The Influence of Cochlear Mechanical Dysfunction, Temporal Processing Deficits, and Age on the Intelligibility of Audible Speech in Noise for Hearing-Impaired Listeners

Peter T. Johannesen[1,2], Patricia Pérez-González[1,2],
Sridhar Kalluri[3], José L. Blanco[1], and Enrique A. Lopez-Poveda[1,2,4]

Abstract

The aim of this study was to assess the relative importance of cochlear mechanical dysfunction, temporal processing deficits, and age on the ability of hearing-impaired listeners to understand speech in noisy backgrounds. Sixty-eight listeners took part in the study. They were provided with linear, frequency-specific amplification to compensate for their audiometric losses, and intelligibility was assessed for speech-shaped noise (SSN) and a time-reversed two-talker masker (R2TM). Behavioral estimates of cochlear gain loss and residual compression were available from a previous study and were used as indicators of cochlear mechanical dysfunction. Temporal processing abilities were assessed using frequency modulation detection thresholds. Age, audiometric thresholds, and the difference between audiometric threshold and cochlear gain loss were also included in the analyses. Stepwise multiple linear regression models were used to assess the relative importance of the various factors for intelligibility. Results showed that (a) cochlear gain loss was unrelated to intelligibility, (b) residual cochlear compression was related to intelligibility in SSN but not in a R2TM, (c) temporal processing was strongly related to intelligibility in a R2TM and much less so in SSN, and (d) age per se impaired intelligibility. In summary, all factors affected intelligibility, but their relative importance varied across maskers.

Hearing-impaired (HI) listeners vary widely in their ability to understand speech in noisy backgrounds, even when the detrimental effect of their hearing loss on intelligibility is compensated for with frequency-specific amplification (e.g., Peters, Moore, & Baer, 1998). The aim of the present study was to shed some light on the relative importance of cochlear mechanical dysfunction, temporal processing deficits, and age on the ability of HI listeners to understand audible speech in noisy backgrounds.

Several factors can affect the ability of HI listeners to understand audible speech in noise (reviewed by Lopez-Poveda, 2014). One of them is outer hair cell (OHC) loss or dysfunction. OHC dysfunction would degrade the representation of the speech spectrum in the mechanical response of the cochlea, particularly in noisy environments, for various reasons. First, OHC dysfunction reduces cochlear frequency selectivity (Robles & Ruggero, 2001). This can smear the cochlear

[1]Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Spain
[2]Instituto de Investigación Biomédica de Salamanca, Universidad de Salamanca, Spain
[3]Starkey Hearing Research Center, Berkeley, CA, USA
[4]Departamento de Cirugía, Facultad de Medicina, Universidad de Salamanca, Spain

Corresponding author:
Enrique A. Lopez-Poveda, Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Calle Pintor Fernando Gallego 1, 37007 Salamanca, Spain.
Email: ealopezpoveda@usal.es

representation of the acoustic spectrum, making it harder for HI listeners to separately perceive the spectral cues of speech from those of interfering sounds (Moore, 2007). Several behavioral studies have supported this idea (Baer & Moore, 1994; Festen & Plomp, 1983; ter Keurs, Festen, & Plomp, 1992). Second, in the healthy cochlea, suppression might facilitate the encoding of speech in noise by enhancing the most salient spectral features of speech against those of the background noise (Deng, Geisler, & Greenberg, 1987; Young, 2008). OHC dysfunction reduces suppression and this might hinder speech-in-noise intelligibility. Third, cochlear mechanical compression might facilitate the understanding of speech in interrupted or fluctuating noise by amplifying the speech in the low-level noise intervals, a phenomenon known as *listening in the dips* (e.g., Gregan, Nelson, & Oxenham, 2013; Rhebergen, Versfeld, & Dreschler, 2009). OHC loss or dysfunction reduces compression (i.e., linearizes cochlear responses; Ruggero, Rich, Robles, & Recio, 1996) and thus could hinder *dip listening* (Gregan et al., 2013). Fourth, medial olivocochlear efferents possibly facilitate the intelligibility of speech in noise by increasing the discriminability of transient sounds in noisy backgrounds (Brown, Ferry, & Meddis, 2010; Guinan, 2010; Kim, Frisina, & Frisina, 2006). Medial olivocochlear efferents exert their action via OHCs, and so OHC dysfunction could reduce the unmasking effects of medial olivocochlear efferents.

The degree of OHC dysfunction could be different across different HI listeners and this might contribute to the wide variability in their ability to understand audible speech in noise. While seemingly reasonable, however, this view is almost certainly only partially correct. First, for HI listeners, there appears to be no significant correlation between residual cochlear compression and the benefit from dip listening (Gregan et al., 2013), which undermines the influence of compression on the intelligibility of suprathreshold speech in noise. Second, at high intensities, cochlear tuning for healthy cochleae is reduced and is only moderately sharper than that of impaired cochleae (Robles & Ruggero, 2001) and yet HI listeners still perform more poorly than do normal-hearing (NH) listeners in speech-in-noise intelligibility tests (reviewed in pp. 205–208 of Moore, 2007).

A second factor that may affect speech-in-noise intelligibility is age. Elderly listeners with normal audiometric thresholds and presumably healthy OHCs have more difficulty in understanding speech in noise than young, NH listeners (Committee on Hearing, Bioacoustics, and Biomechanics, 1988; Kim et al., 2006; Peters et al., 1998), which suggests that age per se or mechanisms other than OHC dysfunction can limit the intelligibility of audible speech.

A third factor that might contribute to the ability of HI listeners to understand suprathreshold speech in noise is temporal processing ability. Several studies support this view. First, for HI listeners, speech-in-noise intelligibility is correlated with their ability to use the information conveyed in the rapid temporal changes that occur in speech, known as *temporal fine structure* (TFS; Lorenzi, Gilbert, Carn, Garnier, & Moore, 2006; Strelcyk & Dau, 2009). Second, both stochastic under-sampling of a noisy speech waveform, as might occur after cochlear synaptopathy or deafferentation (Lopez-Poveda & Barrios, 2013), and temporally jittering the frequency components in speech (Pichora-Fuller, Schneider, Macdonald, Pass, & Brown, 2007) decrease speech-in-noise intelligibility with negligible effects on absolute threshold. Third, Henry and Heinz (2012) showed that the synchronization of auditory nerve discharges to a sound's waveform decreases in noise backgrounds and that the decrease is greater for cochleae with OHC dysfunction than for healthy cochleae. Therefore, it is possible that OHC dysfunction causes temporal processing deficits that manifest in noise. Altogether this suggests that normal processing of speech temporal cues is required for understanding audible speech in noise. Temporal processing abilities could vary across HI listeners and this could contribute to the wide variability in their ability to understand speech in noise.

The present study was aimed at assessing the relative importance of cochlear mechanical dysfunction, temporal processing deficits, and age for understanding audible speech in noisy environments by HI listeners. In a previous study, we reported behaviorally inferred estimates of cochlear mechanical gain loss and residual compression for 68 HI listeners (Johannesen, Perez-Gonzalez, & Lopez-Poveda, 2014). Those estimates were used in the present study as indicators of cochlear dysfunction and the participants from that study were invited back into the laboratory to assess their ability to understand audible speech in various types of noise, as well as their ability to process temporal information. After Moore and Sek (1996), temporal processing ability was assessed using frequency modulation detection thresholds (FMDTs), defined as the minimum detectable excursion in frequency for a pure tone carrier. FMDTs are thought to be dependent on the quality with which frequencies are coded in the phase locking of auditory nerve discharges and on the ability of a listener to discriminate frequencies based on such a code (Moore & Sek, 1996). Speech-in-noise intelligibility was assessed using the speech reception threshold (SRT), defined as the speech-to-noise ratio (SNR) required to understand 50% of the sentences that listeners were presented within a noise background (e.g., Peters et al., 1998). When measuring SRTs, stimuli were linearly amplified in a frequency-specific manner to minimize the effect of reduced audibility on intelligibility. SRTs were measured

for two types of maskers: a steady noise with a speech--shaped long-term spectrum (speech-shaped noise [SSN]) and two-talker masker played in reverse (R2TM). The latter masker was used because it has the same temporal and spectral properties as forward speech and was thus expected to have the same energetic masking properties as speech but without semantic information that may contribute to informational masking (e.g., Hornsby & Ricketts, 2007). Stepwise multiple linear regression models (MLRs) were used to predict SRTs from a linear combination of cochlear dysfunction indicators, FMDTs, and age and to quantify the relative importance of these factors for speech-in-noise intelligibility.

