Predictors of supra-threshold speech-in-noise intelligibility by 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,*]

[1] *Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Salamanca, Spain*

[2] *Instituto de Investigación Biomédica de Salamanca, Universidad de Salamanca, Salamanca, Spain*

[3] *Starkey Hearing Research Center, Berkeley, CA, USA*

[4] *Departamento de Cirugía, Facultad de Medicina, Universidad de Salamanca, Salamanca, Spain*

The aim was to assess the relative importance of cochlear mechanical dysfunction, temporal processing deficits, and age for hearing-impaired listeners to understand supra-threshold speech in noise backgrounds. 68 hearing-aid candidates took part in the study. Intelligibility was assessed for speech-shaped noise (SSN) and reversed two-talker masker (R2TM) backgrounds. Behavioural estimates of cochlear gain loss and residual compression from a previous study 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 thresholds and cochlear gain loss were also included in the analyses. Stepwise multiple linear regression models of intelligibility were designed to assess the relative importance of the various factors for speech intelligibility. Results showed that (1) cochlear gain loss was unrelated to intelligibility; (2) residual cochlear compression was related to intelligibility in SSN but not in R2TM backgrounds; (3) temporal processing was strongly related to intelligibility in R2TM backgrounds and much less so in SSN backgrounds; (4) age *per se* hindered intelligibility. We conclude that all factors affect speech intelligibility but their relative importance varies across masker backgrounds.

INTRODUCTION

Hearing-impaired (HI) people vary widely in their ability to understand speech in noise backgrounds, even when their audiometric loss is compensated with frequency-specific sound amplification (e.g., Moore, 2007). The present study aimed at shedding some light on the relative importance of cochlear mechanical dysfunction, temporal processing deficits, and age as predictors of this variability.

---

*Corresponding author: ealopezpoveda@usal.es

Proceedings of ISAAR 2015: Individual Hearing Loss – Characterization, Modelling, Compensation Strategies. 5th symposium on Auditory and Audiological Research. August 2015, Nyborg, Denmark. Edited by S. Santurette, T. Dau, J. C. Dalsgaard, L. Tranebjærg, and T. Andersen. ISBN: 978-87-990013-5-4. The Danavox Jubilee Foundation, 2015.

Several explanations have been proposed to account for the ability of HI listeners to understand audible speech in noise backgrounds (reviewed by Lopez-Poveda, 2014). One of them is that HI listeners could suffer from outer hair cell (OHC) loss or dysfunction and this 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. This can smear the cochlear representation of the acoustic spectrum, making it harder for HI listeners to separately perceive the spectral cues of speech from those of interfering sounds. Second, in the healthy cochlea, suppression might facilitate the encoding of speech in noise by enhancing the most salient frequency features of the target speech against those of the background noise. OHC dysfunction reduces suppression and this might hinder speech-in-noise intelligibility. Third, cochlear mechanical compression might facilitate the understanding of speech in fluctuating noise by amplifying the speech in the silent noise intervals, a phenomenon known as 'listening in the dips'. OHC loss or dysfunction reduces compression (i.e., linearizes cochlear responses) and thus could hinder dip listening (Gregan *et al.*, 2013).

The view that OHC dysfunction accounts for the ability of HI listeners to understand audible speech in noise 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 role of compression on the intelligibility of supra-threshold speech in noise backgrounds. Second, at high intensities, cochlear tuning is comparable for healthy and impaired cochleae 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). Third, elderly listeners with normal audiometric thresholds and presumably healthy OHCs often have difficulty understanding speech in noise (CHABA, 1988), which suggests that age *per se* or mechanisms other than OHC dysfunction can limit the intelligibility of audible speech.

