DTU Library

Effects of Hearing Loss and Fast-Acting Compression on Amplitude Modulation Perception and Speech Intelligibility

Wiinberg, Alan; Jepsen, Morten Løve; Epp, Bastian; Dau, Torsten

Published in:
Ear and Hearing

Link to article, DOI:
10.1097/AUD.0000000000000589

Publication date:
2018

Document Version
Peer reviewed version

Effects of Hearing Loss and Fast-Acting Compression on Amplitude Modulation Perception and Speech Intelligibility

Alan Wiinberg,[1] Morten Løve Jepsen,[2] Bastian Epp,[1] and Torsten Dau[1]

Objective: The purpose was to investigate the effects of hearing-loss and fast-acting compression on speech intelligibility and two measures of temporal modulation sensitivity.

Design: Twelve adults with normal hearing (NH) and 16 adults with mild to moderately severe sensorineural hearing loss were tested. Amplitude modulation detection and modulation-depth discrimination (MDD) thresholds with sinusoidal carriers of 1 or 5 kHz and modulators in the range from 8 to 256 Hz were used as measures of temporal modulation sensitivity. Speech intelligibility was assessed by obtaining speech reception thresholds in stationary and fluctuating background noise. All thresholds were obtained with and without compression (using a fixed compression ratio of 2:1).

Results: For modulation detection, the thresholds were similar or lower for the group with hearing loss than for the group with NH. In contrast, the MDD thresholds were higher for the group with hearing loss than for the group with NH. Fast-acting compression increased the modulation detection thresholds, while no effect of compression on the MDD thresholds was observed. The speech reception thresholds obtained in stationary noise were slightly increased in the compression condition relative to the linear processing condition, whereas no difference in the speech reception thresholds obtained in fluctuating noise was observed. For the group with NH, individual differences in the MDD thresholds could account for 72% of the variability in the speech reception thresholds obtained in stationary noise, whereas the correlation was insignificant for the hearing-loss group.

Conclusions: Fast-acting compression can restore modulation detection thresholds for listeners with hearing loss to the values observed for listeners with NH. Despite this normalization of the modulation detection thresholds, compression does not seem to provide a benefit for speech intelligibility. Furthermore, fast-acting compression may not be able to restore MDD thresholds to the values observed for listeners with NH, suggesting that the two measures of amplitude modulation sensitivity represent different aspects of temporal processing. For listeners with NH, the ability to discriminate modulation depth was highly correlated with speech intelligibility in stationary noise.

(Ear & Hearing 2018;XX;00–00)

INTRODUCTION

Loudness recruitment is a typical consequence of sensorineural hearing loss and appears to be a consequence of damaged outer hair cells and, thereby, a loss of the level-dependent cochlear amplification the outer hair cells provide (Fowler 1936; Steinberg & Gardner 1937; Moore 2004). It has been proposed that recruitment can be compensated for by mimicking the cochlear amplification with multi-band dynamic range compression (DRC; Villchur 1973; Allen 1996), which provides level-dependent gain such that low-level sounds are amplified more than higher-level sounds. DRC, thus, provides audibility of the low-level portions of the sound while avoiding loudness discomfort at high levels. It has been shown that multiband DRC with compression attack and release time-constants of 10 ms can restore normal loudness perception for listeners with sensorineural hearing loss (LSNHL; Strelcyk et al. 2012). It has also been demonstrated that measures of auditory temporal resolution that are assumed to be affected by recruitment, such as gap detection and forward masking, can at least partly be restored for LSNHL by fast-acting DRC with attack and release times less than 60 ms (e.g., Moore et al. 2001; Brennan et al. 2015; Kowalewski et al. 2015).

Another measure of temporal processing is amplitude modulation detection. Recruitment has been shown to enhance the internal representation of the signal envelope in the auditory system, mainly as a result of reduced amplitude compression of the signal at the level of processing on the basilar membrane (BM) in the impaired auditory system, relative to the processing in the healthy auditory system (Moore et al. 1996). Because of reduced BM compression, listeners with unilateral sensorineural hearing loss (SNHL) perceive a larger modulation depth in the impaired compared with the healthy ear (Moore et al. 1996). Consistent with this result, LSNHL have lower modulation detection thresholds (MDTs), that is, higher sensitivity, than listeners with normal hearing (NH) when using tonal carriers presented at the same low sensation level (e.g., Moore & Glasberg 2001). Since fast-acting DRC reduces the depth of modulation in the temporal envelope, this type of processing might be able to restore the normal internal representation of the envelope in LSNHL. Brennan et al. (2013) obtained MDTs with noise carriers in conditions with and without DRC for LSNHL and found higher MDTs in the conditions with DRC, consistent with the idea that DRC may compensate for recruitment. However, the study did not consider to what extent the DRC processing restored performance to the level found for listeners with NH. Furthermore, since noise carriers contain intrinsic fluctuations that can mask the imposed signal modulation, in contrast to deterministic tonal carriers without any intrinsic fluctuations, MDTs obtained with noise carriers may be dominated by modulation masking effects, whereas MDTs obtained with tonal carriers can only be limited by internal noise (Dau et al. 1997).

Much evidence shows that the intelligibility of speech in background sounds is affected by the auditory processing of the amplitude modulations contained in the speech and the background. Speech intelligibility models, such as the speech transmission index (Houtgast & Steeneken 1985), the spectro-temporal modulation index (Elhilali et al. 2003), and the envelope power spectrum model (Jørgensen & Dau 2011), have modeled the effects of different types of processing channels (such as a room or a nonlinear processor) on the envelope

[1]Hearing Systems Group, Department of Electrical Engineering, Technical University of Denmark, DK-2800 Lyngby, Denmark; and [2]Department of Electronics and Audiology, Widex A/S, DK-3540 Lynge, Denmark.

0196/0202/2018/XXXX-00/0 • Ear & Hearing • Copyright © 2018 Wolters Kluwer Health, Inc. All rights reserved • Printed in the U.S.A.

<zdoi; 10.1097/AUD.0000000000000589>

representations of the signals and their relation to speech intelligibility. However, the link between sensitivity to amplitude modulations and speech intelligibility has been controversial. It has been argued that increased masker modulation fluctuations in the internal representation could reduce speech intelligibility (Kale & Heinz 2010; Schlittenlacher & Moore 2016). MDTs obtained with noise carriers have been shown to be only poorly correlated with the intelligibility of speech in modulated maskers for LSNHL (e.g., Takahashi & Bacon 1992; Feng et al. 2010).

Regarding effects of fast-acting DRC on speech intelligibility, some studies have reported a degradation of speech intelligibility (Drullman & Smoorenburg 1997; Reinhart et al. 2016; Noordhoek & Drullman 1997). For example, Noordhoek and Drullman (1997) found that fast-acting DRC increased the reception threshold (SRT) for speech in stationary noise. The authors argued that fast-acting DRC distorts the temporal envelope of the speech signal and, thereby, affects important speech cues. However, other studies found either no effect of fast-acting DRC on speech intelligibility (Boothroyd et al. 1988; Souza & Turner 1996; Drullman & Smoorenburg 1997; van Buuren et al. 1999) or even improved speech intelligibility (Yund & Buckles 1995; Souza & Turner 1998; Gatehouse et al. 2006). For example, Souza and Turner (1998) found increased speech recognition scores for LSNHL using multiband fast-acting DRC at low speech levels when compared to scores obtained with linear amplification, whereas no improvement was observed at higher speech levels. The authors ascribed this improvement to differences in audibility: the DRC provided more amplification and, thereby, raised the low-level speech stimuli above the audibility threshold by a larger amount than the linear amplification.

