Identification and Discrimination of Sound Textures in Hearing-Impaired and Older Listeners

Oliver Scheuregger[1] [iD], Jens Hjortkjær[1,2] and Torsten Dau[1]

Abstract

Sound textures are a broad class of sounds defined by their homogeneous temporal structure. It has been suggested that sound texture perception is mediated by time-averaged summary statistics measured from early stages of the auditory system. The ability of young normal-hearing (NH) listeners to identify synthetic sound textures increases as the statistics of the synthetic texture approach those of its real-world counterpart. In sound texture discrimination, young NH listeners utilize the fine temporal stimulus information for short-duration stimuli, whereas they switch to a time-averaged statistical representation as the stimulus' duration increases. The present study investigated how younger and older listeners with a sensorineural hearing impairment perform in the corresponding texture identification and discrimination tasks in which the stimuli were amplified to compensate for the individual listeners' loss of audibility. In both hearing impaired (HI) listeners and NH controls, sound texture identification performance increased as the number of statistics imposed during the synthesis stage increased, but hearing impairment was accompanied by a significant reduction in overall identification accuracy. Sound texture discrimination performance was measured across listener groups categorized by age and hearing loss. Sound texture discrimination performance was unaffected by hearing loss at all excerpt durations. The older listeners' sound texture and exemplar discrimination performance decreased for signals of short excerpt duration, with older HI listeners performing better than older NH listeners. The results suggest that the time-averaged statistic representations of sound textures provide listeners with cues which are robust to the effects of age and sensorineural hearing loss.

Introduction

Sound textures (e.g., sounds produced by wind, insect swarms, fire, and water) are a broad class of acoustic stimuli found in everyday listening environments. One unique property of sound textures is that their statistical properties remain relatively constant over time and can thus be efficiently represented via time-averaged statistics. McDermott and Simoncelli (2011) developed a sound texture framework that synthesized textures by analyzing the time-averaged statistics at the output of several processing stages of a biologically plausible auditory model, which were subsequently used to shape a Gaussian noise seed to have matching statistics. The model comprised auditory processing stages derived from both psychophysical and physiological data, which have been shown to be important for the perception of acoustic stimuli, including frequency-selective auditory filters (Glasberg & Moore, 1990;

Patterson, 1976; Patterson et al., 1987), compressive non-linearities (Ruggero, 1992; Yates, 1990; Zilany et al., 2009; Zwicker, 1979), and amplitude-modulation selective filters (Dau et al., 1997; Chi et al., 2005). Behavioral identification tasks with young (NH) listeners found that increasing the number of statistic classes imposed during the synthesis process increased the similarity of the sound texture

[1]Hearing Systems Section, Department of Health Technology, Technical University of Denmark, Kongens Lyngby, Denmark
[2]Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Hvidovre, Kettegård Allé 30, DK-2650 Hvidovre, Denmark

Corresponding Author:
Oliver Scheuregger, Hearing Systems Section, Department of Health Technology, Technical University of Denmark, Kongens Lyngby, Denmark.
Email: oliver.scheuregger@gmail.com

percept to the real-world equivalent. Furthermore, when the auditory model deviated in its biological plausibility, such as applying linearly (instead of logarithmically) spaced auditory filters, the perceptual quality of the sound texture exemplars was reduced. In a subsequent study, McDermott et al. (2013) demonstrated that young NH listeners' ability to discriminate between excerpts of two different sound textures (texture discrimination) improved as the excerpt duration increased, whereas the listeners' ability to discriminate between two unique excerpts taken from the same sound texture (exemplar discrimination) worsened as the excerpt duration increased. The results of both texture and exemplar discrimination suggest that, as the duration of a sound texture increases, the auditory system switches from a fine-grained representation to a time-averaged statistic representation.

The present study investigated the effect of hearing loss and age on sound texture perception. Hearing loss commonly distorts the coding and representation of the sound in the auditory system and affects sound perception, including sound source separation (Bronkhorst & Plomp, 1989, 1992; Ter-Horst, 1993), spectral and temporal resolution (Dubno & Schaefer, 1995; Moore, 1985; Nelson & Freyman, 1987; Reed et al., 2009), and pitch and loudness perception (Arehart & Burns, 1999; Oxenham, 2008; Rosengard et al., 2005) which, in turn, affects music and speech perception, particularly in noisy and reverberant environments (Carhart & Tillman, 1970; Cherry, 1953; Dubno et al., 1984; Duquesnoy, 1983; Eisenberg et al., 1995; Hygge et al., 1992; Moore et al., 1995; Peters et al., 1998; Takahashi & Bacon, 1992). McWalter and Dau (2015) reported that the statistics of sound textures were altered when measured through auditory models which reflected aspects of hearing loss (broader acoustic filters and reduced non-linearities). These alterations were quantified by the coefficient of variance of squared difference. The higher-order marginal moments, correlations and modulation power statistics showed more variation than the other statistic classes. Furthermore, listeners were able to discriminate between sound textures synthesized from the unaltered auditory model and the altered ones.

Previous behavioral studies have typically focused on the perception of environmental sounds, a broader class of sounds found in everyday life of which many, but not all, may also be classified as sound textures. NH adults typically perform well in identification of everyday environmental sounds (Gaver, 1993), as well as more complex sound perception tasks, such as determining the length of objects by their sound when dropped onto a hard surface (Carello et al., 1998), the gender of a walker by the sound of their footsteps (Li et al., 1991), and hand configuration by the sound of applause (Repp, 1987). Identification of gross material categories of impacted objects has been suggested to be relatively robust (Giordano & McAdams, 2006) even when listeners only have access to long-term spectral information as in

texture sounds (Hjortkjær & McAdams, 2016). However, surveys have indicated that hearing loss negatively affects the perception of environmental sounds (Badran, 2001; Cox et al., 2007; Hallberg et al., 2008; Hass-Slavin et al., 2005; Tyler, 1990, 1994; Zhao et al., 1997; Zwolan et al., 1996), as well as listeners' sense of self-awareness, ability to detect danger and overall quality of life (Arlinger, 2003; Hétu et al., 1988; Mulrow et al., 1990; Scherer, 1998). Sound texture perception is therefore an important part of everyday audition, yet little is known regarding (HI) listeners' perception of sound textures.

Older listeners with age-appropriate hearing (pure tone thresholds less than 41 dB HL between 250 and 4000 Hz) have also demonstrated poorer performance than younger listeners in environmental sound identification tasks, in particular when the task complexity increases or the stimulus is distorted or masked (Fabiani et al., 1996; Gygi & Shafiro, 2013). In general, age affects auditory perception and is often confounded with the effect of hearing loss. Age has been demonstrated to degrade the neural representation of sound at the auditory midbrain and cortical levels, particularly with respect to temporal attributes (Hellstrom & Schmiedt, 1990; Presacco et al., 2016; Ross et al., 2010; Sörös et al., 2009). Many psychophysical studies have explored the behavioral effects of age on temporal processing tasks in listeners with normal audiograms, such as gap detection (Lister & Roberts, 2005; Roberts & Lister, 2004), duration detection (Abel et al., 1990; Fitzgibbons & Gordon-Salant, 1994, 1995, 2001, 2004), modulation detection (Dashika et al., 2016; Wallaert et al., 2016), and temporal-fine-structure processing (Füllgrabe, 2013; Füllgrabe & Moore, 2018), in which older listeners have typically demonstrated poorer temporal processing abilities than younger listeners.

While the effects of hearing loss and age have been shown to negatively impact listeners' perception of environmental sounds, little is known about the effects on sound texture perception. Sound textures represent a well-defined subset of environmental sounds and are characterized by temporal homogeneity. The time-averaged statistics of sound textures play an important role in perception and are presumed to be shaped by the auditory periphery. It is hypothesized that distortions to the auditory periphery due to age and/or hearing loss may be reflected in listeners' ability to accurately identify and discriminate sound textures. Here, two of the experiments from McDermott and Simoncelli (2011) and McDermott et al. (2013) were conducted with HI and older listeners to investigate these effects and the results were compared with data from young NH listeners in those reference studies. In the first experiment, sound texture identification performance was measured in mostly older HI listeners to study the listeners' identification sensitivity with varying sound texture statistics. In the second experiment, the effect of hearing loss and age on the listeners' ability to discriminate between sound textures of varying excerpt duration was examined.

Materials and Methods

Sound Texture Synthesis Model

The sound texture synthesis system used in this study was the one developed by McDermott and Simoncelli (2011), and comprised an analysis and a synthesis stage. Here, a condensed description is provided, and further details can be found in McDermott and Simoncelli (2011). In the analysis stage, the real-world sound texture recording is first decomposed into 30 subbands using cosine zero-phase filters whose center frequencies are equally spaced on an equivalent rectangular bandwidth (ERB) scale across the acoustic frequencies 52 to 8844 Hz. Two additional low- and high-pass filters are applied at the extremes of the spectrum to achieve a constant summed squared frequency response across frequency. The decomposition of a broadband signal into separate acoustic frequency subbands reflects the frequency selectivity of the cochlea. Subsequently, the envelopes of each subband are derived by first taking the Hilbert transform of each subband, and then the absolute of the resultant analytic signals. Each subband envelope is compressed using a power law exponent value of 0.3. The envelope extraction represents a rough estimation of the inner hair cells' response to a travelling wave on the basilar membrane, and the compression stage simulates the active non-linear amplitude response of the outer hair cells. The subband envelopes are downsampled to 400 Hz to improve computational efficiency. The marginal moments comprise the mean, variance, skew and kurtosis of the subband envelopes. The cochlear correlation statistics measure the correlations between each subband envelope and its eight neighboring subband envelopes. Each subband envelope is then passed through a modulation filterbank comprising 20 half-cosine filters, logarithmically spaced from 0.5 Hz to 200 Hz with a Q-factor of 2. This stage reflects the auditory system's sensitivity to the slow-varying fluctuations in a signal's temporal envelope. The modulation power is measured across each acoustic-modulation frequency channel. The modulation correlations, C1, capture the correlations between six octave-spaced modulation channels each tuned to the same modulation frequency, as well as its two nearest neighbors, but across all acoustic frequencies. The subset of six octave-spaced modulation filters cover a range of 3 Hz to 100 Hz with a Q-factor of $\sqrt{2}$. The C2 modulation correlations capture the correlations between six octave-spaced modulation channels each tuned to the same acoustic frequency, but across the same six octave-spaced modulation channels as used in the C1 correlations.