Another explanation for the ability of HI listeners to understand speech in noise is that HI listeners may suffer from temporal processing deficits. This view would be consistent with the reported correlation between the reduced speech-in-noise intelligibility of HI listeners and their reduced ability to use the information conveyed in the rapid temporal changes of speech sounds, known as 'temporal fine structure' (Lorenzi *et al.*, 2006; Strelcyk and Dau, 2009). It would also be consistent with evidence that temporally jittering the frequency components in speech, as might occur after auditory neuropathy (Pichora-Fuller *et al.*, 2007), or stochastic undersampling of a noisy speech waveform, as might occur after synaptopathy (Lopez-Poveda and Barrios, 2013), both decrease speech-in-noise intelligibility with negligible reductions in audibility.

The present study aimed at assessing the relative contribution of cochlear mechanical dysfunction, temporal processing deficits, and age to the performance of HI listeners understanding audible speech in noisy environments.

MATERIAL AND METHODS

Subjects

The same 68 subjects (43 males) with symmetrical sensorineural hearing losses of the study of Johannesen *et al.* (2014) participated in the present study. Speech-in-noise intelligibility was assessed in bilaterally listening conditions (see below). Indicators of cochlear mechanical status and temporal processing ability, however, were measured in one ear only. For most cases, the test ear was the ear with better audiometric thresholds in the 2-6 kHz frequency range (30 left ears, 38 right ears).

Indicators of cochlear mechanical dysfunction

OHC dysfunction linearizes cochlear mechanical responses. Johannesen *et al.* (2014) compared behaviourally inferred cochlear input/output curves for each HI listener at each one 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 (HL_{OHC} in dB). It was defined as the contribution of cochlear gain loss to absolute thresholds and was calculated as the difference sound level required for a pure tone at the test frequency to evoke identical mechanical responses in the cochlea of a HI and a NH listener at absolute threshold (see also Lopez-Poveda and Johannesen, 2012). A second variable was inner hair cell (IHC) loss or HL_{IHC}. It was defined as the difference (in dB) between the pure tone threshold (PTT 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 conveniently termed HL_{IHC} (Moore and Glasberg, 1997). A third variable 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.* (2014) for further details.

PTT, HL_{OHC}, HL_{IHC}, and BMCE were taken from Johannesen *et al.* (2014) and were all considered potential predictors of speech-in-noise intelligibility. Note that the four variables had values at each of the five test frequencies.

Johannesen *et al.* (2014) reported that they could not measure input/output curves for listeners and test frequencies where the audiometric loss was too high. Here, these cases were assumed to be indicative of total cochlear gain loss and HL_{OHC} was set equal to the cochlear gain observed for NH listeners (for details, see p. 11 of Johannesen *et al.*, 2014) and BMCE was set equal to 1 dB/dB.

Frequency modulation detection thresholds

Temporal processing ability was assessed using frequency modulation detection thresholds (FMDTs). The experiment was identical to that of Strelcyk and Dau (2009). In short, an FMDT was defined as the minimum detectable excursion in frequency for a tone carrier and was estimated using a three alternative forced choice procedure. In each trial, the three intervals contained a 1500-Hz pure tone with a level of 30 dB above the detection threshold for the tone. The tones in all intervals were amplitude-modulated (AM) with a modulation depth of 6 dB and a time-varying modulation rate. In the target interval (selected at random), the tone's frequency was varied with a rate of 2 Hz and with a maximum frequency excursion. 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% point on the psychometric function (Levitt, 1971). Three FMDTs estimates were obtained and their mean was taken as the threshold.

Speech reception thresholds

Speech-in-noise intelligibility was quantified using the speech reception threshold (SRT), defined as the speech-to-noise ratio (SNR) required to understand 50% of the sentences and was measured using the hearing-in-noise test (HINT) (Nilsson *et al.*, 1994). The background noise was either a steady speech-shaped noise (SSN) or a masker that consisted of two simultaneous talkers (one male and one female) played in reverse (reversed two-talker masker, R2TM). The corresponding SRTs are referred to as $\mathrm{SRT}_{\mathrm{SSN}}$ and $\mathrm{SRT}_{\mathrm{R2TM}}$, respectively.