Since most modulations inherent in speech are well above threshold, that is, above MDT (Edwards 2004; Schlittenlacher & Moore 2016), suprathreshold measures of modulation processing, such as modulation-depth discrimination (MDD) thresholds, might provide stronger links to speech intelligibility performance than MDTs. Schlittenlacher and Moore (2016) observed higher MDD thresholds (with tonal carriers) for LSNHL than for listeners with NH. They argued that recruitment increases the perceived amount of amplitude modulation, called fluctuation strength, such that this sensation "saturates" at lower modulation depths for LSNHL than for listeners with NH. Hence, differences between modulation depths may become less noticeable for LSNHL than for listeners with NH when the fluctuation strength is at ceiling level for the LSNHL and below ceiling level for the listeners with NH. After this argument, fast-acting DRC might be able to compensate for the increased fluctuation strength and thereby restore normal MDD thresholds. Alternatively, it is possible that MDT and MDD thresholds represent two different aspects of temporal processing. Ewert and Dau (2004) demonstrated that MDD thresholds obtained at different reference modulation depths follow Weber's law, that is, thresholds are roughly proportional to the reference modulation depth. Two different "internal noise" processes can be assumed to limit resolution: an "absolute" internal noise that determines the MDT and an internal noise *after* a logarithmic compression that determines the MDD threshold. If recruitment causes an enhancement of the internal representation of the envelope, relative to NH, this may affect the MDT but not necessarily the MDD threshold. Likewise, DRC may compensate for the effect of recruitment and, thus, adjust MDTs back to normal, but this may not affect the MDD threshold. It is possible that aspects of hearing loss other than recruitment, such as suprathreshold deficits because of, for example, inner hair cell loss, can result in increased MDD thresholds, as well as degraded speech intelligibility in some conditions.

The present study addressed the relations between temporal envelope sensitivity (in terms of MDT and MDD thresholds), speech intelligibility, hearing-loss and fast-acting DRC. Two measures of temporal envelope sensitivity were obtained: (i) temporal modulation transfer functions with tonal carriers for which MDTs were obtained as a function of modulation frequency (e.g., Kohlrausch et al. 2000) and (ii) MDD thresholds, for which the just-noticeable increase in modulation depth from a (suprathreshold) standard modulation depth was measured as a function of modulation frequency (e.g., Lee & Bacon 1997). Since previous work suggested that both slow envelope fluctuations (<16 Hz) and fast fluctuations (16–300 Hz) contribute to speech intelligibility in competing-talker conditions (Stone et al. 2008, 2012; Christiansen et al. 2013), MDD thresholds and MDTs were obtained for modulation frequencies in the range from 8 to 256 Hz. Tonal carriers were used since they do not contain intrinsic envelope fluctuations, which may mask, and thereby limit, the detectability of the imposed modulation (e.g., Dau et al. 1997, 1999). For tonal carriers, the imposed modulation introduces spectral sidebands, which may be resolved if they are sufficiently far from the carrier frequency (e.g., Kohlrausch et al. 2000). However, as long as the modulation frequency is within the range where spectral resolution does not play a major role, results obtained with tonal carriers may provide a better measure of the temporal resolution of the auditory system than results obtained with noise carriers. In the first experiment, MDTs and MDD thresholds were obtained with and without DRC for listeners with normal and impaired hearing. In the second experiment, SRTs for speech in both steady and fluctuating noise were obtained with and without DRC for the same groups of listeners.

MATERIALS AND METHODS

Experiment 1: Amplitude Modulation Detection and Modulation-Depth Discrimination

Listeners • Two groups of listeners participated, a group with NH and a group with SNHL. The group with NH consisted of 12 adults (6 males and 6 females). The mean age was 29 years and the range was 21 to 60 years. All had absolute thresholds better than 20 dB HL for the octave frequencies between 0.125 and 8 kHz. The SNHL group consisted of 16 adults (10 males and 6 females) with symmetrical mild to moderately severe sensorineural hearing losses. One listener dropped out of the study without completing the temporal modulation transfer functions measure with the 5 kHz carrier. The mean age was 68 years and the range was 50 to 80 years. The absolute thresholds for the test ear, measured using conventional audiometry (ISO 8253-1:2010), are shown in Figure 1. All listeners (except the first author, who served as one of the listeners with NH) signed an informed consent document and were reimbursed for their efforts. Approval for the study was granted by the Science Ethics Committee of the Capital Region in Denmark ("De Videnskabsetiske Komitéer for Region Hovedstaden").

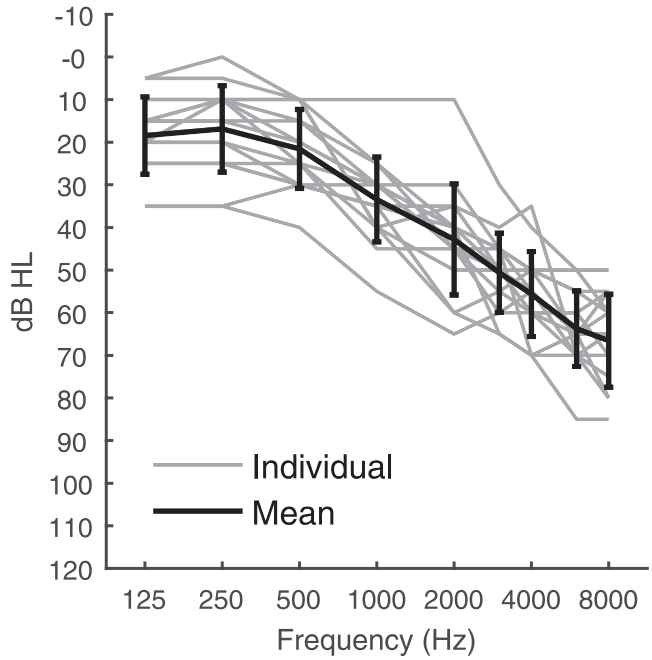

Fig. 1. Individual and mean absolute thresholds for the tested ear of the listeners with sensorineural hearing loss (LSNHL), measured using conventional manual audiometry, and expressed in dB HL. Error bars represent ±1 SD.

Stimuli • The stimulus was a sinusoidally amplitude-modulated tone:

$$s(t) = (1 + m \cos 2\pi f_m t) \sin 2\pi f_c t, \quad (1)$$

where t is time, m is the modulation depth taking values between 0 and 1, and f_c and f_m are the frequency of the carrier and modulator, respectively. The frequency of the carrier was either 1 or 5 kHz for the measurement of the MDTs and 1 kHz for the measurement of the MDD thresholds. The modulation frequencies were 8, 16, 32, 64, 128, and 256 Hz in the unprocessed condition and 8, 16, and 32 Hz in the DRC condition. The stimulus duration was 600 ms, including 50-ms raised-cosine onset and offset ramps. Intervals were separated by 500-ms pauses. The overall level of the stimuli in each interval was kept the same, regardless of the modulation depth and DRC processing. For the listeners with NH, the level of the stimuli was 65 dB SPL. For the LSNHL, the 65 dB SPL stimuli were linearly amplified according to the NAL-R(P) frequency-dependent prescription rule by amounts depending on the individual audiometric thresholds (Byrne et al. 1990). The frequency-dependent amplification was provided using a bank of seven octave-wide bandpass linear-phase, finite-impulse-response filters with center frequencies between 0.125 and 8 kHz.

