better thresholds in the STM condition than the SM condition for their study group (HI; 29-79 yrs.) and a strong correlation between the two (Miller et al., 2018). Bernstein et al. (2013) compared NH and HI participants and found significantly better STM detection thresholds in the NH group. As we use a LP filter on our stimuli, we don't expect differences because the oNH and oHI groups will have similar audiometric thresholds in low frequencies. Several studies showed correlations between spectral ripple tests (especially STM) and SPIN-tests (Bernstein et al., 2016; Bernstein et al., 2013; Davies-Venn et al., 2015; Miller et al., 2018). Therefore, we also expect better results on the STM condition than the SM condition in general; similar thresholds in oNH and oHI participants and correlations with the SPIN tests (especially LP) in this study.
4. For the FFR, we can split up our hypotheses for the ENV and TFS components. In several studies, the ENV component was more influenced by audiometric thresholds than TFS (Ananthakrishnan et al., 2016; Märcher-Rørsted et al., 2022; Ruggles et al., 2012). Age would have an effect on both FFR_Env and FFR_Tfs, but they would not correlate to each other based on the study of Ruggles et al. (2012). One study also showed significant relations between FFR_Env with syllables and SPIN (Mai et al. (2018). As the S(T)M detection thresholds were also correlated with speech understanding (Bernstein et al., 2016; Bernstein et al., 2013; Davies-Venn et al., 2015; Miller et al., 2018), we expect the correlations between FFR (with SM harmonics) and SPIN to be significant. In speech in noise, TFS-cues are less detectable, therefore we expect no correlation with the FFR_Tfs but with the FFR_Env. For speech in quiet, both TFS and ENV cues are conveyed, so both FFR_Tfs and FFR_Env could correlate.
5. Lastly, we could expect that the FFR results correlate to the SM detection threshold, as they include the same spectral modulations. This was the case in the study of Parthasarathy et al. (2020), who also used an FM detection threshold task and FFR measurement with the same stimuli.

Method

Participants

For this study, we recruited 33 middle-aged Dutch-speaking adults from 40 to 60 years old. This population was expected to have cochlear synaptopathy in varying degrees. Three of them were excluded due to measurement issues or because the audiogram did not fit in the criteria for the test groups (see Appendix 1). The remaining 30 participants were split up into two test groups based on their audiograms. The first group consisted of sixteen older normal hearing (oNH) participants, five men and eleven women (mean age = 46.8 yrs.; *SD* = 5.91 yrs.; range 41 to 60 yrs.). For the second group, fourteen patients (six men and eight women) with clinical sloping sensorineural hearing loss (oHI) participated, (mean age = 51.9 yrs.; *SD* = 6.16 yrs.; range 40 to 56 yrs.). Additionally, two young normal hearing participants (yNH), two women (ages 22 and 24 yrs.), were tested as a baseline for the analyzed data.

The study was approved by the ethical commission of the Ghent University Hospital (UZ Ghent). Participants were informed about the experimental procedures according to the ethical guidelines and signed an informed consent. Participants were recruited via flyers, social media and the ear-nose-throat department of UZ Ghent. People that were interested filled in a questionnaire online before the start of the experiment that included the following topics: (i) general sociodemographic questions, (ii) hearing, (iii) tinnitus and (iv) noise exposure and hearing protection. The last section was based on a questionnaire developed by Jokitulppo et al. (2006), resulting in an individual exposure level (Laeq). This questionnaire screened for following exclusion criteria: (i) chronic tinnitus or hyperacusis, (ii) middle-ear pathologies and/or history of middle-ear surgery, (iii) pregnancy and (iv) any known genetic hearing loss. Subjects who participated in the entire study received financial compensation.

Test protocol

The test battery was split up into two sessions of a maximum of two hours, split over two days in order to avoid excessive noise exposure. The first session included tympanometry, tonal audiometry, speech audiometry and an S(T)M detection threshold test. The second session consisted of objective tests, namely OAEs and EEGs. We remind here that since TFS can only be conveyed at low frequencies, most tests were restricted to stimuli in low-frequency regions (< 1500 Hz).

Tympanometry and otoscopy

Conductive hearing losses and middle/outer ear pathologies were excluded at the beginning of the experiment using tympanometry and otoscopy. Tympanometry was performed with the GSI TympStar (Grason-Stadler) in order to exclude conductive hearing losses caused by middle/outer ear pathologies. By use of a leakproof probe, the middle-ear admittance was measured with a 226 Hz, 85 dB sound pressure level (SPL) tone. From the selected participants, 26 (twelve oHI, twelve oNH and two yNH) had 'type A' tympanograms with middle ear pressure (MEP) between -100 and 100 daPa and statical acoustical admittance (SAA) between 0,3 and 1,7 mmho. Six participants (two oHI and four oNH) showed a 'type Ad' tympanogram with an SAA bigger than 1,7 mmho. Otoscopy was performed with two participants. One had bilateral large ear canal volumes (ECV) due to big ear canals, the eardrum showed no perforations. With another participant, we could not get a leak tight seal in the left ear, otoscopy showed no outer or middle ear problems, the right ear was used for further tests to be sure.

Audiometry

Hearing thresholds were measured using the Hughson-Westlake method. Participants were placed in a double-walled sound-attenuating booth and stimuli were presented by an Equinox Interacoustics audiometer and transmitted using Interacoustics TDH-39 headphones for the conventional (half-)octave frequencies (0.125, 0.25, 0.5, 1, 2, 3, 4, 6 and 8 kHz). Also extended high frequencies of 10, 12.5, 14 and 16 kHz were measured using circumoral Sennheiser HAD-200 headphones. If the air-conduction thresholds exceeded 20 dB HL between 0.25 and 4 kHz, bone conduction was measured with an Interacoustics bone vibrator placed on the mastoid in order to compute the air-bone gap (ABG).

If the ABG was 15 dB HL or higher, the participant was excluded from the study because of a possible middle-ear pathology. Middle-aged adults were expected to have mild increased hearing thresholds at high frequencies (8 kHz and higher). Therefore, participants were placed in the oNH-group if their hearing threshold does not exceed 20 dB HL between 0,125 and 4 kHz. The oHI-group included people with a high-frequency sloping hearing loss of more than 20 dB HL at 4000 Hz. For the rest of the experiment, only the best ear (based on the 4 kHz threshold and the pure tone threshold (PTA)) was measured in order to reduce time and prevent crossover of the sound signals used further in the experiment (oHI: seven right and seven left; oNH: nine right and seven left; yNH: both left).

Speech intelligibility

Speech intelligibility was measured using the Flemish Matrix sentence test (Luts et al., 2014), which consists of non-predictable five-word sentences of a closed set (10 names, 10 verbs, 10 numerals, 10 colors and 10 objects), guessing was allowed but not mandatory. The sentences were presented with the Fireface UCX soundcard (RME) and HAD-300 (Sennheiser) headphones in a sound-proof room and were analyzed in MATLAB. Using an automatic script in MATLAB, participants had to click the comprehended words on the word-matrix on a laptop. SPIN was measured in three filter-conditions (for both the speech and noise): BB (no filtering), LP (cutoff 1500 Hz) and HP (cutoff 1650 Hz). SPIQ was only measured in the LP-condition in order to reduce time. For each condition, a different wordlist was examined and the same lists were used for all participants.

For the SPIN, the noise was fixed at an intensity of 75 dB SPL and the signal-to-noise ratio (SNR) started at -4 dB. The noise started 500 ms before and ended 500 ms after each sentence. Depending on the correctness of the answer, the speech intensity (and thus SNR) was adapted using the staircase paradigm in steps of minimum 0.1 dB SNR. For the SPIQ, the speech signal started at 75 dB SPL and the intensity variated depending on the number of correct words. The test was training-dependent, thus only the mean of the last six reversals was used for analysis. The SPIQ resulted in a speech reception threshold (SRT), while the SPIN measured an SNR. The lower the SRT and SNR values, the better the speech intelligibility.

S(T)M detection threshold test

The measurement procedure was almost identical to that described by Miller et al. (2018). The S(T)M detection thresholds were obtained by changing the modulation depth of the envelope of the stimuli with a three-down, one-up adaptive procedure tracking the 79,4% correct point on the psychometric function ((Levitt, 1971) see Figure 1c and d). We used a three-interval two-alternative forced choice task, where the modulated stimulus is presented in the second or third interval with equal probability using a script on MATLAB. Stimuli were presented with HAD-300 (Sennheiser) headphones connected to a Fireface UCX soundcard (RME). All stimuli were presented at 75 dB SPL (rms). Three conditions were examined; (i) a harmonic carrier (f0=110 Hz, in the 300-1500 Hz regions) with a SM envelope at 2 cycle/oct, (ii) a noise carrier (same region 300-1500 Hz) with a SM envelope at 2 cycle/oct, (iii) an harmonic carrier with a STM envelope at 2 cycle/oct and 4 Hz (we also added bands of noise in the 200-300 Hz and 1500-1800 Hz regions at -15 dB of the main level to mask potential temporal off-frequency cues) (see Figure 1a and b).

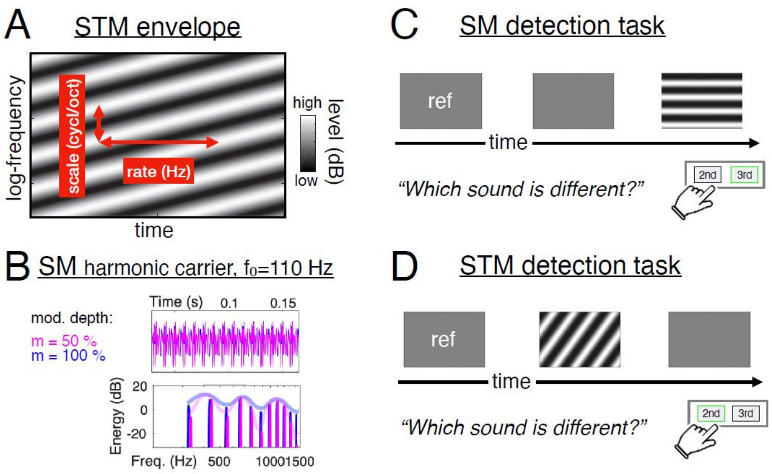


Figure 1: (a) Spectrogram of the spectrotemporal modulated (STM) envelope used in the STM-detection threshold task. (b) Representation of the spectral modulated (SM) harmonic carrier with f_0 = 110Hz used in the SM-detection threshold task and the FFR. (c) Example of the behavioral SM detection task with a three-down, one-up adaptive procedure. (d) Example of the behavioral STM detection task with a three-down, one-up adaptive procedure.

The duration of the stimuli was 500 ms (with 20 ms ramps at onset and offset) and had an interstimulus interval of 300 ms. Two interleaved tracks were collected per condition (a third track was run if the difference between the thresholds measured in the two tracks was greater than 3 dB). If the tracks did not converge because a given listener was unable to reach performance threshold (i.e., >3 incorrect responses at full mod. depth), the tracking stopped, and 40 additional trials were collected at full modulation depth to extrapolate the data and infer the threshold. Participants received training at the beginning of the task (with a stimulus duration of 1000 ms and a modulation depth of 50 or 100%) and got visual trial-by-trial feedback regarding the correctness of their responses. Lower thresholds reflect better discrimination abilities. An important note is our method of determining the S(T)M detection thresholds. As the responses did not converge for some subjects with the adaptive procedure, we developed a method of reconstruction of the psychometric functions to estimate the thresholds because there was a lot of variability in the tracking (see Appendix 2).

OAEs

OAEs were measured to visualize the OHC-integrity, using the Universal Smart Box (Intelligent Hearing System) with Etymotic ER10D probes and analyzed with the corresponding software (SmartDPOAE and SmartEP continuous acquisition module (SEPCAM)) and Spyder. For the DPgram, two tone pairs were presented simultaneously with f_2=1.22*f_1 and f_2 ranging from 500 to 8000 Hz with 2 freqs/octave. The intensities were kept at 65 and 55 dB SPL for L_1 and L_2 respectively and also a measurement with 55 and 40 dB SPL was done. 32 sweeps were taken and the SNR on each frequency was used for further analysis. Also DPOAE input/output thresholds were measured at 1, 2 and 4 kHz as described by Verhulst et al. (2016). Tone pairs were presented at f_2=1.22*f_1. The intensity of L_2 ranged from 35 to 70 dB SPL and L_1 was computed based on the Scissors paradigm of Kummer et al. (1998) (L_1=0.4*L_2+39). For TEOAEs, clicks of 80 µs were produced at 80 dB SPL during 2000 sweeps.

EEG-tests

AEP measurements were executed using the Universal Smart Box (Intelligent Hearing System) with the SEPCAM. Participants were seated in a reclining armchair in a quiet room, watching a silent movie in order to keep them relaxed and awake (Purcell et al., 2004). Ambu Neuroline snap electrodes were connected to the vertex, nose and both mastoids after applying NuPrep® gel with a cotton swab for reducing the skin-impedance to a maximum of 3 kΩ. The stimuli were presented with ER-2 inserts (Etymotic Research) and the ears were covered with earmuffs to minimize the noise intrusion level. All electrical devices other than the measurement equipment were turned off and unplugged.