To measure an SRT, the speech was fixed in level to a nominal value of 65 dB SPL and the masker level was varied adaptively using a one-up, one-down rule. 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 were the speech and the masker would be co-located one meter in front of the listener at eye level (Table 3 in ANSI, 1997). The resulting stimulus was linearly amplified individually for each participant according to the NAL-R rule (Byrne and Dillon, 1986) to account for the potential effect of the audiometric loss on intelligibility. The amplified stimulus was played diotically to the listeners. All other details of the procedure were as in the original HINT test (Nilsson *et al.*, 1994).

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-analogue converted using an RME Fireface 400 sound card with a 24-bit resolution, and were played through Sennheiser HD580 headphones. Data were collected in a sound attenuation booth.

Statistical analyses

Pairwise Pearson correlations were first sought between each of the independent variables (PTT, $\mathrm{HL}_{\mathrm{OHC}}$, $\mathrm{HL}_{\mathrm{IHC}}$, BMCE, FMDT, and age) and each of the dependent variables (aided $\mathrm{SRT}_{\mathrm{SSN}}$ and $\mathrm{SRT}_{\mathrm{R2TM}}$). Prior to the correlation analysis, variables with values at different frequencies were combined into a single value by weighting the value at each test frequency according to the frequency's importance for speech perception (ANSI, 1997) and summing the weighted values across frequencies.

Multiple linear regression (MLR) models were constructed for $\mathrm{SRT}_{\mathrm{SSN}}$ and $\mathrm{SRT}_{\mathrm{R2TM}}$ independently to assess the relative importance of the potential predictors for intelligibility. Sometimes several potential predictors might reflect a common underlying factor, a phenomenon known as co-linearity. To minimize the impact of co-linearity, MLR models were constructed in a stepwise fashion (i.e., by gradually adding new potential predictors to the model in each step). The final model omits co-linear variables.

RESULTS

Raw data

The mean absolute thresholds across listeners for the test ears were 37, 44, 51, 61, and 75 dB HL at 0.5, 1, 2, 4, and 6 kHz, respectively (see also Fig. 1 in Johannesen *et al.*, 2014). The standard deviations were in the range 11 to 20 dB HL across frequencies. High-frequency losses were more frequent than other types of losses.

The listeners' ages ranged from 25 to 82 years, with a mean and a standard deviation of 62 and 14 years, respectively. The 5%, 25%, 50%, 75%, and 95% percentiles of age were 38, 54, 61, 74, and 81 years, respectively.

For most listeners, $\mathrm{SRT_{SSN}}$ were in the range −5 to 1 dB SNR, thus in line with values reported by earlier studies for SSN maskers (Peters *et al.*, 1998; George *et al.*, 2006; Gregan *et al.*, 2013). $\mathrm{SRT_{R2TM}}$ values were in the range −2 to 5 dB SNR and generally higher than $\mathrm{SRT_{SSN}}$ values. This trend and range of values are consistent with the −4 to 2 dB range reported by Festen and Plomp (1990) for $\mathrm{SRT_{R2TM}}$. The present $\mathrm{SRT_{R2TM}}$ 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. This shows that SRTs can be different for different types of fluctuating maskers.

FMDTs for the present participants were in the range 0.7 to 2, in units of $\log_{10}$(Hz), and thus similar to the range of values reported by Strelcyk and Dau (2009) (0.7 to 1.7, when converted to the present units).

Pairwise Pearson's correlations

Table 1 shows squared Pearson's correlation coefficients (R^2 values) for pairs of variables. $\mathrm{HL_{OHC}}$ and $\mathrm{HL_{IHC}}$ were significantly correlated with PTT but were uncorrelated with each other. This supports the idea that the people with similar audiometric losses can suffer from different degrees of mechanical cochlear gain loss (e.g., Plack *et al.*, 2004; Lopez-Poveda and Johannesen, 2012).