In the synthesis stage, the measured statistics are imposed onto a 5-s Gaussian noise seed using an iterative process, until the statistics of the seed match the measured statistics of the input signal. Gaussian noise is chosen to ensure that the fine structure is as random as possible and thus each synthesis yields a novel synthetic texture. The iteration loop operates on the synthetic noise seed and comprises the same acoustic and modulation-frequency decomposition stages and non-linearities. The synthetic texture's statistics at each stage is measured and an error signal is computed corresponding to the difference with the desired real-world texture's statistics. A signal-to-noise ratio (SNR) is computed for each statistic class as the squared statistic values to the squared error of the statistic class. The synthetic texture's statistics are then modified using a gradient descent method and the iteration process is stopped once all statistics classes reach an SNR of 30 dB or higher, or after 60 iteration loops. Only textures whose average SNR across all statistic classes reached 20 dB or higher were included in the present experiments.

Experiment 1: Sound Texture Identification

Procedure. The identification experiment from McDermott and Simoncelli (2011) was replicated in this study. The stimuli were divided into five sound texture classes: 'animals' (e.g., insects in swamp, frogs, seagulls), 'environment' (e.g., pouring coins, wind blowing, fire), 'mechanical' (e.g., bulldozer, train locomotive, jackhammer), 'people' (e.g., babble, crowd noise, laughter) and 'water' (e.g., waterfall, rain in woods, seaside waves). Sound textures within each class possessed similar perceptual and statistical properties such that sound textures within each class were more likely to be confused with each other than sound textures from two different classes. A one-interval five-alternative-forced-choice method was used, in which each trial contained a single stimulus and five labels, one of which had to be chosen. One label was correct, while the other four were incorrect and randomly selected from each of the four remaining sound texture classes. The listeners' performance was measured in terms of percentage correct responses. The same 96 sound textures as in McDermott and Simoncelli (2011) were used, each with a duration of 7 s, a sampling frequency of 20 kHz and 16 bit resolution.

Nine conditions were prepared for each stimulus. Condition 1 represented the synthetic sound texture including the cochlear channel power only; condition 2 comprised only the marginal moments (mean, variance, skew and kurtosis); condition 3 also included cochlear correlations (C); condition 4 used modulation power instead of cochlear correlations; condition 5 included both cochlear correlations and modulation power; condition 6 added the C1 correlations, whereas condition 7 added only the C2 correlations; condition 8 included the full set of statistic classes. Condition 9 represented the original (non-synthesized) real-world sound texture.

All tasks were conducted with MATLAB 2017a on a Macbook Pro 2017 model. The stimuli were presented to both ears diotically at a sound pressure level (SPL) of 70 dB (as also used in McDermott & Simoncelli, 2011) and

the playback system was calibrated using a G.R.A.S IEC 60318–1 Ear Simulator and a Norsonic Nor139 sound level meter. Audio playback was provided via a Focusrite Scarlett 2i2 Universal Serial Bus (USB) audio interface (48 kHz, 24 bit) and Sennheiser HD650 headphones. A soundproof booth (customized IAC 1200-A medical room) was used, with the fan switched off. All listeners conducted the tests without hearing aids but were provided with audibility compensation. This was done using a Cambridge method for loudness equalization (CAMEQ) filter (Moore et al.,

Figure 1. Pure-tone audiometric thresholds for all HI listeners. The mean and standard deviation across listeners are shown by the bold markers and error bars, respectively. Left and right ears are indicated by blue and red lines, respectively. Individual listeners' thresholds (averaged across left and right ears) are indicated by the grey lines.

1999) which brought as much of the stimulus' spectrum above threshold as possible by equalizing, for each ear separately, the average specific loudness of the stimulus across the frequency range from 500 to 8000 Hz. A 10-ms half-Hanning window was applied to the onsets and offsets of the sounds during playback to avoid unwanted spectral cues. Before the session, the listeners read instructions which explained the task and were asked if they had fully understood the task before commencing the test. Each listener completed the identification task in one experimental session, including pauses. Listeners did not receive any training prior to the test, nor was feedback provided during the test, consistent with McDermott and Simoncelli (2011).
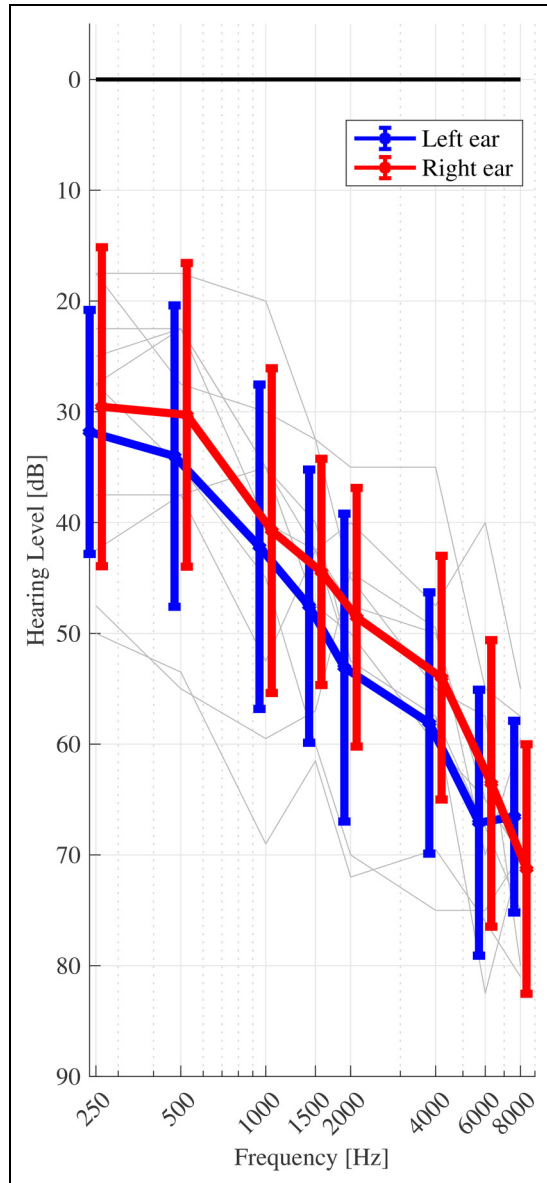
Listeners. Eleven mostly older HI listeners (10 male, age range 20 to 88 years, mean age 70.8 years, s.d. 17.7 years) participated in the experiment. All HI listeners in this study had a sloping, symmetric ($\pm$10 dB) sensorineural hearing loss within the N2 – N5 standard audiograms (Bisgaard et al., 2010). Audiograms were tested at eight frequencies: 250, 500, 1000, 1500, 2000, 4000, 6000 and 8000 Hz). Seven listeners had air-bone gap (ABG) thresholds less than 10 dB at all eight tested frequencies. Four listeners had ABG thresholds less than 10 dB at all but one tested frequency (500–4000 Hz; ABG = 15–20 dB). The listeners received financial compensation for their time and had prior experience in psychoacoustical procedures. All listeners provided informed consent and the experiment was approved by the Science-Ethics Committee for the Capital Region of Denmark (reference H-16036391). **Figure 1** shows all HI listeners' pure tone audiometric thresholds.

Data Analysis. The young NH listeners' scores from McDermott and Simoncelli (2011) were included as controls in the analysis of this experiment, and comprised the mean percent correct of 10 listeners' (one male, mean age 22.2 years, age range not available) identification performance for the same nine conditions as detailed in this experiment. Tests for experimental effects were conducted using mixed linear models, with the fixed effects hearing (NH or HI) and synthesis condition (1 to 9) and a random effect of listener to account for repeated measures. Two sound texture-class confusion patterns were derived, reflecting the listeners' responses versus the (correct) target response. One confusion pattern included the listeners' responses to only synthesis conditions 1 to 3, referred to as 'low-order' textures. The other confusion patterns included the responses to only 'high-order' synthesis conditions 6 to 8. These synthesis conditions were chosen (as opposed to using conditions 1 to 4, and conditions 5 to 8) to place more weight on specifically the 'low' and 'high' order statistics. For each, a fixed effects model was fit with fixed effect confusion group. All t-tests were conducted using Bonferroni adjusted alpha values. The significance levels were consistent throughout this study: n.s: $p \geq 0.05$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Experiment 2: Sound Texture Discrimination

Procedure. Sound texture and exemplar discrimination experiments as described in McDermott et al. (2013) were conducted. The stimulus pool comprised 100 textures in total: 50 pairs for texture discrimination, and 50 individual sound textures for exemplar discrimination. The stimuli were synthesized using the same toolbox as in experiment 1 and the synthesis stage imposed the marginal moments (kurtosis omitted), cochlear correlations, modulation power, and C1 and C2 correlations (synthesis condition 8) onto a 5 s white Gaussian noise seed.