All signals were generated digitally on a PC equipped with a RME UCX Fireface sound card at a sampling rate of 44.1 kHz. The stimuli were presented in a sound-attenuating booth via DT 770 PRO Beyerdynamic headphones to the better ear of the listeners, defined as the ear with the lowest absolute threshold averaged across 500, 1000, and 2000 Hz. The transfer function of each earpiece of the headphones was digitally equalized (1001 point FIR filter) to produce a flat frequency response, measured with an ear simulator (B&K 4153) and a flat plate adaptor as specified in IEC 60318-1 (2009).

Experimental Procedure • MDTs were measured using an adaptive three-interval, three-alternative forced-choice paradigm. A three-down one-up procedure was used to track the 79.4% point on the psychometric function (Levitt 1971). The carrier was unmodulated in two of the intervals, the reference intervals, and amplitude modulated in the other interval, the target interval. In each trial, the intervals were presented in random order, and the listeners had to select the interval containing the modulated carrier. For each carrier frequency, the different modulation frequencies were tested in random order. The thresholds and step sizes were represented by the modulation depth, in decibels: 20 $\log_{10}(m)$. A run started with a modulation depth of −5 dB. The step size was initially 5 dB and was reduced to 2 dB after the first incorrect response. Each run was terminated after seven reversals, and the threshold estimate for that run was computed as the mean value of the modulation depth at the last six reversals. Reported thresholds represent the mean over three runs.

For the measurement of MDD thresholds, the procedure was the same as for the measurement of the MDTs, except that the carrier was modulated with a constant standard modulation depth (m_s) in the reference intervals and modulated with a higher modulation depth in the target interval. The standard modulation depth was −15 dB, which was 10–15 dB above the amplitude MDTs typically found for a 1-kHz carrier for the range of the modulation frequencies studied here (Kohlrausch et al. 2000; Moore & Glasberg 2001). A run started with the target at a modulation depth of −3 dB. The thresholds and step sizes were represented by the ratio of the modulation depth of the target to the modulation depth of the reference: 20 $\log_{10}(m/m_s)$. The step size was initially 2 dB and was reduced to 1 dB after the first incorrect response. The order of the measurements of MDTs and MDD thresholds with and without DRC was randomized for each listener.

Single-Channel Dynamic Range Compression System • The single-channel DRC system was implemented in MATLAB version 2013b (The MathWorks, Inc., Natick, MA). The envelope of the signal was extracted using the Hilbert transform and smoothed using a peak detector (Eq. [8.1] in Kates 2008). The attack and release time-constants, measured according to IEC 60118-2 (1983), were 10 and 60 ms, respectively. The smoothed envelope was converted to decibels. A broken-stick gain function (with linear gain below the compression threshold (20 dB SPL) and 2:1 compression ratio above threshold) was applied to the processed envelope. The resulting sample-wise gain was applied to the input stimulus.

Experiment 2: Perception of Speech in Noise

Stimuli • The Danish version of the hearing in noise test (HINT) was used to measure SRTs (Nielsen & Dau 2011). Two noise maskers provided by the International Collegium for Rehabilitative Audiology (ICRA) were used, ICRA-1 steady speech-shaped noise and ICRA-6 speech-shaped noise with the modulation characteristics of two-talker babble (Dreschler et al. 2001).

Listeners • The same listeners as in experiment 1 participated in experiment 2. However, the listener with SNHL who dropped out during experiment 1 did not participate in experiment 2. The SRT data from another listener with SNHL were excluded from further analysis because of the listener's familiarity with the HINT speech corpus from his work as a clinical audiologist.

Experimental Procedure • The level of the speech was kept constant and the level of the noise masker was varied using an adaptive one-up one-down procedure. The listeners were instructed to verbally repeat the sentences as accurately as possible and to guess if they were uncertain. All of the words in a sentence needed to be correct for the sentence to be scored as correct (Nielsen & Dau 2011). The SRT was measured for each of the noise maskers with and without DRC applied to the mixture of speech and noise. The sentence lists were randomly selected for each of the four conditions (2 Noise Conditions × 2 DRC Conditions). To familiarize the listeners with the task and to reduce possible learning effects, training was conducted before the data collection using two practice lists. Other aspects of the experimental procedure were the same as in experiment 1.

Multi-Band Dynamic Range Compression System • The multiband DRC system used in the second experiment contained seven frequency bands. The input signal was Hanning windowed in time frames of 256 samples, approximately 6 ms in duration, with 75% overlap between frames. Each of the windowed segments was padded with 128 zeros at the beginning and with 128 zeros at the end and transformed to the spectral domain using a 512-point fast Fourier transform (FFT). The power of the resulting frequency bins was combined into seven octave-wide frequency bands with center frequencies between 0.125 and 8 kHz. Other aspects (including smoothing, time-constants, conversion to gain, and the static gain function) of the DRC were the same as for the single-channel system. The resulting band-wise gains were in the frequency domain mapped to the 512 FFT bins using a piecewise cubic interpolation (Kates 2008). The frequency smoothed gains were applied to the bins of the short-time Fourier transformed input stimulus, and an inverse FFT was applied to produce time segments of the compressed stimuli. These time segments were subsequently windowed with a tapered cosine window to avoid aliasing artifacts and combined using an overlap-add method to provide the processed temporal waveform.

Statistical Analysis • Three analyses of variance (ANOVA) assessed the effects of hearing status and modulation frequency on the MDTs for the 1- and 5-kHz carriers and the MDD thresholds. Hearing status (NH versus SNHL) was a between-listener factor, and modulation frequency was a within-listener factor. Two additional ANOVAs assessed the effect of DRC for modulation frequencies between 8 and 32 Hz. The additional factors, DRC (on versus off) and carrier frequency (1 versus 5 kHz) were within-listener factors. The SRT data were analyzed separately for each noise type using a two-factor mixed model ANOVA that included hearing status as a between-listener factors and DRC condition as a within-listener factor. Tukey's Honestly Significant Differences post hoc tests were conducted to test for main effects and interactions using a 5% significance value. When full modulation was not detectable, the threshold was set to a modulation depth corresponding to full modulation.

RESULTS

Amplitude Modulation Detection Thresholds

Figure 2 shows the MDTs as a function of modulation frequency for the carrier frequency of 1 kHz for the listeners with NH (upper panels) and the LSNHL (bottom panels). Figure 2A and C shows the results obtained without DRC, and Figure 2B and D shows the results obtained with DRC. The across-listener variation was larger for the SNHL group (SD = 5.4 dB) than for the NH group (SD = 3.5 dB). The MDT decreased from 8 to 16 Hz and remained constant from 16 to 64 Hz. The increased MDT at 8 Hz may reflect temporal interference between the carrier onset and the onset of modulation, as shown by Sheft and Yost (1990). The MDT decreased with increasing modulation frequency above 64 Hz, reflecting the detection of the spectral sidebands (Kohlrausch et al. 2000). Eleven of the 16 LSNHL showed monotonically decreasing MDT above 64 Hz, similar to the listeners with NH, whereas the MDTs remained roughly constant above 64 Hz for the remaining five LSNHL. This is most likely a consequence of reduced frequency selectivity for these five LSNHL in the frequency region around 1 kHz. To avoid confounding the measure of temporal modulation sensitivity with the ability to detect spectral sidebands, MDTs for the 1-kHz carrier above 64 Hz were excluded from the statistical analysis. There was no effect of hearing status [$F(1,26) = 0.46$, $p = 0.503$] and no interaction between hearing status and modulation frequency [$F(3,78) = 0.81$, $p = 0.503$]. However, there was an effect of modulation frequency [$F(3,78) = 21.40$, $p < 0.001$].