Auditory brainstem response

ABR-responses were measured to obtain an extensive representation of the auditory nerve function. The used parameters are based on the studies of Liberman et al. (2016) and Vasilkov et al. (2021). A click-stimulus of 100 µs was presented with a rate of 11 Hz. Two intensities (75 and 90 dB peSPL (rms)) were measured, with both 3000 alternating sweeps, filtered between 10 and 1500 Hz. Positive and negative waves were pointed on the ABR-curves manually. Both amplitudes and latencies of selected waves were saved as data.

Envelope following response with RAM stimulus

A pure tone of 4 kHz was 100% rectangularly, amplitude-modulated at 110 Hz with a duty cycle of 25% (RAM25), as used in the study of Vasilkov et al. (2021). The stimuli were presented at 75 dB SPL (rms) for 200 ms during 1000 alternating sweeps and the signal was filtered from 10 to 5000 Hz. In order to measure the ENV component, both polarities (condensation and rarefaction) were summed up. The TFS component was analyzed by subtracting both polarities (Aiken & Picton, 2008; Ananthakrishnan et al., 2016). A Fast Fourier transformation (FFT) was performed, resulting in a spectrum with peaks on the modulation frequency (f_1=110 Hz) and the following harmonics (f_n) for the ENV condition, while the TFS shows no measurable peaks (see Figure 2). After a noise-floor correction, the amplitudes of the first four harmonics (f_1 until f_4) were summed up, resulting in an EFR-magnitude (Vasilkov et al., 2021).

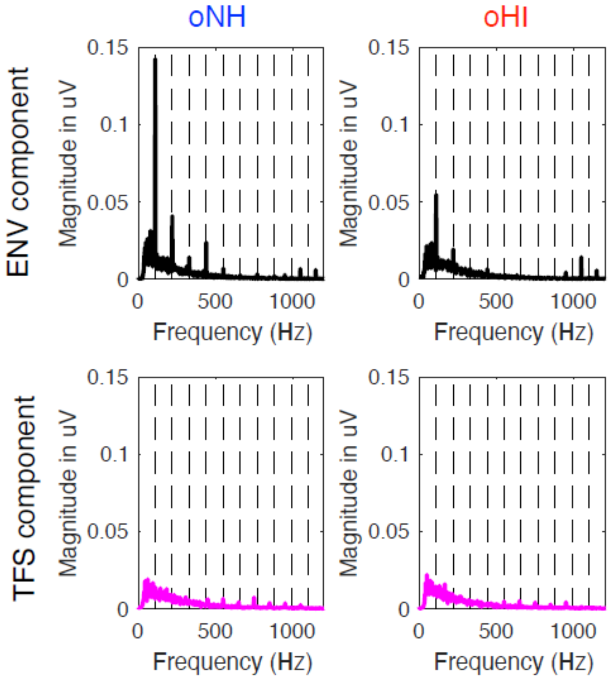


Figure 2: Example of the spectrums of the envelope following response (EFR) after noise correction for a normal hearing (oNH: left) and a hearing impaired (oHI: right) subject. The upper black spectra represent the envelope response with peaks on the harmonics (EFR_Env; sum of the polarities), while the lower pink spectra show the temporal fine structure (EFR_Tfs; subtraction of the polarities).

Frequency following response with SM stimulus

Signals were exactly the same as in the psychoacoustical detection task. We used a harmonic signal with an f0 at 110 Hz, filtered between 300 and 1500 Hz. Two stimuli were used; one with a spectral-modulation depth of 50% and the other one with a spectral-modulation depth of 100%. These stimuli were 400-ms long, had a level 75 dB SPL (rms), and were each presented using 3000 repetitions of alternating polarities. Because there was almost no difference between the responses to these sounds, we used the mean of the two measurements in our analyses. FFR_Env was derived by summing all recorded responses to positive and negative stimulus polarities, while FFR_Tfs was derived by subtracting the responses to the two stimulus polarities (Aiken & Picton, 2008; Ananthakrishnan et al., 2016). A FFT was computed on these responses, showing peaks at f0 (110 Hz) and its harmonics. We then used a bootstrap-based correction algorithm to estimate spectra after noise-floor correction. Finally, an index representing the overall magnitude of the responses (later used for statistical analyses) was computed by summing the magnitude of the harmonics that were significantly above noise-floor; we retained harmonics #1 to #5 for FFR_env and harmonics #3 to #7 for FFR_tfs, since they were those at which the energy was significantly above the noise floor (see Figure 3).

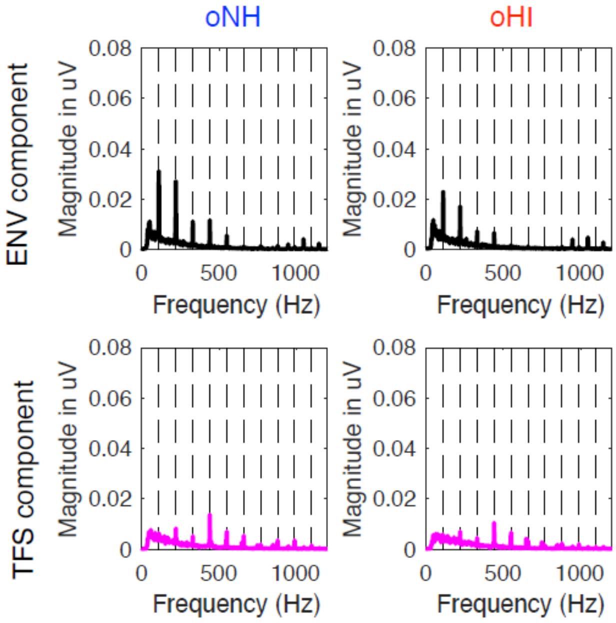


Figure 3: Example of the spectrums of the frequency following response (FFR) after noise correction for a normal hearing (oNH: left) and a hearing impaired (oHI: right) subject. The upper black spectra represent the envelope response with peaks on the harmonics (FFR_Env; sum of the polarities), while the lower pink spectra show the temporal fine structure, which has also peaks on specific harmonics (FFR_Tfs; subtraction of the polarities).

Statistical analysis

Data analysis was performed using IBM SPSS Statistics 28. First, the descriptive parameters of the chosen variables were analyzed, and the Shapiro-Wilk normality test was performed per variable for all participants together and for both oHI and oNH apart. For further analysis, a distinction was made between variables describing hearing loss (age, audiometry, OAE and ABR) and variables describing suprathreshold coding (SPIN, S(T)M detection, EFR and FFR).

In order to consider the influence of reduced audibility, comparisons were made between the oHI and oNH test groups for all variables. yNH-data was only used as a baseline value for the visualization in boxplots. Normal data distribution was distinguished using the Shapiro-Wilk-Test and equal-variance assumption by the Levene-Test. If assumptions were satisfied, the independent student t-test was used. If the normal distribution assumption was not met for two independent samples (which was only the case for EFR_Env), the non-parametric Mann-Whitney U Test was applied. Two-tailed significance level was reached when $p<0.05$.

Correlations were also analyzed; these were distinguished using the Pearson correlation coefficient (r) if both variables are normally distributed, otherwise Spearman's rank correlation coefficient (ρ) was used. Two-tailed significance level was reached when $p<0.05$. We have to notice that a lot of correlations are performed, which results in an increased chance of false positive results. We are aware of this possible problem in this preliminary work. All correlation were measured for the whole test group (all) and for both oNH and oHI groups apart. First, correlations were measured between the suprathreshold variables (SPIN, S(T)M detection and FFR) and the hearing loss variables in order to consider the effect of hair cell loss on the experimental measurements. Afterwards, correlations between the suprathreshold measurements were analyzed.

Results

Parameters describing hearing loss (hypothesis 1)

Table 1: T-values of the student t-test for the variables describing hearing loss (audiometry, OAE and ABR).

		oHI		***oNH***	
	t	*Mean*	*SD*	*Mean*	*SD*
Audiometric threshold 4000 Hz (dB HL)	6.62** *p*<0.001	43.9	18.93	9.1	5.84
Audiometric threshold low frequencies (dB HL)	2.70* *p*=0.012	13.93	7.64	7.81	4.55
High frequency audiometry (dB HL)	7.15** *p*<0.001	62.50	18.13	22.34	12.43
DPgram 4000 Hz (dB)	-1.97 *p*=0.066	-14.56	2.23	-9.23	10.54
ABR latency wave I (ms)	2.27* *p*=0.033	3.90	0.41	3.48	0.50
ABR latency wave V (ms)	2.72* *p* =0.012	8.09	0.46	7.64	0.40
ABR amplitude wave I (µV)	-1.98 *p* =0.060	-0.0025	0.0507	0.0468	0.0691
ABR amplitude wave V (µV)	-2.41* *p* =0.024	0.1777	0.1204	0.2928	0.1257

Significant t-values were indicated with * for *p*<0.05 and ** for *p*<0.01. DP = distortion product; ABR = auditory brainstem response.

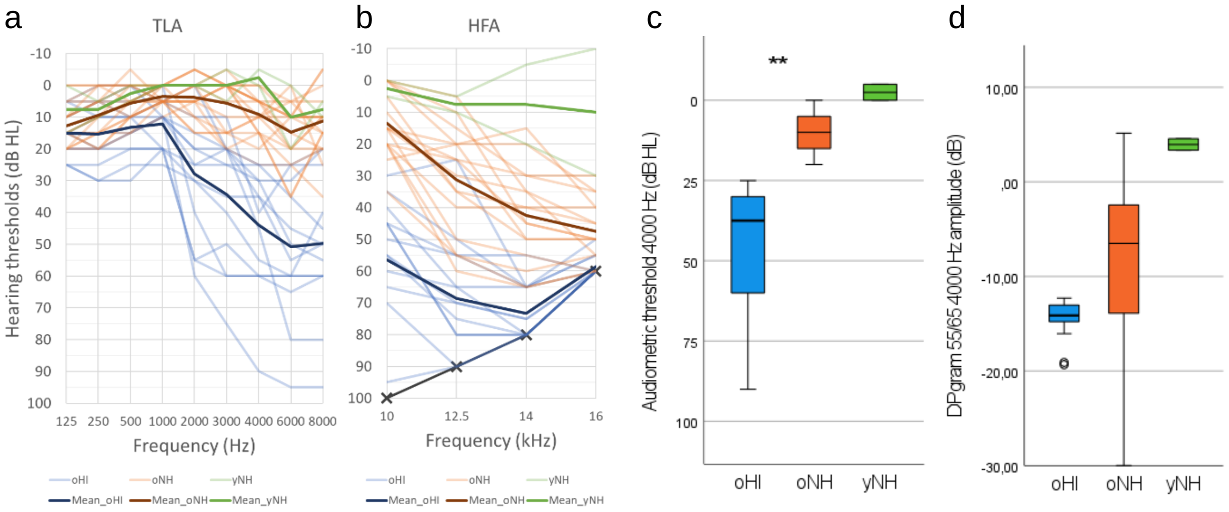


Figure 4: (a) Tonal liminary audiograms and mean thresholds split up in test groups. (b) Extended high frequency audiograms and mean thresholds split up in test groups. (c) Boxplot of the audiometric hearing thresholds on 4000 Hz for each test group. (d) Boxplot of the amplitudes of the DPgram on 4000 Hz for each test group. oHI = older hearing impaired; oNH = older normal hearing; yNH = younger normal hearing.

The tonal audiograms (tonal liminary audiometry (TLA) and high frequency audiometry (HFA)) of the three test groups are plotted in Figure 4(a and b). For the HFA, the output intensity of the audiometer was limited to 100, 90, 80 and 60 dB HL for 10, 12.5, 14 and 16 kHz respectively. If no response was detected at these levels, the threshold was considered as the maximum output intensity. This was the case for one oNH and ten oHI participants and had the most effect on 14 and 16 kHz. Therefore, we used the mean thresholds on

10 and 12.5 kHz instead of 10 to 16 kHz in further analyses based on high-frequency thresholds (see Appendix 3).

The audiometric thresholds and DPgram amplitudes on 4000 Hz are plotted in Figure 4(c and d). The significant group differences for Audio_4000Hz (audiometric thresholds on 4000 Hz), Audio_LF (mean audiometric thresholds from 125 to 1000 Hz) and HFA (mean high frequency audiometric thresholds on 10 and 12.5 kHz) are presented in Table 1. Significant moderate correlations with age were found for Audio_LF (*r*=0.41; *p*=0.027) and Audio_HFA (*r*=0.41; *p*=0.025), confirming our first hypothesis. For Audio_4000Hz, correlations were reversed for the two test groups (HI: *r*=-0.73, *p*=0.003 and NH: *r*=0.55, *p*=0.029).

For the DPgram on 4000 Hz (DP_4000Hz), no significant difference between oHI and oNH groups was found (*t*=-1.97; *p*>0.05). As seen in Figure 4(d) the values for the oHI-group are all below -10 dB and could be interpreted as absent OHC-responses, while the oNH-group shows more spread in DPgram-amplitudes, a trend we described in our first hypothesis. As we also expected, DP_4000Hz significantly correlated with age in the oNH group (*r*=-0.55, *p*=0.027). For the ABR, we expected significant differences in latencies for the first and fifth peaks, but not for their amplitudes. Significant differences were found for the latency in both peaks, but also for the amplitude of peak five (see Figure 5 and Table 1).