A three-interval, two-alternative-forced-choice (odd-one-out) task was used. The listeners were presented with three intervals of equal duration in the order A–reference–B, with a fixed inter-stimulus interval of 400 ms. They were then asked to indicate which interval, A or B, contained the odd-one-out. In the case of sound texture discrimination, one of the intervals, A or B, was randomly assigned a unique excerpt taken from the same synthetic sound texture as the reference, while the other interval was of a different synthetic sound texture entirely. The task thus measures listeners' ability to hear differences *between* two different textures. In exemplar discrimination, three intervals were again presented to the listener; however, this time all three intervals belonged to the same synthetic sound texture, whereby one interval, A or B, was randomly assigned the same excerpt as the reference (i.e., both were physically identical), while the other was a unique excerpt. The task measures listeners' ability to discern fine details between two unique excerpts of the *same* sound texture. All unique excerpts were non-overlapping in time. Within a given trial, all excerpts were of equal duration, but across trials the excerpt duration was varied to be 40, 91, 209, 478, 1093 or 2500 ms. Each discrimination task thus comprised 300 trials (50 sound texture-pairs/textures x 6 excerpt durations) presented in a randomized order. The task order was also randomized and balanced across listeners. The experimental software, hardware, headphone calibration method and level, listening booth, and audibility compensation method were identical to those described in experiment 1.

Listeners. Four listener groups were tested to control for both hearing loss and age. Pilot tests were conducted with a group of four self-reported NH listeners aged between 21 and 28 years whose responses matched those of the 12 young self-reported NH listeners (all female, mean age 21.1 years, s.d. 3.0 years, age range not available) reported by McDermott et al. (2013). The four young NH listeners' results from this study were pooled together with those of McDermott et al. (2013) to form the NH-50 group, which comprised 16 listeners under the age of 50 years (three males, mean age 22.2 years, s.d. 4.2 years). Listener group NH + 50 comprised 15 listeners over the age of 50 years (five males, age range 54 to 76 years, mean age 65.8 years, s.d. 8.1 years).

The inclusion criteria for normal hearing were symmetric audiograms ($\pm$ 10 dB) and a hearing level not exceeding 30 dB HL for all tested frequencies (250, 500, 1000, 1500, 2000, 4000, 6000 and 8000 Hz). Listener group HI + 50 consisted of nine HI listeners over the age of 50 years (eight males, age range 68 to 82 years, mean age 74.6 years, s.d. 4.0 years), of which all also participated in experiment 1. The fourth listener group, HI-50, comprised four listeners under the age of 50 years (three males, mean age 30.3 years, s.d. 13.3 years), of which one also participated in experiment 1. The relatively small number of listeners in this group was, in part, due to a lack of listeners who fit the described criteria for hearing loss. The inclusion criteria for sensorineural hearing loss were identical to those as described in experiment 1. Seven HI + 50 and three HI-50 listeners had ABG thresholds less than 10 dB at all tested frequencies. Two HI + 50 listeners and one HI-50 listener had ABG thresholds less than 10 dB at all but one tested frequency (500–4000 Hz; ABG = 15–20 dB).

The criteria for hearing loss were identical to those described in experiment 1. **Figure 2** shows the individual listeners' pure tone thresholds for the groups HI + 50 (**Figure 2A**), NH + 50 (**Figure 2B**) and HI-50 (**Figure 2C**).

All listeners received compensation for their time and had prior experience with psychoacoustic procedures. All listeners provided informed consent and the experiment was approved by the Science-Ethics Committee for the Capital Region of Denmark (reference H-16036391).

Data Analysis. Due to the relatively small sample size of the HI-50 group, these listeners were omitted from statistical analyses. Therefore, the effect of age was tested on the differences between the NH-50 and NH + 50 listener groups' results, and the effect of hearing loss was tested on the differences between the NH + 50 and HI + 50 listener groups' results. All t-tests were conducted using Bonferroni adjusted alpha values. The significance levels were consistent throughout this study: n.s: $p \geq 0.05$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Results

Experiment 1: Sound Texture Identification

Figure 3 shows the results obtained in the sound texture identification task. The red open markers represent the data for the HI listeners from the present study. For direct comparison, the blue filled markers indicate the results obtained in the control study by McDermott and Simoncelli (2011) with 10 NH listeners. For both listener groups, performance improved as the number of statistic classes imposed during synthesis increased, though the HI listeners' performance was lower overall. The Analysis of Variance (ANOVA) results revealed a significant interaction effect between hearing status and synthesis condition
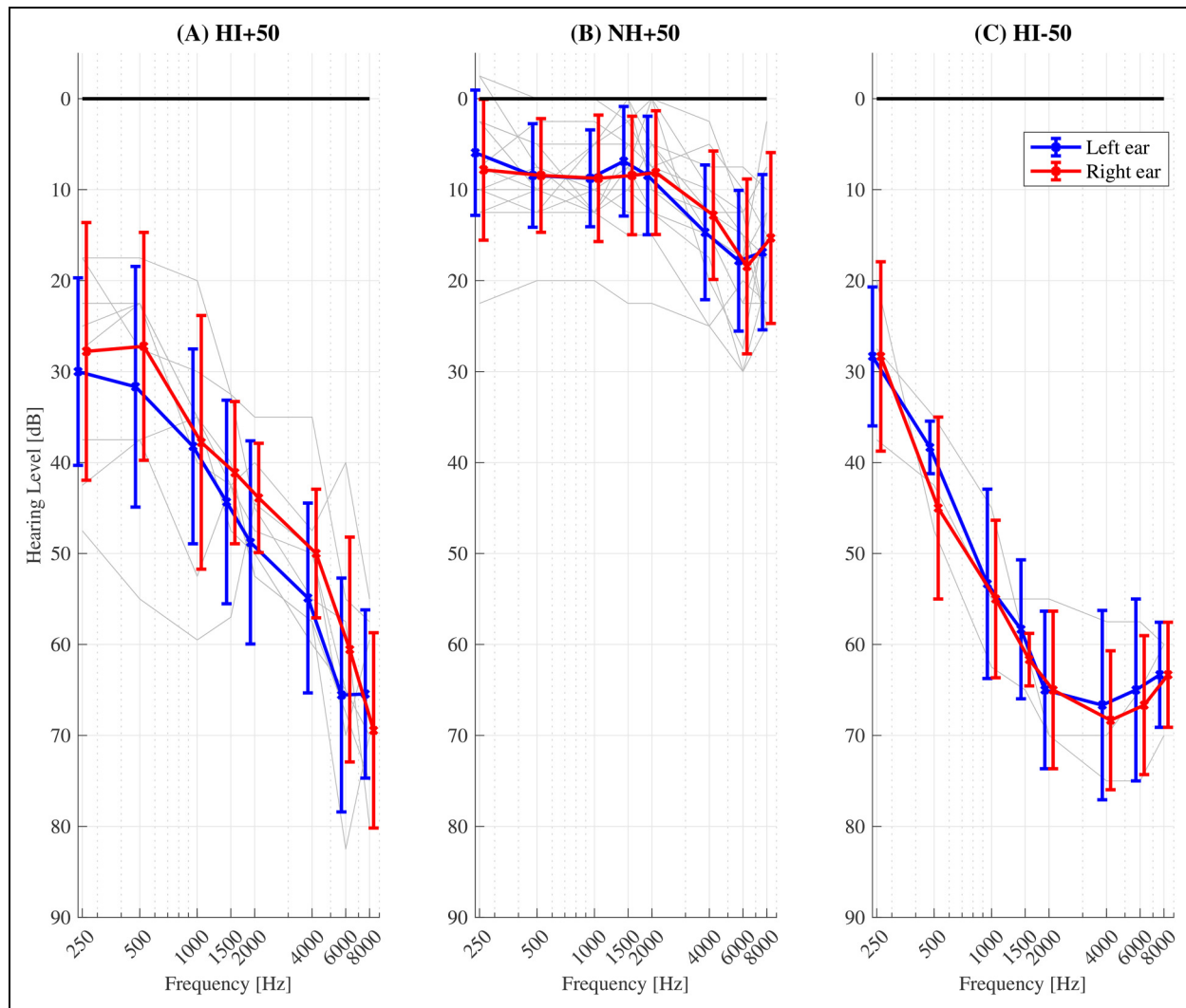
Figure 2. Pure tone audiometric thresholds for the (**A**) HI + 50, (**B**) NH + 50 and (**C**) HI-50 listeners. Audiograms were not obtained for the young NH group. The mean and standard deviation across listeners are shown by the bold markers and error bars, respectively. Left and right ears are indicated by the blue and red lines, respectively. Individual listeners' thresholds (averaged across left and right ears) are indicated by the grey lines.

($F_{8,152}$ = 3.8049, $p < 0.001$). Post hoc t-tests were conducted to compare the two listener groups' responses for each experimental condition (Bonferroni corrected alpha = 0.0056) and found that the HI listeners scored significantly lower than the NH listeners for all synthesis conditions except condition 2.