Figure 3 shows the MDTs as a function of modulation frequency for the carrier frequency of 5 kHz for the listeners with NH (upper panels) and the LSNHL (bottom panels). There was an effect of modulation frequency [$F(5,125) = 44.14$, $p < 0.001$]. Post hoc comparisons revealed that the threshold decreased from 8 to 16 Hz, remained constant from 16 to 128 Hz, and increased between 128 and 256 Hz. There was no effect of hearing status [($F(1,25) = 0.93$, $p = 0.343$]. The interaction between hearing status and modulation frequency was significant [$F(5,125) = 10.43$, $p < 0.001$]. Below 64 Hz, MDT thresholds were significantly lower for the LSNHL than for the listeners with NH ($M = 5.5$ dB, SE = 3 dB). There were no other statistically significant differences. Two of the LSNHL showed much higher MDTs above 32 Hz than the other LSNHL. The increase of the MDT at high modulation frequencies for the 5-kHz carrier reflects a limitation in the temporal processing of

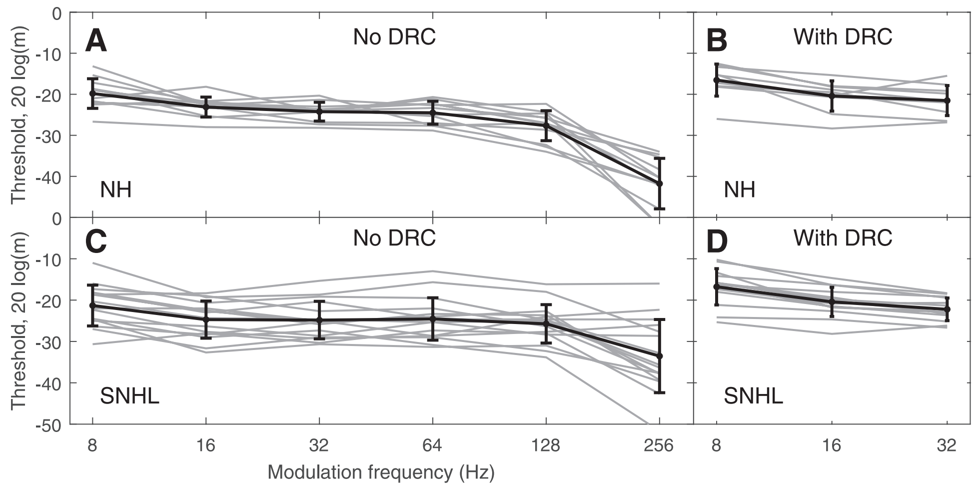

Fig. 2. Individual (gray) and mean (black) modulation detection thresholds (MDTs) for a carrier frequency of 1 kHz. The MDT (20 log *m*) is plotted as a function of the modulation frequency. A and B, the results for the listeners with normal hearing. C and D, the results for the listeners with sensorineural hearing loss (LSNHL). A and C, the results obtained without dynamic range compression (DRC). B and D, the results obtained with DRC. The error bars represents ±1 SD.

fast amplitude modulations in the auditory system (e.g., Ewert & Dau 2000; Kohlrausch et al. 2000).

Panels B and D in Figures 2 and 3 show the MDTs for the listeners with NH and the LSNHL, respectively, when DRC was applied. There were significant effects of DRC [$F(1,19) = 73.28$, $p < 0.001$], hearing status [$F(1,45) = 4.23$, $p = 0.045$], modulation frequency [$F(2,40) = 60.42$, $p < 0.001$], and an interaction between DRC and modulation frequency [$F(2,145) = 6.89$, $p = 0.001$]. No other effects were significant. Figure 4 shows the MDTs with DRC subtracted from the MDTs without DRC as a function of modulation frequency (solid curve). The amount of ΔMDT roughly corresponds to the physical reduction of the modulation depth (dashed curve) which was calculated based on the Hilbert envelopes before and after compression, for an input modulation depth of −15 dB (Stone & Moore 1992).

Modulation Depth Discrimination Thresholds

Figure 5 shows the MDD thresholds as a function of modulation frequency for the carrier frequency of 1 kHz. The results for the listeners with NH are shown in Figure 5A and B, and the results for the LSNHL are shown in Figure 5C and D. The MDD thresholds are expressed as the ratio of the modulation depth of the target to the modulation depth of the reference [$20 \log(m/m_s)$]. Figure 5A and C show the results obtained without DRC, and Figure 5B and D show the results obtained with DRC. The across-listener variation in the MDD thresholds was larger for the SNHL group (SD = 2.1 dB) than for the NH group (SD = 1.3 dB). An ANOVA on the data obtained without DRC revealed an effect of hearing status [$F(1,26) = 5.26$, $p = 0.029$]. The MDD thresholds were significantly lower for the listeners with NH than for the LSNHL (M = 1.6 dB, SE = 0.6 dB). There was an effect of modulation frequency [$F(3,78) = 7.36$, $p < 0.001$] and no interaction between hearing status and modulation frequency [$F(3,78) = 0.45$, $p = 0.717$]. Post hoc comparisons revealed that the MDD threshold decreased from 8 to 16 Hz and remained constant from 16 to 64 Hz. As for the modulation detection experiment, the increased MDD threshold at the lowest modulation frequencies was probably caused by the gating of the carrier (Lee & Bacon 1997).

Figure 5B and D shows the MDD thresholds for the listeners with NH and LSNHL obtained with DRC. There was no

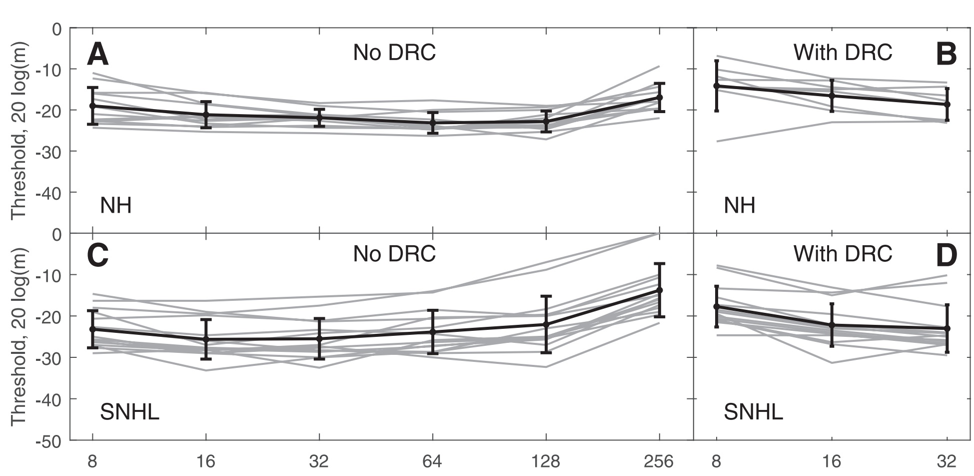

Fig. 3. As Figure 2, but for a carrier frequency of 5 kHz. DRC, dynamic range compression; LSNHL, listeners with sensorineural hearing loss; MDT, modulation detection threshold; NH, normal hearing.

