



OPEN Impaired noise adaptation contributes to speech intelligibility problems in people with hearing loss

Miriam I. Marrufo-Pérez^{1,2}, Milagros J. Fumero^{1,2}, Almudena Eustaquio-Martín^{1,2} & Enrique A. Lopez-Poveda^{1,2,3}✉

Understanding speech in noisy settings is harder for hearing-impaired (HI) people than for normal-hearing (NH) people, even when speech is audible. This is often attributed to hearing loss altering the neural encoding of temporal and/or spectral speech cues. Here, we investigated whether this difference may also be due to an impaired ability to adapt to background noise. For 25 adult hearing-aid users with sensorineural hearing loss, speech reception thresholds (SRTs) were measured for natural and tone-vocoded words embedded in speech-shaped noise (SSN). The stimuli were preceded or not by a 1-second adapting SSN precursor. Adaptation was calculated as the difference in SRT between the two precursor conditions. Corresponding data for 28 NH listeners were taken from a previously published study. SRTs were worse for HI listeners, confirming that hearing loss diminishes access to speech acoustic cues. Furthermore, noise adaptation was negatively correlated with the age-controlled hearing loss both for natural ($\rho = -0.56$, $N = 52$, $p < 0.001$) and vocoded ($\rho = -0.45$, $N = 39$, $p = 0.002$) words. Impaired adaptation contributed up to 10% of the SRT loss in HI listeners. We conclude that HI listeners suffer from poorer speech in noise recognition not only because of impaired access to speech acoustic cues but also because they are less able to adapt to background noise.

Keywords Hearing aids, Adaptation, Speech spectrum, Envelope, Temporal fine structure

Hearing loss affects millions of people worldwide and has a negative impact on many areas, including speech intelligibility. Hearing-impaired (HI) listeners show poorer speech recognition in noisy environments than do normal-hearing (NH) listeners even when speech is presented at suprathreshold levels^{1–8}. This has prompted many studies to investigate the cause of non-audibility related deficits^{9–11} and led to the search for new sound-processing algorithms to improve speech-in-noise intelligibility^{12,13}. To date, however, studies have overlooked that the greater difficulties experienced by HI listeners understanding speech in noisy environments could be partly due to an impaired ability to adapt to background noise. The aim of the present study was to investigate the potential negative impact of hearing loss on the ability of individuals to adapt to noise during speech recognition.

The ability of NH listeners to recognize words or syllables in noisy settings can improve as speech tokens are delayed from the onset of the noise. This has been interpreted as evidence that NH listeners adapt to the background noise^{14–17}. Adaptation to noise occurs in pure tone detection¹⁸ (also known as “overshoot”), amplitude modulation (AM) detection^{19,20}, syllable recognition²¹, and word recognition^{16,17}. This adaptation has a time course of approximately 300–350 ms^{18,22,23} and its benefits can be quite large for NH listeners. For example, for speech presented at a fixed signal-to-noise ratio (SNR), adaptation results in up to ~30% more recognized syllables^{14,15,21}. When the SNR is adaptively varied to estimate the SNR at 50% word recognition, speech reception thresholds (SRTs) can be up to 5 dB better when words are delayed 300 ms from the noise onset than when words and noise start at the same time^{16,22,23}.

Hearing-impaired listeners vary widely in their ability to understand audible speech in noise (or speech in speech). Although previous studies have considered different predictors of this variability^{11,24–29}, it has been neglected that HI listeners may adapt more or less to noise depending on the degree of hearing loss. Evidence exists that adaptation to noise in pure tone or AM detection is less for HI listeners than for NH listeners^{30–35}. In

¹Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Calle Pintor Fernando Gallego 1, 37007 Salamanca, Spain. ²Instituto de Investigación Biomédica, Universidad de Salamanca, Salamanca, Spain.

³Departamento de Cirugía, Facultad de Medicina, Universidad de Salamanca, Salamanca, Spain. ✉email: ealopezpoveda@usal.es

AM detection, HI listeners show overall better AM detection thresholds than NH listeners, and do not benefit from adding an adapting noise (or precursor) before the AM tone, whereas NH listeners do³⁴. The better AM thresholds are thought to be a consequence of HI listeners having more linear basilar membrane (BM) responses because of outer hair cell (OHC) loss, which results in an enhanced AM representation at the BM. The smaller adaptation for HI listeners has been interpreted to be a consequence of medial olivocochlear (MOC) efferent effects having a smaller range of action to linearize BM responses³⁴.

Because speech recognition for HI listeners relies more on envelope (AM) cues than on temporal fine structure (TFS) cues^{4,10,36} and because HI listeners exhibit less than normal adaptation to noise in AM detection, adaptation to noise in speech recognition should be less for HI listeners than for NH listeners, either when speech is natural or processed (vocoded) to preserve envelope cues and discard TFS cues. To our knowledge, only Ben-David et al.²² have investigated adaptation to noise in speech recognition including a group of HI listeners. They showed that older HI listeners do not adapt to multitalker babble noise but adapt to steady noise as younger NH listeners do. However, it is uncertain whether these results are related to age, hearing loss, or the two factors. In addition, because Ben-David et al. aimed at investigating the effects of aging but not of hearing loss on adaptation to noise, listeners performed the experiment using their better ear, and the audiometric thresholds of their HI listeners did not differ much from those of their NH listeners [the thresholds of their HI listeners were ~10 dB HL at low-to-mid frequencies (mean across 0.5, 1, and 2 kHz) and ~32 dB HL at mid-to-high frequencies (mean across 3, 4, 6, and 8 kHz)]. Further research including participants with greater audiometric losses and controlling for the possible effect of age on adaptation is necessary to shed light on the impact of hearing loss on adaptation.

Here, we measured adaptation to noise for NH and HI listeners presented with natural and vocoded words in steady noise. We compared SRTs in noise for NH and HI listeners in different conditions (natural and vocoded words presented with and without an adapting precursor noise) to estimate the contribution of different factors (adaptation loss as well as the impaired access to speech envelope, TFS, or spectral cues) to the SRT loss of HI listeners. HI listeners were provided with frequency-specific, linear amplification to compensate for reduced audibility, and the possible effect of age on adaptation was controlled for statistically.

Materials and methods

Experimental design

A total of 25 HI listeners participated in the study. We measured SRTs (i.e., the SNR at 50% word recognition) for natural (unprocessed) and tone-vocoded words embedded in simultaneous speech-shaped noise (SSN). Words could be preceded or not by a 1-second-long noise precursor with identical spectrum and level as the simultaneous noise. Adaptation to noise was calculated as the improvement in SRT caused by the precursor relative to the no-precursor condition. The stimuli were presented monaurally and amplified using a software-based multichannel hearing-aid (HA), which provided frequency-specific linear amplification according to the National Acoustics Laboratory – Revised (NAL-R) gain prescription³⁷. The results for HI listeners were compared with those of NH listeners published in a previous study¹⁷.