Material and Methods

Subjects

The 68 subjects (43 males) with symmetrical sensorineural hearing losses who participated in the study of Johannesen et al. (2014) participated in the present study. Their ages ranged from 25 to 82 years, with a median of 61 years. The present study was part of a larger hearing-aid study and hence all participants were required to be hearing-aid users or candidates. Speech-in-noise intelligibility was assessed in bilateral listening conditions (see later). Indicators of cochlear mechanical status and temporal processing ability, however, were measured for one ear only. For most participants, the test ear was the ear with better audiometric thresholds in the frequency range of 2 to 6 kHz (30 left ears and 38 right ears). See Johannesen et al. for further details of the subject sample. All procedures were approved by the human experimentation ethical review board of the University of Salamanca. Subjects gave their signed informed consent prior to their inclusion in the study.

Indicators of Cochlear Mechanical Dysfunction

OHC dysfunction linearizes cochlear mechanical responses (Ruggero et al., 1996). Johannesen et al. (2014) compared behaviorally inferred cochlear input/output curves for each HI listener at each of five test frequencies (0.5, 1, 2, 4, and 6 kHz) with corresponding reference input/output curves for NH listeners. They reported three main variables from their analyses. One variable was cochlear mechanical gain loss (also referred to as OHC loss or HL_{OHC}) expressed in decibels (dB). It was defined as the contribution of cochlear gain loss to the audiometric threshold and was inferred from the difference (in dB) between the compression threshold for an individual input/output curve and the reference compression threshold for NH listeners at the corresponding test frequency corrected for compression (see Equation (2) in

Lopez-Poveda & Johannesen, 2012). A second variable was inner hair cell (IHC) loss or HL_{IHC}. It was defined as the difference (in dB) between the audiometric loss (in dB HL) and HL_{OHC}. This difference was reported after earlier studies where the audiometric loss was assumed to be the sum of a cochlear mechanical component, HL_{OHC}, and an additional component of an uncertain nature termed HL_{IHC} (Moore & Glasberg, 1997). A third variable reported by Johannesen et al. was the basilar-membrane compression exponent (BMCE). It was defined as the slope (in dB/dB) of an inferred cochlear input/output curve over its compressive segment. See Johannesen et al. for further details on how these variables were inferred.

HL_{OHC} and BMCE were readily available for most participants and test frequencies from the study of Johannesen et al. (2014) and were used in the present context as indicators of cochlear mechanical dysfunction. Specifically, HL_{OHC} was regarded here as an indicator of cochlear mechanical gain loss and the BMCE as an indicator of residual cochlear compression. For completeness, HL_{IHC} was also used in the present analysis. Note that the three variables had values at each of the five test frequencies.

Johannesen et al. (2014) reported that they could not infer input/output curves for listeners and test frequencies where the audiometric loss was too high. For the present analysis, we assumed that those cases were indicative of total cochlear gain loss. Therefore, for those cases, BMCE was set equal to 1 dB/dB, corresponding to a linear input/output curve, and HL_{OHC} was set equal to the maximum cochlear gain values for NH listeners. These were assumed to be 35.2, 43.5, 42.7, 42.7, and 42.7 dB at 0.5, 1, 2, 4, and 6 kHz, respectively, as reported on p. 11 of Johannesen et al.

Frequency Modulation Detection Thresholds

Temporal processing ability was assessed using FMDTs. The complete justification for this can be found in Moore and Sek (1996). We followed the procedure of Strelcyk and Dau (2009). In short, a FMDT was defined as the minimum detectable excursion in frequency for a tone carrier and was estimated using a three alternative forced choice adaptive procedure. The three intervals contained a pure tone with a frequency of 1500 Hz and a duration of 750 ms, including 50-ms raised cosine onset and offset ramps. The level of the tone was set 30 dB above the absolute threshold for the tone (30 dB sensation level). The tones in all intervals were also sinusoidally amplitude modulated (AM) with a modulation depth of 6 dB and with an instantaneous modulation rate that either increased or decreased linearly with time. The initial and final AM modulation rates were randomized in the interval between 1 and 3 Hz under

the constraint that the modulation rate change was always above 1 Hz. In one interval, chosen at random, the tone's frequency was additionally sinusoidally varied with a rate of 2 Hz and with a certain maximum frequency excursion. The listeners' task was to identify the interval containing the frequency modulation. The use of a low rate of frequency variation (2 Hz), a moderately low-frequency carrier (1500 Hz), and the randomized AM were intended to emphasize the dependency of FMDT on temporal information (*phase locking* to the carrier) rather than on changes in the excitation patterns (Moore & Sek, 1996); the use of a carrier tone frequency of 1500 Hz, thus in the middle of the speech frequency range, was intended to emphasize the dependency of FMDT on the listeners' ability to follow the temporal cues in speech. The listeners were provided with feedback about the correctness of their response. The logarithm of the maximum frequency excursion was varied in successive trials according to an adaptive one-up two-down rule to estimate the 71% correct point in the psychometric function (Levitt, 1971). The initial frequency excursion was set, so the target interval was always easily identified. The initial step size of the frequency excursion was $\log_{10}(1.5)$. This was decreased to $\log_{10}(1.26)$ after four reversals. The adaptive procedure continued until a total of 12 reversals in frequency excursion had occurred. The FMDT was calculated as the mean of the logarithms of the frequency excursions at the last eight reversals. A measurement was discarded if the standard deviation of the logarithm of the frequency excursions at the last eight reversals exceeded 0.15. Three threshold estimates were obtained in this way and their mean was taken as the threshold. If the standard deviation of these three measurements exceeded 0.15, one or more additional threshold estimates were obtained and included in the mean.

Prior to the FMDT task, the absolute threshold of the carrier tone was measured using a three alternative forced choice procedure in which the level of the tone was varied in successive trials according to an adaptive two-down, one-up rule to estimate the 71% correct point in the psychometric function (Levitt, 1971). The duration of the carrier tone was 750 ms, including 50-ms raised cosine onset and offset ramps.

Speech Reception Thresholds

SRTs were measured using the hearing-in-noise test (HINT; Nilsson, Soli, & Sullivan, 1994). Sentences uttered by a male speaker were presented to the listener in the presence of a masker. The sentences were those in the Castilian Spanish version of the HINT (Huarte, 2008). Two different maskers were used. One masker consisted of a steady Gaussian noise filtered in frequency to have the long-term average spectrum of speech

(Table 2 in Byrne et al., 1994). This masker will be referred to as SSN and the corresponding SRT as SRT_{SSN}. The second masker consisted of two simultaneous talkers (one male and one female) played in reverse. This masker will be referred to as time-reversed two-talker masker (R2TM) and the corresponding SRT as SRT_{R2TM}.

To generate the R2TM, the Spanish HINT sentences were uttered by a male and a female native Castilian-Spanish speaker in a sound booth and recorded with a Brüel & Kjaer type 4192 microphone with its amplifier (Brüel & Kjaer Nexus) connected to an RME Fireface 400 sound card. The recorded sentences were segmented and pauses between sentences removed. For each speaker, the sentences were equalized in root mean square amplitude, time reversed, and concatenated. Finally, the concatenated sentences of the male and the female speaker were mixed digitally. A different segment (chosen at random) of the resulting R2TM was used to mask each HINT sentence during an SRT measurement. We note that the long-term spectra of the R2TM, the SSN, and the target speech could have been slightly different because the speakers used to generate the R2TM and the target sentences were different, and the filter used to produce the SSN, though speech shaped, was not based on the long-term spectrum of the specific target sentences.