BMCE was positively correlated with PTT and $\mathrm{HL_{OHC}}$, a result indicative that the greater the audiometric loss or the loss of cochlear gain, the more linear (less compressive) the cochlear input/output curves. The positive correlation between BMCE and PTT appears inconsistent with the study of Johannesen *et al.* (2014) that, based on the same data, reported no correlation between those two variables. 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 weighted averages. Second, BMCE was set here to 1 dB/dB whenever the audiometric loss was so high that a corresponding

			PTT	$\mathrm{HL}_{\mathrm{OHC}}$	$\mathrm{HL}_{\mathrm{IHC}}$	BMCE	FMDT	$\mathrm{SRT}_{\mathrm{SSN}}$	$\mathrm{SRT}_{\mathrm{R2TM}}$
Age	years	R^2	0.01	0.02	0.00	**0.08**	0.00	**0.07**	**0.06**
		p	0.48	0.28	0.63	**0.024**	0.57	**0.032**	**0.039**
PTT	dB HL	R^2	-	**0.63**	**0.30**	**0.13**	0.03	**0.14**	**0.17**
		p	0.088	**10^{-15}**	**$1.4 \cdot 10^{-6}$**	**0.002**	0.13	**0.00144**	**0.00042**
$\mathrm{HL}_{\mathrm{OHC}}$	dB	R^2	-	-	0.01	**0.34**	**0.06**	**0.12**	**0.16**
		p	-	0.25	0.51	**$2.4 \cdot 10^{-7}$**	**0.04**	**0.0046**	**0.00077**
$\mathrm{HL}_{\mathrm{IHC}}$	dB	R^2	-	-	**-**	0.00	0.02	**0.10**	**0.08**
		p	-	-	**0.031**	0.90	0.31	**0.0102**	**0.023**
BMCE	dB/dB	R^2	-	-	-	**-**	0.00	**0.29**	**0.07**
		p	-	-	-	**0.0096**	0.81	**$3.1 \cdot 10^{-6}$**	**0.035**
FMDT	$\log_{10}$(Hz)	R^2	-	-	-	-	-	**0.07**	**0.28**
		p	-	-	-	-	0.26	**0.028**	**$3.4 \cdot 10^{-6}$**
$\mathrm{SRT}_{\mathrm{SSN}}$	dB SNR	R^2	-	-	-	-	-	-	**0.51**
		p	-	-	-	-	-	0.17	**$1.05 \cdot 10^{-11}$**

Table 1. Squared pairwise Pearson correlations (R^2) and significance levels (p) between all potential predictors and aided $\mathrm{SRT}_{\mathrm{SSN}}$ and $\mathrm{SRT}_{\mathrm{R2TM}}$. The p-values in the diagonal indicate the probability for a non-Gaussian distribution of the corresponding variable.

input/output curve could not be measured, something that may have biased and increased the correlation slightly.

Table 1 also shows that FMDTs were not correlated with PTT, $\mathrm{HL}_{\mathrm{IHC}}$, or BMCE and were only slightly positively correlated with $\mathrm{HL}_{\mathrm{OHC}}$. Furthermore, FMDTs were not correlated with age. This suggests that FMDTs were indeed assessing auditory processing aspects unrelated (or only slightly related) to cochlear mechanical dysfunction or age, as was intended.

In addition, Table 1 shows that $\mathrm{SRT}_{\mathrm{SSN}}$ and $\mathrm{SRT}_{\mathrm{R2TM}}$ were significantly and positively correlated with each other. The two SRTs were measured using identical conditions and yet their R^2 (0.51) shows that only 51% of the variance in $\mathrm{SRT}_{\mathrm{SSN}}$ could be explained by the $\mathrm{SRT}_{\mathrm{R2TM}}$. This suggests that different mechanisms and/or deficits mediate speech intelligibility for different masker backgrounds. If the mechanisms or deficits mediating speech intelligibility were identical for the two masker backgrounds, one would expect a higher correlation (higher R^2) between $\mathrm{SRT}_{\mathrm{SSN}}$ and $\mathrm{SRT}_{\mathrm{R2TM}}$ than the one found.