The largest performance improvements, for both listener groups, occurred between synthesis conditions 1 to 5. Post hoc t-tests were conducted to compare the pairwise differences between all experimental conditions (Bonferroni corrected alpha = 0.0014) for each listener group and revealed significant ($p < 0.05$) performance improvements between synthesis conditions 2–3, and 4–5 (NH listeners), and 1–2, 2–3, and 4–5 (HI listeners). For conditions 5 to 8, each additional statistic class resulted in performance increases that

were modest and insignificant for both listener groups. Only the HI listeners' performance increased significantly from condition 8 (synthetic sound textures) to condition 9 (real-world sound textures).

The highest identification performance for both listener groups occurred for the real-world sound texture (condition 9). As for all synthesized sound textures, the performance of the HI listeners was below that of the NH listeners, indicating an inherent deficiency due to hearing loss (beyond pure-tone sensitivity loss) and age that cannot be explained by any distortion resulting from the analysis/synthesis process.

Figure 4 shows the sound texture class confusion patterns for the HI listeners only. The NH listeners' responses to each texture class were not available. The two panels represent the responses to (**A**) synthesis conditions 1 to 3 and (**B**) synthesis
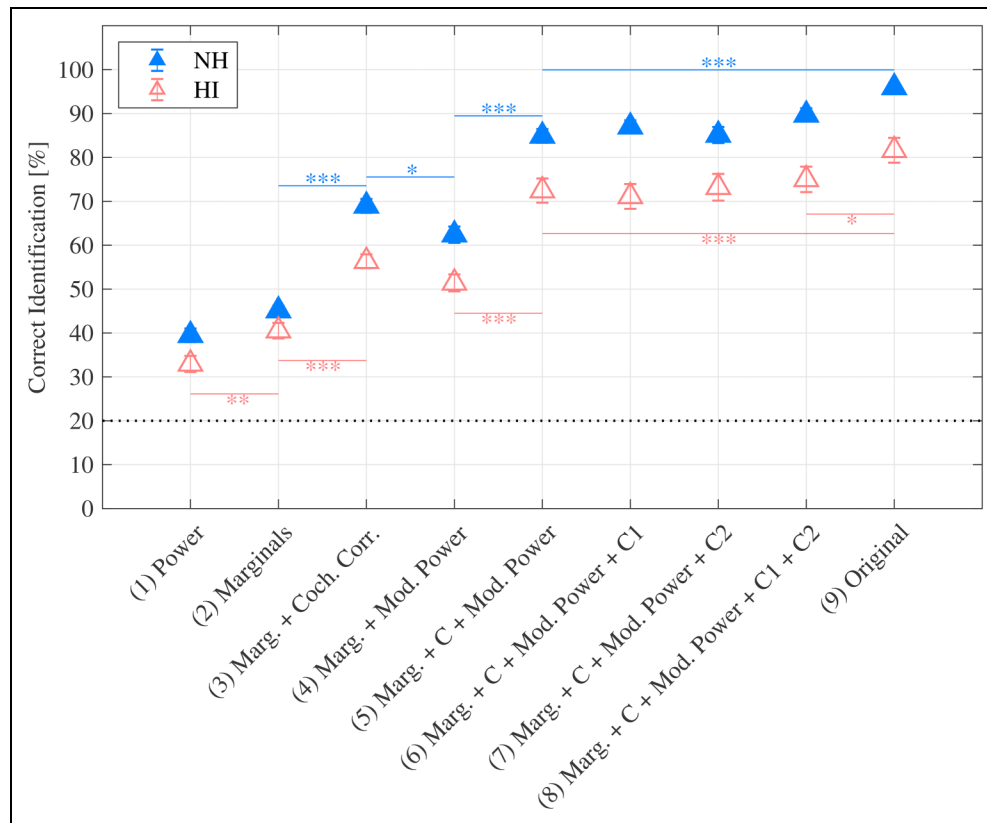
Figure 3. Average sound texture identification performance. The red open markers indicate the HI listeners' performance measured in this study and the blue filled markers indicate the NH listeners' performance taken from McDermott and Simoncelli (2011). Significance of differences (t-tests with Bonferroni corrections) between pairs of synthesis conditions is shown as color-coded lines and asterisks. Error bars represent the standard error of the mean. The black dotted line shows chance level (20%).

conditions 6 to 8 and highlight that the listeners were more biased towards 'water' responses in the low-order synthesis conditions. The ANOVA results revealed a significant main effect of texture class for the low-order confusion rates only ($F_{4,15} = 74.32$, $p < 0.001$), and average confusion rates were 26.4% higher for 'water' textures (compared to 7.6% for the high-order textures). This finding is consistent with McDermott and Simoncelli (2011) in which water sound textures made up the majority of the low-order synthetic sound textures which were identified correctly and incorrectly most often (see **Figure 5B–C**, McDermott and Simoncelli, 2011). The statistics of 'water' sound textures are inherently more similar to those of white noise than are the inherent statistics of the other sound texture classes, which may explain why confusion rates were found to increase in the lower-order synthesis conditions. The results may be relevant, and necessary to consider, for further experiments conducted using low-order synthetic sound textures.

In summary, the results indicated that both listener groups were perceptually sensitive to time-averaged statistics. However, the HI listeners' overall identification performance was lower than reported for the NH listeners.

Experiment 2: Sound Texture Discrimination

Figure 5 shows the results for sound texture discrimination (**A**) and exemplar discrimination (**B**). The results for the young and the old NH listeners are indicated by the filled blue circles (NH-50) and the open blue circles (NH + 50), respectively. The corresponding results for the HI listeners are shown by the filled red circles (HI-50) and the open red circles (HI + 50).

The sound texture discrimination results (**Figure 5A**) indicate that all listener groups' performance improved with longer excerpt duration and approached ceiling levels for excerpts of duration 1093 ms and above. The standard error of all listener groups' responses decreased as performance increased ([40 ms, 2500 ms]: NH-50 [1.91%, 0.37%]; HI-50 [4.80%, 0.50%]; NH + 50 [3.15%, 0.47%]; HI + 50 [3.06%, 1.31%]). No effect of hearing loss or age was found.

Conversely, the exemplar discrimination results (**Figure 5B**) show large listener-group differences, with a significant main effect of age ($F_{1,40} = 20.45$, $p < 0.001$). The NH + 50 listeners performed, on average, the worst and the NH-50 listeners performed the highest. The excerpt
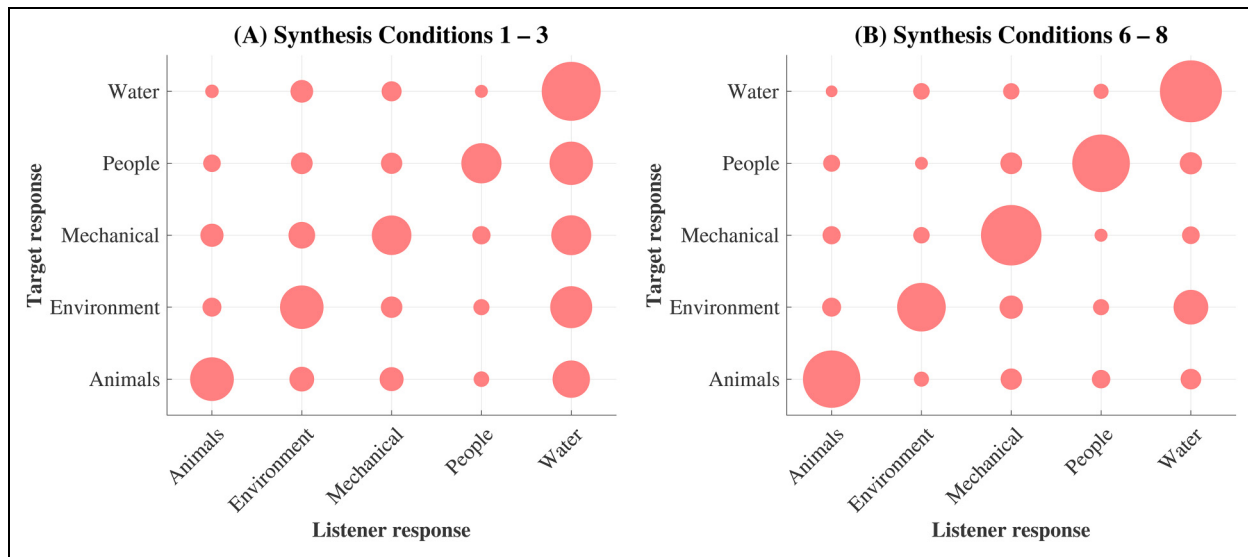
Figure 4. Older HI listeners' confusion patterns. The abscissa represents the listeners' responses for the different sound texture classes; the ordinate represents the target (correct) response. The area of the filled circles is proportional to the percentage of responses averaged across all listeners for each response-target pair. The different panels show the trends for two subsets of the data, indicating: (**A**) only low-order synthesis conditions (1 to 3), and (**B**) only high-order synthesis conditions (6 to 8).

duration at which young NH listeners achieved peak mean performance was between 40 and 91 ms, and the standard error of these listeners' responses decreased as mean performance increased ([2500 ms, 40 ms] = [2.0%, 1.4%]), as also observed in sound texture discrimination. The HI-50 listeners' mean performance was similar to the NH + 50 listeners across most durations, and the increase in mean performance was matched by a decrease in the standard error from 2500 ms (3.9%) to 209 ms (2.2%). Mean performance peaked at 91 ms and was just 2% lower than observed for the NH-50 listeners, though with a small increase in the standard error (2.6%). At 40 ms, the mean performance dropped 6.5% more than observed for the NH-50 listeners and the standard error of the responses was more than twice the value at 91 ms: ([91 ms, 40 ms] = [2.6%, 5.5%]).