To measure an SRT, the speech was fixed in level at 65 dB SPL and the masker level was varied adaptively using a one-up, one-down rule to find the SNR (in dB) at which the listener correctly identified 50% of the sentences (i.e., to find the SRT in dB SNR units). After setting the levels of the speech and the masker, the two sounds were mixed digitally and filtered to simulate a free-field listening condition where the speech and the masker were colocated 1 m away in front of the listener at eye level (i.e., $0°$ azimuth and $0°$ elevation) and had a spectrum according to Table 3 in American National Standards Institute (ANSI, 1997). The filtering included corrections for the frequency response of the headphones. The resulting stimulus was linearly amplified individually for each participant according to the National Acoustics Laboratory Revised (NAL-R) rule (Byrne & Dillon, 1986) and played diotically to the listeners. The masker started 500 ms before and ended 250 ms after the target sentence. Twenty sentences were played to the listeners for each SRT measurement. In the adaptive procedure, the masker level was varied with a step size of 4 dB for the first 5 sentences and of 2 dB for the last 15 sentences. The SRT was calculated as the mean SNR used for the last 15 sentences. For each masker, three SRT estimates were obtained and the mean was taken as the final result. All other details of the procedure were as for the original HINT test (Nilsson et al., 1994). Feedback on the correctness of the listener's response was not provided.

Stimuli and Apparatus

For all measurements, stimuli were digitally generated or stored as digital files with a sampling rate of 44100 Hz. They were digital to analog converted using an RME Fireface 400 sound card with a 24-bit resolution and were played through Sennheiser HD-580 headphones. Subjects sat in a double-wall sound-attenuating booth during data collection.

Statistical Analyses

Audiometric pure-tone thresholds (PTT, in dB HL), HL_{OHC}, HL_{IHC}, BMCE, FMDTs, and age were used as potential predictors (independent variables) of the aided SRT_{SSN} and SRT_{R2TM} (dependent variables). Pairwise Pearson correlations were first calculated between each of the six independent variables and each of the two dependent variables. Statistically significant correlations were regarded as indicative that the independent variable could be a potential predictor of the dependent variable. This type of analysis, however, does not reveal the relative importance of the identified predictors. Indeed, sometimes several potential predictors might reflect a common underlying factor, a phenomenon known as colinearity. In such cases, often only one of the colinear predictors is sufficient to explain the variance in the dependent variable. To better assess the relative importance of potential predictors while minimizing the impact of colinearity, we conducted a stepwise MLR analysis. In a MLR model, it is assumed that the dependent variable may be expressed as a linear combination of independent variables. If the MLR model is constructed in a stepwise fashion (i.e., by gradually adding new potential predictors to the model in each step), the final model omits colinear variables and, most importantly, provides information about the relative importance of the various predictors. Here, we used MLR models to predict the SRT_{SSN} and SRT_{R2TM} independently. As is common practice, the variance explained by the models was adjusted for the number of predictors used in the model (i.e., the explained variance was reduced more as more predictors were included in the model; Theil & Goldberger, 1961).

Unlike FMDTs, which were measured for one carrier frequency only, PTT, HL_{OHC}, HL_{IHC}, and BMCE were available for each of the five test frequencies. Because the correlation and MLR analyses required that the independent variables be single valued, multivalued variables were combined into a single value by weighting the value at each test frequency according to the importance of that frequency for speech recognition (speech-intelligibility index (SII) weightings; ANSI, 1997) and summing the SII-weighted values across all frequencies. The weights were 0.18, 0.25, 0.28, 0.23, and 0.06 for the test frequencies 0.5, 1, 2, 4, and 6 kHz, respectively (from Tables 3 and 4 of ANSI, 1997). The implications of this approach are addressed in the Discussion section.

Results

Raw Data

Figure 1 shows distributions of absolute thresholds for the test ears and for each test frequency. Note that high-frequency losses were more frequent than other type of losses. The 5, 25, 50, 75, and 95th percentiles of age were 38, 54, 61, 74, and 81 years, respectively. The mean age was 62 years and the standard deviation was 14 years. Although the distribution of the ages was slightly skewed toward higher ages, it was not significantly different from Gaussian ($p = .40$, Chi-squared test).
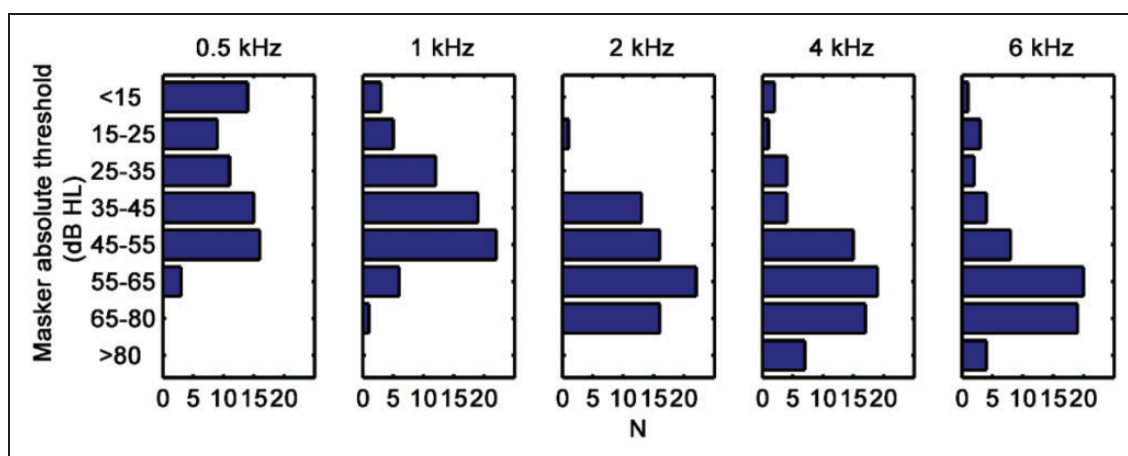

Figure 1. Distribution of hearing losses in categories for the test ears of all participants for each of the test frequencies. Replotted with permission from Johannesen et al. (2014).

For most listeners, SRT_{SSN} values were in the range -5 to $1\,dB$ SNR (Figure 2), thus in line with values reported by earlier studies for SSN maskers (George, Festen, & Houtgast, 2006; Gregan et al., 2013; Peters et al., 1998). SRT_{R2TM} values were in the range -2 to $5\,dB$ SNR and generally higher than SRT_{SSN} values (Figure 2). This trend and range of values are consistent with those reported elsewhere for HI listeners for a R2TM (e.g., Festen & Plomp, 1990 reported SRT_{R2TM} values from -4 to $2\,dB$ SNR). The present SRT_{R2TM} values were about 3, 5, and $5\,dB$ higher than the SRTs for interrupted or modulated noise backgrounds reported by George et al. (2006), Peters et al. (1998), and Gregan et al. (2013), respectively. The fact that SRTs differ for different types of fluctuating maskers is consistent with previous studies (e.g., Festen & Plomp, 1990).

FMDTs for the present participants were in the range of 0.7 to 2 (in units of $log_{10}(Hz)$; Figure 2) and thus similar to the range of values reported by Strelcyk and Dau (2009; 0.7–1.7, when converted to the present units). The participants in the study of Strelcyk and Dau had almost normal audiometric thresholds at frequencies $\leqslant 1$ kHz while the present listeners typically had greater hearing losses over that frequency range (Figure 1), which might explain the higher upper limit in the present FMDTs.

Pairwise Pearson Correlations

Table 1 shows squared Pearson correlation coefficients (R^2 values) for pairs of variables. HL_{OHC} and HL_{IHC} were significantly correlated with PTT but were uncorrelated with each other. This supports the idea put forward elsewhere that listeners with similar audiometric losses can suffer from different degrees of mechanical cochlear gain loss (e.g., Johannesen et al., 2014; Lopez-Poveda & Johannesen, 2012; Lopez-Poveda, Johannesen, & Merchán, 2009; Moore & Glasberg, 1997; Plack, Drga, & Lopez-Poveda, 2004).

BMCE was positively correlated with PTT and HL_{OHC}, indicating that the greater the audiometric loss or the loss of cochlear gain, the more linear (greater BMCE) the cochlear input/output curves. The positive correlation between BMCE and PTT appears inconsistent with earlier studies that reported no correlation between those two variables (Johannesen et al., 2014; Plack et al., 2004). Indeed, Johannesen et al. (2014) reported no correlation between BMCE and audiometric loss based on the same data as were used here. Differences in the data analyses might explain this discrepancy. First, the cited studies based their conclusions on frequency-by-frequency correlation analyses, whereas the present result is based on across-frequency SII-weighted averages. Second, BMCE was set here to

$1\,dB/dB$ whenever the audiometric loss was so high that a corresponding input/output curve could not be measured, something that possibly biased and increased the correlation slightly.