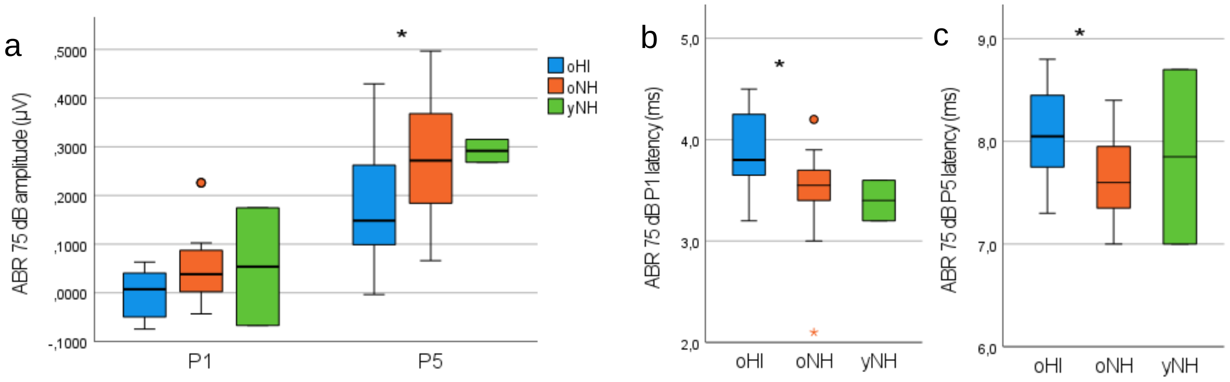


Figure 5: Boxplots of auditory brainstem responses wave I (P1) and wave V (P1), split up in test groups. (a) Amplitudes of the ABR peaks. (b) Latencies of wave I. (c) Latencies of wave V. oHI = older hearing impaired; oNH = older normal hearing; yNH = younger normal hearing.

Parameters describing suprathreshold coding

Table 2: T-values of the student t-test for the variables describing suprathreshold coding (SPIN, S(T)M detection thresholds and FFR) and U-values of the Mann-Whitney U test for the EFR.

		oHI		***oNH***	
	t/U	*Mean*	*SD*	*Mean*	*SD*
SPIN broad band (dB SNR)	3.08** *p*=0.008	-5.439	2.327	-7.403	0.565
SPIN low pass (dB SNR)	-1.62 *p*=0.116	-2.964	2.248	-1.813	1.628
SPIN high pass (dB SNR)	2.33* *p*=0.037	0.950	6.729	-3.450	1.180
SM harmonics (dB)	-1.78 *p*=0.086	-14.242	3.638	-11.951	3.671
SM noise (dB)	-0.11 *p*=0.914	-12.544	4.134	-12.389	3.646
STM harmonics (dB)	-0.86 *p*=0.398	-9.092	4.429	-7.670	4.615
EFR envelope (µV)	191** *p*<0.001	0.0879	0.0622	0.2213	0.1646
FFR envelope (µV)	-2.61* *p*=0.014	0.0555	0.0255	0.0863	0.0372
FFR temporal fine structure (µV)	-1.06 *p*=0.302	0.0269	0.0102	0.0331	0.0209

Significant t-values were indicated with * for *p*<0.05 and ** for *p*<0.01. SPIN = speech in noise; SM = spectral modulated; STM = spectrotemporal modulated; EFR = envelope following response; FFR = frequency following response.

Speech-in-noise intelligibility thresholds (hypothesis 2)

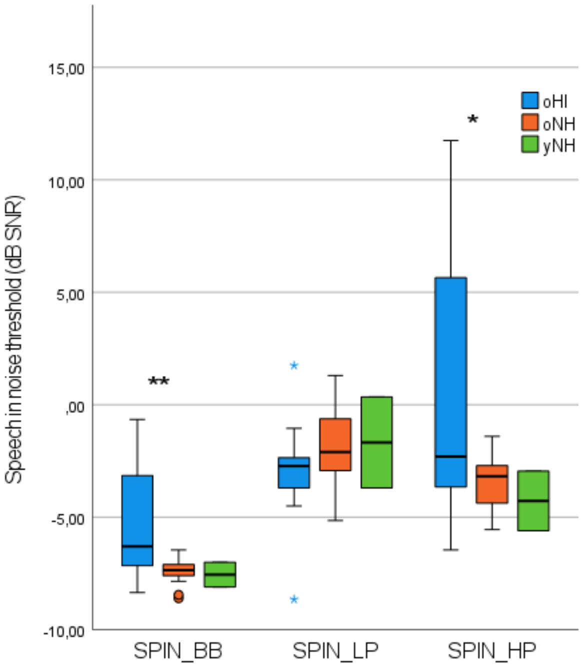


*Figure 6: Boxplots of the speech in noise (SPIN) thresholds for each filter condition (broad band (BB), low pass (LP) and high pass (HP)), split up for the test groups. oHI = older hearing impaired; oNH = older normal hearing; yNH = younger normal hearing. Lower thresholds represent better speech in noise intelligibility. → Sign: SPIN_BB**, SPIN_HP**

Because SPIQ_LP strongly correlates with Audio_LF ($r=0.87$; $p<0.001$), the audiometric thresholds will possibly be, next to cognitive factors, the only parameter influencing the SPIQ_LP results (see Appendix 4). Therefore, we will only use Audio-LF and the SPIN-conditions for further analysis. The results of the SPIN measurements are visualized in Figure 6.

Contrary to the findings of Garrett et al. (2020), significant differences across the two groups were found for SPIN_BB ($t=3.08$; $p=0.008$) and SPIN_HP ($t=2.33$; $p=0.037$). Significant correlations were found between these thresholds and the audiometric thresholds at 4000Hz in the oHI group (SPIN_BB ($r=0.72$; $p=0.003$) and SPIN_HP ($r=0.66$; $p=0.015$)). DP_4000Hz correlated significantly to SPIN_HP ($r=0.61$; $p=0.013$). SPIN_LP showed no significant differences between oHI and oNH test groups ($t=-1.62$; $p>0.05$) and no correlation with Audio_4000Hz or Audio_LF ($p>0.05$). These results suggest that thresholds in SPIN_BB and SPIN_HP conditions are influenced by hearing loss in the participants we selected (with high frequent slopes), whereas thresholds in SPIN_LP are not. A negative correlation was found between SPIN_BB and age, but only in oHI participants ($r=-0.56$; $p=0.024$). Audio_4000Hz could be a confounding factor as a (probably false) negative correlation was also seen between Audio_4000Hz and age.

Also contrary to the study of Garrett et al. (2020) with the German Matrix sentence-test, scores on SPIN_BB were significantly better than SPIN_LP ($t=-9.28$; $p<0.001$) and SPIN_HP ($t=-7.87$; $p<0.001$). SPIN_BB correlated significantly with SPIN_HP ($\rho=0.54$; $p=0.002$), especially for the oHI-group ($\rho=0.89$; $p<0.001$), not for the oNH-group ($\rho=-0.20$; $p>0.05$). For SPIN_LP, nonsignificant correlations were found with SPIN_BB for each test group apart and with SPIN_HP only for the oHI ($p>0.05$, see

Table 3).

Table 3: Correlation coefficients of speech in noise with age, audiometry, OAE, ABR and between the different SPIN-conditions.

	SPIN broad band			*SPIN low pass*			*SPIN high pass*		
	All	*oHI*	*oNH*	*All*	*oHI*	*oNH*	*All*	*oHI*	*oNH*
Age	-0.10	-0.56*	-0.23	-0.12	-0.14	-0.15	-0.09	-0.34	-0.19
Audio_4000Hz	0.52**	0.72**	-0.35	-0.28	-0.06	0.02	0.31	0.66*	-0.34
Audio_LF	0.02	-0.18	-0.34	-0.30	-0.35	0.11	0.12	-0.31	0.29
HFA	0.49**	0.17	-0.01	-0.19	0.04	0.17	0.34	0.46	-0.30
DP_4000Hz	-0.24	0.01	-0.07	0.16	-0.22	0.17	0.21	0.21	0.61*
ABR latency wave I	-0.08	0.01	-0.03	0.05	-0.28	0.10	-0.04	0.11	-0.12
ABR latency wave V	-0.29	-0.25	0.08	0.33	0.27	0.21	-0.22	-0.25	-0.03
ABR amplitude wave I	0.12	-0.09	0.12	-0.33	-0.25	-0.37	-0.02	-0.04	-0.02
ABR amplitude wave V	0.31	0.29	-0.10	-0.05	0.29	-0.12	0.09	0.45	-0.16
SPIN broad band	1.00	1.00	1.00	0.21	0.46	0.48	0.54**	0.89**	-0.20
SPIN low pass				1.00	1.00	1.00	0.26	0.50	0.15
SPIN high pass							1.00	1.00	1.00

Significant correlations were indicated with * for $p<0.05$ and ** for $p<0.01$. SPIN = speech in noise; oHI = older hearing impaired; oNH = older normal hearing; Audio_4000Hz = audiometric thresholds on 4000 Hz; Audio_LF = audiometric thresholds low frequencies; HFA = high frequency audiometry; DP_4000Hz = distortion product otoacoustic emissions on 4000 Hz; ABR = auditory brainstem response.

S(T)M detection thresholds and relations to SPIN (hypothesis 3)

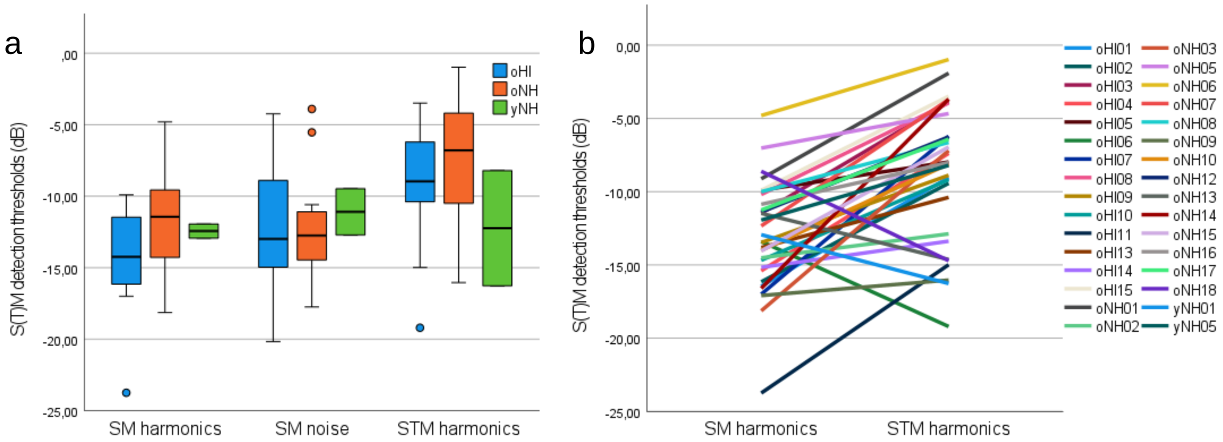


Figure 7: (a) Boxplots of the spectral (temporal) modulation (S(T)M) detection thresholds for each condition (SM harmonics, SM noise and STM harmonics), split up for the test groups. Lower thresholds mean better detection ability. (b) Line graphs of each participant between SM harmonics and STM harmonics. oHI = older hearing impaired; oNH = older normal hearing; yNH = younger normal hearing.

A significant difference was found between SM_harm and STM_harm (*t*=-5.86; *p*<0.001), participants showed better thresholds for SM_harm than STM_harm, contrary to the findings of Miller et al. (2018). This was not the case for two oNH, one oHI and one yNH participant (see Figure 7b). Also, significant moderate correlations were found between the three conditions (see Table 4). SM_harm and STM_harm, shown in Figure 8, were also correlated in the study of Miller et al. (2018). As these correlations could be influenced by confounding factors (described by Borjigin et al. (2022)), also the difference between SM_harm and SM noise and between STM_harm and SM_harm was calculated in order to reduce the influence of nonsensory factors. The subtraction of SM_harm and SM_noise correlated with SM_harm (*r*=0.43; *p*=0.018) as also did the subtraction of STM_harm and SM_harm with STM_harm (*r*=0.64: *p*<0.001) and with SM_harm (*r*=-0.41; *p*=0.026) (see Appendix 5). We also compared the S(T)M detection thresholds and SPIN results between participants who performed STM first with the ones who first did the SPIN test, in order to check the effect of concentration/attention. Only for SM_noise (*t*=2.10; *p*=0.046) and SPIN_LP (*t*=2.75; *p*=0.010) significant differences were found, where the participants who did the SPIN test first, scored better on both SM_noise and SPIN_LP, so these differences are probably found by chance.