Participants

Twenty-five HI listeners (9 men) participated in the experiment (mean age = 61.5 years; standard deviation (SD) = 9.3 years). The participants were native Spanish speakers, except for one of them, who was a native French speaker but had a high level of proficiency in Spanish.

Air- and bone-conduction hearing thresholds were measured using a clinical audiometer (Interacoustics AD229e). Air-conduction thresholds were measured in the two ears at octave frequencies between 0.25 and 8.0 kHz, as well as at 6.0 kHz. All but two participants had symmetrical hearing loss (mean air-conduction threshold difference between the two ears < 15 dB at 0.5, 1, and 2 kHz, and < 30 dB at 4 and 6 kHz³⁸). For participants with symmetrical hearing loss, the test ear was chosen at random (14 left, 9 right ears). The two participants with asymmetrical hearing loss were tested in the ear with the worst air-conduction thresholds (the right ear in the two cases). The audiogram for the test ear of the 25 HI participants is plotted in Fig. 1A. For further analyses, we computed the arithmetic mean of the air conduction thresholds at 0.5, 1, 2, and 4 kHz, hereinafter referred as the pure-tone average (PTA). The PTA thresholds ranged from 17.5 to 82.5 dB HL [mean = 43.7 dB HL; SD = 18.1 dB].

To confirm that participants had sensorineural hearing loss, we also measured their bone-conduction hearing thresholds at octave frequencies between 0.5 and 4.0 kHz for the participants with air-conduction PTA > 30 dB HL in the tested ear. The exception was one participant with an air-conduction PTA of 82.5 dB HL, for whom we did not measure bone-conduction thresholds. For participants with symmetrical hearing loss, bone-conduction thresholds were measured in one ear without masking the contralateral ear. The measured ear was chosen at random because we assumed that the interaural attenuation in bone conduction was zero or close to zero³⁹. For the two participants with asymmetrical hearing loss, thresholds were measured in the test ear with masking in the contralateral ear. For all participants, their bone conduction PTAs ranged from 21.3 to 65 dB HL (mean = 40.8 dB HL; SD = 14.9 dB). This indicates that they had sensorineural hearing loss.

The results for HI listeners were compared with the results for NH listeners from a previous study¹⁷. The air-conduction audiograms for the test ear of the 28 NH listeners are shown in Fig. 1B. The PTA thresholds for this group ranged from 0 to 21.3 dB HL (mean = 6.9 dB HL; SD = 5 dB). NH listeners were younger (mean age = 26.6 years; SD = 8.3; min = 17; max = 56) than HI listeners were (mean age = 61.5 years; SD = 9.3; min = 39; max = 82). For that reason, the correlation between hearing loss and adaptation to noise was controlled for age as described later.

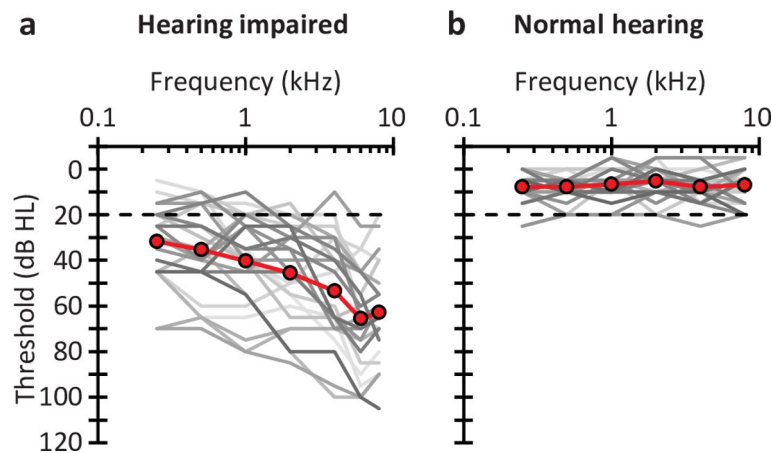


Fig. 1. Audiometric thresholds for the study participants. **(A)** Air-conduction thresholds for the test ear of the 25 HI listeners. The line with symbols depicts the group mean audiogram; lines without symbols depict the audiograms of individual participants. The dashed line at 20 dB HL depicts the maximum level typically accepted for a person to be regarded as NH. **(B)** As panel (A) but for NH listeners ($N=28$) (data from Marrufo-Pérez et al.¹⁷).

Participants were volunteers and not paid for their time. All of them signed an informed consent form before they could participate in the study. Methods were approved by the Ethics Committee of the University of Salamanca (Spain) and performed in accordance with the Declaration of Helsinki.

Stimuli

Hearing-impaired listeners

SRTs for noise-masked words were measured monaurally in the presence and in the absence of a noise precursor. Both the simultaneous and precursor noises were steady, had a speech-shaped spectrum, and had a level of 65 dB SPL before HA processing. In the condition without precursor, the simultaneous masking noise started 50 ms before the word onset and finished 50 ms after the word offset. The precursor noise had a duration of 1 s. Both the precursor and simultaneous noise had 20-ms raised-cosine onset and offset ramps. The time gap between the precursor noise and the simultaneous masking noise was 0 ms. Different precursor and simultaneous noises were generated for each word presentation (i.e., noises were not ‘frozen’). Adaptation to noise was calculated as the improvement in SRT with the precursor relative to the no-precursor condition.

SRTs were measured for natural⁴⁰ and tone-vocoded disyllabic words (one participant was tested only with natural words). The vocoder was applied to the words but not to the noise. The vocoder included a high-pass preemphasis filter (first-order Butterworth filter with a 3-dB cutoff frequency of 1.2 kHz); a bank of 12, sixth-order Butterworth band-pass filters whose 3-dB cutoff frequencies followed a modified logarithmic distribution between 100 and 8500 Hz (center frequencies from channel #1 to #12 equal to 149, 262, 408, 601, 855, 1191, 1638, 2234, 3029, 4091, 5511, and 7412 Hz); and envelope extraction via full-wave rectification and low-pass filtering (fourth-order Butterworth low-pass filter with a 3-dB cutoff frequency of 400 Hz). The envelope for each frequency channel was used to modulate a sinusoidal carrier at the channel center frequency, and the modulated signals were filtered again through the corresponding filter in the bank and samplewise added to obtain the vocoded speech.

Stimuli (speech and noise) were processed through a 12-channel, software-based HA implemented in the frequency domain in Matlab (The MathWorks, v2017a). The HA has been fully described elsewhere¹³ but was set with linear amplification according to the NAL-R gain prescription³⁷.

Normal-hearing listeners

The results of the present study for HI listeners were compared with the results for NH listeners from a previous study¹⁷. The stimuli used in the two studies were similar except that a precursor noise 1.0 s in duration was used for HI listeners while for NH listeners words were delayed in the noise by 1.5 s. The difference in the preceding noise duration is not expected to affect the amount of adaptation because the time course of adaptation is ~350 ms and SRTs in noise do not improve further for noise-word onset delays beyond 600 ms. This time course is independent of age or hearing status²².

