Modeling auditory signal processing in hearing-impaired listeners

Morten L. Jepsen and Torsten Dau

Centre for Applied Hearing Research, Ørsted•DTU, Technical University of Denmark, DK-2800 Lyngby, Denmark

Recently, an auditory signal processing model was developed which can simulate psychoacoustic data from a large variety of conditions related to spectral and temporal masking in normal-hearing listeners (Jepsen *et al.*, 2008). The model includes the dual-resonance non-linear (DRNL) filterbank suggested by Lopez-Poveda and Meddis (2001) to simulate the non-linear cochlear signal processing, and is otherwise similar to the modulation filterbank model by Dau *et al.* (1997). In the present study, the model parameters were modified to simulate cochlear hearing impairment. The modifications of the model were based on individual data from notched-noise masking and forward masking and were associated with changes of the parameters of the DRNL stage of the model. Data from a pure-tone audiogram were used to further reduce listener sensitivity in connection with an assumed loss of inner hair cells. In addition, intensity discrimination experiments and a modulation depth discrimination experiment were performed to estimate potential retro-cochlear (central) limitations of the processing of supra-threshold stimuli. The model helps understanding the perceptual consequences of hearing impairment in individual listeners and might be useful for the evaluation of hearing-aid signal processing.

INTRODUCTION

In the processing model of Dau *et al.* (1997), a linear basilar-membrane (BM) filtering stage, the gammatone filterbank, was used to simulate a variety of detection and masking data. However, the model could not account for effects associated with non-linearities, such as compression and suppression, observed in the intact cochlea. Jepsen *et al.* (2008) developed a substantially modified version of the original model that includes the dual-resonance nonlinear (DRNL) filter at its cochlea stage (Lopez-Poveda and Meddis, 2001). The Jepsen *et al.* model was shown to account for a variety of spectral- and temporal masking aspects and modulation detection for normal-hearing (NH) listeners. In particular, the new model was more successful than the original model in several conditions where cochlear nonlinearity is crucial.

In the present study, the idea was to modify the Jepsen *et al.* model in order to account for masking data from individual listeners with hearing-impairment (HI) of cochlear origin (sensori-neural). This common type of impairment is often associated with the loss or deficit of outer hair-cell function and, thus, a loss or reduction of compression. It is assumed here that the hearing impairment can be modeled primarily by reducing the amount of cochlear compression and the sensitivity at the hair cell stage while keeping all subsequent processing stages identical to the model for NH. Lopez-Pov-

Auditory signal processing in hearing-impaired listeners. 1st International Symposium on Auditory and Audiological Research (ISAAR 2007). T. Dau, J. M. Buchholz, J. M. Harte, T. U. Christiansen (Eds.). ISBN: 87-990013-1-4. Print: Centertryk A/S.

eda and Meddis (2001) showed how the parameters of the DRNL filter could be modified to simulate moderate and severe hearing loss. Several psychoacoustic experiments were conducted in order to derive model parameters that can characterize an individual hearing loss. The experiments were: forward masking to estimate the amount of compression (Oxenham and Plack, 1997), notched-noise masking to estimate frequency selectivity (Rosen *et al.*, 1998), and intensity discrimination and modulation-depth discrimination (Ewert and Dau, 2004) in order to estimate the variance of the internal noise at a central stage of processing in the model. Additionally, information from a pure-tone audiogram was used. In this study, data were obtained from two listeners with severe high-frequency hearing loss. These preliminary results will show how a model of auditory perception can be modified to characterize an individual hearing-loss. This could potentially be useful for hearing-aid development, evaluation of hearing-aid processing and more generally, for a better understanding of the perceptual consequences of a cochlear hearing loss.

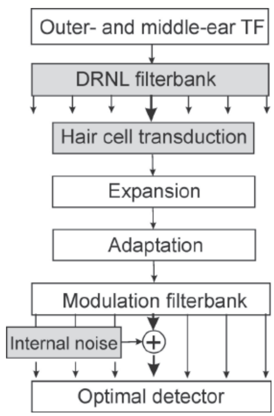

Fig. 1: The auditory processing model comprising stages of the outer- and middle-ear, DRNL unit, hair cell transduction, expansion, adaptation, modulation filterbank and the optimal detector.