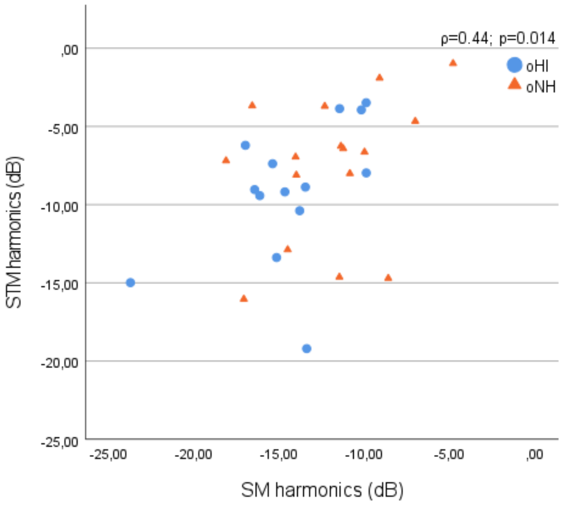


Figure 8: Scatter plot of thresholds with STM harmonics in relation to thresholds with SM harmonics for oHI (blue dots) and oNH (orange dots) participants. SM = spectral modulated; STM = spectrotemporal modulated; oHI = older hearing impaired; oNH = older normal hearing.

Confirming our hypothesis based on the similar audiometric thresholds in the low frequencies, no significant differences between oHI and oNH were found for all S(T)M detection threshold tests (*p*>0.05, see Table 4 and Figure 7). No correlations were found between our markers and audiometric thresholds (*p*>0.05), except between the thresholds on 4000 Hz and SM_noise for oNH (*r*=0.55; *p*=0.026) and between HFA and SM_harm for all subjects (*r*=-0.39; *p*=0.036). The last correlation is negative, which means that subjects with worse and higher hearing thresholds scored better on the SM_harm detection task. For the oNH-group, DP_4000Hz significantly correlates with all three test conditions (SM_harm: *r*=-0.66 (*p*=0.006); SM_noise: *r*=-0.62 (*p*=0.010); STM_harm: *r*=-0.63 (*p*=0.008)) (see Figure 9). This would mean that OHC-loss on 4000 Hz affected the S(T)M detection thresholds.

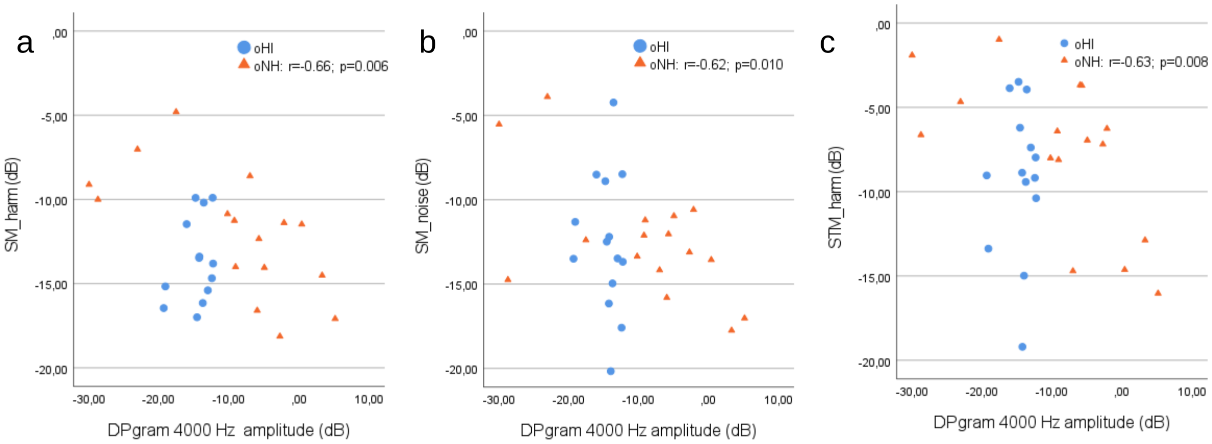


Figure 9: Scatter plots of the S(T)M detection thresholds ((a) SM_harm, (b) SM_noise and (c) STM_harm) in relation to the DPgram amplitude on 4000 Hz for oHI (blue dots) and oNH (orange dots) participants. SM = spectral modulated; STM = spectrotemporal modulated; DP = distortion product; oHI = older hearing impaired; oNH = older normal hearing.

SPIN_BB shows no correlations with the S(T)M detection thresholds. Although there is a trend in the oHI group, as in the study of Miller et al. (2018), but these are not significant (*p*>0.05; SM_harm *r*=0.41; SM_noise *r*=0.39; STM_harm *r*=0.36). Also for SPIN_LP and SPIN_HP, no correlations were found for the whole group (p>0.05). Trends of positive correlations can be seen in the oHI group with SM_noise and STM_harm but these are again not significant (see Table 4 and Appendix 5). Significant negative correlations are seen in the oNH group between SPIN_LP and SM_harm (*r*=-0.53; *p*=0.034), between SPIN_HP and SM_harm (*r*=-0.51; *p*=0.042), and between SPIN_HP and STM_harm (*r*=-0.60; *p*=0.014) (see Appendix 5). This would mean that participants who are better at detecting spectral harmonics score worse at the filtered SPIN tests. Also, no significant correlations were found between the subtraction of the different SM-conditions and SPIN (*p*>0.05).

Table 4: Correlation coefficients of the S(T)M-detection threshold test with age, audiometry, OAE, ABR, SPIN and between the different S(T)M-conditions.

	SM harmonics			***SM noise***			***STM harmonics***		
	All	*oHI*	*oNH*	*All*	*oHI*	*oNH*	*All*	*oHI*	*oNH*
Age	-0.05	-0.04	0.29	-0.02	-0.15	0.17	0.06	0.22	0.17
Audio_4000Hz	-0.10	0.11	0.38	0.11	0.12	0.55*	-0.22	-0.20	0.10
Audio_LF	-0.07	0.23	-0.11	0.16	0.21	0.16	-0.02	0.27	-0.24
HFA	-0.39*	-0.42	0.01	-0.05	-0.25	0.20	-0.05	0.11	0.19
DP_4000Hz	-0.30	0.19	-0.66**	-0.40*	-0.16	-0.62*	-0.37*	-0.15	-0.63**
ABR latency wave I	-0.26	0.07	-0.46	-0.11	0.04	-0.17	-0.23	-0.35	-0.14
ABR latency wave V	-0.15	-0.17	0.18	-0.01	0.26	-0.22	-0.26	-0.17	-0.22
ABR amplitude wave I	0.11	-0.42	0.19	-0.18	-0.53	-0.05	0.04	-0.01	-0.01
ABR amplitude wave V	-0.10	-0.38	-0.23	-0.05	-0.25	0.05	0.11	-0.17	0.19
SPIN broad band	-0.12	0.41	-0.16	0.21	0.39	0.05	0.07	0.36	0.33
SPIN low pass	0.02	0.29	-0.53*	0.29	0.51	0.02	0.25	0.44	-0.04
SPIN high pass	-0.20	0.18	-0.51*	0.09	0.30	-0.23	-0.17	0.39	-0.60*
EFR envelope	0.13	0.24	-0.28	-0.05	0.04	-0.14	0.24	0.46	-0.02
SM harmonics	1.00	1.00	1.00	0.63**	0.80**	0.51*	0.44*	0.51	0.35
SM noise				1.00	1.00	1.00	0.54**	0.53	0.56*
STM harmonics							1.00	1.00	1.00

Significant correlations were indicated with * for *p*<0.05 and ** for *p*<0.01. SM = spectral modulated; STM = spectrotemporal modulated; oHI = older hearing impaired; oNH = older normal hearing; Audio_4000Hz = audiometric thresholds on 4000 Hz; Audio_LF = audiometric thresholds low frequencies; HFA = high frequency audiometry; DP_4000Hz = distortion product otoacoustic emissions on 4000 Hz; ABR = auditory brainstem response; SPIN = speech in noise; EFR = envelope following response.

EFR and FFR results, their potential relationships and relations to SPIN (hypothesis 4)

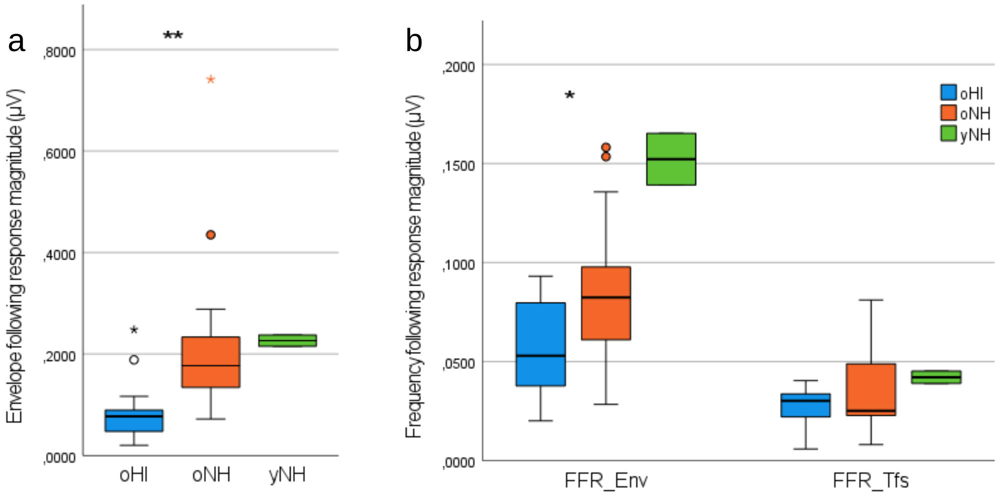


Figure 10: Boxplots of the amplitudes of (a) the envelope following response (EFR) and (b) frequency following response envelope (FFR_Env) and frequency following response temporal fine structure (FFR_Tfs), split up for the test groups. oHI = older hearing impaired; oNH = older normal hearing; yNH = younger normal hearing. → Sign: EFR_Env and FFR_Env

Differences between oHI and oNH were significant for EFR_Env (*U*=919; *p*<0.001) and FFR_Env (*t*=-2.61; *p*=0.014), not for FFR_Tfs (*p*>0.05) (see Figure 10). In addition, FFR_Env correlated with the audiometric thresholds at 4000 Hz (*ρ*=-0.49; *p*=0.006, see Figure 11a), while no correlation is found with the audiometric thresholds in low frequencies (*p*>0.05), confirming our expectations in hypothesis 4. This might suggest that FFR_Env is influenced by high frequency hearing loss, despite the low-frequency content of the stimulus. For the oNH group, FFR_Tfs shows nonsignificant correlations with audiometric thresholds on 4000 Hz (*r*=-0.39; *p*>0.05) and DPgram amplitudes on 4000 Hz (*r*=0.41; *p*>0.05) and significant correlations with low-frequency audiometry thresholds (*r*=-0.51; *p*=0.044, see Figure 11b). As seen in Table 5, also the correlations with ABR show an effect of neural hearing loss on FFR_Env and less (no significant correlations) on FFR_Tfs.

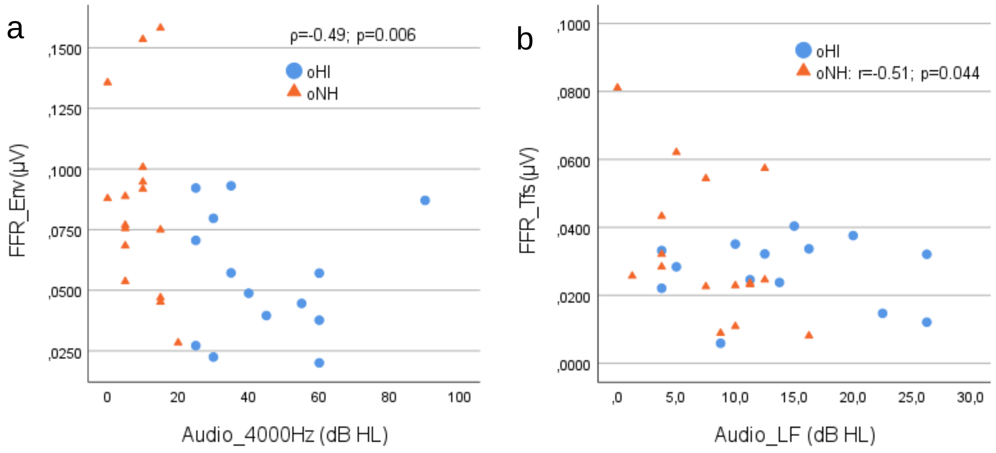


Figure 11: Scatter plots of the significant correlations between the frequency following response (FFR) magnitudes in relation to audiometric thresholds and age. (a) FFR_Env with audiometric thresholds on 4000 Hz (Audio_4000Hz) and (b) FFR_Tfs with audiometric thresholds in the low frequencies (mean 125 to 1000 Hz; Audio_LF) all for oHI (blue dots) and oNH (orange dots) participants. Env = envelope; Tfs = temporal fine structure; Audio_LF = audiometric thresholds low frequencies; Audio_4000Hz = audiometric thresholds on 4000 Hz; oHI = older hearing impaired; oNH = older normal hearing.