Procedure

Phonetically balanced word lists⁴⁰ were used to measure the SRTs. Before starting an SRT measurement, the experimenter chose one of the ten available lists at random. Each list contained 25 disyllabic words. Words from a given list were presented in random order across test conditions to minimize the possibility that participants remembered the words. To measure the SRT, the noise level was fixed. The speech level varied adaptively using a one-down, one-up adaptive rule, i.e., it decreased after a correct response and increased after an incorrect response. The SRT was thus defined as the SNR yielding 50% correct word recognition in the psychometric function⁴¹. The initial SNR was 10 dB. The speech level changed in 4-dB steps between words 1 and 8 and in 2-dB steps between words 9 and 25. The SRT was calculated as the mean of the SNRs for the final 15 words (the SNR for the 16th word was calculated and used in the SRT estimate but not actually presented). Feedback was not given to the participants on the correctness of their responses.

The SRTs with and without precursor were measured in pairs but in random order. An SRT was discarded immediately after the measurement, and a new SRT was measured when the SD within the measurement was greater than 3 dB. Three SRTs were obtained for each condition, but if the experimenter deemed that the across-measures SD was high (no objective criteria were used), a fourth pair of SRTs was obtained. The mean of the three or four SRTs was taken as the final SRT.

During the measurements, participants were seated in a double-wall sound attenuating booth and the presentation of each word was controlled by the experimenter, who was sitting outside the booth without visual interaction with the participant. Participants were asked to repeat each word into a microphone after hearing them or, if they did not recognize the word, they were asked to indicate so. The experimenter received the participant's response through headphones and marked the word as correct or incorrect in the software. The word was scored as correct when the participant correctly repeated all its phonemes. A sound cue (1-kHz pure tone with 500 ms duration) was presented 500 ms before the stimulus onset to warn the listener about the stimulus presentation and to focus his/her attention on the speech recognition task. The level of the cue was 83 dB SPL for HI listeners and 66 dB SPL for NH listeners.

Apparatus

The stimuli were presented through custom-made Matlab software (The MathWorks, v2017a) and played via an RME Fireface UCX soundcard at a sampling rate of 44.1 kHz and with 24-bit resolution. For NH listeners, stimuli were presented using insert earphones (ER2, Etymotic Research). The sound pressure levels (SPLs) were calibrated by inserting the earphones into a Zwislocki DB-100 coupler connected to a sound level meter (Brüel Kjaer, mod. 2238). For HI listeners, stimuli were presented using circumaural headphones (Sennheiser HD580), which provided higher output levels than did the insert earphones and prevented clipping from occurring when the hearing loss was severe. The SPLs were calibrated by placing the headphones on a KEMAR head (Knowles Electronics) equipped with a Zwislocki (Knowles Electronics DB-100) artificial ear connected to a sound level meter (Brüel & Kjaer 2238). For the two transducers, calibration was performed at 1 kHz and the obtained sensitivity was used at all other frequencies.

Statistical analyses

All the statistical analyses were performed using IBM SPSS Statistics, version 28. The Shapiro-Wilk test of normality revealed that the data did not conform to a Gaussian distribution. Hence, nonparametric statistical tests were used to compare SRTs or adaptation to noise across participant groups (NH and HI listeners) or across conditions (no precursor/precursor) as well as to correlate different measures. We hypothesized that (1) SRTs would be better with than without the precursor noise, and (2) adaptation would be less for HI listeners than for NH listeners. Because these hypotheses were directional, we report one-tailed significance levels for all tests. An effect was regarded as statistically significant when the null hypotheses could be rejected with 95% confidence ($p \leq 0.05$). No corrections for multiple comparisons were applied.

Results

Figure 2A shows the SRTs for NH listeners presented with natural and vocoded words, without and with precursor. Figure 2B shows the corresponding data for HI listeners. Let us first analyze the SRTs measured without the noise precursor to assess the effect of hearing loss on intelligibility independent of adaptation. In that case, SRTs were overall worse (higher) for HI listeners than for NH listeners, both for natural (Mann-Whitney $U = 32.5$, $p < 0.001$) and vocoded words ($U = 119.0$, $p = 0.039$). In addition, the SRTs worsened as hearing loss increased both for natural (Spearman Rho $\rho = 0.74$, $N = 52$, $p < 0.001$) and vocoded ($\rho = 0.36$, $N = 39$, $p = 0.012$) words (Fig. 3A). Because stimuli were presented to HI listeners through a HA with frequency-specific linear amplification³⁷, the worse SRTs for HI listeners were unlikely related to audibility deficits. Instead, the results suggest that HI listeners were less able to use speech-related acoustic cues as hearing loss increases. The specific cues will be discussed later.

To investigate whether listeners showed adaptation to noise, Fig. 2A, B compares SRTs without and with the noise precursor. For NH listeners, the Wilcoxon signed rank test showed that the noise precursor improved SRTs for both natural ($N = 27$; $Z = -2.73$; $p = 0.003$) and vocoded ($N = 15$; $Z = -3.0$; $p = 0.002$) words, although adaptation was greater for vocoded words ($N = 14$; $Z = -2.2$; $p = 0.013$) (Fig. 2C). HI listeners did not show statistically significant adaptation for natural words ($N = 25$; $Z = -1.5$; $p = 0.061$) but they did show adaptation for vocoded words ($N = 24$; $Z = -2.3$; $p = 0.010$) (Fig. 2D). The amount of adaptation to noise was low overall and not significantly different for the two listener groups when listeners were presented with natural words (0.5 dB for NH; 0.4 dB for HI; Mann-Whitney $U = 332.0$; $p = 0.460$). In contrast, adaptation was greater for NH listeners (1.7 dB) than for HI listeners (0.7 dB) when they were presented with vocoded words ($U = 113.0$; $p = 0.027$).