The HI + 50 listeners' peak mean performance occurred at both 209 (78.9%) and 91 ms (78.2%) and was lower at 40 ms (65.8%), while for the NH + 50 listeners, peak performance occurred at 209 ms (70.1%), and decreased at both 91 ms (63.9%) and 40 ms (61.9%). The standard error increased consistently across both older listener groups' responses as the excerpt duration decreased from 2500 ms (HI + 50 = 3.4%; NH + 50 = 2.7%) to 40 ms (HI + 50 = 4.2%; NH + 50 = 3.7%). The HI + 50 listeners scored higher than the NH + 50 listeners at all excerpt durations, with greatest differences found at 91 ms and 209 ms: HI + 50 scored 14.3% and 8.8% higher than NH + 50, respectively.

Figure 6 shows the differences in texture (**A**) and exemplar (**B**) discrimination performance between the hearing

status groups (solid line) and the age groups (dashed line). Due to the relatively small sample size of the HI-50 group, this group was omitted: differences in hearing status groups were computed between NH + 50 and HI + 50, and differences between age groups were computed between NH-50 and NH + 50.

Figure 6A shows that the differences in sound texture discrimination performance between the hearing status groups were close to 0% for all excerpt durations. T-tests were conducted to compare the differences between the NH and HI listeners' responses at each excerpt duration (Bonferroni corrected alpha = 0.0083) and indicated that the differences failed to reach significance at all excerpt durations. The same t-tests were also conducted to test for the differences between the younger and older groups and found that the differences were also close to 0% for excerpt durations of 209 ms and above ($p > 0.05$) but increased to 8.4% at 91 ms ($p < 0.01$), and 9.3% at 40 ms ($p < 0.001$).

The duration-dependent effect of age is apparent in **Figure 6B** (dashed line). T-tests that evaluated the differences between the younger and older groups at each excerpt duration (Bonferroni corrected alpha = 0.0083) reached significance at all excerpt durations, and the greatest differences were found at both 40 ms and 91 ms (29.6%, $p < 0.001$). The same t-tests were conducted to test for the differences between the hearing status groups (solid line) and indicated that the HI + 50 listeners performed better than the NH + 50 listeners at all excerpt durations (differences reached significance only at 91 ms: $p < 0.05$), despite overall poor mean performance, relative to the younger listeners (**Figure 5B**).
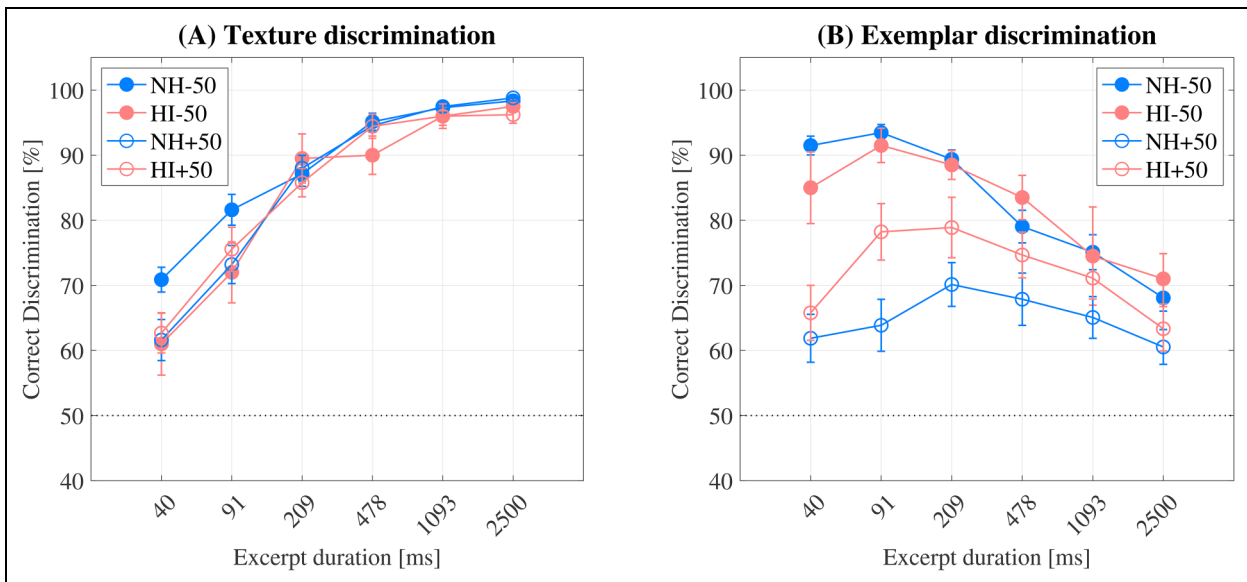
Figure 5. (**A**) sound texture and (**B**) exemplar discrimination performance of all four listener groups: NH-50 (filled blue circles), HI-50 (filled red circles), NH + 50 (open blue circles) and HI + 50 (open red circles). Error bars indicate the standard error of the mean. The black dotted line indicates chance level (50%).

Texture discrimination performance was found to be robust to both age and sensorineural hearing loss, whereas exemplar discrimination performance – in particular at the shorter excerpt durations – was significantly affected by age. Furthermore, the older HI listeners' exemplar discrimination performance was better than the older NH listeners – a finding that was also more pronounced at the shorter excerpt durations.

Discussion

This study investigated the effects of sensorineural hearing loss and age on the perception of sound textures. Two experiments, originally described in McDermott and Simoncelli (2011) and McDermott et al. (2013), were conducted using different listener profiles.

Experiment 1 comprised a sound texture identification task conducted on mostly older HI listeners whose results indicated three findings similar to those previously reported for younger NH listeners: (i) listeners were perceptually sensitive to time-averaged statistics, (ii) their performance improved as the statistics of the synthetic sound texture approached those of the real-world equivalent, and (iii) largest performance improvements occurred when correlations between frequency channels were included in the synthesis stage. The HI listeners, however, showed an overall performance deficit compared to the previously reported NH listeners for all synthesis conditions. Thus, while the sound texture statistics were perceptually important to both listener groups for correct identification, the NH listeners were able to identify

sound textures with a higher level of accuracy for all synthetic and real-world conditions.

In the second experiment, two different sound texture discrimination tasks were conducted on four listener groups categorized by both age (younger/older than 50 years) and sensorineural hearing loss (NH/HI). In sound texture discrimination, the listeners' ability to hear differences between two different sound textures was measured, while in exemplar discrimination, the listeners' ability to hear differences between two unique excerpts of the same sound texture was measured. All listener groups' sound texture discrimination performance improved as the excerpt duration increased, and performance was accompanied by a decrease in the variability of the responses. In exemplar discrimination, the younger listeners improved monotonically as the excerpt duration decreased, indicating accurate perception of the fine-grain differences between sound textures down to durations as short as 40 ms – an ability not significantly affected by sensorineural hearing loss. The older listeners, on the other hand, reached much lower performance levels than both younger listener groups over a narrower range of short excerpt durations, indicating an overall poorer representation of the sounds' fine-grain structure. The variability of the older listeners' responses increased with decreasing excerpt duration, which suggests that many older listeners, with or without sensorineural hearing loss, were severely limited in their ability to perform the task. Furthermore, the older HI listeners performed better than the older NH listeners at all measured excerpt durations.

The results from this study indicate that listeners' sensitivity to the time-averaged statistic representations of sound
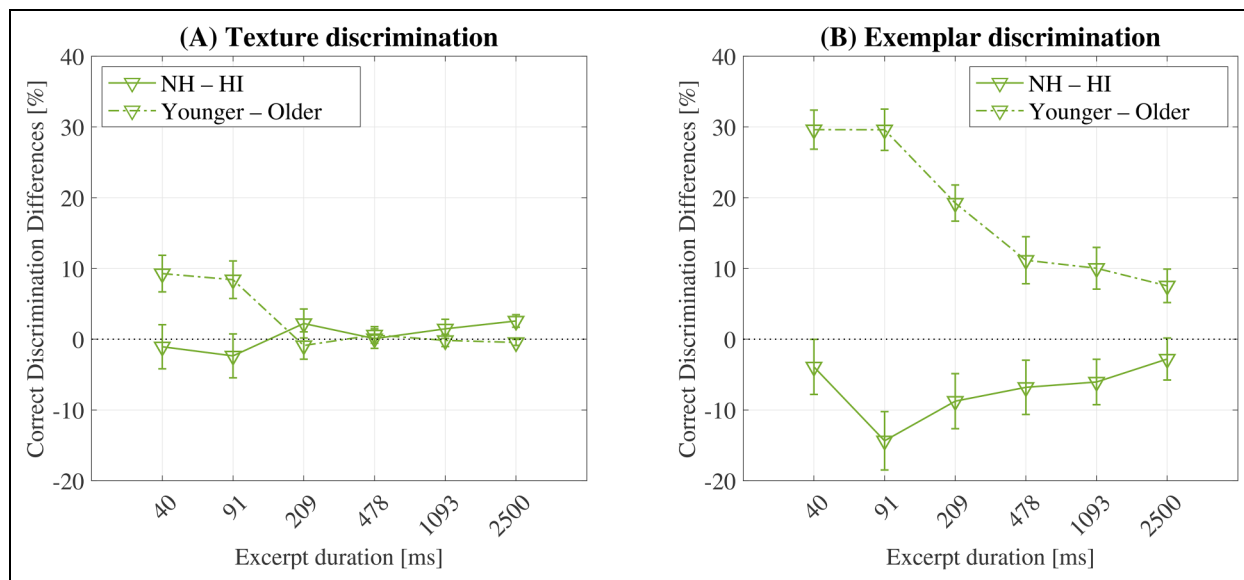
Figure 6. Differences in texture (**A**) and exemplar (**B**) discrimination performance for listeners grouped by hearing (solid line) and age (dash-dotted line). For the differences between hearing groups, positive values indicate that the NH listeners scored higher than the HI listeners, and for the differences between age groups, positive values indicate that the younger listeners scored higher than the older listeners. The differences were computed across balanced groups only. Error bars represent the standard error of the mean. The black dotted line indicates no differences (0%).