Whereas we expected correlations with age for both FFR_Tfs and FFR_Env, only FFR_Tfs correlated significantly with age for both test groups (*ρ*=-0.43; *p*=0.017), while this effect was not seen with FFR_Env (*ρ*=-0.17; *p*>0.05) (see Appendix 5). FFR_Env and FFR_Tfs significantly correlate with each other (*ρ*=0.47; *p*=0.009), which is contrary to the hypothesis. FFR_Env strongly correlated with EFR_Env (*ρ*=0.75; *p*<0.001, see Figure 12), while FFR_Tfs did not (*ρ*=-0.14; *p*>0.05). This would mean that the ability to detect envelopes of temporal or spectral modulations relies on a similar mechanism and that the FFR_Tfs is not affected by cochlear synaptopathy. Significant correlations with EFR_Env and other parameters can be seen in Table 5, but are less important for our research question.

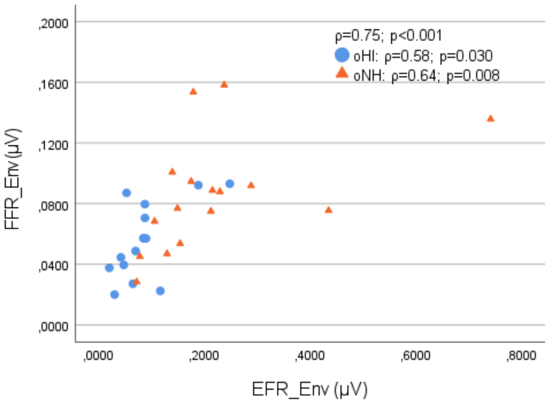


Figure 12: Scatter plot of the frequency following response magnitude envelope (FFR_Env) in relation to the envelope following response (EFR_Env), for oHI (blue dots) and oNH (orange dots) participants. Env = envelope; Tfs = temporal fine structure; oHI = older hearing impaired; oNH = older normal hearing.

Significant correlations between FFR_Env and SPIN were not found (*p*>0.05, see Table 5). For FFR_Tfs, there is a significant correlation with SPIN_LP (*ρ*=0.39; *p*=0.033) and nonsignificant correlations with SPIN_BB when test groups were split up (*p*>0.05; oHI: *r*=0.33; oNH: *r*=0.46). These are positive correlations, which means that participants with better FFR_Tfs, scored worse on the SPIN_LP and SPIN_BB (see Figure 13). In our hypothesis, we expected that FFR_Tfs and FFR_Env would correlate with SPIQ_LP, as TFS cues are used. As the variation in SPIQ_LP could be declared by the low-frequency audiometry (see Appendix 4), the correlation between FFR_Tfs and Audio_LF confirms our hypothesis.

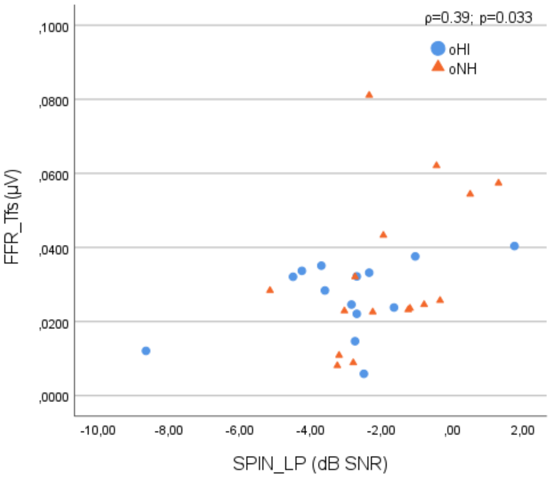


Figure 13: Scatter plot of the frequency following response magnitude for temporal fine structure (FFR_Tfs) in relation to speech in noise low pass (SPIN_LP), for oHI (blue dots) and oNH (orange dots) participants. oHI = older hearing impaired; oNH = older normal hearing.

Table 5: Correlation coefficients of the frequency following response (FFR) with age, audiometry, OAE, ABR, SPIN, S(T)M detection thresholds and between FFR_Env and FFR_Tfs.

	FFR envelope			***FFR temporal fine structure***			***EFR envelope***		
	All	*oHI*	*oNH*	*All*	*oHI*	*oNH*	*All*	*oHI*	*oNH*
Age	-0.17	0.24	-0.25	-0.43*	-0.44	-0.56*	-0.18	0.57*	-0.32
Audio_4000Hz	-0.49**	-0.19	-0.26	-0.11	0.13	-0.39	-0.73**	-0.59*	-0.47
Audio_LF	-0.24	0.14	-0.25	-0.20	-0.08	-0.51*	-0.13	0.59*	-0.33
HFA	-0.28	0.03	0.26	-0.03	-0.14	0.11	-0.60**	0.05	-0.53*
DP_4000Hz	0.24	-0.18	0.13	0.21	-0.34	0.41	0.50**	-0.22	0.37
ABR latency wave I	-0.41*	-0.29	-0.20	0.08	0.16	0.03	0.19	0.24	-0.26
ABR latency wave V	-0.72**	-0.63*	-0.68**	-0.10	0.29	-0.13	-0.44*	-0.17	-0.50
ABR amplitude wave I	0.13	0.51	-0.24	0.11	0.23	0.11	0.48*	0.11	0.61*
ABR amplitude wave V	0.57**	-0.03	0.71**	0.31	0.20	0.20	-0.68**	-0.36	-0.62*
SPIN broad band	-0.25	0.01	0.12	0.19	0.33	0.46	-0.39*	-0.32	0.09
SPIN low pass	0.23	-0.05	0.27	0.39*	0.44	0.45	0.34	0.16	0.31
SPIN high pass	-0.08	-0.03	<0.01	-0.05	0.09	0.10	-0.10	-0.29	0.25
SM harmonics	-0.02	0.03	-0.32	-0.25	0.01	-0.45	0.13	0.24	-0.28
SM noise	-0.13	-0.36	-0.01	-0.08	-0.06	-0.36	-0.05	0.04	-0.14
STM harmonics	0.20	0.14	0.15	0.10	-0.11	-0.01	0.24	0.46	-0.02
FFR envelope	1.00	1.00	1.00	0.47**	0.40	0.42	0.75**	0.58*	0.64**
FFR temporal fine structure				1.00	1.00	1.00	0.14	0.10	0.33
EFR envelope							1.00	1.00	1.00

Significant correlations were indicated with * for *p*<0.05 and ** for *p*<0.01. FFR = frequency following response; oHI = older hearing impaired; oNH = older normal hearing; Audio_4000Hz = audiometric thresholds on 4000 Hz; Audio_LF = audiometric thresholds low frequencies; HFA = high frequency audiometry; DP_4000Hz = distortion product otoacoustic emissions on 4000 Hz; ABR = auditory brainstem response; SPIN = speech in noise; EFR = envelope following response; SM = spectral modulated; STM = spectrotemporal modulated.

Relationships between FFR and psychophysical detection thresholds of SM stimuli (hypothesis 5)

Between FFR and SM detection thresholds, no significant correlations were found (*p*>0.05), contrary to our hypothesis and the findings of Parthasarathy et al. (2020). Because the redundant same stimuli were used, a correlation between SM_harm and FFR was expected. SM_harm shows nonsignificant correlations with both FFR_Env (*r*=-0.32) and FFR_Tfs (*r*=-0.45) when only the oNH group was taken into count. Also, SM_noise shows nonsignificant correlations with FFR_Tfs for the oNH group (*r*=-0.36) and with FFR_Env for the oHI group (*r*=-0.36). The correlations are also visualized in Appendix 5. Also, no correlations were found with the subtraction of SM_harm and SM_noise and the two FFR_conditions (*p*>0.05).

Discussion

S(T)M detection threshold test

We hypothesized that the ability to detect spectral modulations would not show a significant difference between oNH and oHI due to similar hearing thresholds in the low frequencies. The S(T)M detection threshold task was split up in three conditions (SM_harm, SM_noise and STM_harm). In none of them, significant differences between oNH and oHI participants were found and only the SM_noise condition showed a correlation with the audiometric thresholds on 4000 Hz for the oNH group. Bernstein et al. (2013) found significant better STM-detection thresholds in their NH than HI group for a broad band noise with low temporal rates (4 or 12 Hz) and high spectral densities (2 or 4 c/o). These findings were explained with the reasoning that the ability to use TFS cues was affected in the HI group, which could be seen in the low rate condition, but that both HI and NH participants would have difficulties detecting the TFS cues in the high rate condition because both had to rely on temporal cues. We only measured the low temporal rate condition in the low frequency range and found no differences, what would mean that the ability to use TFS with low temporal rates and high spectral rates would be affected by hearing thresholds (high frequencies in their study and low frequencies in ours). Another difference with our study, is that we made a distinction between oNH and yNH, while these were taken together in the study of Bernstein et al. (2013) (NH: 24-60 yrs.; HI: 70-87 yrs.).

In the boxplot of our S(T)M detection results (Figure 7) could be seen that one baseline-yNH scored better for the STM_harm task in comparison to the SM_harm thresholds, what could mean that age has an impact on the detection thresholds. As this is only the case for the condition with temporal modulations, we could hypothesize that the temporal ENV could also explain the differences found by Bernstein et al. (2013). The differences in coding of the temporal cues could be related to cochlear synaptopathy. A whole yNH group should be included in order to confirm or refute this hypothesis.

Reverse on the study of Bernstein et al. (2013), we could see in Figure 7 that the oHI group seems to had better detection thresholds than the oNH group for the SM-harm condition (although this was not statistically significant), which was not the case for the SM_noise condition. Such enhanced thresholds in hearing impaired participants was also seen in the study of Anderson et al. (2013), where the FFR_Env showed bigger responses for HI participants, which was not the case for FFR_Tfs. There were two possible explanations proposed: (i) change of levels of neurotransmitters in the inferior colliculus or (ii) greater broadband activity. Nevertheless, these interpretations should be taken with caution, as the FFR in the study of Anderson et al. (2013) was presented at sensation level, to compensate for hearing loss. They also concluded that the spectral ENV response was more sensitive to amplification, while TFS was less, what could also explain the differences exhibited in their study. In our study, oNH and oHI were presented to the same stimulus intensity (75 dB SPL), making high-frequency hearing loss the only distinguishing factor.

While the audiometric thresholds show no correlations with the S(T)M detection thresholds, the DPgrams on 4000 Hz did correlate with all three conditions. This would mean that OHC-loss in the high frequencies has an impact on the spectral modulation detection in the low frequencies. Bernstein et al. (2013) also found correlations between the STM detection thresholds based on low-frequency sounds and frequency selectivity in the high frequencies (which is a typical consequence of OHC-loss). In a study of Wu et al. (2019) on temporal bones, a similar degeneration of OHCs was seen across all audiometric frequencies (low and high) by ageing, which is in contrast to the high-frequency threshold shift on the audiogram. This would mean that the OAEs on 4000 Hz could also represent the OHC loss in low frequencies, which would impact the S(T)M detection thresholds.

Miller et al. (2018) investigated the S(T)M detection threshold in participants with hearing loss. They found significant better thresholds for the STM-condition (moving ripples) than the SM-condition (static ripples), while we saw the opposite for our oHI and oNH group (29 – 79 yrs.). What we should take in account, is again the age-factor. Miller et al. (2018) used both yHI and oHI participants in their study without comparing the STM results and correlations in both groups. Moreover, the study of Miller et al. (2018) only used participants with higher audiometric thresholds (mean: 56.6 dB on 4 kHz) who used hearing aids, while we had only two hearing aid users in our HI group. Further studies should investigate whether hearing aid use would have an impact on ability to detect spectral or temporal modulations.

With our findings (worse STM than SM detection in all older participants), we could hypothesize that the inclusion of a temporal envelope, made it more difficult to detect TFS information for the older group, while this is not the case for one of the yNH participants. Clinard and Cotter (2015) also concluded that dynamic frequencies rather that static ones are influenced by age-effects, also in low frequencies and that oNH adults have degraded phase-locked neural representations of dynamic frequencies. A possible explanation could be cochlear synaptopathy, nevertheless, no significant correlation between STM_harm and EFR was found. A possible explanation could be related to the fact that the EFR had a high frequency carrier (4000 Hz), while the STM had a low-frequency carrier. Also the effect of OHC_loss we saw on the STM-detection thresholds and not on the EFR response (see Vasilkov et al. (2021)), could be consistent with this absence of a correlation. Another factor could be concentration, as the STM detection threshold is a behavioral task, while EFR is purely objective. Since the STM-stimulus closely resembles the speech signal, a lot of factors could interact for the final response (effects of audiometric hearing loss, OHC-loss, age…). Again, a larger sample of yNH listeners could add more information about age-related influences on SM and STM detection.

Finally, Miller et al. (2018) found significant correlations between STM and SM ripples in their HI-group, what would mean that their detection is driven by the same mechanisms. This finding could also be seen in our older group between STM_harm and SM_harm, but only when the whole group was considered, not when oHI and oNH individuals are considered apart. When we take a look at Figure 8, the correlation can be seen especially for subjects with worse STM detection thresholds (> -10 dB). Subjects with better results on STM (< -10 dB), showed wider spread in the SM detections thresholds. This could mean that the temporal cue dominated when the participant was able to detect it. Also between the subtraction of STM_harm and SM_harm, where the temporal aspect is the only factor that is different, a correlations was found with both STM_harm (meaning that there is less influence by nonsensory factors (see below)) and SM_harm (suggesting that the two variables are possibly driven by the same underlying mechanisms).