Potential predictors of speech-in-noise intelligibility

Table 1 shows that $\mathrm{SRT}_{\mathrm{SSN}}$ and $\mathrm{SRT}_{\mathrm{R2TM}}$ were significantly correlated with all of the independent variables and hence in principle they could all be contributing to the measured SRTs. The correlations (Table 1) show that PTT explained slightly more $\mathrm{SRT}_{\mathrm{R2TM}}$ variance ($R^2 = 0.17$) than $\mathrm{SRT}_{\mathrm{SSN}}$ ($R^2 = 0.14$) variance. This trend and values are consistent with those reported by Peters *et al.* (1998).

Table 1 suggests that PTT, HL_{OHC}, and HL_{IHC} had only a mild influence on aided SRTs, as the largest amount of variance explained by any of these three predictors on any of the two SRTs was 17% (Table 1). For both SRT_{SSN} and SRT_{R2TM}, HL_{OHC} and HL_{IHC} predicted less variance than the 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 more information than the PTT alone about supra-threshold speech-in-noise intelligibility deficits.

In addition, Table 1 reveals that BMCE predicted 29% of SRT_{SSN} variance but only 7% of SRT_{R2TM} variance, while FMDTs predicted 28% of the SRT_{R2TM} variance but only 7% of the SRT_{SSN} variance. 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-masker backgrounds.

Stepwise multiple linear regression models

Stepwise MLR models for SRT_{SSN} and SRT_{R2TM} are shown in Table 2.

Priority	Predictor	Coefficient	*t*-value	*p*	Adj. accum. R^2
			SRT_{SSN}		
n/a	Intercept	-7.5	-8.0	$3.5\cdot10^{-11}$	-
1	BMCE	4.25	5.0	$5.6\cdot10^{-6}$	0.28
2	HL_{IHC}	0.097	3.3	0.0017	0.37
3	Age	0.023	2.1	0.038	0.41
4	FMDT	0.90	2.0	0.045	0.44
			SRT_{R2TM}		
n/a	Intercept	-7.1	-5.5	$7.0\cdot10^{-7}$	-
1	FMDT	2.24	4.8	$1.25\cdot10^{-5}$	0.27
2	PTT	0.061	3.5	0.008	0.38
3	Age	0.032	2.9	0.0047	0.45

Table 2. Stepwise MLR models of aided SRT_{SSN} and SRT_{R2TM}. Columns indicate the predictor's priority order and name, the regression coefficient, the *t*-value, the corresponding probability for a significant contribution *(p*), and the adjusted accumulated proportion of total variance explained (Adj. accum. R^2), respectively. The priority order is established according to how much the corresponding predictor contributed to the predicted variance.

The top part of Table 2 shows that the most significant predictor of SRT_{SSN} was BMCE, which explained 28% of the SRT_{SSN} variance. Additional predictors were HL_{IHC}, age, and FMDT which contributed an additional 9, 4, and 3% to the predicted variance, respectively. The model predicted a total of 44% of the SRT_{SSN} variance. PTT and HL_{OHC} were not significant additional predictors.

The MLR model for aided $\mathrm{SRT}_{\mathrm{R2TM}}$ was strikingly different than the model for $\mathrm{SRT}_{\mathrm{SSN}}$ (compare the top and bottom parts of Table 2). The most significant predictor of $\mathrm{SRT}_{\mathrm{R2TM}}$ was FMDT, which explained 27% of the $\mathrm{SRT}_{\mathrm{R2TM}}$ variance. Additional predictors were PTT and age, which contributed an additional 11 and 7% to the model predicted variance, respectively. Altogether, the model accounted for 45% of the $\mathrm{SRT}_{\mathrm{R2TM}}$ variance. Neither $\mathrm{HL}_{\mathrm{OHC}}$, or BMCE, or the $\mathrm{HL}_{\mathrm{IHC}}$ were significant predictors of $\mathrm{SRT}_{\mathrm{R2TM}}$.