textures is robust to the effects of both age and sensorineural hearing loss. Although the older HI listeners were less able to accurately identify sound textures when synthesized with few statistics, high performance was nonetheless achieved as the number of statistics imposed during the synthesis stage increased, or when presented with the real-world sound texture. Furthermore, sound texture discrimination performance was unaffected by hearing loss. This is consistent with other studies (Ballas & Barnes, 1988; Harris et al., 2017) using environmental sounds in which accurate perception was observed in older listeners with normal hearing or (relative to this study) very mild hearing loss, but only when the task complexity was low and the stimulus quality was high (e.g., Gygi and Shafiro, 2013). The finding here might therefore be, at least partially, attributed to the simplicity of the experimental tasks.

Two interesting observations were that (i) the HI listeners performed worse than the NH listeners in texture identification, including the real-world textures, and (ii) the older HI listeners performed better than the older NH listeners in exemplar discrimination. Observation (i) suggests that the statistical representation of sound textures in HI listeners may differ from that in NH listeners, and is reflected in perception. Changes to the sound texture synthesis model which reflect aspects of hearing loss (e.g., broader auditory filters) have previously been shown to alter the statistical representation of sound textures, and to be perceptually discriminable (McWalter & Dau, 2015). These changes in the statistical representation may be reflected by an increase in the

variability of the measured statistics of the sound texture: i.e., statistics measured across an impaired auditory system may not converge at the same rate or to the same degree as those measured across an unimpaired auditory system. McWalter and McDermott (2018) reported that young NH listeners integrate sound texture statistics over a time window on the order a few seconds. It is possible that older HI listeners require a longer integration window. Sound textures whose statistics are represented with greater variability may then pose HI listeners with greater difficulty in tasks such as identification where an accurate representation of the sound texture is necessary, while in texture discrimination, where differences between two representations are required for good performance, the effect may be less detrimental to HI listeners' ability to perform the task.

Observation (ii) implies that the older HI listeners were better than the older NH listeners at discerning the fine temporal details of sound textures. There is evidence that HI listeners perform better than NH listeners in intensity discrimination (Jerger, 1962) and first-order amplitude modulation detection thresholds (Füllgrabe et al., 2003), possibly due to a loss of cochlear compression resulting in an expansion of the listener's internal representation of the signal's temporal envelope. This may in turn offer HI listeners with additional discrimination cues.

Alternatively, observation (ii) may reflect differences between the older HI and older NH listeners' experience with psychoacoustic experiments. Due to their hearing loss, the older HI listeners have participated in numerous

psychoacoustic experiments in the past, while the older NH listeners were less experienced. Exemplar discrimination relies on the listener's ability to focus on very short sound excerpts and make judgements on often quite subtle differences. In contrast, texture discrimination is a simpler task in that it relies on making judgements in response to longer excerpts of sounds. Therefore, the fact that the older HI listeners scored higher in exemplar discrimination than the older NH listeners, while texture discrimination performance was equal, may rather be explained by experience than hearing loss.

The results from exemplar discrimination showed that, despite demonstrating robust statistical representations of sound textures, the older listeners' ability to access the fine temporal details of sound textures was severely limited. Factors such as task complexity, signal quality and loss of high frequency sensitivity, which have been shown to negatively affect older listeners' perception of environmental sounds (Dick et al., 2015; Gygi & Shafiro, 2013), do not explain the findings in this study: the task was simple, the textures were synthesized with the full set of statistics, and the experiments were conducted in quiet with audibility compensation up to 8 kHz. Non-auditory cognitive decline has been linked to deficits in auditory processing tasks (for a review, see Aydelott et al., 2010). For example, Füllgrabe et al. (2015) conducted various cognitive tests on both younger and older NH listeners, and also measured their speech in noise performance. Cognitive abilities were positively correlated with speech in noise performance, and tended to be lower in the older listeners. Pichora-Fuller (2003) found that cognitive stressors, such as memory load, may exacerbate older listeners' ability to understand spoken language. A more recent study (Strelcyk et al., 2019) measured older HI listeners' sensitivity to interaural phase and time differences as well as their cognitive abilities (Trail Making Test) and reported a strong correlation between the auditory and non-auditory test scores.

Additionally, the older listeners' poorer performance in the exemplar discrimination task may be linked to an age-related decline in temporal resolution. Evidence from psychophysical studies typically support the view that auditory temporal processing deteriorates with age, resulting in a loss of sensitivity to the temporal fine structure (Füllgrabe et al., 2015; Füllgrabe & Moore, 2018; Hopkins & Moore, 2011). Neural deafferentation and temporal asynchrony of neural activity have been discussed as age-related pathologies (Makary et al., 2011; Pichora-Fuller & Schneider, 1991, 1992; Schneider, 1997) which distort the temporal encoding of acoustic stimuli and their internal neural representation at cochlear and/or retro-cochlear levels of processing (Lopez-Poveda & Barrios, 2013). Typically, models have been employed to simulate neural deafferentation and asynchrony, and the effects on listeners' perception have been measured. For example, Pichora-Fuller et al. (2007) simulated temporal asynchrony by temporally jittering speech signals, and reported a significant reduction in word identification performance when younger adults were presented with jittered speech as opposed to the clean speech. Marmel et al. (2015) used stochastic undersampling to model the effects of neural deafferentation and found that listeners' duration discrimination performance systematically decreased with undersampling. However, a more recent study (Oxenham, 2016) based on signal detection theory suggests that even a synapse loss of 50% is unlikely to result in any measurable effect in perceptual tasks, such as signal detection in quiet and discrimination of intensity, frequency and interaural time differences.

Further work is required to explore the underlying factors that may account for the observed effects of sensorineural hearing loss and age on sound texture identification and on the discrimination of short texture excerpts. Firstly, more data on young HI listeners' sound texture identification and discrimination performance would be beneficial to help extricate the two effects. Likewise, data on older NH listeners' sound texture identification performance would provide a fully balanced data set. With respect to hearing loss, one line of enquiry would be to test if changes to the auditory model which reflect aspects of hearing loss result in a more variable representation (increased variability of the measured statistics) of the sound texture. To test if the poor exemplar discrimination performance of the older listeners is confined to short excerpts of sound textures, or is observed for a broader range of sounds, the exemplar discrimination experiment could be repeated with brief excerpts of other sounds, such as speech. Additionally, more complex listening tests, such as using sound textures to mask target signals, may expose further differences between the listener profiles and offer more ecologically valid insights into sound texture perception. Regarding the effect of age, further work may involve incorporating models of neural degeneration as well as correlations between measures of temporal resolution and individual listeners' exemplar discrimination performance.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ORCID iD

Oliver Scheuregger 🔟 https://orcid.org/0000-0001-8061-7139

References

Abel S. M., Krever E. M., & Alberti P. W. (1990). Auditory detection, discrimination and speech processing in ageing, noise-sensitive and hearing-impaired listeners. *Scandinavian Audiology*, *19*(1), 43–54. https://doi.org/10.3109/01050399009070751

Arehart K. H., & Burns E. M. (1999). A comparison of monotic and dichotic complex-tone pitch perception in listeners with hearing loss. *The Journal of the Acoustical Society of America*, *106*(2), 993–997. https://doi.org/10.1121/1.427111

Arlinger S. (2003). Negative consequences of uncorrected hearing loss—a review. *In International Journal of Audiology*, *42*(sup2), 17–20. https://doi.org/10.3109/14992020309074639

Aydelott J., Leech R., & Crinion J. (2010). Normal adult aging and the contextual influences affecting speech and meaningful sound perception. *Trends in Amplification*, *14*(4), 218–232. https://doi.org/10.1177/1084713810393751

Badran O. E. (2001). Difficulties perceived by hearing aid candidates and users. *Indian Journal of Otology*, *7*(2), 53–56.