Recently, Borjigin et al. (2022), studied the effect of extraneous factors on both the behavioral and electrophysiological TFS-tests. Nonsensory factors (attention, concentration…) would account for more than half of the variance seen in behavioral thresholds, making it not clear whether the SM-detection threshold is actually sensitive for TFS-cues, as stated in on other studies (Bernstein et al., 2016; Bernstein et al., 2013; Miller et al., 2018). They subtracted the influence of nonsensory factors (lapse rate in catch trials) from their behavioral scores and found no correlations anymore with the TFS. We did a similar thing by subtracting SM_harm with SM_noise, which are assumed to be affected by the same cognitive factors and envelope components, but would differ in the amount of TFS cues. As we still found correlations with SM_harm, we could assume that there is no big effect of cognitive factors. Also, no significant order-effect was found for the S(T)M detection thresholds and SPIN test. The same was found between STM_harm and SM_harm. Another way Borjigin et al. (2022) used to consider the influence of attention on their TFS measurement, was including another behavioral task that should not be affected by TFS-processing (an AM detection test), where they found significant correlations, probably due to nonsensory factors. In our study, we did not found correlations between the S(T)M detection thresholds and SPIN test, despite that they were performed at the same day, on the same place, with the same equipment. This would suggest that there was no influence of cognitive factors on both tests (nevertheless, it is also possible that only one of the two was influenced by cognition). Another factor that could influence the results and correlations was the use of estimated thresholds as not all responses converged (see above).

Another TFS-sensitive test that we could have included to check the sensitivity of our SM detection threshold task, is the TFS1 (Sęk & Moore, 2012). Nevertheless, we had not enough time for each participants to include this test in our protocol. Another way to test for TFS processing capacities would be by measuring FM thresholds with low-frequency carriers (Borjigin et al., 2022; Parthasarathy et al., 2020).

FFR

In this study, significant group differences between oNH and oHI were found for the FFR_Env and not for FFR_Tfs. Also the correlations with the audiometric thresholds on 4000 Hz confirms other studies where the ENV was more affected by high-frequency hearing loss based on the audiogram than the TFS for FFR with syllable-stimuli (Ananthakrishnan et al., 2016; Märcher-Rørsted et al., 2022; Ruggles et al., 2012). As earlier discussed, Anderson et al. (2013) found the opposite (enhanced FFR_Env for HI). It should be taken into account that they presented the stimuli at sensation level in order to compensate for hearing loss. Ananthakrishnan et al. (2016) did the same thing and both studies showed that the FFR_Env was more affected by increased stimulus intensity than FFR_Tfs. This would again mean that the spectral ENV is more affected by audibility, while the TFS detection would better be explained by the neural representation. While correlations with audiometric thresholds on 4000 Hz were found, neither FFR_Env or FFR_Tfs were correlated with DPgrams on 4000 Hz. This would mean that not the OHC-loss influenced the differences in FFR_Env, but possibly loss of high-frequent IHCs or synapses of the ANFs.

The fact that high frequency audiometric thresholds can affect the FFR_Env, despite a low-frequency carrier, is also discussed in other studies (Ananthakrishnan et al., 2016; Märcher-Rørsted et al., 2022). Ananthakrishnan et al. (2016) speaks of a disruption in the temporal pattern of neural activity that results from a peripheral hearing loss, which then has adverse effects on neural timing (and synchronization of neural activity) that is cumulative along the auditory neuroaxis. This could refer to cochlear synaptopathy, which affects the synapses in the apex of the basilar membrane and could influence the travelling wave and the neural timing on the basis, where low frequencies are detected. Märcher-Rørsted et al. (2022) also searched for an explanation of differences in pure-tone FFRs between oNH and yNH, despite similar audiograms in the low frequencies. They hypothesize that post-synaptical effects of ANF-loss (more at brainstem level) would declare their findings.

A strong correlation was found between FFR_Env and EFR_Env and a similar trend could be seen between the test groups in Figure 10 (small difference between oHI and oNH, bigger difference with yNH group). This could mean that both envelope recordings measure a similar mechanism, despite other frequencies of the stimulus. The problems with envelope coding in low frequencies, could again be explained by an effect of cochlear synaptopathy in the apex of the basilar membrane that affects the temporal pattern, or by central mechanisms for phase locking that are disturbed as described by Ananthakrishnan et al. (2016). When we compare the EFR with the FFR response, we notice that the FFR magnitude is much lower, despite that the two EEG measures were based on stimuli at the same SPL, this could more easily be corrupted by noise and show less correlations. Compared to the findings of Garrett et al. (2020), only a weak correlation was found between EFR_Env and SPIN_BB. Nevertheless, they only found correlations when also yNH participants were included in the analyzes, not for the older participants only (as in our study group).

While no correlations were found between FFR_Tfs and high frequency audiometric thresholds, significant correlations were found between FFR_Tfs and low-frequency audiometry and age, particularly for the oNH group. Füllgrabe et al. (2015) had the same findings when comparing their yNH group with an oNH group (difference for TFS, not for ENV). Contrary, in our boxplots (Figure 10), the difference with the two yNH participants seems much larger for the FFR_Env, than the FFR_Tfs. The difference in results could be explained by the fact that Füllgrabe et al. (2015) did not use an objective FFR, but a psychoacoustic test. A study that better leans to ours, is the one from Ruggles et al. (2012). They found correlations in age for both FFR_Env and FFR_Tfs for their NH group, including younger and older participants. When splitting up the yNH and oNH participants, they concluded that the yNH more rely on the ENV cues for spatial attention, why their older participants rely on TFS cues. Also Mai et al. (2018) stated that the FFR_Tfs reflects the processing of pitch information. Contrary, Ananthakrishnan et al. (2016) concluded in their study that changes in the FFR (both TFS and ENV) would be more affected by peripheral hearing loss and less by age-related changes. With this reasoning in mind, it would be the audiometric thresholds in the low frequencies (and thus IHC loss or high-SR fibers) that affect FFR_Tfs coding. As the audiogram deteriorates with age, also for low frequencies (correlated significantly), this could be a confounding factor of the correlation we found with age in the oNH group. Also the fact that no significant differences were found between the oNH and oHI group could be declared by the small difference between low-frequency thresholds.

A correlation between FFR_Env and FFR_Tfs was not found in the NH group of Ruggles et al. (2012) suggesting that these would represent other mechanisms. In our study, there was a correlation when both oNH and oHI groups were taken together, but not apart. This would mean that a similar mechanism impacts both factors in older participants. In the study of Ruggles et al. (2012), also younger participants were included, what could have impact on the FFR_Env, but not on FFR_Tfs, resulting in the correlation between the two, as they would rely more on ENV cues. We have to take in account that EEG-related factors could also be a confounding factor of these correlations (noise, head circumference, gender, impedances…), as also discussed in the paper of Borjigin et al. (2022).

Despite that the S(T)M detection threshold task and FFR measurements rely on the exact same stimuli, presented at the same intensity, the results did not show correlations between these measurements. In addition, this would mean that the 50% and 100% modulated FFRs could not explain the small threshold differences in modulation depths of the psychoacoustic measurement and vise-versa. Also a cognitive factor (attention, concentration…) could still play a role in the SM-detection, while the FFR was an objective test. The SM-harm detection threshold correlated with OHC-damage (DPgrams) at 4000Hz, but not with audiometric thresholds (IHC-damage or high-SR ANFs) at 4000 Hz or EFR (cochlear synaptopathy; low-SR fibers) on 4000Hz. Contrary, FFR_Env did not correlate with OHC-loss at 4000 Hz, but did with cochlear synaptopathy and FFR_Tfs with the IHCs/high-SR fibers in low frequencies. This could mean that robust detecting of spectral modulations in the 50% and 100% modulated FFR_Env, more depends on IHC and ANF functioning, while the detection of small spectral differences would more rely on the frequency-selectivity based on OHC-functioning. Parthasarathy et al. (2020) showed a correlation between their SM detection thresholds an objective FFR with the same stimuli. Nevertheless, it is not clear which age-group is used in the study. Based on the HFA (where the mean of the 23 participants is better than 20 dB HL until 16 kHz), we could suggest that these participants are younger than our oNH test group and/or have better hearing thresholds. As discussed earlier, younger participants would rely more on ENV ques than older subjects, who would rely on the TFS cues (Ruggles et al., 2012), maybe the spectral ENV cues were dominant in the study Parthasarathy et al. (2020). Again, the effect of age could be analyzed by adding a yNH test group to our study.

Relation with speech in noise intelligibility

In our results, SPIN_BB and especially SPIN_HP are affected by high-frequency hearing loss and SPIN_HP also by OHC-loss on 4000 Hz. Also Bernstein et al. (2016) concluded that audibility partly explains variance in SPIN with hearing aids. Contrary, Garrett et al. (2020) showed no significant differences between oNH and oHI participants on SPIN without hearing aids. Their oHI participants had lower hearing thresholds on 4000 Hz (max. 50 dB), what could explain why they found no significant effects of hearing loss. This would mean that SPIN can be influenced by both IHC- and OHC-loss.

Contrary to other studies (Bernstein et al., 2016; Bernstein et al., 2013; Davies-Venn et al., 2015; Miller et al., 2018), we found no significant positive correlations between S(T)M detection thresholds and SPIN performance, only some nonsignificant trends were seen. It should be noted that two of the studies measured SPIN with hearing aids (Bernstein et al., 2016; Miller et al., 2018) and that Bernstein et al. (2013) compensated for hearing loss by presenting the sentences at higher intensity (92 dB HL). Davies-Venn et al. (2015) concluded that STM could predict SPIQ and SPIN in amplified and unamplified conditions, but also in the unamplified conditions, the intensity level was based on audiometric thresholds. Moreover, they did not take the age-factor into account, while the test groups had large ranges of ages (HI: 21 – 89 yrs. and NH: 23 – 31 yrs.). We did find negative correlations for the oNH group, especially with SM_harm, what would mean that participants with worse SPIN scores, score better on the SM_harm detection test. This could have the same explanation as why the oHI group seemed to score better at the SM_harm task than the oNH group (levels of neurotransmitters or greater broadband activity).

Also for the FFR, no correlations with speech intelligibility were found. Mai et al. (2018) showed a correlation between FFR_Env and SPIN, but also in this study, SPIN was adapted based on audiometry. The SPIN results should be compensated for hearing thresholds in order to make conclusions about the correlations with S(T)M detection thresholds and FFR.

Clinical applications and future studies

In our findings, the ENV seems to provide the most information about the hearing status of older participants. Literature shows evidence that cochlear synaptopathy would affect that envelope coding. Further research should be conducted to compare the effects of low frequent and high frequent modulations, in order to understand more underlying pathophysiology of the FFR. The TFS component seems to be less useful for the clinical practice, nevertheless, a correlation with audiometric thresholds in the low frequencies was found. If the FFR_Tfs could tell us more about IHC/high-rate ANFs, we would expect correlations with higher audiometric thresholds if a high frequent harmonic would be presented. If this is the case, the FFR with harmonics could give information about low- and high-SR ANFs with the FFR_Env and FFR_Tfs respectively. An advantage of using harmonics instead of syllables (see Aiken and Picton (2008), Ruggles et al. (2012) and Ananthakrishnan et al. (2016)), is that the ENV and TFS components could be more split up. Aiken and Picton (2008) mentioned an overlap between ENV and TFS cues over their formants of syllables due to cochlear nonlinearity.

In contrast to previous studies, the SM detection thresholds seem to be less useful in order to predict speech intelligibility. A pitfall of our study is that SPIN and S(T)M detection thresholds are measured at the same intensity for all participants. Tree participants with the worse high-frequency thresholds mentioned that the speech was not hearable at the start level of SPIN_HP, which could mean that our start intensity (71 dB HL) was too low to have trustable results. Compensation of reduced audibility could be done by giving the start intensity at sensation level, based on the tonal audiogram or by giving all HI participants the same, higher intensity. Also the S(T)M detection threshold task and EEG could be presented at sensation level. Another aspect that is already discussed in the literature, is the type of noise used for the SPIN test. Several studies showed that TFS cues are more important in noise that consist of multiple speakers instead of a simple white noise (Bharadwaj et al., 2019; Ruggles et al., 2012; Zeng et al., 2005). Also, the use of speech digits instead of the sentences of the Matrix test, could result in fewer cognitive factors that influence our results (Parthasarathy et al., 2020; Ruggles et al., 2012).

Further studies should consist of larger sample sizes and add a yNH and if possible yHI test group. Younger participants could give information about which parameters would be affected by cochlear synaptopathy or other age-related changes. Especially for the SM_harm and STM_harm, a larger yNH group would be interesting to see if they score better in the STM_harm condition. If this is the case, the addition of a temporal envelope would have negative effects on TFS coding in older participants, which could be important for the optimization of hearing aids.