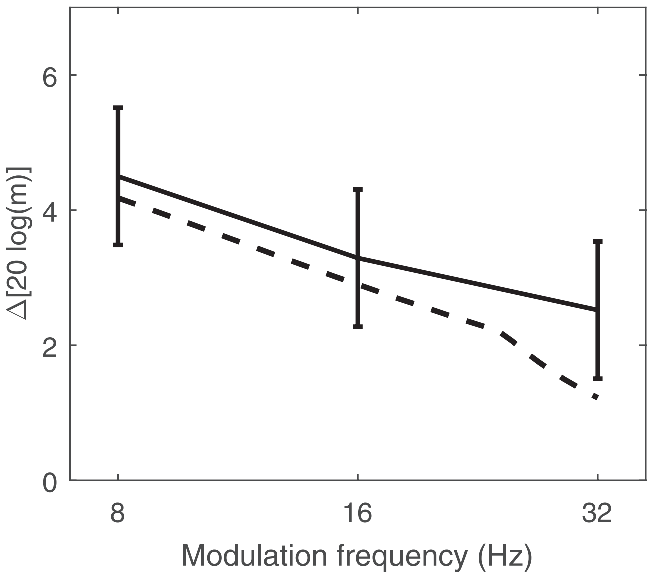

Fig. 4. The effect of dynamic range compression (DRC) on the modulation detection thresholds (MDTs; solid curve) averaged across listeners and carrier frequencies. The error bars represent ±2 standard errors of the mean. For comparison, the dashed curve shows the physical reduction of the modulation depth.

effect of DRC [$F(1,50) = 0.79$, $p = 0.38$], an effect of hearing status [$F(1,54) = 16.62$, $p = 0.44$], an effect of modulation frequency [$F(2,110) = 41.54$, $p < 0.001$], and no significant second or third order interactions (DRC and modulation frequency [$F(2,53) = 0.24$, $p = 0.78$], modulation frequency and hearing status [$F(2,108) = 0.8$, $p = 0.45$], DRC and hearing status [$F(1,52) = 0.10$, $p = 0.75$], DRC, hearing status, and modulation frequency [$F(1,104) = 0.82$, $p = 0.44$].

Speech Reception Thresholds

Figure 6 shows the average SRTs for the listeners with NH (open symbols) and the LSNHL (closed symbols) obtained in the two noise conditions with and without multiband DRC processing. The listeners with NH showed, on average, lower SRTs than the LSNHL in all conditions. The SRT difference between the stationary noise (ICRA-1) and the speech-modulated noise (ICRA-6), the masking release, can be regarded as a measure of the benefit from the dips in the modulated noise. The masking release was smaller for the LSNHL (0.9 dB) than for the listeners with NH (3.3 dB); five of the LSNHL showed a negative masking release. The variation of SRTs across listeners was larger for the SNHL group than for the NH group, especially for the speech-modulated noise.

For the stationary noise masker, there was an effect of DRC [$F(1,25) = 5.76$, $p = 0.02$] and hearing status [$F(1,25) = 48.11$, $p < 0.001$] but no interaction [$F(1,25) = 1.31$, $p = 0.26$]. A post-hoc analysis showed that the SRT was 0.5 dB higher (SE = 0.2 dB) for the DRC-on condition than for the DRC-off condition. For the speech-modulated noise masker, there was an effect of hearing status [$F(1,25) = 36.64$, $p < 0.001$] but no effect of DRC [$F(1,25) = 0.04$, $p = 0.83$] and no interaction [$F(1,25) = 0.43$, $p = 0.51$].

Regression Analyses

Correlations between the various measures were calculated separately for the NH and the SNHL groups and are shown in Table 1. The MDT and MDD values were based on the mean thresholds across modulation frequencies. For the MDTs at 1 kHz, only the thresholds for modulation frequencies below 128 Hz were considered to compute the mean values. For the listeners with NH, the only significant correlation was between the SRT obtained with the stationary noise masker and MDD thresholds. For the LSNHL, no significant correlation between the SRTs and the MDD thresholds was observed. Figure 7 shows a scatterplot of the SRT for the stationary noise masker against MDD thresholds. For the listeners with NH, the SRT increased with increasing MDD thresholds. For both groups, there was no significant correlation between MDD thresholds and MDTs for the 1-kHz carrier.

DISCUSSION

This study investigated modulation detection, modulation depth discrimination, and speech intelligibility in noise for listeners with NH and LSNHL in conditions with and without

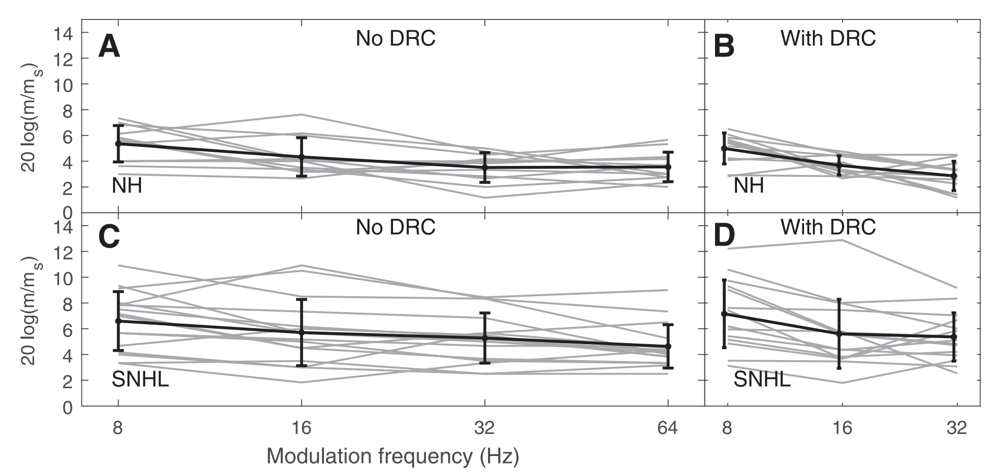

Fig. 5. Individual and mean modulation-depth discrimination (MDD) thresholds for the 1-kHz carrier. The modulation discrimination threshold (20 log(m/m_s)) is plotted as a function of the modulation frequency. A and B, the results for the listeners with normal hearing (NH). C and D, the results for the listeners with sensorineural hearing loss (LSNHL). A and C, the results obtained without dynamic range compression (DRC). B and D, The results obtained with DRC. The error bars represent ±1 SD.