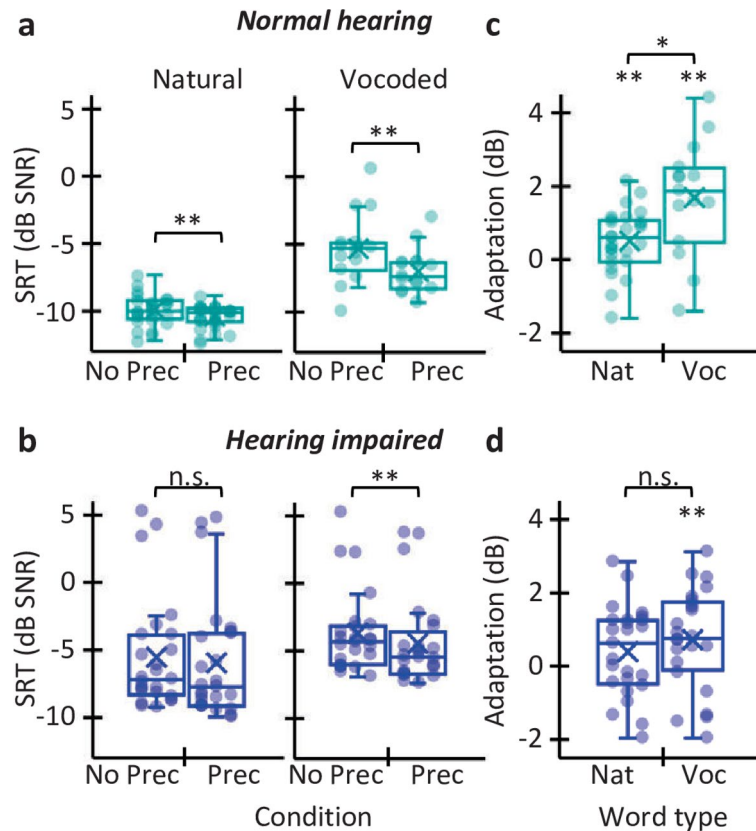


Fig. 2. SRTs and adaptation to noise for NH and HI listeners. **(A)** SRTs for NH listeners presented with natural ($N=27$) or vocoded ($N=15$) words in steady noise. SRTs are shown without ('No Prec') and with ('Prec') steady noise precursor. Bottom, middle, and top lines in each box plot indicate the 25th, 50th (median), and 75th percentiles, respectively. Crosses represent mean scores. Dots illustrate individual scores. **(B)** As panel (A) but for HI listeners. The sample size was 25 and 24 participants for natural and vocoded words, respectively. **(C)** Adaptation to noise for NH listeners for natural and vocoded words. Positive values indicate better SRTs with than without a precursor. **(D)** As panel (C) but for HI listeners. * $p < 0.05$; ** $p \leq 0.01$.

Because HI listeners did not adapt to noise with natural words, and showed less than normal adaptation with vocoded words, the results show that hearing loss impairs adaptation to noise. To investigate this further, Fig. 3A, B shows the relationship between the PTA threshold and the SRTs measured with and without the noise precursor. Clearly, the effect of the precursor on SRT decreases as the PTA increases, indicating that hearing loss impairs adaptation to noise. [Similar results were found when the PTA threshold was calculated by averaging audiometric thresholds across all audiometric frequencies (0.25–8 kHz) (not shown)]. Figure 3C, D shows the amount of adaptation as a function of PTA. The correlation was not statistically significant for natural words (Spearman Rho $\rho = -0.19$, $N = 52$, $p = 0.089$) but it was significant for vocoded words ($\rho = -0.41$, $N = 39$, $p = 0.005$). For the two types of words, listeners no longer benefited from adaptation when the PTA is 70 dB HL or higher (i.e., the regression line approaches zero).

The relationship between PTA and adaptation might, however, be reflecting the effect of a mixture of variables because PTA is correlated with age ($\rho = 0.705$, $N = 53$, $p < 0.001$) as well as with the baseline SRT measured without the precursor (SRT_{NoPrec}) (natural: $\rho = 0.743$, $N = 52$, $p < 0.001$; vocoded: $\rho = 0.361$, $N = 39$, $p = 0.012$). Age²² and the baseline SRT_{NoPrec} can affect the amount of adaptation. To investigate the specific impact of hearing loss on adaptation, we performed a semipartial correlation between PTA and adaptation to noise while controlling for the effects of age and SRT_{NoPrec} on the PTA. Spearman's rank-order semipartial correlation revealed a correlation between PTA and adaptation to noise for natural ($\rho = -0.56$, $N = 52$, $p < 0.001$) and vocoded words ($\rho = -0.45$, $N = 39$, $p = 0.002$) (Fig. 3E, F). This finding suggests that hearing loss per se, rather than age or baseline intelligibility, degrades adaptation to noise.

Note that our sample includes 25 HI listeners but only four of them had severe or close to severe hearing loss (PTA ≥ 68 dB HL) (Fig. 3A–B). Three of those listeners showed significantly worse SRTs than the rest of the sample (Fig. 2B) and one might think that the relationships shown in Fig. 3A–B may disappear when the SRTs (~ 4 dB SNR) from those listeners are excluded. This, however, was not the case (supplementary Fig. S1). Moreover, the Spearman's rank-order semipartial correlation between PTA and adaptation to noise remains significant for natural ($\rho = -0.49$, $N = 49$, $p < 0.001$) and vocoded words ($\rho = -0.31$, $N = 36$, $p = 0.033$) when the data for those three HI listeners are omitted (supplementary Fig. S1E, F). This corroborates that the results described above are not biased by those three participants.