Table 1 also shows that FMDTs were not correlated with PTT, HL_{IHC}, or BMCE and were only slightly positively correlated with HL_{OHC}. Furthermore, FMDTs were not correlated with age. This suggests that FMDTs were assessing auditory processing aspects unrelated (or only slightly related) to cochlear mechanical dysfunction or age.

Potential Predictors of Speech-in-Noise Intelligibility

Table 1 shows that all of the independent variables used in the present study were significantly correlated with SRT_{SSN} and SRT_{R2TM}, except for BMCE which was significantly correlated with SRT_{SSN} only. Therefore, virtually all of them could in principle be related to the measured SRTs. Figure 2 shows scatter plots of SRT_{SSN} (left) and SRT_{R2TM} (right) against each of the six predictor variables (one variable per row), together with linear regression functions (fitted by least squares) and corresponding statistics. The plots suggest a linear relationship between each of the predictors and the aided SRTs. PTT explained slightly more SRT_{R2TM} variance ($R^2 = .17$) than SRT_{SSN} variance ($R^2 = .14$; compare Figure 2(C) and 2(D)). This trend and values are consistent with those reported by Peters et al. (1998), who found R^2 values in the range of 0.07 to 0.11 for SSN and 0.11 to 0.25 for fluctuating maskers, although they did not specifically use R2TMs (see their Table IV).

Figure 2 suggests that PTT, HL_{OHC}, and HL_{IHC} had only a small influence on aided SRTs, as the largest amount of variance explained by any of these three predictors for either of the two SRTs was 17%; this was the variance in SRT_{R2TM} predicted by PTT (Figure 2(D)). For both SRT_{SSN} and SRT_{R2TM}, HL_{OHC} and HL_{IHC} predicted less variance than PTT, which suggests that specific knowledge about the proportion of the PTT that is due to cochlear mechanical gain loss (HL_{OHC}) or other uncertain factors (HL_{IHC}) does not provide much more information than the PTT alone about aided speech-in-noise intelligibility deficits.

BMCE predicted 23% of SRT_{SSN} variance (Figure 2(I)) but was not a significant predictor of SRT_{R2TM} (Figure 2(J)), while FMDTs predicted 28% of the SRT_{R2TM} variance (Figure 2(L)) but only 7% of the SRT_{SSN} variance (Figure 2(K)). This suggests that residual cochlear compression could be more important than temporal processing abilities for understanding speech in steady noise backgrounds while temporal processing abilities could be more important for understanding speech in fluctuating backgrounds.
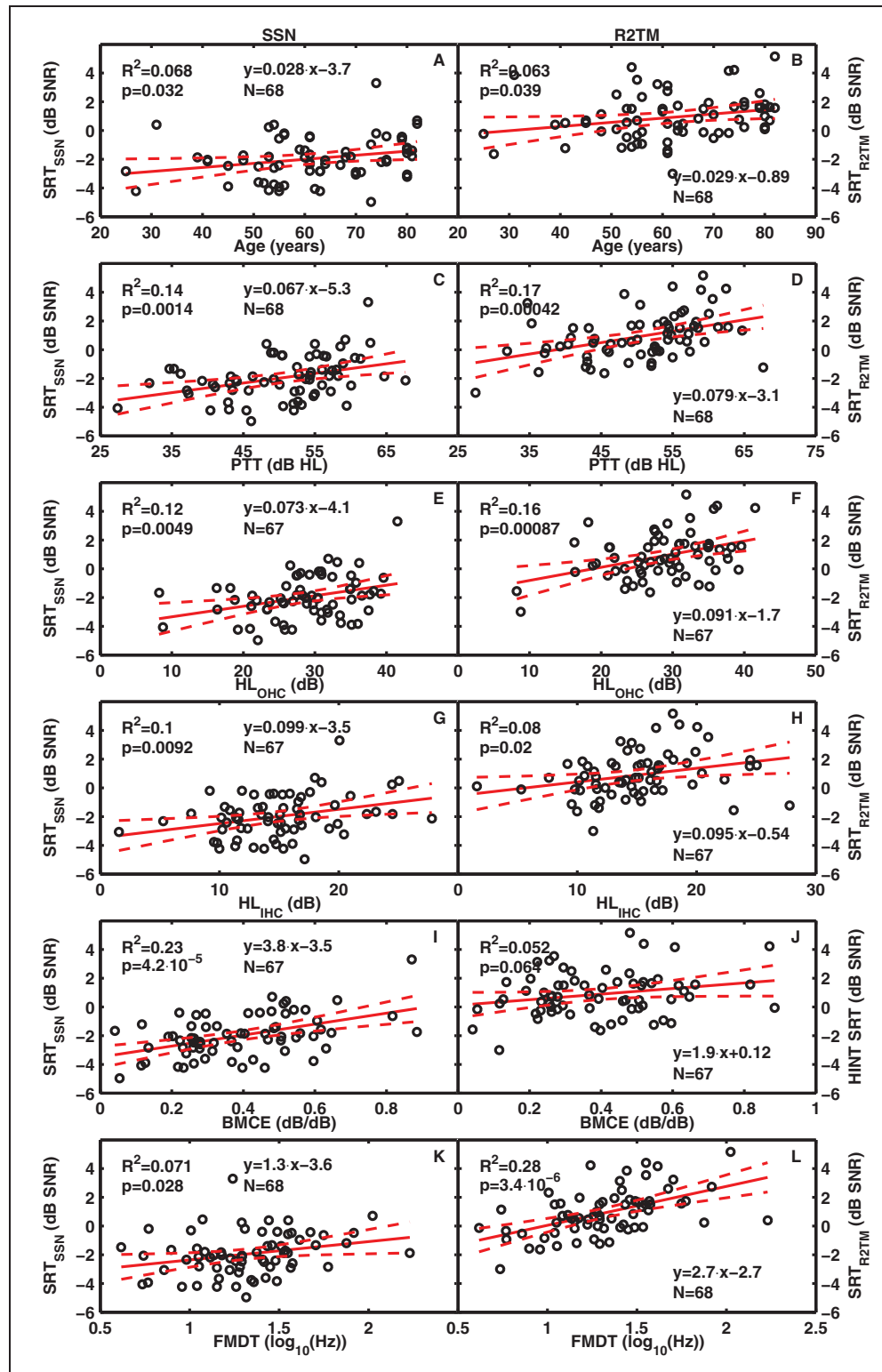
Figure 2. Aided speech reception thresholds for SSN (left column) and R2TM (right column) against across-frequency SII-weighted predictor variables. Each row is for a different predictor (age, PTT, HL$_{OHC}$, HL$_{IHC}$, BMCE, and FMDT) as indicated in the abscissa of each panel. Solid lines depict linear regression lines; dashed lines depict the 5 and 95% confidence interval of the regression line. The upper left inset in each panel informs of the proportion of variance of aided HINT SRTs (R^2) explained by the different predictors and the probability (p) for the value to occur by chance. The lower inset presents the regression equation and the number of cases (N). *Note.* SRT = speech reception threshold; SSN = speech-shaped noise; R2TM = time-reversed two-talker masker; SII = speech-intelligibility index; HLOHC: contribution of cochlear gain loss to the audiometric loss; HLIHC: contribution of inner hair cell dysfunction to the audiometry loss; BMCE = basilar-membrane compression exponent; FMDT = frequency modulation detection threshold; HINT = hearing-in-noise test.

Table 1. Squared Pairwise Pearson Correlations (R^2) and Significance Levels (p) Between All Potential Predictors and Aided HINT SRTs for SSN and R2TM.