THE MODEL

The model (Fig. 1) has an overall structure similar to the models proposed in Dau *et al.* (1997) and Jepsen *et al.* (2008). The first stages consist of outer- and middle ear filters and the DRNL filterbank that simulates BM processing. The DRNL filter has two paths; one linear and one nonlinear path both with a bandpass characteristic. The linear path is dominant at high levels (> 70 dB SPL) while the nonlinear path is dominant at low and medium levels and has a compressive non-linear gain. For further details on the DRNL parameters, the reader is referred to Lopez-Poveda and Meddis (2001). The transformation of the mechanical BM oscillations into inner hair-cell (IHC) receptor potentials is simulated roughly by half-wave rectification and low-pass filtering at 1-kHz. At this point, an estimated loss of IHC can be simulated by a linear attenuation if desired. The signal is then transformed into an intensity-like representation, by applying a squaring expansion. This step is motivated by findings showing that the auditory-nerve (AN) spike rate exhibits a square-law behaviour near AN threshold as a function of stimulus level (Muller *et al.*, 1991). The adaptation stage in the model simulates adaptive properties of the auditory periphery. As in the original model, the effects of adaptation are realized by a chain of five feedback loops in series with different time constants. The output of the entire stage approaches a logarithmic compression for stationary signals. For input variations that are rapid, compared with the time constants of the low-pass filters, the transformation through the adaptation loops is more linear, leading to a higher sensitivity for fast temporal variations. The output of the adaptation stage is filtered by a 1st-order low-pass filter at 150 Hz, motivated by results from modulation detection data with sinusoidal carriers (e.g., Kohlrausch *et al.*, 2000). The low-pass filter is followed by a modulation filterbank as proposed in Dau *et al.* (1997). The lowest modulation filter is a 2nd order Butterworth lowpass filter at 2.5 Hz. For frequencies above 5 Hz there is an array of bandpass filters with a quality factor of Q = 2. Modulation filters with a centre frequency above 10 Hz only output the Hilbert envelope of the modulation filters. Internal noise is added in order to limit the resolution of the model. The decision device is realized as an optimal detector. The model was calibrated by adjusting the variance of the internal noise such that the model satisfies Weber's law when considering an intensity discrimination task using deterministic stimuli. The stages shown in gray in Fig. 1 reflect where parameters can be modified to simulate a hearing loss. These comprise the DRNL filterbank, the hair-cell transduction stage and the variance of the internal noise. The DRNL is modified if there is evidence for a loss of BM compression and the sensitivity is reduced in the hair-cell stage if it is estimated that there is an IHC loss. The internal noise is adjusted according to the limitation in the listener's sensitivity in discrimination tasks which may reflect a retro-cochlear or cognitive component of the hearing loss.

EXPERIMENTS

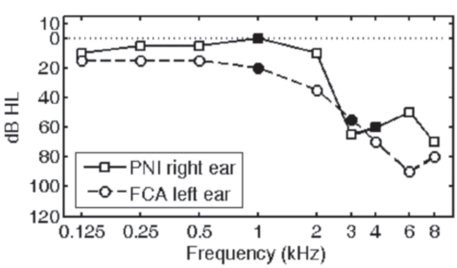

Fig. 2: Pure-tone audiograms of the hearing-impaired subjects PNI (squares) and FCA (circles).