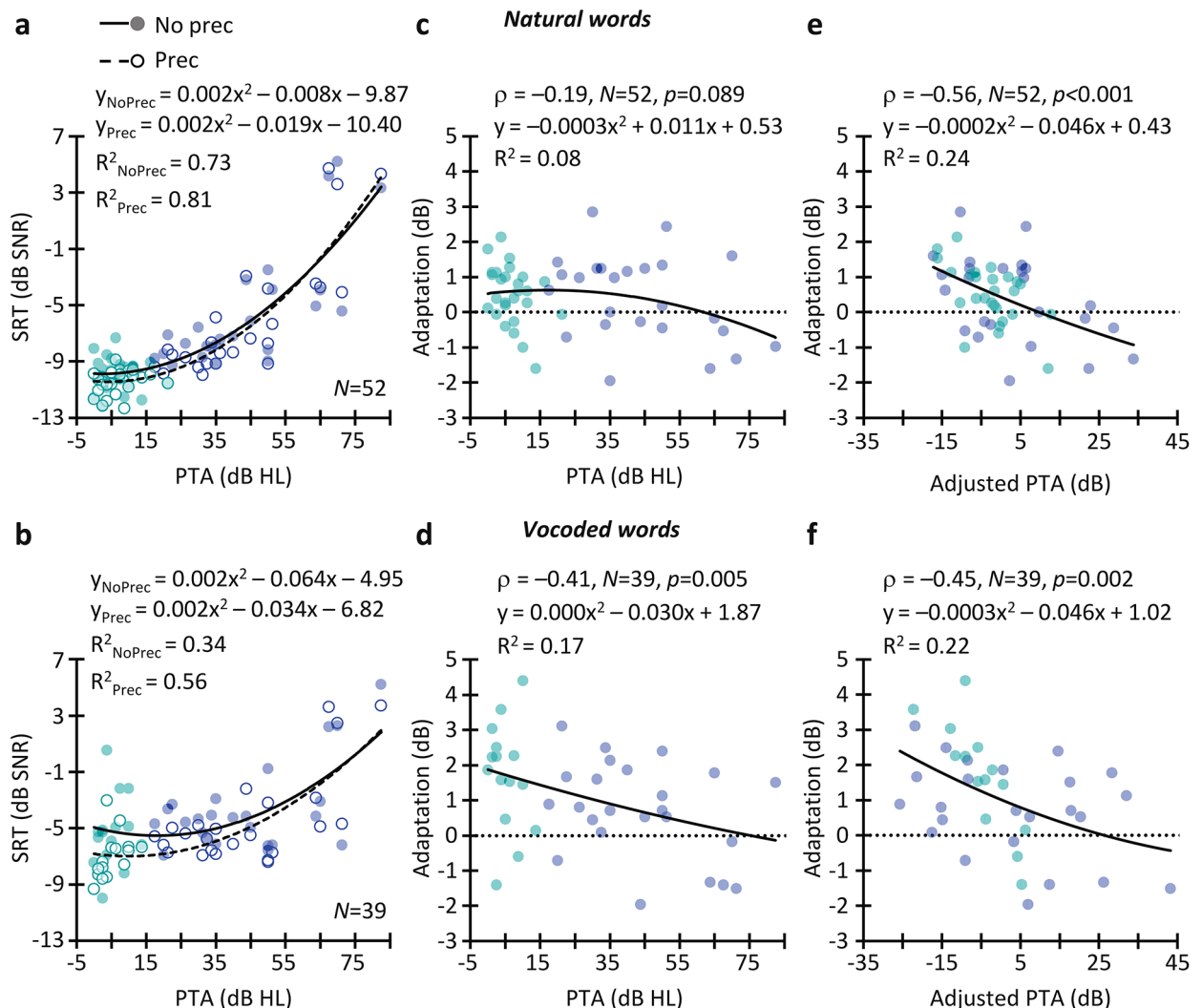


Fig. 3. Relationship between hearing loss and adaptation to noise. **(A)** SRTs for natural words measured without (filled dots; solid line) and with a noise precursor (empty dots; dashed line) as a function of PTA (0.5–4 kHz). Green dots depict data for NH listeners, and blue dots depict data for HI listeners. Lines indicate quadratic regression fits to the data. **(B)** As A but for vocoded words. **(C)** Relationship between adaptation to noise for natural words and PTA. **(D)** As C but for vocoded words. **(E)** Relationship between adaptation to noise for natural words and PTA when the effects of age and SRT_{NoPrec} were removed from the PTA. **(F)** As E but for vocoded words.

Discussion

The present study aimed to investigate the impact of hearing loss on the ability to adapt to noise during speech recognition. We found that, in aided conditions, the SRT for words in competition with SSN worsens with increasing hearing loss, both for natural and vocoded words. We also found that adaptation to noise decreases as hearing loss increases, both for natural and vocoded words. These relationships held when the effect of age on PTA was partialled out, which demonstrates that the loss of adaptation related to hearing loss and not to ageing.

The loss of adaptation to noise adds to the loss of speech information

Previous studies have shown that hearing loss impairs the encoding of and thus the access to speech TFS^{4,10,36,42,43} and envelope information^{42,44,45}. It also degrades the resolution of the long-term speech spectrum in the auditory system because of poor frequency selectivity^{2,46–50}. In addition, hearing loss at frequencies higher than those commonly measured in the audiogram (8–20 kHz) contributes to impaired speech recognition in broadband noise⁵¹. Here, we demonstrate that hearing loss not only impairs access to speech acoustic cues, but also impairs the ability of listeners to adapt to noise.

Figure 4 illustrates the contribution of various factors to the loss of intelligibility for the present HI listeners with the greatest hearing loss (supplementary Fig. S2 shows similar data for participants with smaller losses). Panel A shows the SRTs for the six NH listeners with the best PTA thresholds (from 0 to 3 dB HL; mean = 1.7 dB HL) and for the six HI listeners with the worst PTA thresholds (from 64 to 83 dB HL; mean = 70.0 dB HL).

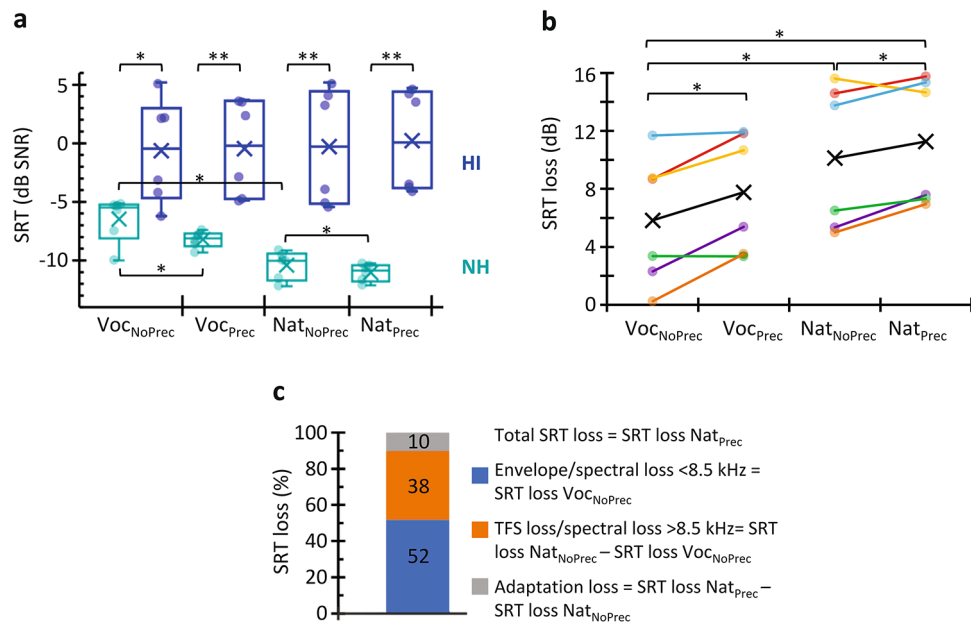


Fig. 4. The contribution of different factors to the loss of intelligibility for HI listeners with 64 to 83 dB hearing loss. **(A)** SRTs for the six listeners with the best (NH, green) and the worst PTA thresholds (HI, blue) in the four experimental conditions of the study. * $p < 0.05$, ** $p \leq 0.01$ for Mann-Whitney test (across participants) or Wilcoxon test (across condition). **(B)** Difference between the SRTs for HI listeners and the mean SRT for NH listeners in the respective condition. Each color line depicts the results for one HI listener. Black line with crosses depicts the mean. **(C)** Contribution of the loss of speech cues and adaptation to noise to the total SRT loss.

