Modeling human auditory evoked brainstem responses based on nonlinear cochlear processing

Harte, James; Rønne, Filip Munch; Dau, Torsten

Published in: Proceedings of the 20th International Congress on Acoustics

Publication date: 2010

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Modeling human auditory evoked brainstem responses based on nonlinear cochlear processing

James M. Harte (1), Filip Rønne (1) and Torsten Dau (1)

(1) Centre for Applied Hearing Research, Department of Electrical Engineering, Technical University of Denmark, DK

PACS: 43.64.Ri, 43.64.Bt

ABSTRACT

The aim of this study was to accurately simulate auditory evoked potentials (AEPs) from various classical stimuli such as clicks and tones, often used in research and clinical diagnostics. In an approach similar to Dau (2003), a model was developed for the generation of auditory brainstem responses (ABR) to transient sounds and frequency following responses (FFR) to tones. The model includes important cochlear processing stages (Zilany and Bruce, 2006) such as basilar-membrane (BM) tuning and compression, inner hair-cell (IHC) transduction, and IHC auditory-nerve (AN) synapse adaptation. To generate AEPs recorded at remote locations, a convolution was made on an empirically obtained elementary unit waveform with the instantaneous discharge rate function for the corresponding AN unit. AEPs to click-trains, as well as to tone pulses at various frequencies, were both modelled and recorded at different stimulation levels and repetition rates. The observed nonlinearities in the recorded potential patterns, with respect to ABR wave V latencies and amplitudes, could be largely accounted for by level-dependent BM processing as well as effects of short-term neural adaptation. The present study provides further evidence for the importance of cochlear tuning and AN adaptation on AEP patterns, and provides a useful basis for the study of more complex stimuli including speech.

INTRODUCTION

For sounds which convey information, such as speech and music, much of the information is carried in the changes in the stimulus, rather than in the parts of the sound which are relatively stable. Through the last decades both psychoacoustic and physiological studies have investigated how the auditory system analyses the temporal modulations of sounds. When various sounds are presented to human subjects, it is possible to record auditory evoked potentials (AEPs) on the surface of the human scalp. Auditory evoked potentials are the summed response from many remotely located neurons recorded via scalp electrodes. They can be recorded from all levels of the auditory pathway, from the auditory nerve, the brainstem up to the cortex. They are typically grouped in terms of time of occurrence after stimulus offset and thus are known as: auditory brainstem responses (ABRs) recorded between 1 and 7 ms after stimulus offset; middle latency responses (MLRs) recorded in the interval 15-50 ms after acoustic stimulus; and auditory late response (ALR) recorded in the interval 75-200 ms after stimulus.

Hearing deficiencies often lead to difficulties in understanding speech, especially in noisy and reverberant environments. Auditory evoked potentials are a powerful tool used to diagnose and assess classical hearing deficiencies. This has led to a trend in the literature of assessing and investigating speech and complex speech-like stimuli with AEPs, i.e. Aiken and Picton (2008); Akhoun et al. (2008); Chandrasekaran and Kraus (2010); Lalor and Foxe(2010) to name a few. AEPs are relatively well understood for basic stimuli, i.e. transients, tone bursts and tones. However, for more complex stimuli, which include amplitude and frequency modulations as well as sharp on-set and off-set transients, it is still relatively poorly understood how the various neurophysiological processing along the auditory pathway gives rise to the AEP recorded at surface electrodes. A clearer understanding of how the underlying neurophysiology in the auditory system leads to surface-recorded scalp potentials could help to assess hearing impairment, or indeed how well this has been compensated for with an auditory prosthestic (Aiken and Picton, 2008), such as a hearing aid or cochlear implant.

The long term goal of this study is to model and simulate speech-evoked and complex (non-speech) sound evoked AEPs originating in the auditory nerve and brainstem, based on current knowledge of neural auditory signal processing. Dau (2003) developed a model for the generation of early AEPs, including auditory brainstem responses (ABR) to transient sounds like clicks and frequency following responses (FFR) to tones. Both of these AEPs are generated by neurons in the auditory nerve (AN) and subsequent stages along the auditory brainstem. The model included important cochlear processing stages such as basilar-membrane (BM) filtering with a compressive feedback loop, inner hair-cell (IHC) transduction, and IHC-AN synapse adaptation. The instantaneous AN discharge rate from the model was convolved with an empirically obtained elementary unit waveform, to simulate AEPs.