		Age	PTT	HL_{OHC}	HL_{IHC}	BMCE	FMDT	SRT_{SSN}	SRT_{R2TM}
Age (years)	R^2	–	.01	.02	.00	.02	.00	**.07**	**.06**
	p	.40	.48	.28	.62	.22	.57	**.032**	**.039**
PTT (dB HL)	R^2	–	–	**.63**	**.30**	**.15**	.03	**.14**	**.17**
	p	–	.09	**$8.9 \cdot \times 10^{-16}$**	**$1.4 \cdot \times 10^{-6}$**	**1.0×10^{-3}**	.13	**1.4×10^{-3}**	**4.2×10^{-4}**
HL_{OHC} (dB)	R^2	–	–	–	.01	**.39**	**.06**	**.12**	**.16**
	p	–	–	.25	.50	**1.5×10^{-8}**	**.04**	**4.9×10^{-3}**	**8.7×10^{-4}**
HL_{IHC} (dB)	R^2	–	–	–	–	.00	.02	**.10**	**.08**
	p	–	–	–	**.03**	.76	.30	**9.2×10^{-3}**	**.020**
BMCE (dB/dB)	R^2	–	–	–	–	–	.01	**.23**	.05
	p	–	–	–	–	**.04**	.38	**4.2×10^{-5}**	.064
FMDT ($\log_{10}$(Hz))	R^2	–	–	–	–	–	–	**.07**	**.28**
	p	–	–	–	–	–	.26	**.028**	**3.4×10^{-6}**
SRT_{SSN} (dB SNR)	R^2	–	–	–	–	–	–	–	**.51**
	p	–	–	–	–	–	–	.17	**1.1×10^{-11}**
SRT_{R2TM} (dB SNR)	R^2	–	–	–	–	–	–	–	–
	p	–	–	–	–	–	–	–	.31

Note. SRT = speech reception threshold; SSN = speech-shaped noise; R2TM = time-reversed two-talker masker; HLOHC: contribution of cochlear gain loss to the audiometric loss; HLIHC: contribution of inner hair cell dysfunction to the audiometry loss; BMCE = basilar-membrane compression exponent; FMDT = frequency modulation detection threshold; HINT = hearing-in-noise test. The *p*-values in the diagonal indicate the probability for a Gaussian distribution of the corresponding variable. Statistical significance ($p < .05$) is indicated with bold font.

Stepwise MLR Models

Stepwise MLR models for SRT_{SSN} and SRT_{R2TM} are shown in Table 2. Each model includes only those predictors whose contribution to the predicted variance was statistically significant. For each model, the predictors' priority order was established according to how much the corresponding predictor contributed to the predicted variance (higher priority was given to larger contributions).

The top part of Table 2 shows that, in the MLR model for aided SRT_{SSN}, the most significant predictor was cochlear compression (BMCE), which explained 22% of the SRT_{SSN} variance, followed by HL_{IHC} and age, which explained 10% and 7% more of the predicted variance, respectively. The model predicted a total of 39% of the SRT_{SSN} variance. Including FMDT, PTT, or HL_{OHC} as additional predictors did not increase the variance predicted by the model. Also, despite the correlation between HL_{OHC} and BMCE ($R^2 = .39$, Table 1), these two variables could not be interchanged in the MLR model. In other words, HL_{OHC} alone explained less variance that BMCE alone. Indeed, BMCE remained as a significant predictor of SRT_{SSN} when HL_{OHC} was included as the first predictor in the stepwise approach but HL_{OHC} became a nonsignificant predictor as soon as BMCE was included in the model.

The MLR model for aided SRT_{R2TM} was strikingly different from the model for SRT_{SSN} (compare the top and bottom parts of Table 2). The most significant predictor of SRT_{R2TM} was FMDT, which explained 27% of the SRT_{R2TM} variance, followed by PTT and age, which explained 10% and 2% more of the variance, respectively. Altogether, the model accounted for 39% of the SRT_{R2TM} variance. Neither HL_{OHC} or BMCE, the two indicators of cochlear mechanical dysfunction, was found to be a significant predictor of SRT_{R2TM}. HL_{IHC} did not increase the variance predicted by this model.

The Role of Audibility

Reduced audibility decreases speech-in-noise intelligibility (e.g., Peters et al., 1998). Here, we tried to minimize the effects of reduced audibility on the SRTs by individually amplifying the target sentences (and the noise) used in the SRT measurements according to the NAL-R linear amplification rule (Byrne & Dillon, 1986). The NAL-R amplification rule, however, compensates at most for half of the audiometric loss. Therefore, it is possible that the low-level portions of the amplified speech spectrum might have been below absolute threshold, something that might have reduced the intelligibility. We attempted to verify that this was not the case by using the SII (ANSI, 1997).

Table 2. Stepwise MLR Models of Aided SRT_{SSN} and SRT_{R2TM}.

Priority	Predictor	Coefficient	*t*-value	*p*	Accum. R^2
SRT_{SSN}					
n/a	Intercept	−7.7	−8.0	4.1×10^{-11}	–
1	BMCE	3.46	4.5	3.2×10^{-5}	.22
2	HL_{IHC}	0.104	3.4	1.0×10^{-3}	.32
3	Age	0.031	2.9	5.8×10^{-3}	.39
SRT_{R2TM}					
n/a	Intercept	−6.6	−5.5	7.0×10^{-7}	–
1	FMDT	2.34	4.7	1.3×10^{-5}	.27
2	PTT	0.060	3.2	1.9×10^{-3}	.37
3	Age	0.022	2.0	4.9×10^{-2}	.39

Note. MLR = multiple linear regression; SRT = speech reception threshold; SSN = speech-shaped noise; R2TM = time-reversed two-talker masker; HLIHC: contribution of inner hair cell dysfunction to the audiometry loss; BMCE = basilar-membrane compression exponent; FMDT = frequency modulation detection threshold. Columns indicate the predictor's priority order and name, the regression coefficient, the *t*-value, and corresponding probability for a significant contribution (*p*), and the accumulated proportion of total variance explained (Accum. R^2), respectively. The priority order is established according to how much the corresponding predictor contributed to the predicted variance (higher priority is given to larger contributions). The accumulated R^2 is the predicted variance adjusted for the number of variables included in the regression model.

The SII is an estimate of the proportion of the speech spectrum that is above the absolute threshold *and* above the background noise (ANSI, 1997). Here, however, the SII was calculated using only the listeners' absolute thresholds, the speech spectrum, and the NAL-R amplification but disregarding the maskers; that is, here, the SII indicated the proportion of the speech spectrum that was above absolute threshold. The rationale was that if the full speech spectrum were audible, then performance deficits in noise would be due to the presence of the noise rather than to reduced audibility and would thus reflect suprathreshold deficits. The resulting SII will be referred to as SII_Q to emphasize that it corresponds to the value in quiet. In all other aspects, the SII calculations conformed to ANSI (1997) for 1/3 octave bands. Our analysis was identical to that of Peters et al. (1998).

As shown in Table 3, 95% of the participants had SII_Q values above 0.52. An SII value of 0.52 corresponds to an intelligibility of almost 90% for NH listeners (e.g., see Figure 3 in Eisenberg, Dirks, Takayanagi, & Martinez, 1998). This suggests that audibility could have affected SRT_{SSN} or SRT_{R2TM} slightly. To further assess the influence of reduced audibility on the present SRTs, new MLR models of SRT_{SSN} and SRT_{R2TM} that included the SII_Q as a potential predictor were explored. The resulting models were identical to those reported in Table 2 and the SII_Q was not a significant predictor in the final MLR model for SRT_{SSN} (*p* = .51) or in the model for SRT_{R2TM} (*p* = .75). Therefore, it is concluded

Table 3. Distribution of SII_Q Values for Individually NAL-R Amplified Speech.

	Percentile (%)						
	0	5	25	50	75	95	100
SII_Q	0.37	0.52	0.58	0.66	0.74	0.86	0.88

Note. SII = speech-intelligibility index.

that reduced audibility was unlikely to have a substantial influence on the present SRTs.

Discussion

The aim of the present study was to assess the relative importance of cochlear mechanical dysfunction, temporal processing deficits, and age for the ability of HI listeners to understand individually amplified speech in SSN (SRT_{SSN}) and a R2TM (SRT_{R2TM}). The main findings were:

1. For the present sample of HI listeners, age, PTT, BMCE, and FMDTs were virtually uncorrelated with each other (Table 1) and yet they were significant predictors of SRTs in noisy backgrounds (Table 2).
2. Residual cochlear compression (BMCE) was the most important single predictor of SRT_{SSN}, while FMDT was the most important single predictor of SRT_{R2TM} (Figure 2, Table 2).
3. Cochlear mechanical gain loss (HL_{OHC}) was correlated with SRT_{SSN} and SRT_{R2TM} (Table 1) but did not improve the MLR models of SRT_{SSN} or SRT_{R2TM} once the previously mentioned predictors were included in the models.
4. Age was a significant predictor of SRT_{SSN} and SRT_{R2TM}, and it was independent of FMDTs and virtually independent of BMCE (Table 1).