Ballas J. A., & Barnes M. E. (1988). Everyday sound perception and aging. *Proceedings of the Human Factors Society Annual Meeting*, *32*(3), 194–197. https://doi.org/10.1177/154193128803200305

Bisgaard N., Vlaming M. S. M. G., & Dahlquist M. (2010). Standard audiograms for the IEC 60118-15 measurement procedure. *Trends in Amplification*, *14*(2), 113–120. https://doi.org/10.1177/1084713810379609

Bronkhorst A. W., & Plomp R. (1989). Binaural speech intelligibility in noise for hearing-impaired listeners. *Journal of the Acoustical Society of America*, *86*(4), 1374–1383. https://doi.org/10.1121/1.398697

Bronkhorst A. W., & Plomp R. (1992). Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing. *The Journal of the Acoustical Society of America*, *92*(6), 3132–3139. https://doi.org/10.1121/1.404209

Carello C., Anderson K. L., & Kunkler-Peck A. J. (1998). Perception of object length by sound. *Psychological Science*, *9*(3), 211–214. https://doi.org/10.1111/1467-9280.00040

Carhart R., & Tillman T. W. (1970). Interaction of competing speech signals With hearing losses. *Archives of Otolaryngology*, *91*(3), 273–279. https://doi.org/10.1001/archotol.1970.00770040379010

Cherry E. C. (1953). Some experiments on the recognition of speech, with One and with Two ears. *Journal of the Acoustical Society of America*, *25*(5), 975–979. https://doi.org/10.1121/1.1907229

Chi T., Ru P., & Shamma S. A. (2005). Multiresolution spectrotemporal analysis of complex sounds. *The Journal of the Acoustical Society of America*, *118*(2), 887–906. https://doi.org/10.1121/1.1945807

Cox R. M., Alexander G. C., & Gray G. A. (2007). Personality, hearing problems, and amplification characteristics: Contributions to self-report hearing aid outcomes. *Ear and Hearing*, *28*(2), 141–162. https://doi.org/10.1097/AUD.0b013e31803126a4

Dashika G. M., Theruvan N. B., Bhat J. S., & Nambi P. M. A. (2016). Effect of age and physical exercise on amplitude modulation detection thresholds. *International Journal of Pharma and Bio Sciences*, *7*(4), B467–B472. https://doi.org/10.22376/ijpbs.2016.7.4.b467-472

Dau T., Kollmeier B., & Kohlrausch A. (1997). Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. *The Journal of the Acoustical Society of America*, *102*(5), 2892–2905. https://doi.org/10.1121/1.420344

Dick F., Krishnan S., Leech R., & Saygin A. P. (2015). Environmental sounds. In *Neurobiology of language* (pp. 1121–1138). Elsevier Inc. https://doi.org/10.1016/B978-0-12-407794-2.00089-4.

Dubno J. R., Dirks D. D., & Morgan D. E. (1984). Effects of age and mild hearing loss on speech recognition in noise. *Journal of the Acoustical Society of America*, *76*(1), 87–96. https://doi.org/10.1121/1.391011

Dubno J. R., & Schaefer A. B. (1995). Frequency selectivity and consonant recognition for hearing-impaired and normal-hearing listeners with equivalent masked thresholds. *Journal of the Acoustical Society of America*, *97*(2), 1165–1174. https://doi.org/10.1121/1.413057

Duquesnoy A. J. (1983). Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *Journal of the Acoustical Society of America*, *74*(3), 739–743. https://doi.org/10.1121/1.389859

Eisenberg L. S., Dirks D. D., & Bell T. S. (1995). Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing. *Journal of Speech and Hearing Research*, *38*(1), 222–233. https://doi.org/10.1044/jshr.3801.222

Fabiani M., Kazmerski V. A., Cycowicz Y. M., & Friedman D. (1996). Naming norms for brief environmental sounds: Effects of age and dementia. *Psychophysiology*, *33*(4), 462–475. https://doi.org/10.1111/J.1469-8986.1996.TB01072.X

Fitzgibbons P. J., & Gordon-Salant S. (1994). Age effects on measures of auditory duration discrimination. *Journal of Speech and Hearing Research*, *37*(3), 662–670. https://doi.org/10.1044/jshr.3703.662

Fitzgibbons P. J., & Gordon-Salant S. (1995). Age effects on duration discrimination with simple and complex stimuli. *Journal of the Acoustical Society of America*, *98*(6), 3140–3145. https://doi.org/10.1121/1.413803

Fitzgibbons P. J., & Gordon-Salant S. (2001). Aging and temporal discrimination in auditory sequences. *The Journal of the Acoustical Society of America*, *109*(6), 2955–2963. https://doi.org/10.1121/1.1371760

Fitzgibbons P. J., & Gordon-Salant S. (2004). Age effects on discrimination of timing in auditory sequences. *The Journal of the Acoustical Society of America*, *116*(2), 1126–1134. https://doi.org/10.1121/1.1765192

Füllgrabe C. (2013). Age-dependent changes in temporal-fine-structure processing in the absence of peripheral hearing loss. *American Journal of Audiology*, *22*(2), 313–315. https://doi.org/10.1044/1059-0889(2013/12-0070)

Füllgrabe C., Meyer B., & Lorenzi C. (2003). Effect of cochlear damage on the detection of complex temporal envelopes. *Hearing Research*, *178*(1–2), 35–43. https://doi.org/10.1016/S0378-5955(03)00027-3

Füllgrabe C., & Moore B. C. J. (2018). The association between the processing of binaural temporal-fine-structure information and audiometric threshold and Age: A meta-analysis. *Trends in Hearing*, *22*, 1–14. https://doi.org/10.1177/2331216518797259

Füllgrabe C., Moore B. C. J., & 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*, *7*(JAN), 1–25. https://doi.org/10.3389/fnagi.2014.00347

Gaver W. W. (1993). How Do We hear in the world?: Explorations in ecological acoustics. *Ecological Psychology*, *5*(4), 285–313. https://doi.org/10.1207/s15326969eco0504_2

Giordano B. L., & McAdams S. (2006). Material identification of real impact sounds: Effects of size variation in steel, glass, wood, and plexiglass plates. *The Journal of the Acoustical Society of America*, *119*(2), 1171. https://doi.org/10.1121/1.2149839

Glasberg B. R., & Moore B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, *47*(1–2), 103–138. https://doi.org/10.1016/0378-5955(90)90170-T

Gygi B., & Shafiro V. (2013). Auditory and cognitive effects of aging on perception of environmental sounds in natural auditory scenes. *Journal of Speech, Language, and Hearing Research*, *56*(5), 1373–1388. https://doi.org/10.1044/1092-4388(2013/12-0283)

Hallberg L. R. M., Hallberg U., & Kramer S. E. (2008). Self-reported hearing difficulties, communication strategies and psychological general well-being (quality of life) in patients with acquired hearing impairment. *Disability and Rehabilitation*, *30*(3), 203–212. https://doi.org/10.1080/09638280701228073

Harris M. S., Boyce L., Pisoni D. B., Shafiro V., & Moberly A. C. (2017). The relationship between environmental sound awareness and speech recognition skills in experienced cochlear implant users. *Otology and Neurotology*, *38*(9), e308–e314. https://doi.org/10.1097/MAO.0000000000001514

Hass-Slavin L., McColl M. A., & Pickett W. (2005). Challenges and strategies related to hearing loss among dairy farmers. *Journal of Rural Health*, *21*(4), 329–336. https://doi.org/10.1111/j.1748-0361.2005.tb00103.x

Hellstrom L. I., & Schmiedt R. A. (1990). Compound action potential input/output functions in young and quiet-aged gerbils. *Hearing Research*, *50*(1–2), 163–174. https://doi.org/10.1016/0378-5955(90)90042-N

Hétu R., Riverin L., Lalande N., Getty L., & St-Cyr C. (1988). Qualitative analysis of the handicap associated with occupational hearing loss. *British Journal of Audiology*, *22*(4), 251–264. https://doi.org/10.3109/03005368809076462

Hjortkjær J., & McAdams S. (2016). Spectral and temporal cues for perception of material and action categories in impacted sound sources. *The Journal of the Acoustical Society of America*, *140*(1), 409–420. https://doi.org/10.1121/1.4955181

Hopkins K., & Moore B. C. J. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *The Journal of the Acoustical Society of America*, *130*(1), 334–349. https://doi.org/10.1121/1.3585848

Hygge S., Ronnberg J., Larsby B., & Arlinger S. (1992). Normal-hearing and hearing-impaired subjects' ability to just follow conversation in competing speech, reversed speech, and noise backgrounds. *Journal of Speech and Hearing Research*, *35*(1), 208–215. https://doi.org/10.1044/jshr.3501.208

Jerger J. (1962). The sisi test. *International Journal of Audiology*, *1*(2), 246–247. https://doi.org/10.3109/05384916209074055

Li X., Logan R. J., & Pastore R. E. (1991). Perception of acoustic source characteristics: Walking sounds. *Journal of the Acoustical Society of America*, *90*(6), 3036–3049. https://doi.org/10.1121/1.401778

Lister J. J., & Roberts R. A. (2005). Effects of age and hearing loss on gap detection and the precedence effect: Narrow-band stimuli. *Journal of Speech, Language, and Hearing Research*, *48*(2), 482–493. https://doi.org/10.1044/1092-4388(2005/033)

Lopez-Poveda E. A., & Barrios P. (2013). Perception of stochastically undersampled sound waveforms: A model of auditory deafferentation. *Frontiers in Neuroscience*, *7*(7 JUL), 1–13. https://doi.org/10.3389/fnins.2013.00124

Makary C. A., Shin J., Kujawa S. G., Liberman M. C., & Merchant S. N. (2011). Age-related primary cochlear neuronal degeneration in human temporal bones. *JARO - Journal of the Association for Research in Otolaryngology*, *12*(6), 711–717. https://doi.org/10.1007/s10162-011-0283-2

Marmel F., Rodríguez-Mendoza M. A., & Lopez-Poveda E. A. (2015). Stochastic undersampling steepens auditory threshold/duration functions: Implications for understanding auditory deafferentation and aging. *Frontiers in Aging Neuroscience*, *7*(APR), 1–13. https://doi.org/10.3389/fnagi.2015.00063

McDermott J. H., Schemitsch M., & Simoncelli E. P. (2013). Summary statistics in auditory perception. *Nature Neuroscience*, *16*(4), 493–498. https://doi.org/10.1038/nn.3347

McDermott J. H., & Simoncelli E. P. (2011). Sound texture perception via statistics of the auditory periphery: Evidence from sound synthesis. *Neuron*, *71*(5), 926–940. https://doi.org/10.1016/j.neuron.2011.06.032

McWalter R., & Dau T. (2015). Statistical representation of sound textures in the impaired auditory system. *Proceedings of ISAAR 2015: Individual Hearing Loss – Characterization, Modelling, Compensation Strategies. 5th Symposium on Auditory and Audiological Research*, August.