In the present paper, the Dau (2003) model is extended to include current advances in AN modeling (Zilany and Bruce, 2006) and is humanised. The original Dau (2003) model, used the Heinz et al. (2001) AN model fitted to experimental cat AN data. Here, the Zilany and Bruce (2006) AN model will be adapted for humans by ensuring the model has appropriate thresholds, tuning curves, BM travelling wave latencies etc., based on current state-of-the-art knowledge derived from both behavioural and objective measures where possible. This paper will present a comparison of the model output with basic transient, tone-burst and click-train data, in an attempt to build up stimulus complexity towards the final goal of speech. Thus it is possible to challenge the model with relatively basic stimuli, before increasing complexity. This paper focuses on the role of basilar membrane tuning and the adaptation mechanism of the AN model and looks at the consequences for AEPs generated.
Neural adaptation is the phenomenon where neural output is reduced due to prolonged or repeated stimulation, in each stage of the auditory pathway. The role of adaptation in AEPs, and more specifically ABRs, is important because in clinical practice it is highly desirable to get accurate recordings of ABRs quickly, particularly from uncooperative subjects, i.e., neonates. Any morphological differences, such as amplitude and latency, from normative data caused by stimulus rate adaptation could interfere with diagnosis. The desire for quicker acquisition time has led to the use of rapid rates of stimulation via so-called pseudo-random binary sequences or maximum length sequences, see Burkard et al. (1990); Jewett et al. (2004) for a few examples. The response to these pseudo-random pulse trains needs to be deconvolved to obtain an estimate of the ABR. The higher rate of the sequence leads to typically smaller ABR amplitudes. This is believed to be a result of neural adaptation.

METHODS

Model for AEP generation

Melcher and Kiang (1996) investigated the relationship between brainstem cell populations and the auditory brainstem response (ABR) in cats. They described the generation of ABRs as a summation of individual cell potentials, \( v_i \), in response to a given stimulus, \( s \),

\[
ABR(t, x_1, x_2) = \sum_i v_i(t, x_1, x_2, s)
\]  

(1)

where \( x_1 \) and \( x_2 \) are the locations of the electrodes on the scalp. The potential, \( v_i \), produced by the individual nerve cell in response to a given stimulus, can be determined by a convolution between the instantaneous firing rate of the \( p \)th cell, \( r_i(t, s) \), and the unitary response, \( u_i(t, x_1, x_2) \). This latter function is defined as the potential produced between the electrode positions on the scalp, \( x_1 \) and \( x_2 \), each time the cell discharges:

\[
v_i = u_i(t, x_1, x_2) \ast r_i(t, s)
\]  

(2)

where \( \ast \) denotes the convolution operation. To obtain an ABR with this method, all cells need to be considered individually. To avoid this, Melcher and Kiang (1996) suggested the use of a cell population potential, \( V \), instead. These can be based on the physio-anatomical type of the cell, \( p \),

\[
ABR(t, x_1, x_2, s) = \sum_p V_p(t, x_1, x_2, s)
\]  

(3)

where \( P \) is the number of different cell populations. It is reasonable to assume that all cells of a given population have the same unitary response, \( r \), as they are all of the same type. Each AN fiber is tuned to a specific CF, where the number of fibers \( N \) is the total number of cells. This implies that the different contributions from the different populations to the scalp potential are assumed to be reflected in the shape of the corresponding individual unitary responses, \( u_p(\cdot) \). Taking advantage of linear super-position, this can be further simplified to:

\[
ABR(t, x_1, x_2, s) = \sum_{p=1}^{N_p} r_p(t, s) \ast u_p(t, x_1, x_2)
\]  

(4)

The summed instantaneous firing rate is then convolved with a unitary response to create the modelled ABR

\[
ABR(t, x_1, x_2, s) = \sum_{p=1}^{N_p} r_p(t, s) \ast u_p(t, x_1, x_2)
\]  

(5)

where \( N \) is the number of different cell populations. It is reasonable to assume that all cells of a given population have the same unitary response, \( u \), as they are all of the same type. Each AN fiber is tuned to a specific CF, where the number of fibers \( N \) is the total number of cells. This implies that the different contributions from the different populations to the scalp potential are assumed to be reflected in the shape of the corresponding individual unitary responses, \( u_p(\cdot) \). Taking advantage of linear super-position, this can be further simplified to:

\[
ABR(t, x_1, x_2, s) = \sum_{p=1}^{N_p} r_p(t, s) \ast u_p(t, x_1, x_2)
\]  

(6)

where \( R(t, s) \) is the summed discharge rate function, and \( U(\cdot) \) is the unitary response function summed across the different neural populations. Thus the generation mechanism of ABR is determined as the sum of instantaneous firing from all cells, convolved with a unitary response summed over all cell types that is dependent on the electrode location on the scalp but is assumed to be independent of efferent influence and stimulus.