Conclusion

The SM detection thresholds showed less useful results for use in the clinical practice as these did not correlate with SPIN and would be affected by OHC-loss. Nevertheless, a compensation for audiometric thresholds for the SPIN test could result in significant correlations and give more information about underlying mechanisms influencing speech intelligibility. Adding a temporal modulation to the SM harmonics made the detection of spectral modulations possibly more difficult in older participants with and without hearing loss.

The FFR of a harmonic carrier carrying spectral modulations seems to be a promising objective diagnostic test to differentiate between ENV and TFS coding in participants with and without hearing loss. The envelope component would be affected by cochlear synaptopathy, while the TFS component could be more important to represent threshold detection and is possibly impacted by IHC loss in the measured frequency-range. Possibly, younger participants rely more on envelope cues, therefore, a yNH test group should be added to give more information about age-related deficits.

References

Aiken, S. J., & Picton, T. W. (2008). Envelope and spectral frequency-following responses to vowel sounds. *Hearing Research*, *245*(1-2), 35-47.

Ananthakrishnan, S., Krishnan, A., & Bartlett, E. (2016). Human frequency following response: neural representation of envelope and temporal fine structure in listeners with normal hearing and sensorineural hearing loss. *Ear and hearing*, *37*(2), e91.

Anderson, S., Parbery-Clark, A., White-Schwoch, T., Drehobl, S., & Kraus, N. (2013). Effects of hearing loss on the subcortical representation of speech cues. *The Journal of the Acoustical Society of America*, *133*(5), 3030-3038.

Aronoff, J. M., & Landsberger, D. M. (2013). The development of a modified spectral ripple test. *The Journal of the Acoustical Society of America*, *134*(2), EL217-EL222.

Bernstein, J. G., Danielsson, H., Hällgren, M., Stenfelt, S., Rönnberg, J., & Lunner, T. (2016). Spectrotemporal modulation sensitivity as a predictor of speech-reception performance in noise with hearing aids. *Trends in hearing*, *20*, 2331216516670387.

Bernstein, J. G., Mehraei, G., Shamma, S., Gallun, F. J., Theodoroff, S. M., & Leek, M. R. (2013). Spectrotemporal modulation sensitivity as a predictor of speech intelligibility for hearing-impaired listeners. *Journal of the American Academy of Audiology*, *24*(4), 293-306.

Bharadwaj, H. M., Mai, A. R., Simpson, J. M., Choi, I., Heinz, M. G., & Shinn-Cunningham, B. G. (2019). Non-Invasive Assays of Cochlear Synaptopathy–Candidates and Considerations. *Neuroscience*, *407*, 53-66.

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*, *8*, 26.

Borjigin, A., Hustedt-Mai, A. R., & Bharadwaj, H. M. (2022). Individualized Assays of Temporal Coding in the Ascending Human Auditory System. *Eneuro*, *9*(2).

Bramhall, N., Beach, E. F., Epp, B., Le Prell, C. G., Lopez-Poveda, E. A., Plack, C. J., Schaette, R., Verhulst, S., & Canlon, B. (2019). The search for noise-induced cochlear synaptopathy in humans: Mission impossible? *Hearing Research*, *377*, 88-103.

Clinard, C. G., & Cotter, C. M. (2015). Neural representation of dynamic frequency is degraded in older adults. *Hearing Research*, *323*, 91-98.

Davies-Venn, E., Nelson, P., & Souza, P. (2015). Comparing auditory filter bandwidths, spectral ripple modulation detection, spectral ripple discrimination, and speech recognition: Normal and impaired hearing. *The Journal of the Acoustical Society of America*, *138*(1), 492-503.

Elliott, T. M., & Theunissen, F. E. (2009). The modulation transfer function for speech intelligibility. *PLoS comput biol*, *5*(3), e1000302.

Füllgrabe, C., Moore, B. C., & Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Frontiers in aging neuroscience*, *6*, 347.

Furman, A. C., Kujawa, S. G., & Liberman, M. C. (2013). Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. *Journal of neurophysiology*.

Garrett, M., Vasilkov, V., Mauermann, M., Wilson, J. L., Henry, K. S., & Verhulst, S. (2020). Speech-in-noise intelligibility difficulties with age: the role of cochlear synaptopathy. *bioRxiv*.

Jokitulppo, J., Toivonen, M., & Bjoörk, E. (2006). Estimated leisure-time noise exposure, hearing thresholds, and hearing symptoms of Finnish conscripts. *Military medicine*, *171*(2), 112-116.

Joris, P., Schreiner, C., & Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiological reviews*, *84*(2), 541-577.

Kujawa, S. G., & Liberman, M. C. (2009). Adding insult to injury: cochlear nerve degeneration after “temporary” noise-induced hearing loss. *Journal of Neuroscience*, *29*(45), 14077-14085.

Kummer, P., Janssen, T., & Arnold, W. (1998). The level and growth behavior of the 2 f1− f2 distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss. *The Journal of the Acoustical Society of America*, *103*(6), 3431-3444.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2B), 467-477.

Liberman, M. C., Epstein, M. J., Cleveland, S. S., Wang, H., & Maison, S. F. (2016). Toward a differential diagnosis of hidden hearing loss in humans. *PloS one*, *11*(9), e0162726.

Luts, H., Jansen, S., Dreschler, W., & Wouters, J. (2014). Development and normative data for the Flemish/Dutch Matrix test.

Mai, G., Tuomainen, J., & Howell, P. (2018). Relationship between speech-evoked neural responses and perception of speech in noise in older adults. *The Journal of the Acoustical Society of America*, *143*(3), 1333-1345.

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. *Hearing Research*, *414*, 108411.

Miller, C. W., Bernstein, J. G., Zhang, X., Wu, Y.-H., Bentler, R. A., & Tremblay, K. (2018). The effects of static and moving spectral ripple sensitivity on unaided and aided speech perception in noise. *Journal of Speech, Language, and Hearing Research*, *61*(12), 3113-3126.

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

Parthasarathy, A., & Kujawa, S. G. (2018). Synaptopathy in the aging cochlea: Characterizing early-neural deficits in auditory temporal envelope processing. *Journal of Neuroscience*, *38*(32), 7108-7119.

Plack, C. J., Barker, D., & Prendergast, G. (2014). Perceptual consequences of “hidden” hearing loss. *Trends in hearing*, *18*, 2331216514550621.

Purcell, D. W., John, S. M., Schneider, B. A., & Picton, T. W. (2004). Human temporal auditory acuity as assessed by envelope following responses. *The Journal of the Acoustical Society of America*, *116*(6), 3581-3593.

Ruggles, D., Bharadwaj, H., & Shinn-Cunningham, B. G. (2012). Why middle-aged listeners have trouble hearing in everyday settings. *Current Biology*, *22*(15), 1417-1422.

Sabin, A. T., Eddins, D. A., & Wright, B. A. (2012). Perceptual learning evidence for tuning to spectrotemporal modulation in the human auditory system. *Journal of Neuroscience*, *32*(19), 6542-6549.

Sęk, A., & Moore, B. C. (2012). Implementation of two tests for measuring sensitivity to temporal fine structure. *International Journal of Audiology*, *51*(1), 58-63.

Sergeyenko, Y., Lall, K., Liberman, M. C., & Kujawa, S. G. (2013). Age-related cochlear synaptopathy: an early-onset contributor to auditory functional decline. *Journal of Neuroscience*, *33*(34), 13686-13694.

Shaheen, L. A., Valero, M. D., & Liberman, M. C. (2015). Towards a diagnosis of cochlear neuropathy with envelope following responses. *Journal of the Association for Research in Otolaryngology*, *16*(6), 727-745.

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. *Hearing Research*, *400*, 108132.

Verhulst, S., Ernst, F., Garrett, M., & Vasilkov, V. (2018). Suprathreshold psychoacoustics and envelope-following response relations: Normal-hearing, synaptopathy and cochlear gain loss. *Acta Acustica united with Acustica*, *104*(5), 800-803.

Verhulst, S., Jagadeesh, A., Mauermann, M., & Ernst, F. (2016). Individual differences in auditory brainstem response wave characteristics: relations to different aspects of peripheral hearing loss. *Trends in hearing*, *20*, 2331216516672186.

Verhulst, S., & Warzybok, A. (2018). Contributions of Low-and High-Frequency Sensorineural Hearing Deficits to Speech Intelligibility in Noise. *bioRxiv*, 358127.

Wu, P., Liberman, L., Bennett, K., De Gruttola, V., O'malley, J., & Liberman, M. (2019). Primary neural degeneration in the human cochlea: evidence for hidden hearing loss in the aging ear. *Neuroscience*, *407*, 8-20.

Zeng, F.-G., Nie, K., Stickney, G. S., Kong, Y.-Y., Vongphoe, M., Bhargave, A., Wei, C., & Cao, K. (2005). Speech recognition with amplitude and frequency modulations. *Proceedings of the National Academy of Sciences*, *102*(7), 2293-2298.

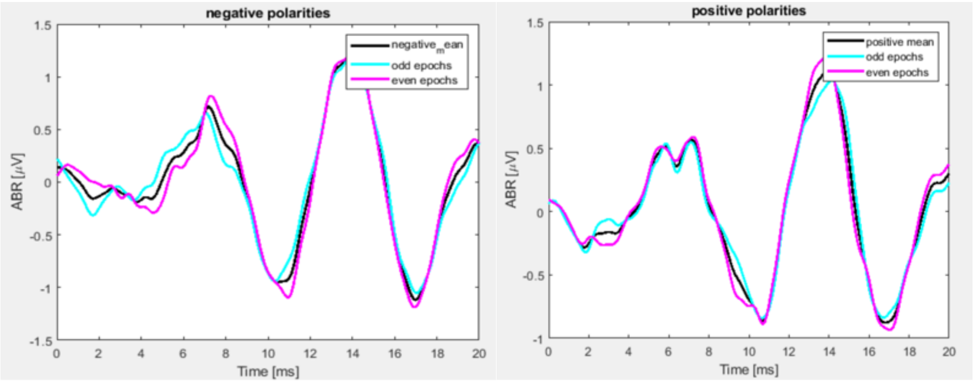
Appendices

Appendix 1

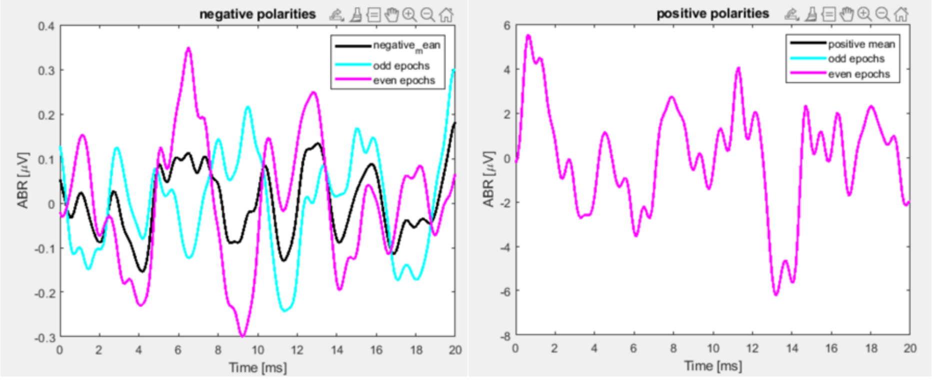
Excluded participants

oNH04, oNH11 and oHI12 were excluded

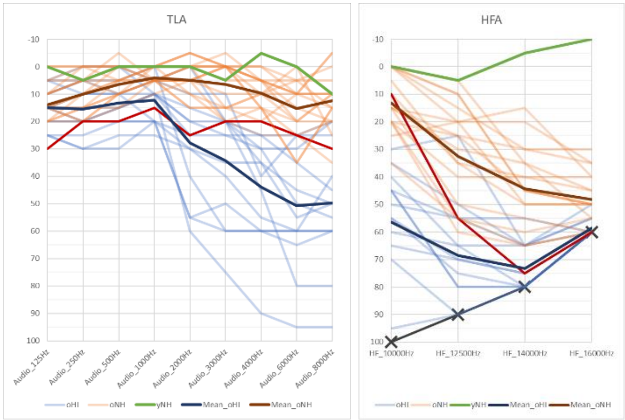
- *oNH04*: electrode came loose during measurement
  ABR:



- *oHI12*: electrode fell out of the equipment during measurement and/or wrong settings were set (rarefaction instead of alternating polarity)
  ABR:



- *oNH11*: the audiogram did not fit in both test groups (threshold on 4000Hz ≤ 20 dB HL but exceeds 20 dB HL on 125 and 2000 Hz).