Six listeners were chosen for each group as a compromise between showing the most extreme differences and not including only the participants with the worst and the best PTA. Figure 4A shows that NH listeners had better SRTs for natural than for vocoded speech and even better SRTs when they had the opportunity to adapt to the noise. In contrast, HI listeners exhibited worse SRTs than NH listeners ($p < 0.05$) and virtually constant SRTs across conditions ($p \geq 0.147$). Figure 4B shows the SRT loss, which was calculated as the difference between the SRTs for individual HI listeners and the mean SRT for NH listeners in each condition. The largest SRT loss occurred for natural words with precursor (mean = 11.3 dB; Fig. 4B) presumably because in this condition HI listeners suffered from impaired access to all the possible speech cues (envelope, TFS, spectrum) plus a loss of adaptation. For convenience, we will refer to the largest SRT loss as ‘total’ SRT loss (Fig. 4C). To shed further light on the contribution of each factor to the total SRT loss we interpreted the results as follows. First, because the vocoder preserved only the speech envelope and spectral information below 8.5 kHz (the highest cutoff frequency of the filter bank in the vocoder), the worse SRTs for vocoded words for HI listeners than for NH listeners (Voc_{NoPrec} in Fig. 4A, B) reveal an impaired ability of HI listeners to use envelope and/or spectral information below 8.5 kHz. This impairment explained 52% of the total SRT loss (Fig. 4C). Second, because the recognition of natural speech depends on all the same cues as the recognition of vocoded speech plus TFS³⁶ and high-frequency (> 8.5 kHz) information⁵¹, the difference between natural and vocoded words reveals the ability to use TFS and high-frequency spectral information. Because NH but not HI listeners showed better SRTs for natural words than for vocoded words (Voc_{NoPrec} vs. Nat_{NoPrec} in Fig. 4A), our results suggest that NH but not HI listeners benefit from adding TFS or high-frequency spectral information to speech. The lack of TFS and high-frequency spectral speech information explained 38% of the total SRT loss (Fig. 4C). Lastly, the improvement in SRT when a precursor is present shows the benefit of adaptation to noise. NH but not HI listeners showed adaptation (Fig. 4A), and the loss of adaptation explained the remaining 10% of the total SRT loss (Fig. 4C). Taken together, the present results show that HI listeners have difficulties recognizing audible speech in noisy conditions not only because they are less able to use/encode speech acoustic cues but also because they are less able to adapt to noisy backgrounds.

On the reduced adaptation to noise for HI listeners

We found that hearing loss impairs adaptation to noise (Figs. 3 and 4). Multiple mechanisms can underlie this result^{35,52}. Among them may be a decrease in the MOC reflex (MOCR)-mediated adjustment of the cochlear gain in the damaged cochlea. MOC fibers are reflexively activated with a time course of 277 ± 62 ms⁵³. MOC efferents terminate upon OHCs, and their activation inhibits the gain of the cochlear amplifier, linearizing BM responses and reducing compression⁵⁴. Jennings et al.³⁴ reasoned that HI listeners do not show adaptation in AM detection because they have more linear BM responses and hence less MOCR-mediated BM linearization. A smaller adaptation to noise in speech recognition might also occur if the smaller MOCR-mediated BM linearization produces less enhancement of the speech envelope at the output of the BM responses in HI listeners. Some

studies, however, have shown that adaptation to noise in speech recognition or AM detection can occur without MOCR effects^{16,20}.

Another proposed mechanism for adaptation to noise is a shift in the dynamic range of auditory neurons toward the most common level in the noise preceding the word^{16,17,20,21,35}. When auditory neurons are presented with low-varying-level noise, neurons shift their dynamic ranges toward the most common noise level so long as the level in question is above the neuron's threshold. This adaptation increases a neuron's sensitivity to changes in sound level^{55,56}. The improvement in sensitivity to level changes, however, is smaller when the variance in the stimulus level is large^{55,57}. Because audiometric hearing loss is often associated with the loss or dysfunction of inner hair cells (IHCs) and OHCs^{58,59}, it is possible that the reduced auditory peripheral compression due to OHC loss causes the IHC receptor potential representation to fluctuate more than normal. This could result in neurons not having a prevailing level to adapt to, thus in less adaptation to noise for HI listeners than for NH listeners¹⁷. However, whether the inherent fluctuations of the steady noise at the output of a linear BM response are fluctuating enough to impair dynamic range adaptation to sound level statistics has yet to be determined.

A third potential mechanism is related to the disruption of the cues used by listeners to segregate streams, i.e., to segregate the speech and noise stimuli. Jennings et al.⁶⁰ measured pure tone detection thresholds for short tones presented 2 or 197 ms after the onset of a 400 ms narrowband noise masker with a flattened or inherent fluctuating temporal envelope. They found that when the probe was delayed in the noise, detection thresholds improved for the tones presented in the flattened noise but worsened for the tones presented in the fluctuating noise. Jennings et al. reasoned that listeners rely on a temporal envelope-based cue to detect the probe and that the fluctuations in the preceding noise disrupt this cue. It has been shown that the amplitude fluctuations of steady noise can also impair speech recognition when presented simultaneously with the speech, presumably because it is difficult to distinguish the noise fluctuations from the amplitude fluctuations that convey speech information^{61,62}. However, it is uncertain whether the precursor fluctuations can hinder speech recognition. If they did, the more linear BM responses of HI listeners would enhance the precursor noise fluctuations, making it harder for HI listeners to distinguish between noise and speech fluctuations, and resulting in less adaptation for HI listeners than for NH listeners. This, however, does not seem to be the case here because although SRTs were sometimes worse with the precursor, the worsening occurred for NH and HI listeners, so it did not seem to be related to the hearing loss (points below zero in Fig. 3C, D).

Implications

The present study shows that adaptation to noise in speech recognition decreases with increasing hearing loss, even when the potential confounding effect of age is factored out. This finding is relevant in at least two ways. First, as explained earlier, HI listeners find it harder to recognize audible speech in noisy settings than NH listeners^{1–5,8}. Impaired access to speech acoustic cues (envelope, TFS, spectrum) has been shown to contribute to impaired intelligibility. The present study reveals that the impact of hearing loss on speech recognition can be underestimated if adaptation is disregarded (Fig. 4). Second, research on the factors that can affect speech-in-noise intelligibility (for both NH and HI listeners) is often conducted disregarding adaptation or, more generally, temporal effects. Indeed, the speech-to-noise onset delay varies widely across studies (e.g., 500 ms in Johannesen et al.²⁴; 3 s in Souza et al.²⁶), sometimes it is not even reported^{25,27–29}, and most times it is not justified. The present study suggests that the relevance of some factors may differ depending on the time when the speech is presented relative to the noise onset. For instance, cochlear mechanical dysfunction does not predict intelligibility in speech maskers for HI listeners¹¹, but it could predict the SRT impairment related to a loss of adaptation to noise.

Conclusions

We measured aided SRTs for words in competition with a steady SSN for listeners with different degrees of sensorineural hearing loss. Words were preceded or not by 1-second-long precursor adapting noise. We found that SRTs worsen and that adaptation to noise in speech recognition decreases with increasing hearing loss when speech is both natural and vocoded. Overall, the present results suggest that HI listeners exhibit impaired intelligibility in noisy environments not only because they are less able to use speech acoustic cues but also because they are less able to adapt to background noise. Impaired adaptation can contribute up to 10% of the SRT loss in HI listeners.