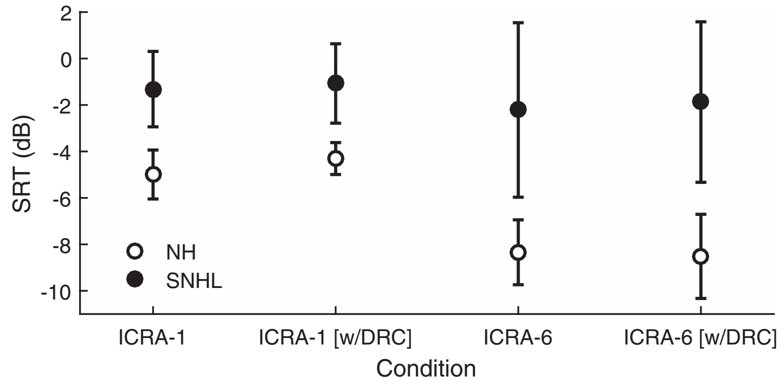

Fig. 6. Mean speech reception thresholds (SRTs) for the listeners with normal hearing (NH) and listeners with sensorineural hearing loss (LSNHL) for the two noise conditions with and without dynamic range compression (DRC). International Collegium for Rehabilitative Audiology (ICRA)-1 is steady speech-shaped noise and ICRA-6 is speech-shaped noise with the modulation characteristics of two-talker babble. The error bars represent ±1 SD.

fast-acting DRC. In the conditions without DRC, the LSNHL showed lower MDTs (for the 5-kHz carrier) than the listeners with NH, consistent with the idea that the internal representation of the signal envelope in the impaired auditory system is enhanced because of a loss of compression at the level of BM processing (e.g., Moore et al. 1996; Kale & Heinz 2010). DRC was found to increase MDTs for the SNHL group, such that they were close to the corresponding MDTs for the NH group. Thus, DRC processing can "restore" the MDTs of LSNHL to those for listeners with NH.

In contrast, MDD thresholds were *higher* for the LSNHL than for the listeners with NH, and there was no effect of DRC on the MDD thresholds. Hence, the modulation depth reduction introduced by DRC did not restore the MDD thresholds of the LSNHL to those for the listeners with NH. This result is not consistent with the hypothesis that recruitment increases the perceived amount of amplitude modulation and, therefore, can lead to a saturation of this sensation such that the differences between modulation depths are less noticeable for LSNHL than for listeners with NH (Schlittenlacher & Moore 2016). The perceived amount of amplitude modulation approaches an asymptotic value ("saturation") when the modulation depth is large but still well below full modulation (Fastl 1983). DRC should have lowered the perceived amount of amplitude modulation away from saturation, such that lower MDD thresholds for the LSNHL were expected when DRC was applied. This was clearly not the case. Instead, the results seem to support the idea that two internal noise sources limit temporal processing: an "absolute" noise term that limits MDTs (in the case of deterministic carriers, like the tonal carriers considered here) and a multiplicative noise term that limits suprathreshold modulation processing. Modeling is not the focus of the present study, and no specific model framework and quantitative predictions are presented here. However, a potential process is a constant-amplitude internal noise after a logarithmic compression of the modulation amplitude, as in the envelope power spectrum model (Ewert & Dau 2000). In this model, the MDD threshold is determined by the ratio of the signal modulation power and the internal noise power. This ratio may be considered as an "efficiency factor" for suprathreshold processing, conceptually similar to the efficiency factor represented in the classical power spectrum model of masking in the audio-frequency domain (Patterson & Moore 1986). Thus, an increased MDD threshold would be accounted for by an increased signal to noise ratio in such a model, without affecting the MDT. Likewise, an enhancement of the internal representation of the envelope in the LSNHL because of a loss of BM compression, relative to NH, may affect the MDT but not the MDD threshold.

TABLE 1. Correlation tables for the NH and the SNHL groups

1. Listeners with NH, N = 12					
SRT_{stat}	—				
SRT_{fluc}	−0.04	—			
MDD1k	−0.85*	−0.19	—		
MDT1k	0.01	0.30	−0.07	—	
MDT5k	−0.14	0.20	−0.04	0.10	—
	SRT_{stat}	SRT_{fluc}	MDD1k	MDT1k	MDT5k
2. Listeners with SNHL, N = 14					
SRT_{stat}	—				
SRT_{fluc}	0.91*	—			
MDD1k	−0.01	−0.19	—		
MDT1k	0.18	0.34	0.37	—	
MDT5k	−0.08	0.08	0.03	0.28	—
	SRT_{stat}	SRT_{fluc}	MDD1k	MDT1k	MDT5k

SRTs for fluctuating noise and stationary noise are denoted by SRT_{fluc} and SRT_{stat}, respectively. The mean MDD across modulation frequency, obtained for the 1-kHz carrier, is denoted by MDD1k. The mean MDTs across modulation frequency for the 1-kHz and 5-kHz carriers are denoted by MDT1k and MDT5k, respectively.

$^*p < 0.001$.

LSNHL, listeners with sensorineural hearing loss; MDD, modulation-depth discrimination; MDT, modulation detection thresholds; N, the number of listeners; SRT, speech reception threshold.

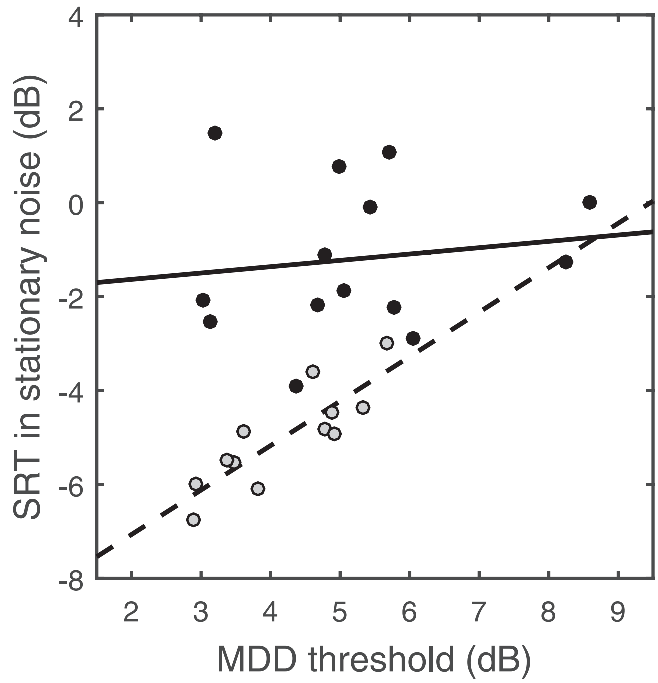

Fig. 7. Scatterplot of the speech reception threshold (SRT) in steady noise against the modulation-depth discrimination (MDD) threshold. Each symbol represents one listener. The gray and black symbols show results for the listeners with normal hearing (NH) and listeners with sensorineural hearing loss (LSNHL), respectively. The dashed and solid lines represent the regression lines for the two groups.

This approach is conceptually similar to the concept of simulating the hearing threshold (or absolute detection threshold) versus the intensity discrimination threshold via different internal noise processes (e.g., Dau et al. 1996a,b; Glasberg et al. 2001).

It is not clear which physiological processes underlie decreased "efficiency" in terms of suprathreshold amplitude modulation processing. Increased MDD thresholds could be a consequence of a reduced population of auditory-nerve fibers and synaptic elements in the inner hair cells (IHCs). Lopez-Poveda and Barrios (2013) proposed that the conversion of the envelope waveform to electrical discharges in the auditory nerve by the inner hair cells resembles a stochastic sampling of the continuous waveform. The ability to process the envelope modulation may depend on the fidelity of the sampling and thereby on whether the population of intact auditory nerve fibers and synaptic elements in the IHCs is sufficient to capture differences in modulation depth in terms of aggregated discharges. A reduced population of auditory nerve fibers and synaptic elements in the IHCs may lead to higher thresholds in the MDD task. For the MDTs, this effect might be balanced by the loss of compression at the level of BM processing.