Two hearing-impaired listeners participated in the experiments (PNI, 72 years and FCA, 76 years). Both had a sloping sensorineural hearing loss. Measurements were performed on one ear; the pure-tone audiograms of the measured ears are shown in Fig. 2. All experiments (except the audiogram measurement) used a three alternative forced-choice (3-AFC) method with a 1-up-2-down adaptive tracking scheme leading to thresholds corresponding to the 70.7% point on the psychometric function. Subjects received 30 to 45 min. training sessions before each new experiment until no systematic improvements in threshold could be observed. Thresholds were calculated as the mean threshold from two to four runs. Measurements including training lasted 12 h for each subject

The stimuli in the forward-masking experiment were similar to the stimuli used in Oxenham and Plack (1997) and Rosengaard *et al.* (2005). A short 10-ms probe signal was presented at frequencies (f_p) of 1, 3 or 4 kHz. There were two forward masker conditions; the on-frequency condition where the masker frequency (f_m) was equal to f_p and the off-frequency condition where f_m was 0.55 x f_p. The duration of the masker was 110 ms including 5-ms squared-cosine ramps. The probe was presented 2 ms after the masker offset at fixed levels depending on the hearing loss at the specific probe frequency. The level of the masker just necessary to mask the probe was measured. The notched-noise experiment followed the constant signal level paradigm of Rosen *et al.* (1998). Masked thresholds were measured at 1, 2 or 2.5 kHz using signal levels of either 50 or 60 dB SPL. Five symmetric and two asymmetric conditions were considered. The masker and signal durations were 550 ms and 440 ms, respectively. Squared-cosine ramps with a duration of 50 ms were used for the signal and the masker. The masker levels necessary to just mask the signal were measured. In the intensity discrimination task, just noticeable differences (JNDs) were measured both using stochastic broadband noise and pure tones. The stimulus durations were 500 ms including 50-ms squared-cosine ramps. Two noise bandwidths were considered: one ranging from 20 Hz to 8 kHz and one ranging from 1 kHz to 8 kHz. Tonal JNDs were measured at 1, 3 or 4 kHz. At 1 kHz, the measurements were performed using standard levels at 60 and 80 dB SPL, respectively, and for the 3 and 4 kHz tone, the standard levels were 80 and 90 dB SPL due to the subject's elevated absolute threshold. Modulation depth discrimination thresholds were measured for a tonal carrier, $c(t)$, at 1 kHz, similar as in Ewert and Dau (2004). A sinusoidal amplitude modulation (AM) of 16 Hz (fmod) was imposed. The signal had a duration of 500 ms including 50-ms squared-cosine ramps. The stimulus level was 65 dB SPL. The equation describing the stimulus $s(t)$ was:

$$s(t) = [1 + m_s(1 + m_{inc})0.5 \times \sin(2\pi f_{mod} t)]\, c(t) \qquad \text{(Eq. 1)}$$

where ms is the standard modulation depth and minc is the relative AM increment. Level cues were eliminated by scaling each stimulus by a factor of $(1+m^2/2)^{0.5}$. Increment detection thresholds were measured at standard modulation depths ranging from -23 dB to -3 dB in 5 dB steps.

RESULTS

Data from the forward masking experiment are presented as growth-of-masking (GOM) functions and were used to estimate the maximal amount of compression on the BM according to the paradigm of Rosengaard *et al.* (2005). In the on-frequency condition, the masker and the signal are assumed to be subjected to similar amounts of compression. In the off-frequency condition the signal is processed compressively, while the masker is processed linearly. These conditions are valid for the range of signal and masker levels used here. The paradigm suggests that the amount of BM compression can be calculated as the ratio of the slopes of GOM functions in the off- and on-frequency conditions. Figure 3 shows the data from the two hearing-impaired subjects (PNI and FCA). The dynamic range of the signal levels is limited by the subjects' absolute threshold of the probe and the maximal masker level that was allowed here (98 dB SPL). Straight lines were fitted to the data and the ratio of the slopes, reflecting the amount of BM compression, was calculated. For subject PNI the amount of compression was 0.51 and 0.83 for 1 and 4 kHz, respectively. For normal-hearing listeners, compression has been estimated to be 0.48 and 0.28 at these two frequencies (Rosengaard *et al.*, 2005). This indicates that PNI shows a near normal amount of compression at 1 kHz and reduced BM compression at 4 kHz. The estimated compression in subject FCA was 0.97 and 1.14 for 1 and 3 kHz; this subject thus shows presumably no compression at both frequencies. The data from the notched-noise experiments were used to fit 'roex' filter parameters as suggested in Rosen *et al.* (1998). The estimated auditory filter shapes are shown in Fig. 5. A widely used measure of auditory-filter width is the equivalent rectangular bandwidths (ERBs) which, for normal hearing listeners, are 133, 241 and 295 Hz at 1, 2 and 2.5 kHz, respectively. Subject PNI has a near normal auditory-filter