Data availability

All data are provided in the manuscript figures and supplementary information files. The actual data values will be provided by the corresponding author (E.A.L.P.) upon reasonable request.

Received: 7 May 2024; Accepted: 15 November 2024

Published online: 20 November 2024

References

- Summers, V., Makashay, M. J., Theodoroff, S. M. & Leek, M. R. Suprathreshold auditory processing and speech perception in noise: Hearing-impaired and normal-hearing listeners. *J. Am. Acad. Audiol.* **24**, 274–292 (2013).
- Bağkent, D. Speech recognition in normal hearing and sensorineural hearing loss as a function of the number of spectral channels. *J. Acoust. Soc. Am.* **120**, 2908–2925 (2006).
- Duquesnoy, A. J. & Plomp, R. The effect of a hearing aid on the speech-reception threshold of hearing-impaired listeners in quiet and in noise. *J. Acoust. Soc. Am.* **73**, 2166–2173 (1983).
- Hopkins, K., Moore, B. C. J. & Stone, M. A. Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. *J. Acoust. Soc. Am.* **123**, 1140–1153 (2008).
- Moore, B. C. J., Peters, R. W. & Stone, M. A. Benefits of linear amplification and multichannel compression for speech comprehension in backgrounds with spectral and temporal dips. *J. Acoust. Soc. Am.* **105**, 400–411 (1999).

6. Plomp, R. Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J. Acoust. Soc. Am.* **63**, 533–549 (1978).
7. Plomp, R. A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired. *J. Speech Hear. Res.* **29**, 146–154 (1986).
8. Lopez-Poveda, E. A. Why do I hear but not understand? Stochastic undersampling as a model of degraded neural encoding of speech. *Front. Neurosci.* **8**, 1–7 (2014).
9. Baer, T. & Moore, B. C. J. Effects of spectral smearing on the intelligibility of sentences in noise. *J. Acoust. Soc. Am.* **94**, 1229–1241 (1993).
10. Lorenzi, C., Gilbert, G., Carn, H., Garnier, S. & Moore, B. C. Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc. Natl. Acad. Sci.* **103**, 18866–18869 (2006).
11. Lopez-Poveda, E. A. et al. Predictors of hearing-aid outcomes. *Trends Hear.* **21**, 2331216517730526 (2017).
12. Lopez-Poveda, E. A., Eustaquio-Martín, A. & San-Victoriano, F. M. Binaural pre-processing for contralateral sound field attenuation and improved speech-in-noise recognition. *Hear. Res.* **418**, 108469 (2022).
13. San-Victoriano, F. M., Eustaquio-Martín, A. & Lopez-Poveda, E. A. Binaural pre-processing for contralateral sound field attenuation can improve speech-in-noise intelligibility for bilateral hearing-aid users. *Hear. Res.* **432**, 108743 (2023).
14. Cervera, T. & Ainsworth, W. A. Effects of preceding noise on the perception of voiced plosives. *Acta Acust United Acust.* **91**, 132–144 (2005).
15. Cervera, T. & Gonzalez-Alvarez, J. Temporal effects of preceding band-pass and band-stop noise on the recognition of voiced stops. *Acta Acust United Acust.* **93**, 1036–1045 (2007).
16. Marrufo-Pérez, M. I., Eustaquio-Martín, A. & Lopez-Poveda, E. A. Adaptation to noise in human speech recognition unrelated to the medial olivocochlear reflex. *J. Neurosci.* **38**, 4138–4145 (2018).
17. Marrufo-Pérez, M. I., Sturla-Carreto, D. P., Eustaquio-Martín, A. & Lopez-Poveda, E. A. Adaptation to noise in human speech recognition depends on noise-level statistics and fast dynamic-range compression. *J. Neurosci.* **40**, 6613–6623 (2020).
18. Zwicker, E. Temporal effects in simultaneous masking and loudness. *J. Acoust. Soc. Am.* **38**, 132–141 (1965).
19. Marrufo-Pérez, M. I., Eustaquio-Martín, A. & López-Bascuas, L. E. & Lopez-Poveda, E. A. Temporal effects on monaural amplitude-modulation sensitivity in ipsilateral, contralateral and bilateral noise. *J. Assoc. Res. Otolaryngol.* **19**, 147–161 (2018).
20. Marrufo-Pérez, M. I. et al. Adaptation to noise in amplitude modulation detection without the medial olivocochlear reflex. *Hear. Res.* **377**, 133–141 (2019).
21. Ainsworth, W. A. & Meyer, G. F. Recognition of plosive syllables in noise: Comparison of an auditory model with human performance. *J. Acoust. Soc. Am.* **96**, 687–694 (1994).
22. Ben-David, B. M., Tse, V. Y. & Schneider, B. A. Does it take older adults longer than younger adults to perceptually segregate a speech target from a background masker? *Hear. Res.* **290**, 55–63 (2012).
23. Ben-David, B. M., Avivi-Reich, M. & Schneider, B. A. Does the degree of linguistic experience (native versus nonnative) modulate the degree to which listeners can benefit from a delay between the onset of the maskers and the onset of the target speech? *Hear. Res.* **341**, 9–18 (2016).
24. Johannsen, P. T., Pérez-González, P., Kalluri, S., Blanco, J. L. & Lopez-Poveda, E. A. The influence of cochlear mechanical dysfunction, temporal processing deficits, and age on the intelligibility of audible speech in noise for hearing-impaired listeners. *Trends Hear.* **20**, 2331216516641055 (2016).
25. Saunders, G. H. & Forsline, A. The performance-perceptual test (PPT) and its relationship to aided reported handicap and hearing aid satisfaction. *Ear Hear.* **27**, 229–242 (2006).
26. Souza, P. et al. Understanding variability in individual response to hearing aid signal processing in wearable hearing aids. *Ear Hear.* **40**, 1280–1292 (2019).
27. Tognola, G., Mainardi, A., Vincenti, V. & Cuda, D. Benefit of hearing aid use in the elderly: The impact of age, cognition and hearing impairment. *Acta Otorhinolaryngol. Ital.* **39**, 409–418 (2019).
28. Wu, M., Cañete, O. M., Schmidt, J. H., Fereczkowski, M. & Neher, T. Influence of three auditory profiles on aided speech perception in different noise scenarios. *Trends Hear.* **25**, 23312165211023709 (2021).
29. Wu, M., Christiansen, S., Fereczkowski, M. & Neher, T. Revisiting auditory profiling: Can cognitive factors improve the prediction of aided speech-in-noise outcome? *Trends Hear.* **26**, 23312165221113889 (2022).
30. Bacon, S. P. & Takahashi, G. A. Overshoot in normal-hearing and hearing-impaired subjects. *J. Acoust. Soc. Am.* **91**, 2865–2871 (1992).
31. Carlyon, R. P. & Sloan, E. P. The overshoot effect and sensory hearing impairment. *J. Acoust. Soc. Am.* **82**, 1078–1081 (1987).
32. Strickland, E. A. & Krishnan, L. A. The temporal effect in listeners with mild to moderate cochlear hearing impairment. *J. Acoust. Soc. Am.* **118**, 3211–3217 (2005).
33. Jennings, S. G., Ahlstrom, J. B. & Dubno, J. R. Effects of age and hearing loss on overshoot. *J. Acoust. Soc. Am.* **140**, 2481–2493 (2016).
34. Jennings, S. G., Chen, J., Fultz, S. E., Ahlstrom, J. B. & Dubno, J. R. Amplitude modulation detection with a short-duration carrier: Effects of a precursor and hearing loss. *J. Acoust. Soc. Am.* **143**, 2232–2243 (2018).
35. Marrufo-Pérez, M. I. & Lopez-Poveda, E. A. Adaptation to noise in normal and impaired hearing. *J. Acoust. Soc. Am.* **151**, 1741–1753 (2022).
36. Moore, B. C. J. The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *J. Assoc. Res. Otolaryngol.* **9**, 399–406 (2008).
37. Byrne, D. & Dillon, H. The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. *Ear Hear.* **7**, 257–265 (1986).
38. AAO-HNS. Academy responses to the FDA request for comment on hearing aid regulations. *Am. Acad. Otolaryngol.-Head Neck Surg. Bull.* **16–17**, 26–28 (1993).
39. Hood, D. The principles and practice of bone conduction audiometry: A review of the present position. *Laryngoscope* **70**, 1211–1228 (1960).
40. Cárdenas, M. & Marrero, V. *Cuaderno de logaudiometría* (Universidad Nacional de Educación a Distancia, 1994).
41. Levitt, H. Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* **49**, 467–677 (1971).
42. Kale, S. & Heinz, M. G. Envelope coding in auditory nerve fibers following noise-induced hearing loss. *J. Assoc. Res. Otolaryngol.* **11**, 657–673 (2010).
43. Moore, B. C. J. The importance of temporal fine structure for the intelligibility of speech in complex backgrounds. *Proc. Intl Symp. Auditory Audiol. Res.* **3**, 21–32 (2011).
44. Füllgrabe, C., Meyer, B. & Lorenzi, C. Effect of cochlear damage on the detection of complex temporal envelopes. *Hear. Res.* **178**, 35–43 (2003).
45. Özmeral, E. J., Eddins, A. C. & Eddins, D. A. How do age and hearing loss impact spectral envelope perception? *J. Speech Lang. Hear. Res.* **61**, 2376–2385 (2018).
46. Leek, M. R. & Summers, V. Reduced frequency selectivity and the preservation of spectral contrast in noise. *J. Acoust. Soc. Am.* **100**, 1796–1806 (1996).
47. Henry, B. A., Turner, C. W. & Behrens, A. Spectral peak resolution and speech recognition in quiet: normal hearing, hearing impaired, and cochlear implant listeners. *J. Acoust. Soc. Am.* **118**, 1111–1121 (2005).
48. Noordhoek, I. M., Houtgast, T. & Festen, J. M. Measuring the threshold for speech reception by adaptive variation of the signal bandwidth. II. Hearing-impaired listeners. *J. Acoust. Soc. Am.* **107**, 1685–1696 (2000).