The absence of a correlation between age and PTT in the present sample of HI listeners was surprising, given the well-established relationship between those two variables (reviewed by Gordon-Salant, Frisina, Popper, & Fay, 2010). One possible explanation is that our participants were required to be hearing-aid candidates (something necessary for an aspect of the study not reported here) while having mild-to-moderate audiometric losses in the frequency range from 0.5 to 6 kHz, something necessary to infer HL_{OHC} estimates using behavioral masking methods (Johannesen et al., 2014). Thus, it is possible that their hearing losses spanned a narrower range than would be observed across the same age span in a random sample. Our across-frequency SII-weighted averaging of audiometric thresholds

(see Methods section) may have contributed to the lack of correlation between age and PTT.

The absence of a correlation between age and FMDTs is not unprecedented (Strelcyk & Dau, 2009) but was nevertheless unexpected, given that several studies have reported lower (better) FMDTs for younger than for middle-aged or older NH listeners (e.g., Grose & Mamo, 2012; Grose, Mamo, Buss, & Hall, 2015; He, Mills, & Dubno, 2007; Hopkins & Moore, 2011; Moore, Vickers, & Mehta, 2012). The number of synapses between IHCs and auditory nerve fibers is known to decrease gradually with increasing age, even in cochleae with normal IHC and OHC counts and thus presumably normal PTTs (Makary, Shin, Kujawa, Liberman, & Merchant, 2011). Insofar as audiometric loss can be caused by noise exposure and noise exposure decreases the number of afferent synapses (Kujawa & Liberman, 2009), audiometric loss is also thought to be associated with a reduction in the number of synapses. A reduced synapse count (or synaptopathy) is thought to impair auditory temporal processing (Lopez-Poveda, 2014; Lopez-Poveda & Barrios, 2013; Plack, Barker, & Prendergast, 2014; Shaheen, Valero, & Liberman, 2015). The absence of a correlation between age and FMDTs (Table 2) suggests either that our participants did not suffer from synaptopathy (something unlikely given the wide range of their ages and audiometric losses, Figure 1) or that FMDTs reflect temporal processing abilities not directly (or not solely) related to synaptopathy. On the other hand, the effect of age on FMDTs is greater for low- (500 Hz) than for high-frequency (4 kHz) carrier tones (He et al., 2007) and the coding of frequency modulation detection cues in the temporal discharge pattern of auditory nerve responses (phase locking) deteriorates gradually with increasing frequency (Johnson, 1980; Palmer & Russell, 1986). Therefore, another explanation for the present lack of correlation between age and FMDTs might be that the frequency of carrier tone used here was perhaps too high (1500 Hz) to unveil any relationship between age and FMDTs. Another explanation might be that both cochlear mechanical dysfunction and aging can independently impair frequency-modulation detection cues but that in the present case (because all listeners had moderate to moderately severe sensorineural audiometric loss), the detrimental effect of cochlear dysfunction on those cues dominated over the (possibly more modest) effect of aging.

The finding that age, PTT, FMDT, and BMCE were correlated with speech-in-noise intelligibility was expected (Table 1) for the reasons reviewed in the Introduction section. A significant, though incidental aspect of the present study is, however, that for the present group of HI listeners those factors were uncorrelated or barely correlated with each other (Table 1) and yet they all affected intelligibility in different proportions for the two types of maskers (Table 2).

Of the two indicators used here to characterize cochlear mechanical dysfunction, cochlear gain loss (HL_{OHC}) was correlated with speech intelligibility in SSN and a R2TM, while residual compression (BMCE) was correlated with speech intelligibility in SSN. The two indicators (HL_{OHC} and BMCE) were correlated with each other (Table 1). However, HL_{OHC} did not remain as a significant predictor of intelligibility for either of the two maskers when other variables were included in the MLR model, while BMCE was the most significant predictor of intelligibility only for SSN (Table 2). The present estimates of HL_{OHC} and BMCE are indirect and based on numerous assumptions (Johannesen et al., 2014; see also Pérez-González, Johannesen, & Lopez-Poveda, 2014). Assuming nonetheless that these estimates are reasonable, the present findings suggest that cochlear mechanical gain loss and residual compression are not equivalent predictors of the impact of cochlear mechanical dysfunction on intelligibility in SSN. This further suggests that residual compression might be more significant than cochlear gain loss, perhaps because the impact of HL_{OHC} on intelligibility may be compensated with linear amplification but the impact of BMCE may not.

The importance of compression for intelligibility in SSN appears inconsistent with the findings reported in some studies. For example, Noordhoek, Houtgast, and Festen (2001) found little influence of residual compression on the intelligibility of narrow-band speech centered on 1 kHz. Similarly, Summers, Makashay, Theodoroff, and Leek (2013) reported that compression was not clearly associated with understanding intense speech (at a fixed level of 92 dB SPL) in steady noise. This inconsistency may be partly due to methodological differences across studies. First, Summers et al. assessed intelligibility using the percentage of sentences identified correctly for a fixed SNR rather than the SRT (in dB SNR). Second, Summers et al. reported correlations between intelligibility and estimates of compression at single frequencies, while we have reported correlations between SRTs and across-frequency SII-weighted averages of compression. Finally, and perhaps most importantly, although Summers et al. inferred compression from temporal masking curves (TMCs), as we did, they did not take into account important precautions that are necessary to infer accurate compression estimates using the TMC method. This method is based on the assumption that cochlear compression may be inferred from comparisons of the slopes of TMCs unaffected by compression (linear references) with those of TMCs affected by compression. Summers et al. used different linear reference TMCs for different test frequencies and their linear references were TMCs for a masker frequency equal to

0.55 times the probe frequency. It has been shown that this almost certainly underestimates compression, particularly at lower frequencies (e.g., Lopez-Poveda & Alves-Pinto, 2008; Lopez-Poveda, Plack, & Meddis, 2003; Pérez-González et al., 2014). As a result, Summers et al. almost certainly underestimated compression, particularly for their NH listeners, something that might have contributed to hiding differences in compression across listeners with different audiometric thresholds.

For the present sample of HI listeners, BMCE was the most significant predictor of aided speech intelligibility in SSN, while FMDT was the most significant predictor of aided intelligibility in a R2TM (Table 2). Hopkins and Moore (2011) investigated the effects of age and cochlear hearing loss on TFS sensitivity, frequency selectivity, and speech reception in noise for a sample of NH and HI listeners. They reported that once absolute threshold was partialled out, TFS sensitivity was the only significant predictor of speech intelligibility in modulated noise while auditory filter bandwidth was the only significant predictor of intelligibility in steady SSN (see their Table IV). The present results appear broadly consistent with those of Hopkins and Moore considering (a) that FMDT may be thought to index the quality of the internal representation of the TFS at the outputs of the cochlear filters (Moore & Sek, 1996) and (b) that auditory filter bandwidth is a psychoacoustical correlate of cochlear frequency selectivity (e.g., Evans, 2001; Shera, Guinan, & Oxenham, 2002) and thus an indicator of cochlear dysfunction.

Of course, peripheral compression (BMCE) and auditory filter bandwidth are both indirect and different behavioral indicators of cochlear dysfunction but are related. A behavioral study has shown that the auditory filter bandwidth increases as the compression decreases (Moore, Vickers, Plack, & Oxenham, 1999) and physiological studies have shown that OHC dysfunction reduces both cochlear frequency selectivity and compression (Ruggero et al., 1996). In the light of this evidence, the mechanism behind intelligibility in SSN for the present HI listeners might be poorer spectral separation of the speech cues due to increased filter bandwidth (Moore, 2007). It seems puzzling, however, that the same mechanism is not at least slightly involved in the intelligibility of speech in modulated noise (Hopkins & Moore, 2011) or R2TMs (present results).

The present finding that TFS sensitivity (as assessed by FMDT) is less important for intelligibility in steady than in fluctuating maskers appears consistent with the results of Hopkins and Moore (2009), who showed that SRTs improved less with increasing TFS information for steady than for fluctuating backgrounds. Perhaps good TFS sensitivity is unhelpful or unnecessary to understand speech in backgrounds where the TFS of the background is uninformative (as would be the case for SSN) but is required to separate a TFS-rich signal from a TFS-rich masker (as would be the case for a R2TM). If this were the case, intelligibility would be better when the TFS information of the signal and the masker are well represented internally.