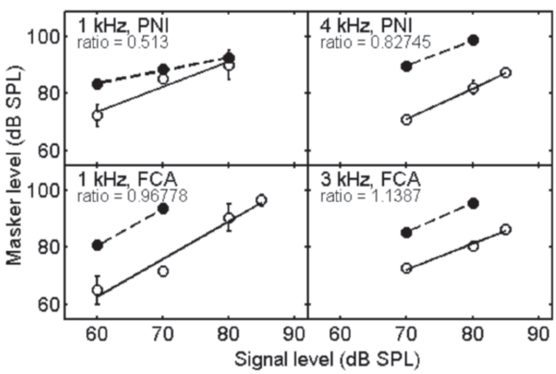

Fig. 3: Growth of masking functions showing the masker level just necessary to mask a signal at a given level. Data from the on- and off-frequency conditions are indicated by open and closed symbols, respectively. bandwidth (ERB = 178 Hz) at 1 kHz and a broadened filter (ERB = 452 Hz) at 2.5 kHz. FCA has broader filters at 1 and 2 kHz with ERBs of 352 and 567 Hz, respectively. It was not possible to obtain notched-noise data at 3 or 4 kHz due to the subjects' elevated pure-tone thresholds at this frequency.

Figure 4 (left panels) shows the data from the intensity discrimination experiments. For normal-hearing listeners, the JNDs typically lie in the range from 0.5 to 1 dB depending on the stimulus condition. Subject PNI has JNDs similar to normal hearing listeners even though slightly higher at the lowest standard level. A comparison of the JNDs obtained in the two noise conditions reveals that this listener does not benefit from the low-frequency content (0.1 – 1.0 kHz) of the broadband noise in this task. Subject FCA seems to have difficulties in this task by showing elevated JNDs in all conditions (1.7 to 3.5 dB). These results suggest that the hearing loss is of cochlear origin for PNI, while FCA may have an additional hearing loss component, since elevated JNDs can be expected in the case of a more centrally located hearing loss. Figure 4 (right panels) shows the data from the modulation depth discrimination task. Open circles indicate the data from the hearing-impaired subjects while filled symbols indicate data from normal hearing listeners (Ewert and Dau, 2004). Both subjects show performance close to NH at the highest standard modulation depths (-13 to -3 dB) while they are less sensitive at the two lowest standard modulation depths. This may reflect a limited sensitivity in this task, but the data could also reflect that more training is necessary in this rather difficult task.

DERIVING MODEL PARAMETERS

In the model for normal hearing (Jepsen *et al.*, 2008), cochlear compression was simulated using the DRNL filterbank. In this processing stage, it is possible to reduce the amount of simulated cochlear compression according to the results of the forward masking experiment. There are numerous parameters in the DRNL filter, but the compressive characteristics can be modified using four parameters; *a, b, c,* that control the non-linear gain function in the non-linear path, and g, the gain in the linear path. Here, these parameters were chosen such that the estimated BM compression in the measured level ranges is reflected in the input/output-function of the corresponding DRNL filter. After the fit to the amount of compression it was investigated how the