49. Stelmachowicz, P. G., Jesteadt, W., Gorga, M. P. & Mott, J. Speech perception ability and psychophysical tuning curves in hearing-impaired listeners. *J. Acoust. Soc. Am.* **77**, 620–627 (1985).
50. Turner, C. W., Chi, S. L. & Flock, S. Limiting spectral resolution in speech for listeners with sensorineural hearing loss. *J. Speech Lang. Hear. Res.* **42**, 773–784 (1999).
51. Zadeh, L. M. et al. Extended high-frequency hearing enhances speech perception in noise. *Proc. Natl. Acad. Sci.* **116**, 23753–23759 (2019).
52. Willmore, B. D. B. & King, A. J. Adaptation in auditory processing. *Physiol. Rev.* **103**, 1025–1058 (2023).
53. Backus, B. C. & Guinan, J. J. Time-course of the human medial olivocochlear reflex. *J. Acoust. Soc. Am.* **119**, 2889–2904 (2006).
54. Murugasu, E. & Russell, I. J. The effect of efferent stimulation on basilar membrane displacement in the basal turn of the guinea pig cochlea. *J. Neurosci.* **16**, 325–332 (1996).
55. Dean, I., Harper, N. S. & McAlpine, D. Neural population coding of sound level adapts to stimulus statistics. *Nat. Neurosci.* **8**, 1684–1689 (2005).
56. Wen, B., Wang, G. I., Dean, I. & Delgutte, B. Dynamic range adaptation to sound level statistics in the auditory nerve. *J. Neurosci.* **29**, 13797–13808 (2009).
57. Rabinowitz, N. C., Willmore, B. D. B., Schnupp, J. W. H. & King, A. J. Contrast gain control in auditory cortex. *Neuron* **70**, 1178–1191 (2011).
58. Liberman, M. C. & Dodds, L. W. Single-neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves. *Hear. Res.* **16**, 55–74 (1984).
59. Wu, P. Z., O'Malley, J. T., de Gruttola, V. & Liberman, M. C. Age-related hearing loss is dominated by damage to inner ear sensory cells, not the cellular battery that powers them. *J. Neurosci.* **40**, 6357–6366 (2020).
60. Jennings, S. G., Sivas, K. & Stone, C. Effects of masker envelope fluctuations on the temporal effect. *J. Assoc. Res. Otolaryngol.* **19**, 717–727 (2018).
61. Stone, M. A., Füllgrabe, C., Mackinnon, R. C. & Moore, B. C. J. The importance for speech intelligibility of random fluctuations in steady background noise. *J. Acoust. Soc. Am.* **130**, 2874–2881 (2011).
62. Stone, M. A., Füllgrabe, C. & Moore, B. C. J. notionally steady background noise acts primarily as a modulation masker of speech. *J. Acoust. Soc. Am.* **132**, 317–326 (2012).

Acknowledgements

This work was supported by the Spanish Ministry of Science and Innovation (grant PID2019-108985GB-I00) and the European Regional Development Fund.

Author contributions

M.I.M.P. and E.A.L.P. designed the research; M.I.M.P. and M.J.F. performed the research; A.E.M. contributed research tools; M.I.M.P. analyzed the data; M.I.M.P. and E.A.L.P. wrote the paper. E.A.L.P. secured funding and supervised the research. All authors read and reviewed the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-80131-9>.

Correspondence and requests for materials should be addressed to E.A.L.-P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024