The role of audibility

Reduced audibility decreases speech-in-noise intelligibility (e.g., Peters *et al.*, 1998). Although NAL-R amplification was provided, audibility might still have been reduced and could have affected the SRTs. To discard this possibility, we calculated the speech intelligibility index (SII) (ANSI, 1997). The SII indicates the proportion of the speech spectrum that is above the absolute threshold and above the background noise (ANSI, 1997). Here, however, we calculated an SII taking into account only the absolute thresholds, the speech spectrum, and the NAL-R amplification while the background noise was disregarded (i.e., here, the SII informed of the proportion of the speech spectrum that was above absolute threshold). In all other aspects, our SII calculations conformed to ANSI (1997). The rationale behind this approach is that if the full speech spectrum were audible, then performance deficits in a masker background would be due to the presence of the masker (Peters *et al.* 1998) rather than to reduced audibility, and would thus reflect supra-threshold deficits.

For 95% of the participants, the SII values were above 0.52, a value that corresponds to an intelligibility of almost 90% for NH listeners (see, e.g., Fig. 3 in Eisenberg *et al.*, 1998). The high SII values indicate that it is unlikely that audibility affected $\mathrm{SRT}_{\mathrm{SSN}}$ or $\mathrm{SRT}_{\mathrm{R2TM}}$. To further rule out the influence of reduced audibility, new MLR models of $\mathrm{SRT}_{\mathrm{SSN}}$ and $\mathrm{SRT}_{\mathrm{R2TM}}$ were explored including the SII as a potential predictor. The resulting models in this case were identical to those of Table 2 and the SII did not become a significant predictor in any of the final MLR models. Therefore, it is unlikely that reduced audibility have influenced 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 understanding audible speech in noise backgrounds. 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 aided SRT in noise backgrounds (Table 2).
2) Residual cochlear compression (BMCE) was the most important single predictor of aided $\mathrm{SRT}_{\mathrm{SSN}}$, while FMDT was the most important single predictor of aided $\mathrm{SRT}_{\mathrm{R2TM}}$ (Table 2).
3) Cochlear mechanical gain loss ($\mathrm{HL}_{\mathrm{OHC}}$) was correlated with aided $\mathrm{SRT}_{\mathrm{SSN}}$ and $\mathrm{SRT}_{\mathrm{R2TM}}$ (Table 1) but did not increase the variance explained by 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).

For the present sample, age, PTT, FMDT, and (virtually) BMCE were uncorrelated with each other. This result was incidental. Given the well-established relationship between age and PTT (reviewed by Gordon-Salant *et al.*, 2010), the absence of a correlation between those two variables was surprising. One possible explanation is that our participants were required to be hearing aid candidates (something necessary for a different 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 behavioural 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 weighted-averaging of audiometric thresholds (see Methods) may have contributed to wash out any correlation between age and PTT.

The absence of a correlation between age or PTTs with FMDTs was unexpected. The number of synapses between IHC and auditory nerve fibres is known to decrease gradually with increasing age, even in cochleae with normal IHC and OHC counts and thus presumably normal PTT (Makary *et al.*, 2011). Insofar as hearing impairment can be caused by noise exposure and noise exposure decreases the number of afferent synapses (Kujawa and Liberman, 2009), hearing impairment is also thought to be associated with a reduced number of synapses. A reduced synapse count (or synaptopathy) is thought to impair auditory temporal processing (Lopez-Poveda and Barrios, 2013). The absence of a correlation between age and FMDTs or between FMDTs and PTT (Table 2) suggests that either our participants did not suffer from synaptopathy (unlikely given the wide age range) or that FMDTs reflect temporal processing abilities not directly (or not solely) related to synaptopathy.

The finding that age, PTT, FMDT, and BMCE are correlated with supra-threshold speech-in-noise intelligibility (Table 1) was expected for the reasons reviewed in the Introduction. A significant though incidental aspect of the present study is, however, that for the present sample those factors were uncorrelated or poorly correlated with each other (Table 1) and yet they affected intelligibility in different proportions for different types of masker backgrounds (Table 2).