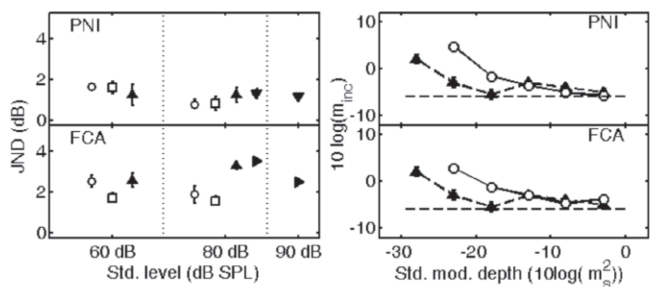

Fig. 4: Left panels show the data from the JND measurements. Open circles indicate JNDs with noise (20 Hz – 8 kHz), open squares (1 – 8 kHz), upward triangles tonal JNDs at 1 kHz, downward at 4 kHz and right-pointing indicate 3 kHz. Right panels show data from the modulation depth discrimination task. Closed triangles represent data from NH listeners. Open circles indicate data from HI listeners at a carrier frequency of 1 kHz.

DRNL filter bandwidths could account for the measured auditory filter bandwidths estimated in the notched-noise experiment. A reduced BM compression should result in a larger filter bandwidth. The top panels in Fig. 5 show how the estimated amount of compression was used to fit the DRNL input/output functions. For PNI, the function is identical to that for normal hearing at 1 kHz, while it is completely linear at 4 kHz. For FCA, the I/O-function is linear both at 1 and 3 kHz. The bottom panels show the corresponding derived auditory filters, represented as iso- intensity curves. The dotted curves indicate the filters obtained with the NH model, while the solid curves show the DRNL filters for the HI model. The dashed curves indicate the estimated roex auditory filters derived from the notched-noise data. It appears that these are in good agreement with the derived DRNL filters for both subjects. For FCA, the DRNL filter width at 2 kHz is narrower on the low frequency side compared to the estimated roex filter. However, in general, the DRNL filters do account for the asymmetry of the estimated roex filters.

The horizontal offset between NH and HI input/output functions reflects that the simulated listeners' sensitivity is reduced by up to 40 dB due to the loss of compression associated with the loss of outer hair-cells. To reduce the subjects' sensitivity further in order to be in agreement with the pure-tone sensitivity according to the audiogram there should be a frequency-dependent linear attenuation stage after the DRNL stage. This linear attenuation can be associated with the loss of inner hair-cells leading to reduced transmission capability. As an example, subject FCA has a pure-tone threshold of 55 dB SPL at 3 kHz. Loss of compression accounts for 40 dB while the remaining 15 dB could be accounted for by such a linear attenuation.

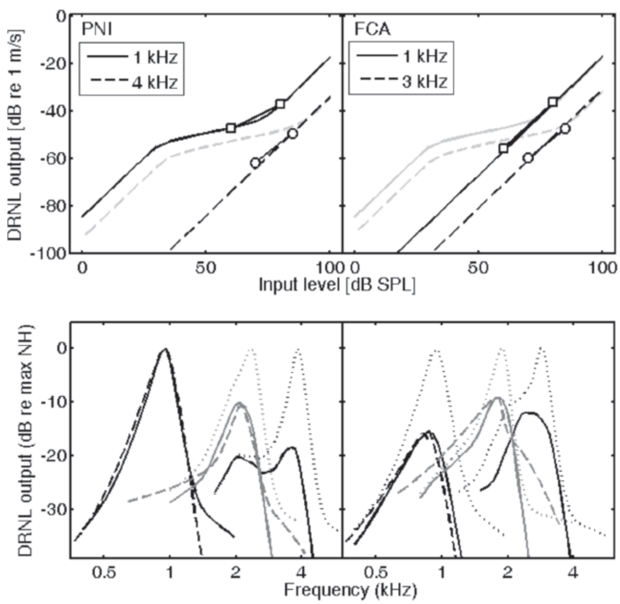

Fig. 5: Top panels: Input/output functions of the DRNL filter fitted to the slopes derived from the forward masking experiments (marked with symbols) Gray lines indicate the NH functions. Bottom: Iso-intensity response curves of the DRNL filter at different center frequencies. Dotted curves indicate filters for NH, solid curves indicate filters for HI listener and dashed curves indicate estimated filter shapes based on the notched-noise data.