The finding that age remained as a significant predictor of intelligibility after the effects of BMCE on SRT_{SNN} and of FMDT on SRT_{R2TM} were partialled out suggests that age per se affects intelligibility in noise. This result is consistent with that of Füllgrabe, Moore, and Stone (2015), who showed that for audiometrically matched young and old listeners, age was a significant contributor for intelligibility in various types of maskers also after the effect of TFS sensitivity was accounted for.

The present conclusions are based on MLR models where all of the predictors were the SII-weighted sum of parameters across frequency. One might wonder whether conclusions would hold if the across-frequency weightings had been different for different parameters. For example, cochlear filter bandwidths tend to increase with increasing HL_{OHC} or BMCE (e.g., Moore et al., 1999) and broader filters can interact with TFS processing (Hopkins & Moore, 2011). Such an interaction could affect intelligibility and would be greater at low frequencies, where phase locking is more significant, than at high frequencies (Hopkins & Moore, 2010). This suggests that predictors of cochlear dysfunction (HL_{OHC} and BMCE) might need to have more weight at low frequencies than suggested by the SII and conclusions might be different in this case. This, however, was unlikely. Alternative MLR models were constructed using the unweighted mean predictors at low (0.5, 1, and 2 kHz) and at high (4 and 6 kHz) frequencies simultaneously. The alternative models (Table 4) suggested (a) that intelligibility depends mostly on the parameter values at low frequencies (something not surprising considering that most subjects had moderate-to-severe audiometric losses at high frequencies) and (b) that the relative importance of the predictors differed from that in the SII-weighted models (Table 2). Nonetheless, the predictors in these alternative models were the same as in the models with SII-weighted parameters across frequency (compare the predictors in Tables 2 and 4).

Speech intelligibility is often analyzed using the SII, which, when calculated taking into account the background noise (ANSI, 1997), provides a measure of the proportion of the speech spectrum that is audible and *above* the noise. The SII is often combined with an additional *proficiency* index that accounts for the effect on intelligibility of factors unrelated to audibility but related to the speaker's enunciation, the type of speech material, and the experience of the listeners with that particular speaker (e.g., Fletcher & Galt, 1950; Studebaker, McDaniel, & Sherbecoe, 1995). Using such a SII/

Table 4. As Table 2 but for MLR Models Where Separate Unweighted Means of the Predictors' Values at 0.5, 1, and 2 kHz (LF), and 4 and 6 kHz (HF) Were Simultaneously Available as Predictors, While SII-Weighted Means Across All Frequencies Were Not.

Priority	Predictor	Coefficient	t-value	p	Accum. R^2
SRT_{SSN}					
n/a	Intercept	−7.73	−7.8	7.9×10^{-11}	—
1	HL_{IHC} (HF)	0.0347	2.9	5.4×10^{-3}	.13
2	BMCE (LF)	3.81	4.1	1.2×10^{-4}	.21
3	HL_{IHC} (LF)	0.113	3.6	6.5×10^{-4}	.28
4	Age	0.037	3.3	1.6×10^{-3}	.38
SRT_{R2TM}					
n/a	Intercept	−6.80	−6.5	1.6×10^{-8}	—
1	PTT (LF)	0.0812	4.5	2.7×10^{-5}	.29
2	Age	0.033	3.1	2.7×10^{-3}	.43
3	FMDT	1.59	3.1	2.8×10^{-3}	.50

Note. MLR = multiple linear regression; SRT = speech reception threshold; SSN = speech-shaped noise; SII = speech-intelligibility index; R2TM = time-reversed two-talker masker; HL = audiometric loss; IHC = inner hair cell; BMCE = basilar-membrane compression exponent; FMDT = frequency modulation detection threshold.

proficiency model, intelligibility for HI listeners has been reasonably well predicted/explained for quiet and noise backgrounds and for aided and unaided conditions (e.g. Humes, 2002; Pavlovic & Studebaker, 1984; Scollie, 2008; Woods, Kalluri, Pentony, & Nooraei, 2013). Here, we did not opt for such an approach for several reasons. First, the SII is not sufficiently validated for fluctuating backgrounds like the R2TM employed here (a revised SII has been proposed that works also for fluctuating noise; Rhebergen, Versfeld, & Dreschler, 2006). Second, in the present study, audibility was largely achieved by providing individualized amplification (see the Results section); hence only the effects of proficiency would remain if the present SRTs were to be explained using the SII/proficiency model. Third, and most importantly, the proficiency index would inform of the magnitude of intelligibility deficits unrelated to audibility while we intended to assess the relative contributions of the physiological causes for those deficits.

Conclusions

1. Estimated cochlear gain loss is unrelated to the ability to understand speech in steady noise.
2. Residual cochlear compression is related to speech understanding in speech-shaped steady noise but not in a time-reversed two-talker masker..
3. Auditory temporal processing ability, as estimated by frequency-modulation detection thresholds, is related to good speech understanding in a time-reversed

two-talker masker but has only minor importance for intelligibility in steady noise.
4. Age per se reduces the intelligibility of speech in any of the two maskers tested here, regardless of absolute thresholds, cochlear mechanical dysfunction, or temporal processing deficits.

References

American National Standards Institute. (1997). *S3.5 methods for calculation of the speech intelligibility index*. New York, NY: Author.

Baer, T., & Moore, B. C. J. (1994). Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech. *Journal of the Acoustical Society of America*, *95*, 2277–2280.

Brown, G. J., Ferry, R. T., & Meddis, R. (2010). A computer model of auditory efferent suppression: Implications for the recognition of speech in noise. *Journal of the Acoustical Society of America*, *127*, 943–954.

Byrne, D., & Dillon, H. (1986). The national acoustic laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. *Ear and Hearing*, *7*, 257–265.

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., ... Ludvidsen, C. (1994). An international comparison of long-term average speech spectra. *Journal of the Acoustical Society of America*, *96*, 2108–2120.

Committee on Hearing, Bioacoustics, and Biomechanics. (1988). Speech understanding and aging. Working group on speech understanding and aging. *Journal of the Acoustical Society of America*, *83*, 859–895.

Deng, L., Geisler, C. D., & Greenberg, S. (1987). Responses of auditory-nerve fibers to multiple-tone complexes. *Journal of the Acoustical Society of America*, *82*, 1989–2000.

Eisenberg, L. S., Dirks, D. D., Takayanagi, S., & Martinez, A. S. (1998). Subjective judgements of clarity and intelligibility for filtered stimuli with equivalent speech intelligibility index predictions. *Journal of Speech, Language, and Hearing Research*, *41*, 327–339.

Evans, E. F. (2001). Latest comparison between physiological and behavioral frequency selectivity. In A. J. M. Houtsma, A. Kohlraush, V. F. Prijs, R. Schoonhoven, & J. Breebaart (Eds.), *Physiological and psychophysical bases of auditory function* (pp. 382–387). Maastricht, The Netherlands: Shaker.

Festen, J. M., & Plomp, R. (1983). Relations between auditory functions in impaired hearing. *Journal of the Acoustical Society of America*, 73, 652–662.

Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America*, 88, 1725–1736.

Fletcher, H., & Galt, R. (1950). Perception of speech and its relation to telephony. *Journal of the Acoustical Society of America*, 22, 89–151.

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*, 6, 347.

George, E. L., Festen, J. M., & Houtgast, T. (2006). Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 120, 2295–2311.

Gordon-Salant, S., Frisina, R. D., Popper, A. N., & Fay, R. R. (2010). *The aging auditory system*. New York, NY: Springer.

Gregan, M. J., Nelson, P. B., & Oxenham, A. J. (2013). Behavioral measures of cochlear compression and temporal resolution as predictors of speech masking release in hearing-impaired listeners. *Journal of the Acoustical Society of America*, 134, 2895–2912.

Grose, J. H., & Mamo, S. K. (2012). Frequency modulation detection as a measure of temporal processing: Age-related monaural and binaural effects. *Hearing Research*, 294, 49–54.

Grose, J. H., Mamo, S. K., Buss, E., & Hall, J. W. (2015). Temporal processing deficits in middle age. *American Journal of. Audiology*, 24, 91–93.

Guinan, J. J. (2010). Cochlear efferent innervation and function. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 18, 447–453.