The two indicators of cochlear mechanical dysfunction (HL_{OHC} and BMCE) were correlated with speech intelligibility in the two noise backgrounds, and they were correlated with each other (Table 1). However, HL_{OHC} did not remain as a significant predictor of intelligibility in neither of the two masker backgrounds when other variables were included in the MLR models, while BMCE became the most significant predictor of intelligibility only in SSN backgrounds (Table 2). The estimates of cochlear gain loss (HL_{OHC}) and residual compression (BMCE) are indirect and based on numerous assumptions (Johannesen *et al.*, 2014). Assuming that these estimates are reasonable, the present finding suggests that cochlear mechanical gain loss and residual compression are not equivalent predictors of the impact of cochlear mechanical dysfunction on the intelligibility of speech in SSN. The finding further suggests that residual compression is more significant than cochlear gain loss, perhaps because the impact of HL_{OHC} on intelligibility may be compensated for with linear amplification but the impact of BMCE may not.

The importance of compression for understanding supra-threshold speech in SSN appears inconsistent with the findings of Summers *et al.* (2013) who reported that compression was not clearly associated with understanding loud speech (at a fixed level of 92 dB SPL) in a steady noise background. This inconsistency may be partly due to methodological differences across studies. First, Summers *et al.* (2013) assessed intelligibility using the percentage of sentences identified correctly for a fixed SNR rather than the SRT (in dB SNR). Second, Summers *et al.* (2013) reported correlations between intelligibility and estimates of compression at single frequencies while we are reporting correlations between SRTs and across-frequency weighted average of compression. Lastly, Summers *et al.* (2013) did not take into account important precautions regarding inference of compression estimates using the temporal masking curve (TMC) method. This method is based on the assumption that cochlear compression may be inferred from comparisons of the slope of TMCs unaffected by compression (linear references) with that of TMCs affected by compression. Summers *et al.* (2013) 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. This almost certainly underestimates compression (e.g., Lopez-Poveda *et al.*, 2003; Lopez-Poveda and Alves-Pinto, 2008), particularly at lower frequencies and for NH listeners, something that might have contributed to 'hiding' differences in compression across listeners with different audiometric thresholds in the data of Summers *et al.* (2013).

Residual compression (BMCE) was the best single predictor of supra-threshold speech intelligibility in a SSN background while FMDT became the most significant predictor in a R2TM background (Table 2). The reason is uncertain, though it seems reasonable that temporal processing ability be more important for intelligibility in fluctuating than in steady masker backgrounds.

CONCLUSIONS

1) Cochlear gain loss is unrelated to understanding audible speech in noise.
2) Residual cochlear compression is related to speech understanding in speech-shaped steady noise but not in reversed two-talker masker backgrounds.
3) Auditory temporal processing ability is strongly related to speech understanding in fluctuating masker backgrounds but has relatively minor importance in a steady noise background.
4) Age hinders the intelligibility of supra-threshold speech in any of the two masker backgrounds tested here, regardless of absolute thresholds, cochlear mechanical dysfunction, or temporal processing deficits.

ACKNOWLEDGEMENTS

We thank William S. Woods, Olaf Strelcyk, and Brent Edwards for useful suggestions. Work supported by Starkey Ltd. (USA), by Junta de Castilla y León and MINECO (ref. BFU2012-39544-C02).

REFERENCES

ANSI (**1997**). *S3.5 Methods for Calculation of the Speech Intelligibility Index* (American National Standards Institute, New York).

Byrne D. and Dillon H. (**1986**). “The National Acoustic Laboratories’ (NAL) new procedure for selecting the gain and frequency response of a hearing aid,” Ear Hearing, **7**, 257-265.

CHABA (**1988**). “Speech understanding and aging. Working Group on Speech Understanding and Aging. Committee on Hearing, Bioacoustics, and Biomechanics, Commission on Behavioral and Social Sciences and Education, National Research Council,” J. Acoust. Soc. Am., **83**, 859-895.