McWalter R., & McDermott J. H. (2018). Adaptive and selective time averaging of auditory scenes. *Current Biology*, *28*(9), 1405–1418.e10. https://doi.org/10.1016/J.CUB.2018.03.049

Moore B. C. J. (1985). Frequency selectivity and temporal resolution in normal and hearing-impaired listeners. *British Journal of Audiology*, *19*(3), 189–201. https://doi.org/10.3109/03005368509078973

Moore B. C. J., Glasberg B. R., & Stone M. A. (1999). Use of a loudness model for hearing aid fitting: III. A general method for deriving initial fittings for hearing aids with multi-channel compression. *British Journal of Audiology*, *33*(4), 241–258. https://doi.org/10.3109/03005369909090105

Moore B. C. J., Glasberg B. R., & Vickers D. A. (1995). Simulation of the effects of loudness recruitment on the intelligibility of speech in noise. *British Journal of Audiology*, *29*(3), 131–143. https://doi.org/10.3109/03005369509086590

Mulrow C. D., Aguilar C., Endicott J. E., Tuley M. R., Velez R., Charlip W. S., Rhodes M. C., Hill J. A., & DeNino L. A. (1990). Quality-of-life changes and hearing impairment. A randomized trial. *Annals of Internal Medicine*, *113*(3), 188–194. https://doi.org/10.7326/0003-4819-113-3-188

Nelson D. A., & Freyman R. L. (1987). Temporal resolution in sensorineural hearing-impaired listeners. *Journal of the Acoustical Society of America*, *81*(3), 709–720. https://doi.org/10.1121/1.395131

Oxenham A. J. (2008). Pitch perception and auditory stream segregation: Implications for hearing loss and cochlear implants. *Trends in Amplification*, *12*(4), 316–331. https://doi.org/10.1177/1084713808325881

Oxenham A. J. (2016). Predicting the perceptual consequences of hidden hearing loss. *Trends in Hearing*, *20*, https://doi.org/10.1177/2331216516686768

Patterson R. D. (1976). Auditory filter shapes derived with noise stimuli. *The Journal of the Acoustical Society of America*, *59*(3), 640–654. https://doi.org/10.1121/1.380914

Patterson R. D., Nimmo-smith I., Holdsworth J., & Rice P. (1987). *An efficient auditory filterbank based on the gammatone function (Part A : The Auditory Filterbank) AN EFFICIENT AUDITORY FIL TERBANK BASED ON THE GAMMATONE FUNCTION Roy Patterson and Ian Nimmo-Smith MQC Applied PsycholotjY Unit 15 Chaucer Qoad Cambride. December.*

Peters R. W., Moore B. C. J., & Baer T. (1998). Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *The Journal of the Acoustical Society of America*, *103*(1), 577–587. https://doi.org/10.1121/1.421128

Pichora-Fuller M. K. (2003). Cognitive aging and auditory information processing. *International Journal of Audiology*, *42*(SUPPL. 2). https://doi.org/10.3109/14992020309074641

Pichora-Fuller M. K., & Schneider B. A. (1991). Masking-level differences in the elderly - a comparison of antiphasic and time-delay dichotic conditions. *Journal of Speech and Hearing Research*, *34*(6), 1410–1422. https://doi.org/10.1044/jshr.3406.1410

Pichora-Fuller M. K., & Schneider B. A. (1992). The effect of interaural delay of the masker on masking-level differences in young and old adults. *Journal of the Acoustical Society of America*, *91*(4), 2129–2135. https://doi.org/10.1121/1.403673

Pichora-Fuller M. K., Schneider B. A., MacDonald E., Pass H. E., & Brown S. (2007). Temporal jitter disrupts speech intelligibility: A simulation of auditory aging. *Hearing Research*, *223*(1–2), 114–121. https://doi.org/10.1016/j.heares.2006.10.009

Presacco A., Simon J. Z., & Anderson S. (2016). Evidence of degraded representation of speech in noise,in the aging midbrain and cortex. *Journal of Neurophysiology*, *116*(5), 2346–2355. https://doi.org/10.1152/jn.00372.2016

Reed C. M., Braida L. D., & Zurek P. M. (2009). Review article: Review of the literature on temporal resolution in listeners With cochlear hearing impairment: A critical assessment of the role of suprathreshold deficits. *Trends in Amplification*, *13*(1), 4–43. https://doi.org/10.1177/1084713808325412

Repp B. H. (1987). The sound of two hands clapping: An exploratory study. *The Journal of the Acoustical Society of America*, *81*(4), 1100–1109. https://doi.org/10.1121/1.394630

Roberts R. A., & Lister J. J. (2004). Effects of age and hearing loss on gap detection and the precedence effect: Broadband stimuli. *Journal of Speech, Language, and Hearing Research*, *47*(5), 965–978. https://doi.org/10.1044/1092-4388(2004/071)

Rosengard P. S., Oxenham A. J., & Braida L. D. (2005). Comparing different estimates of cochlear compression in listeners with normal and impaired hearing. *The Journal of the Acoustical Society of America*, *117*(5), 3028–3041. https://doi.org/10.1121/1.1883367

Ross B., Schneider B. A., Snyder J. S., & Alain C. (2010). Biological markers of auditory gap detection in young, middle-aged, and older adults. *PLoS ONE*, *5*(4), e10101. https://doi.org/10.1371/journal.pone.0010101

Ruggero M. A. (1992). Responses to sound of the basilar membrane of the mammalian cochlea. *Current Opinion in Neurobiology*, *2*(4), 449–456. https://doi.org/10.1016/0959-4388(92)90179-O

Scherer M. J. (1998). Characteristics associated with marginal hearing loss and subjective well-being among a sample of older adults. *Journal of Rehabilitation Research and Development*, *35*(4), 420–426.

Schneider B. (1997). Psychoacoustics and aging: Implications for everyday listening. *Journal of Speech-Language Pathology and Audiology*, *21*(2), 111–124.

Sörös P., Teismann I. K., Manemann E., & Lütkenhöner B. (2009). Auditory temporal processing in healthy aging: A magnetoencephalographic study. *BMC Neuroscience*, *10*, 1–9. https://doi.org/10.1186/1471-2202-10-34

Strelcyk O., Zahorik P., Shehorn J., Patro C., & Derleth R. P. (2019). Sensitivity to interaural phase in older hearing-impaired listeners correlates With nonauditory trail making scores and With a spatial auditory task of unrelated peripheral origin. *Trends in Hearing*, *23*, https://doi.org/10.1177/2331216519864499

Takahashi G. A., & Bacon S. P. (1992). Modulation detection, modulation masking, and speech understanding in noise in the elderly. *Journal of Speech and Hearing Research*, *35*(6), 1410–1421. https://doi.org/10.1044/jshr.3506.1410

Ter-Horst K. (1993). Ability of hearing-impaired listeners to benefit from separation of speech and noise. *Australian Journal of Audiology*, *15*(2), 71–84.

Tyler R. S. (1990). Advantages and disadvantages reported by some of the better cochlear-implant patients. *American Journal of Otology*, *11*(4), 282–289.

Tyler R. S. (1994). Advantages of disadvantages expected and reported by cochlear implant patients. *American Journal of Otology*, *15*(4), 523–531.

Wallaert N., Moore B. C. J., & Lorenzi C. (2016). Comparing the effects of age on amplitude modulation and frequency modulation detection. *The Journal of the Acoustical Society of America*, *139*(6), 3088–3096. https://doi.org/10.1121/1.4953019

Yates G. K. (1990). Basilar membrane nonlinearity and its influence on auditory nerve rate-intensity functions. *Hearing Research*, *50*(1–2), 145–162. https://doi.org/10.1016/0378-5955(90)90041-M

Zhao F., Stephens S. D. G., Sim S. W., & Meredith R. (1997). The use of qualitative questionnaires in patients having and being considered for cochlear implants. *Clinical Otolaryngology and Allied Sciences*, *22*(3), 254–259. https://doi.org/10.1046/j.1365-2273.1997.00036.x

Zilany M. S. A., Bruce I. C., Nelson P. C., & Carney L. H. (2009). A phenomenological model of the synapse between the inner hair cell and auditory nerve: Long-term adaptation with power-law dynamics. *The Journal of the Acoustical Society of America*, *126*(5), 2390–2412. https://doi.org/10.1121/1.3238250

Zwicker E. (1979). A model describing nonlinearities in hearing by active processes with saturation at 40 dB. *Biological Cybernetics*, *35*(4), 243–250. https://doi.org/10.1007/BF00344207

Zwolan T. A., Kileny P. R., & Telian S. A. (1996). Self-report of cochlear implant use and satisfaction by prelingually deafened adults. *Ear and Hearing*, *17*(3), 198–210. https://doi.org/10.1097/00003446-199606000-00003