MDTs for listeners with NH (using tonal carriers) have been shown to decrease with increasing sound pressure level (Kohlrausch et al. 2000), and in fact, MDTs for listeners with NH and LSNHL have been found to be similar when measured at the same sound pressure level (Moore & Glasberg 2001; Grose et al. 2016). Hence, the lower MDTs obtained for the LSNHL in the present study could be, at least to some extent, a consequence of the higher presentation level and might not be an effect of hearing loss per se. In contrast, MDD thresholds are not affected by presentation level as shown by Schlittenlacher and Moore (2016). No correlation between age and the measures of temporal envelope sensitivity were found for either of the listener groups. This is also consistent with Schlittenlacher and Moore (2016) who did not find any clear effect of age on MDD thresholds. For MDTs, significant effects of age have been reported, with thresholds being higher for older listeners (He et al. 2008; Füllgrabe et al. 2015).

The DRC processing used in the present study adversely affected the intelligibility of speech in stationary noise when the effect of audibility was controlled for, even though the effect was small. This suggests that the effect of restoring MDTs back to normal may not lead to improved speech perception performance, which is consistent with earlier findings showing that MDTs are only weakly correlated with speech intelligibility for LSNHL (e.g.; Takahashi & Bacon 1992; Feng et al. 2010). In contrast, interestingly, the observed strong correlation between speech intelligibility in stationary noise and MDD thresholds for the listeners with NH suggests that MDD thresholds are linked to speech intelligibility performance. This is consistent with the hypothesis that since most modulations inherent in speech are well above the MDT, suprathreshold measures of modulation processing, such as MDD thresholds, might provide stronger links to speech perception than MDTs. However, no such correlation was observed for the SNHL. The reason might be that the increased MDD thresholds only reflect one out of several deficits (such as, e.g., reduced frequency selectivity, deficits of temporal fine structure coding, as well as cognitive decline), limiting the intelligibility of speech in noisy sound environments. More generally, the LSNHL may have been more heterogeneous (despite similar audiograms across the LSNHL) than the group with NH. Such individual differences across measures of auditory and cognitive functions were not assessed in the present study. Further investigations, including auditory modeling, may help to clarify the relation between MDD processing, other auditory components characterizing suprathreshold listening, and speech intelligibility in noise for listeners with NH, LSNHL, and aided LSNHL.

CONCLUSIONS

This study investigated the effects of fast-acting DRC on listeners with NH and LSNHL' speech intelligibility and two measures of temporal modulation sensitivity (MDTs and MDD thresholds). The LSNHL showed lower MDTs (for the 5-kHz carrier) than the listeners with NH. MDTs for the LSNHL were more similar to those for the listeners with NH when using fast-acting DRC. This is consistent with the idea that the internal representation of the signal envelope in the impaired auditory system is enhanced because of a loss of compression in the cochlea. In contrast, MDD thresholds were higher for the LSNHL than for the listeners with NH, and fast-acting DRC had no effect on the MDD thresholds. This result is not consistent with the hypothesis that the differences between modulation depths are less noticeable for LSNHL than for listeners with NH because the (enhanced) envelope fluctuations are processed at or near ceiling level. These results suggest that MDTs and MDD thresholds represent two independent aspects of envelope processing. There was no beneficial effect of restoring MDTs to the level observed in the listeners with NH on speech perception. For the listeners with NH, the ability to discriminate modulation depth was highly correlated with speech intelligibility in stationary noise.

ACKNOWLEDGMENTS

We thank Nicoline Thorup and Pernille Holtegaard for their assistance in recruiting the listeners with hearing loss. We thank the Audiological Department at Bispebjerg Hospital for providing support through their facilities and staff. Many thanks to Brian Moore and two anonymous reviewers for their very helpful feedback on earlier versions of this paper.

This project was carried at the Centre for Applied Hearing Research (CAHR) supported by Widex, Oticon, GN ReSound and the Technical University of Denmark.

The authors have no conflicts of interest to disclose.

Address for correspondence: Alan Wiinberg, Department of Electrical Engineering, Technical University of Denmark, DK-2800 Lyngby, Denmark. E-mail: alwiin@elektro.dtu.dk

Received January 18, 2017; accepted February 24, 2018.

REFERENCES

Allen, J. B. (1996). Derecruitment by multiband compression in hearing aids. In Kollmeier B. (Ed.), *Psychoacoustics, Speech and Hearing Aids* (pp. 141–152). Singapore: World Scientific.

Boothroyd, A., Springer, N., Smith, L., et al. (1988). Amplitude compression and profound hearing loss. *J Speech Hear Res*, *31*, 362–376.

Brennan, M. A., Gallun, F. J., Souza, P. E., et al. (2013). Temporal resolution with a prescriptive fitting formula. *Am J Audiol*, *22*, 216–225.

Brennan, M. A., McCreery, R. W., Jesteadt, W. (2015). The influence of hearing-aid compression on forward-masked thresholds for adults with hearing loss. *J Acoust Soc Am*, *138*, 2589–2597.

Byrne, D., Parkinson, A., Newall, P. (1990). Hearing aid gain and frequency response requirements for the severely/profoundly hearing impaired. *Ear Hear*, *11*, 40–49.

Christiansen, C., MacDonald, E. N., Dau, T. (2013). Contribution of envelope periodicity to release from speech-on-speech masking. *J Acoust Soc Am*, *134*, 2197–2204.

Dau, T., Kollmeier, B., Kohlrausch, A. (1997). Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration. *J Acoust Soc Am*, *102*(5 Pt 1), 2906–2919.

Dau, T., Püschel, D., Kohlrausch, A. (1996a). A quantitative model of the "effective" signal processing in the auditory system. I. Model structure. *J Acoust Soc Am*, *99*, 3615–3622.

Dau, T., Püschel, D., Kohlrausch, A. (1996b). A quantitative model of the "effective" signal processing in the auditory system. II. Simulations and measurements. *J Acoust Soc Am*, *99*, 3623–3631.

Dau, T., Verhey, J., Kohlrausch, A. (1999). Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers. *J Acoust Soc Am*, *106*, 2752–2760.

Dreschler, W. A., Verschuure, H., Ludvigsen, C., et al. (2001). ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. International Collegium for Rehabilitative Audiology. *Audiology*, *40*, 148–157.

Drullman, R., & Smoorenburg, G. F. (1997). Audio-visual perception of compressed speech by profoundly hearing-impaired subjects. *Audiology*, *36*, 165–177.

Edwards, B. (2004). Hearing aids and hearing impairment. In S. Greenberg, A. N. Popper, R. R. Fay (Eds.), *Speech Processing in the Auditory System* (pp. 339–421). New York, NY: Springer-Verlag.

Elhilali, M., Chi, T., Shamma, S.A. (2003). A spectro-temporal modulation index (STMI) for assessment of speech intelligibility. *Speech Commun*, *41*, 331–348.

Ewert, S. D., & Dau, T. (2000). Characterizing frequency selectivity for envelope fluctuations. *J Acoust Soc Am*, *108*(3 Pt 1), 1181–1196.

Ewert, S. D., & Dau, T. (2004). External and internal limitations in amplitude-modulation processing. *J Acoust Soc Am*, *116*, 478–490.