He, N. J., Mills, J. H., & Dubno, J. R. (2007). Frequency modulation detection: Effects of age, psychophysical method, and modulation waveform. *Journal of the Acoustical Society of America*, 122, 467–477.

Henry, K. S., & Heinz, M. G. (2012). Diminished temporal coding with sensorineural hearing loss emerges in background noise. *Nature Neuroscience*, 15, 1362–1364.

Hopkins, K., & Moore, B. C. J. (2009). The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *Journal of the Acoustical Society of America*, 125, 442–446.

Hopkins, K., & Moore, B. C. J. (2010). The importance of temporal fine structure information in speech at different spectral regions for normal-hearing and hearing-impaired subjects. *Journal of the Acoustical Society of America*, 127, 1595–1608.

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. *Journal of the Acoustical Society of America*, 130, 334–349.

Hornsby, B. W. Y., & Ricketts, T. A. (2007). Directional benefit in the presence of speech and speech like maskers. *Journal of the American Academy of Audiology*, 18, 5–16.

Huarte, A. (2008). The Castilian Spanish hearing in noise test. *International Journal of Audiology*, 47, 369–370.

Humes, L. E. (2002). Factors underlying the speech-recognition performance of elderly hearing-aid wearers. *Journal of the Acoustical Society of America*, 112, 1112–1132.

Johannesen, P. T., Perez-Gonzalez, P., & Lopez-Poveda, E. A. (2014). Across-frequency behavioral estimates of the contribution of inner and outer hair cell dysfunction to individualized audiometric loss. *Frontiers in Neuroscience*, 8, 214.

Johnson, D. H. (1980). The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *Journal of the Acoustical Society of America*, 68, 1115–1122.

Kim, S. H., Frisina, R. D., & Frisina, D. R. (2006). Effects of age on speech understanding in normal hearing listeners: Relationship between the auditory efferent system and speech intelligibility in noise. *Speech Communication*, 48, 855–862.

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

Levitt, H. (1971). Transformed up-down methods in psycho-acoustics. *Journal of the Acoustical Society of America*, 49(Suppl. 2): 467–477.

Lopez-Poveda, E. A. (2014). Why do I hear but not understand? Stochastic undersampling as a model of degraded neural encoding of speech. *Frontiers in Neuroscience*, 8, 348.

Lopez-Poveda, E. A., & Alves-Pinto, A. (2008). A variant temporal-masking-curve method for inferring peripheral auditory compression. *Journal of the Acoustical Society of America*, 123, 1544–1554.

Lopez-Poveda, E. A., & Barrios, P. (2013). Perception of stochastically undersampled sound waveforms: A model of auditory deafferentation. *Frontiers in Neuroscience*, 7, 124.

Lopez-Poveda, E. A., & Johannesen, P. T. (2012). Behavioral estimates of the contribution of inner and outer hair cell dysfunction to individualized audiometric loss. *Journal of the Association for Research in Otolaryngology*, 13, 485–504.

Lopez-Poveda, E. A., Johannesen, P. T., & Merchán, M. A. (2009). Estimation of the degree of inner and outer hair cell dysfunction from distortion product otoacoustic emission input/output functions. *Audiological Medicine*, 7, 22–28.

Lopez-Poveda, E. A., Plack, C. J., & Meddis, R. (2003). Cochlear nonlinearity between 500 and 8000 Hz in listeners with normal hearing. *Journal of the Acoustical Society of America*, 113, 951–960.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Science of the United States of America*, 103, 18866–18869.

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. *Journal of the Association for Research in Otolaryngology*, 12, 711–717.

Moore, B. C. J. (2007). *Cochlear hearing loss*. Chichester, England: John Wiley.

Moore, B. C. J., & Glasberg, B. R. (1997). A model of loudness perception applied to cochlear hearing loss. *Auditory Neuroscience*, 3, 289–311.

Moore, B. C. J., & Sek, A. (1996). Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking. *Journal of the Acoustical Society of America*, 100, 2320–2331.

Moore, B. C. J., Vickers, D. A., & Mehta, A. (2012). The effects of age on temporal fine structure sensitivity in monaural and binaural conditions. *International Journal of Audiology*, 51, 715–721.

Moore, B. C. J., Vickers, D. A., Plack, C. J., & Oxenham, A. J. (1999). Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism. *Journal of the Acoustical Society of America*, 106, 2761–2778.

Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America*, 95, 1085–1099.

Noordhoek, I. M., Houtgast, T., & Festen, J. M. (2001). Relations between intelligibility of narrow-band speech and auditory functions, both in the 1-kHz frequency region. *Journal of the Acoustical Society of America*, 109, 1197–1212.

Palmer, A. R., & Russell, I. J. (1986). Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hearing Research*, 24(1), 1–15.

Pavlovic, C. V., & Studebaker, G. A. (1984). An evaluation of some assumptions underlying the articulation index. *Journal of the Acoustical Society of America*, 75, 1606–1612.

Pérez-González, P., Johannesen, P. T., & Lopez-Poveda, E. A. (2014). Forward-masking recovery and the assumptions of the temporal masking curve method of inferring cochlear compression. *Trends in Hearing*, 18, 1–14.

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. *Journal of the Acoustical Society of America*, 103, 577–587.

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, 114–121.

Plack, C. J., Barker, D., & Prendergast, G. (2014). Perceptual consequences of "hidden" hearing loss. *Trends in Hearing*, 18, 1–11.

Plack, C. J., Drga, V., & Lopez-Poveda, E. A. (2004). Inferred basilar-membrane response functions for listeners with mild to moderate sensorineural hearing loss. *Journal of the Acoustical Society of America*, 115, 1684–1695.

Rhebergen, K. S., Versfeld, N. J., & Dreschler, W. A. (2006). Extended speech intelligibility index for the prediction of the speech reception threshold in fluctuating noise. *Journal of the Acoustical Society of America*, 120, 3988–3997.

Rhebergen, K. S., Versfeld, N. J., & Dreschler, W. A. (2009). The dynamic range of speech, compression, and its effect on the speech reception threshold in stationary and interrupted noise. *Journal of the Acoustical Society of America*, 126, 3236–3245.

Robles, L., & Ruggero, M. A. (2001). Mechanics of the mammalian cochlea. *Physiological Reviews*, 81, 1305–1352.

Ruggero, M. A., Rich, N. C., Robles, L., & Recio, A. (1996). The effects of acoustic trauma, other cochlear injury and death on basilar-membrane responses to sound. In A. Axelson, H. Borchgrevink, P. -A. Hellström, D. Henderson, R. P. Hamernik, & R. J. Salvi (Eds.), *Scientific basis of noise-induced hearing loss* (pp. 23–35). New York, NY: Medical Publishers.

Scollie, S. D. (2008). Children's speech recognition scores: The speech intelligibility index and proficiency factors for age and hearing level. *Ear and Hearing*, 29, 543–556.

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

Shera, C. A., Guinan, J. J., & Oxenham, A. J. (2002). Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. *Proceedings of the National Academy of Science*, 99, 3318–3323.

Strelcyk, O., & Dau, T. (2009). Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *Journal of the Acoustical Society of America*, 125, 3328–3345.

Studebaker, G. A., McDaniel, D. M., & Sherbecoe, R. L. (1995). Evaluating relative speech recognition performance using the proficiency factor and rationalized arcsine differences. *Journal of the American Academy of Audiology*, 6, 173–182.

Summers, V., Makashay, M. J., Theodoroff, S. M., & Leek, M. R. (2013). Suprathreshold auditory processing and speech perception in noise: Hearing-impaired and normal-hearing listeners. *Journal of the American Academy of Audiology*, 24, 274–292.

ter, Keurs, M., Festen, J. M., & Plomp, R. (1992). Effect of spectral envelope smearing on speech reception. I. *Journal of the Acoustical Society of America*, 91, 2872–2880.

Theil, H., & Goldberger, A. S. (1961). On pure and mixed statistical estimation in economics. *International Economic Review*, 2, 65–78.

Woods, W. S., Kalluri, S., Pentony, S., & Nooraei, N. (2013). Predicting the effect of hearing loss and audibility on amplified speech reception in a multi-talker listening scenario. *Journal of the Acoustical Society of America*, 133, 4268–4278.

Young, E. D. (2008). Neural representation of spectral and temporal information in speech. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363, 923–945.