When the results from the intensity discrimination tasks or the modulation depth discrimination task show degraded performance, as it was the case for subjects FCA of the present study, the variance of the internal noise in the model can be increased in order to limit the resolution of the model, thus leading to higher JNDs. The internal noise can be adjusted according to the results of the tonal JND measurements presented here. Such simulations were not performed in the present study.

DISCUSSION

In this study, preliminary results have shown how data from psychoacoustic experiments can be used to modify parameters of an auditory processing model to characterize individual hearing impairment. Forward masking was used to estimate the amount of compression at different frequencies, and notched-noise masking data were used to investigate if the DRNL filter width was reasonable when reducing or removing compression. Interestingly, it was found that subject FCA did not show any compression at 1 kHz, even though his audiogram only indicates a 20-dB hearing loss. Further information about the compressive behavior could be included if one could estimate the breakpoint on the BM input/output function by a similar forward masking paradigm. This would be most relevant when dealing with moderate hearing loss where BM compression still exists. The pulsation threshold paradigm (Plack and Oxenham, 2000) can do this, but large training effects and standard deviations do not favour this. Only few frequencies were investigated here, and future investigations should consider how to interpolate information about compression and filter bandwidths across frequency. It appears that the loss of compression (outer hair-cell function) within the model cannot fully account for the reduced audibility. Therefore, an additional linear attenuation stage should be included and this might be associated with a loss of inner hair-cell function. It should also be investigated how such frequency-dependent linear attenuation values could be estimated. The intensity discrimination and modulation-depth discrimination tasks were included to investigate more central processing performance. It was shown that even with these two subjects, reflecting similar audiograms, only FCA had difficulties in the intensity discrimination task. The subjects performed similarly in the modulation depth discrimination task, both having higher thresholds than normal for the lowest standard modulation depths.

The approach of the present study could be useful when characterizing and modeling individual hearing losses. Future work will try to establish a reliable battery of psychoacoustic tests which are efficient and sufficient. A battery of psychoacoustic validation experiments should also be conducted on the same subjects. These should cover several aspects of spectral and temporal masking, such as spectral masking, forward masking with noise maskers and modulation detection and more (e.g., Jepsen *et al.*, 2008). Data should be measured on several subjects to establish a clear idea about the influence of the model parameters used to characterize the hearing loss. If an auditory processing model can appropriately account for the individual hearing loss, this would help understanding the perceptual consequences of the hearing impairment. Secondly, such a model could be used in applications dealing with objective measures of audio quality or speech intelligibility. Such applications would be very useful in e.g. development of hearing-aid algorithms in order to save time and resources on subjective testing.

REFERENCES

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997**). “Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers,” J. Acoust. Soc. Am. **102**, 2892-2905

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

Jepsen, M. L., Ewert, S. D., and Dau, T. (**2007**), “Modeling spectral and temporal masking in the human auditory system,” J. Acoust. Soc. Am., accepted.

Kohlrausch, A., Fassel, R., and 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.

Lopez-Poveda, E. A., and Meddis, R. (**2001**). “A human nonlinear cochlear filterbank,” J. Acoust. Soc. Am. **110**, 3107–3118.

Muller M., Robertson D., and Yates G. K. (**1991**). “Rate-versus-level functions of primary auditory nerve fibres: evidence for square law behaviour of all fibre categories in the guinea pig,” Hear. Res. **55**, 50–56.

Oxenham, A., and Plack, C. (**1997**). “A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing,” J. Acoust. Soc. Am. **101**, 3666-3675

Plack C. J., and Oxenham A. J. (**2000**). “Basilar-membrane nonlinearity estimated by pulsation threshold,” J. Acoust. Soc. Am. **107**, 501–507.

Rosen, S., Baker, R. J., and Darling, A. (**1998**). “Auditory filter nonlinearity at 2 kHz in normal hearing listeners,” J. Acoust. Soc. Am. **103**, 2539-2550.

Rosengard, P.S., Oxenham, A.J., and Braida, L.D. (**2005**). “Comparing different estimates of cochlear compression in listeners with normal and impaired hearing,” J. Acoust. Soc. Am. **117**, 3028-3041