Fastl, H. (1983). Fluctuation strength of modulated tones and broad-band noise. In R. Klinke and R. Hartmann (Eds.), *Hearing—Physical Bases and Psychophysics* (pp. 282–288). Berlin, Germany: Springer.

Feng, Y., Yin, S., Kiefte, M., et al. (2010). Temporal resolution in regions of normal hearing and speech perception in noise for adults with sloping high-frequency hearing loss. *Ear Hear*, *31*, 115–125.

Fowler, E. P. (1936). A method for the early detection of otosclerosis: A study of sounds well above threshold. *Arch Otolaryngol*, *24*, 731–741.

Füllgrabe, C., Moore, B. C., Stone, M. A. (2014). Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition. *Front Aging Neurosci*, *6*, 347.

Gatehouse, S., Naylor, G., Elberling, C. (2006). Linear and nonlinear hearing aid fittings–1. Patterns of benefit. *Int J Audiol*, *45*, 130–152.

Glasberg, B. R., Moore, B. C., Peters, R. W. (2001). The influence of external and internal noise on the detection of increments and decrements in the level of sinusoids. *Hear Res*, *155*, 41–53.

Grose, J. H., Porter, H. L., Buss, E., et al. (2016). Cochlear hearing loss and the detection of sinusoidal versus random amplitude modulation. *J Acoust Soc Am*, *140*, EL184.

He, N. J., Mills, J. H., Ahlstrom, J. B., et al. (2008). Age-related differences in the temporal modulation transfer function with pure-tone carriers. *J Acoust Soc Am*, *124*, 3841–3849.

Houtgast, T., & Steeneken, H. J. M. (1985). A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *J Acoust Soc Am*, *77*, 1069–1077.

International Electrotechnical Commission (1983). *Hearing Aids. Part 2: Hearing Aids With Automatic Gain Control Circuits. IEC 60118-2-1983*. Geneva: IEC.

International Organization for Standardization (2010). ISO 8253-1:2010. *Audiometric Test Methods—Pure Tone Audiometry*. Geneva, Switzerland: ISO.

Jørgensen, S., & Dau, T. (2011). Predicting speech intelligibility based on the signal-to-noise envelope power ratio after modulation-frequency selective processing. *J Acoust Soc Am*, *130*, 1475–1487.

Kale, S., & Heinz, M. G. (2010). Envelope coding in auditory nerve fibers following noise-induced hearing loss. *J Assoc Res Otolaryngol*, *11*, 657–673.

Kates, J. M. (2008). Digital Hearing Aids (Plural, San Diego, CA). Kohlrausch, A., Fassel, R., Dau, T. (2000). The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. *J Acoust Soc Am*, *108*, 723–734.

Kowalewski, B., Macdonald, E., Strelcyk, O., et al. (2015). Auditory-model based assessment of the effects of hearing loss and hearing-aid compression on spectral and temporal resolution. *Proc Int Symp Audit Audiol Res*, *5*(December), 173–180.

Lee, J., & Bacon, S. P. (1997). Amplitude modulation depth discrimination of a sinusoidal carrier: Effect of stimulus duration. *J Acoust Soc Am*, *101*, 3688–3693.

Levitt, H. (1971). Transformed up δown methods in psychoacoustics. *J Acoust Soc Am*, *49*(2B), 467–477.

Lopez-Poveda, E. A., & Barrios, P. (2013). Perception of stochastically undersampled sound waveforms: A model of auditory deafferentation. *Front Neurosci*, *7*, 124.
Moore, B. C. (2004). Testing the concept of softness imperception: loudness near threshold for hearing-impaired ears. *J Acoust Soc Am*, *115*, 3103–3111.
Moore, B. C., & Glasberg, B. R. (2001). Temporal modulation transfer functions obtained using sinusoidal carriers with normally hearing and hearing-impaired listeners. *J Acoust Soc Am*, *110*, 1067–1073.
Moore, B. C., Glasberg, B. R., Alcántara, J. I., et al. (2001). Effects of slow- and fast-acting compression on the detection of gaps in narrow bands of noise. *Br J Audiol*, *35*, 365–374.
Moore, B. C. J., Wojtczak, M., Vickers, D.A. (1996). Effect of loudness recruitment on the perception of amplitude modulation. *J Acoust Soc Am*, *100*, 481–489.
Nielsen, J. B., & Dau, T. (2011). The Danish hearing in noise test. *Int J Audiol*, *50*, 202–208.
Noordhoek, I. M., & Drullman, R. (1997). Effect of reducing temporal intensity modulations on sentence intelligibility. *J Acoust Soc Am*, *101*, 498–502.
R. D. Patterson and B. C. J. Moore (1986). Auditory filters and excitation patterns as representations of frequency resolution. In B.C.J. Moore, (Ed.), *Frequency Selectivity in Hearing*, Academic, London.
Reinhart, P. N., Souza, P. E., Srinivasan, N. K., et al. (2016). Effects of Reverberation and Compression on Consonant Identification in Individuals with Hearing Impairment. *Ear Hear*, *37*, 144–152.
Schlittenlacher, J., & Moore, B. C. (2016). Discrimination of amplitude-modulation depth by subjects with normal and impaired hearing. *J Acoust Soc Am*, *140*, 3487.
Sheft, S., & Yost, W. A. (1990). Temporal integration in amplitude modulation detection. *J Acoust Soc Am*, *88*, 796–805.
Souza, P. E., & Turner, C. W. (1996). Effect of single-channel compression on temporal speech information. *J Speech Hear Res*, *39*, 901–911.
Souza, P. E., & Turner, C. W. (1998). Multichannel compression, temporal cues, and audibility. *J Speech Lang Hear Res*, *41*, 315–326.
Steinberg, J., & Gardner, M. (1937). The dependence of hearing impairment on sound intensity. *J Acoust Soc Am*, *9*, 11–23.
Stone, M. A., & Moore, B. C. (1992). Syllabic compression: Effective compression ratios for signals modulated at different rates. *Br J Audiol*, *26*, 351–361.
Stone, M. A., Anton, K., Moore, B. C. (2012). Use of high-rate envelope speech cues and their perceptually relevant dynamic range for the hearing impaired. *J Acoust Soc Am*, *132*, 1141–1151.
Stone, M. A., Füllgrabe, C., Moore, B. C. (2008). Benefit of high-rate envelope cues in vocoder processing: Effect of number of channels and spectral region. *J Acoust Soc Am*, *124*, 2272–2282.
Strelcyk, O., Nooraei, N., Kalluri, S., et al. (2012). Restoration of loudness summation and differential loudness growth in hearing-impaired listeners. *J Acoust Soc Am*, *132*, 2557–2568.
Takahashi, G. A., & Bacon, S. P. (1992). Modulation detection, modulation masking, and speech understanding in noise in the elderly. *J Speech Hear Res*, *35*, 1410–1421.
van Buuren, R. A., Festen, J. M., Houtgast, T. (1999). Compression and expansion of the temporal envelope: Evaluation of speech intelligibility and sound quality. *J Acoust Soc Am*, *105*, 2903–2913.
Villchur, E. (1973). Signal processing to improve speech intelligibility in perceptive deafness. *J Acoust Soc Am*, *53*, 1646–1657.
Yund, E. W., & Buckles, K. M. (1995). Multichannel compression hearing aids: effect of number of channels on speech discrimination in noise. *J Acoust Soc Am*, *97*, 1206–1223.