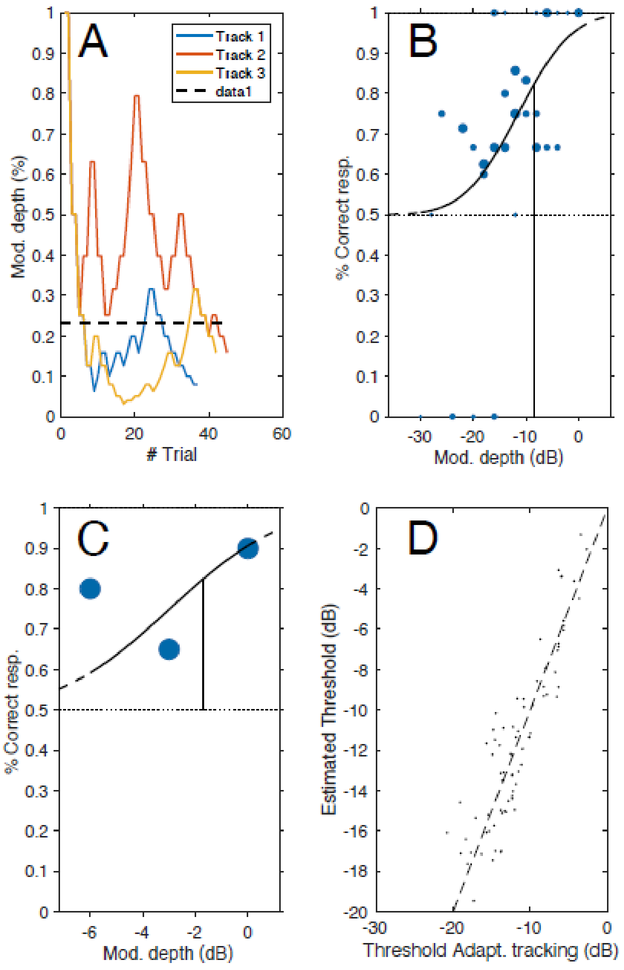


Appendix 2

Estimation of behavioral detection thresholds

We used a procedure to reconstruct the whole psychometric functions, as the results of some participants did not converge.

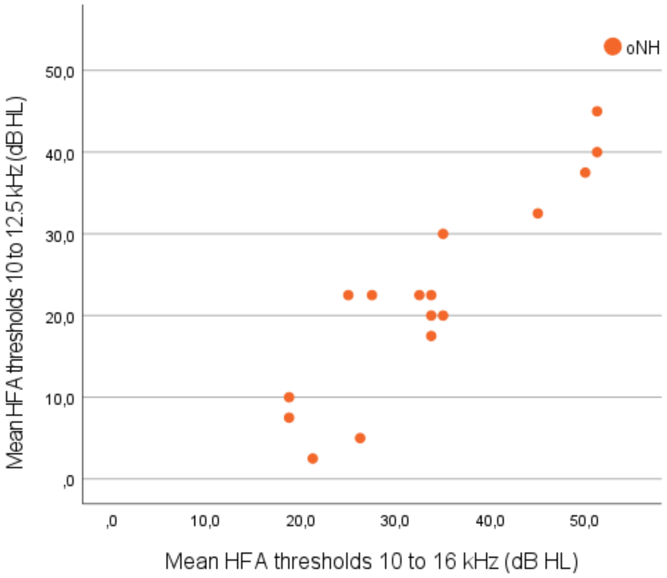
(A) Illustration of the convergence problems we had with the psychoacoustical tracking procedure for some individuals. (B) Tracks in panel A. (C) For an individual that was measured using a constant stimuli procedure with only 3 mod. depths because the adaptive procedure was even not able to go down, and for which we are nevertheless able to obtain a threshold. This allows us to estimate a threshold for every individual and condition, and avoid missing values. (D) Comparison of the two procedures (mean of the last 5 reversals vs. thresholds estimated from the reconstructed psychometric function) on the conditions where we could extract both, these are very close and strongly correlated, which is a good demonstration that we can use either one or the other for the stats.



Appendix 3

Mean high frequency audiometry (HFA)

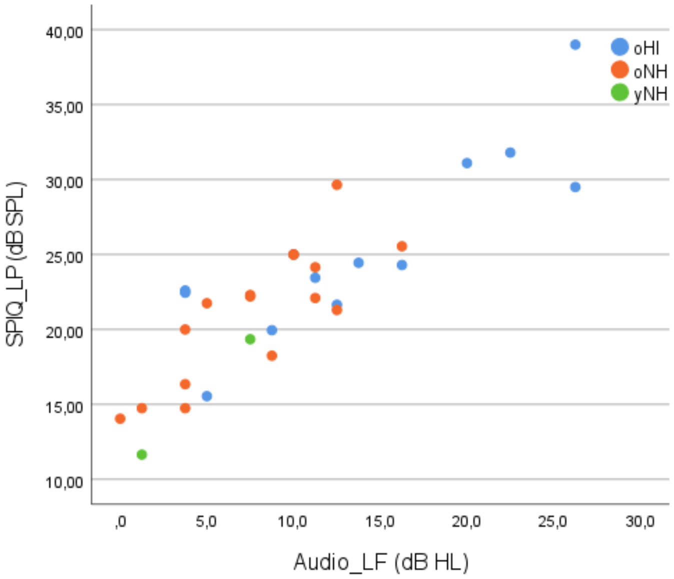
The mean of 10 and 12.5 kHz will be used for the analysis, because 14 and 16 kHz are not trustworthy because of output limits of the audiometer. For participants where the limit is not reached (oNH, except one), both means correlated significantly (r=0.91; p<0.001).



Appendix 4

SPIQ_LP and Audio_LF

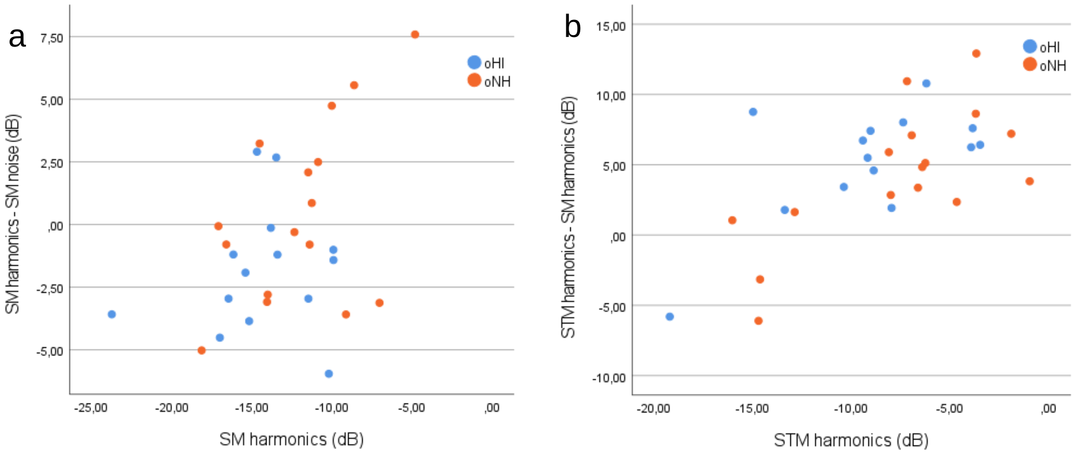
Scatter plot of the speech in quiet low pass (SPIQ_LP) in relation to the mean audiometric thresholds of the low frequencies (125 to 1000 Hz; Audio_LP), for oHI (blue dots), oNH (orange dots) participants and yNH (green dots). oHI = older hearing impaired; oNH = older normal hearing; yNH = younger normal hearing. The correlation is significant (r=0.87; p<0.001), only Audio_LF will be used for the further analyzes.



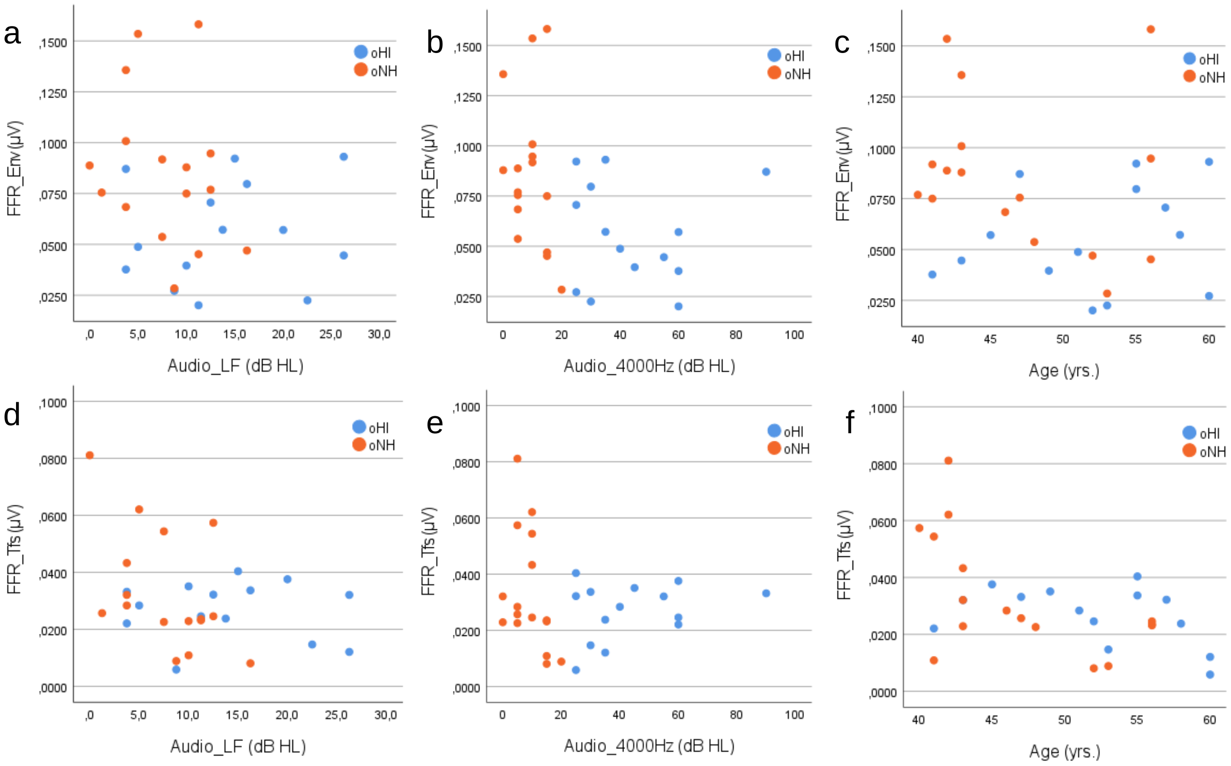
Appendix 5

Additional scatter plots

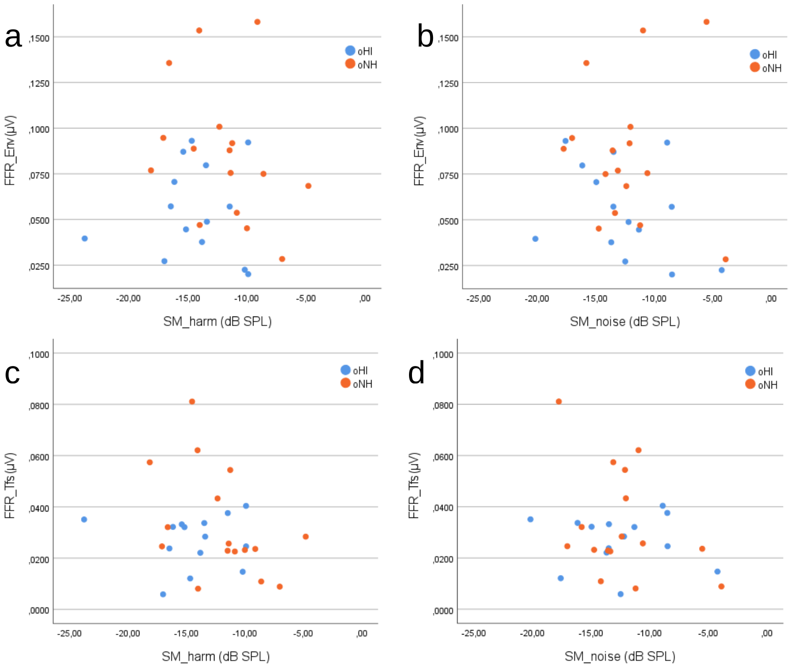
Scatter plots of (a) the subtraction of SM_harm and SM_noise in relation to SM_harm and (b) the subtraction of STM_harm and SM_harm with STM_harm.



Scatter plots of the frequency following response (FFR) magnitudes in relation to (a and d) audiometric thresholds in the low frequencies (mean 125 to 1000 Hz), (b and e) audiometric thresholds on 4000 Hz and (c and f) age, for oHI (blue dots) and oNH (orange dots) participants. Env = envelope; Tfs = temporal fine structure; Audio_LF = audiometric thresholds low frequencies; Audio_4000Hz = audiometric thresholds on 4000 Hz; oHI = older hearing impaired; oNH = older normal hearing.



Scatter plots of the frequency following response (FFR) magnitudes in relation to (a and c) SM harmonics and (b and d) SM noise, for oHI (blue dots) and oNH (orange dots) participants. Env = envelope; Tfs = temporal fine structure; SM = spectral modulated; STM = spectrotemporal modulated; oHI = older hearing impaired; oNH = older normal hearing.



Scatter plots of the speech in noise (SPIN; low pass (LP) and broad band (BB)) thresholds in relation to (a and c) SM harmonics and (b and d) STM harmonics, for oHI (blue dots) and oNH (orange dots) participants. SM = spectral modulated; STM = spectrotemporal modulated; oHI = older hearing impaired; oNH = older normal hearing.