Eisenberg L.S., Dirks D.D., Takayanagi S., and Martinez A.S. (**1998**). “Subjective judgements of clarity and intelligibility for filtered stimuli with equivalent speech intelligibility index predictions,” J. Speech Lang. Hear. Res., **41**, 327-339.

Festen J.M. and Plomp, R. (**1990**). “Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing,” J. Acoust. Soc. Am., **88**, 1725-1736.

George E.L., Festen J.M., and Houtgast T. (**2006**). “Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners,” J. Acoust. Soc. Am., **120**, 2295-2311.

Gordon-Salant, S., Frisina R.D., Popper A.N., and Fay R.R. (**2010**). *The Aging Auditory System* (Springer, New York).

Gregan M.J., Nelson P.B., and Oxenham A.J. (**2013**). “Behavioral measures of cochlear compression and temporal resolution as predictors of speech masking release in hearing-impaired listeners,” J. Acoust. Soc. Am., **134**, 2895-2912.

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

Kujawa S.G. and Liberman M.C. (**2009**). “Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss,” J. Neurosci., **29**, 14077-14085.

Levitt H. (**1971**). “Transformed up-down methods in psychoacoustics,” J. Acoust. Soc. Am., **49**, 467-477.

Lopez-Poveda, E.A., Plack, C.J., and Meddis, R. (**2003**). “Cochlear nonlinearity between 500 and 8000 Hz in listeners with normal hearing,” J. Acoust. Soc. Am., **113**, 951-960.

Lopez-Poveda, E.A. and Alves-Pinto A. (**2008**). “A variant temporal-masking-curve method for inferring peripheral auditory compression,” J. Acoust. Soc. Am., **123**, 1544-1554.

Lopez-Poveda E.A. and Johannesen, P.T. (**2012**). "Behavioral estimates of the contribution of inner and outer hair cell dysfunction to individualized audiometric loss," J. Assoc. Res. Otolaryngol., **13**, 485-504.

Lopez-Poveda, E.A. and Barrios, P. (**2013**). "Perception of stochastically undersampled sound waveforms: a model of auditory deafferentation," Front. Neurosci., **7**, 124.

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

Lorenzi C., Gilbert G., Carn H., Garnier S., and Moore B.C.J. (**2006**). "Speech perception problems of the hearing impaired reflect inability to use temporal fine structure," Proc. Natl. Acad. Sci. USA, **103**, 18866-18869.

Makary C.A., Shin, J., Kujawa, S.G., Liberman, M.C., and Merchant, S.N. (**2011**). "Age-related primary cochlear neuronal degeneration in human temporal bones," J. Assoc. Res. Otolaryngol., **12**, 711-717.

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

Moore, B.C.J. (**2007**). *Cochlear Hearing Loss* (John Wiley & Sons, Ltd., Chichester).

Nilsson M., Soli S.D., and Sullivan J.A. (**1994**). "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise," J. Acoust. Soc. Am., **95**, 1085-1099.

Peters, R.W., Moore, B.C.J., and Baer, T. (**1998**). "Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people," J. Acoust. Soc. Am., **103**, 577-587.

Pichora-Fuller, M.K., Schneider, B.A., MacDonald, E., Pass, H.E., and Brown, S. (**2007**). "Temporal jitter disrupts speech intelligibility: a simulation of auditory aging," Hear. Res., **223**, 114-121.

Plack, C.J., Drga, V., and Lopez-Poveda, E.A. (**2004**). "Inferred basilar-membrane response functions for listeners with mild to moderate sensorineural hearing loss," J. Acoust. Soc. Am., **115**, 1684-1695.

Strelcyk, O. and Dau, T. (**2009**). "Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing," J. Acoust. Soc. Am., **125**, 3328-3345.

Summers, V., Makashay, M.J., Theodoroff, S.M., and Leek, M.R. (**2013**). "Suprathreshold auditory processing and speech perception in noise: Hearing-impaired and normal-hearing listeners," J. Am. Acad. Audiol., **24**, 274-292.