The structure of the ABR model is shown in figure 1, where a parallel bank of AN fibers are individually modelled by the AN model. The summed instantaneous firing rate is then convolved with a unitary response to create the modelled ABR

\[
\text{ABR}(t, x_1, x_2, s) = \sum_{p=1}^{N_p} r_p(t, s) \ast u_p(t, x_1, x_2)
\]  

(5)

where \( N \) is the number of different cell populations. It is reasonable to assume that all cells of a given population have the same unitary response, \( u \), as they are all of the same type. Each AN fiber is tuned to a specific CF, where the number of fibers \( N \) is the total number of cells. This implies that the different contributions from the different populations to the scalp potential are assumed to be reflected in the shape of the corresponding individual unitary responses, \( u_p(\cdot) \). Taking advantage of linear super-position, this can be further simplified to:

\[
\text{ABR}(t, x_1, x_2, s) = \sum_{p=1}^{N_p} r_p(t, s) \ast u_p(t, x_1, x_2)
\]  

(6)

where \( R(t, s) \) is the summed discharge rate function, and \( U(\cdot) \) is the unitary response function summed across the different neural populations. Thus the generation mechanism of ABR is determined as the sum of instantaneous firing from all cells, convolved with a unitary response summed over all cell types that is dependent on the electrode location on the scalp but is assumed to be independent of efferent influence and stimulus.

The AN model from Zilany and Bruce (2006) is shown schematically in figure 2. The input to the AN model is the instantaneous pressure waveform of the stimulus in units of Pa. The output of the AN model of Zilany and Bruce (2006) is the spike times in response to the stimulus pressure. The model has a number of key functional stages: a middle ear filter; a feed-forward control path representing the active mechanism; a primary signal-path C1 filter representing the basilar membrane (BM) filtering adapted by the control path; a parallel-path C2 filter for high-level stimuli; an inner-hair cell (IHC) section followed by a synapse model and a stochastic AN spike discharge generator. In the schematic figure 2, the following abbreviations are used: outer hair cell (OHC), low-pass (LP)
filter, static nonlinearity (NL), characteristic frequency (CF) and inverting nonlinearity (INV). $C_{\text{OHC}}$ and $C_{\text{IHC}}$ are scaling constants that indicate OHC and IHC status, respectively. The bold and gray lines in the filter functions represent the tunings at low and high sound pressure levels, respectively. The wideband C2 filter shape is fixed and is the same as the broadest possible C1 filter. The bold and gray lines in the stage following the C1 filter (C1 transduction function) indicate the nonlinearity in the IHC in normal and impaired (scaled down according to $C_{\text{IHC}}$) conditions, respectively.

In the present study, the spikes/s output from the synapse model are used, rather than the stochastic output from the spike generator. Otherwise the simulations would need to be re-run many times to determine the averaged output from the ABR model.

**A humanized AN model**

The Zilany and Bruce (2007, 2006) AN model was fitted to cat AN data, and has thus been modified to better model human AN response here. The following changes to the original cat AN model were implemented by Bruce and co-workers:

\[
K_{\text{CF}} = 0.91 \cdot \min \left\{ 4000, \ 10^{0.1f_c/10^{0.4}} \right\} 
\]

where $x$ is the distance on the BM apex in mm, and the constants are: $A = 165.4$, $a = 0.06$ and $k = 1$.

Two additional changes were made to the Zilany and Bruce (2006) model by the present paper authors. In Zilany and Bruce (2006), the synapse gain, which describes the relationship of the inner hair cell potential to the synaptic release rate, varies as a function of CF to ensure that the model thresholds match empirical data from cats. Without such physiological data available, human behavioural monaural absolute thresholds Killion (1978) were used to fit the model. Thus the synapse gain function from Zilany and Bruce (2006) was changed to be

\[
f_c = A (10^{0.1f_c} - k)
\]

Cochlear tuning is a level-dependent property, where increasing excitation level results in broader BM filters. This is implemented in the original Zilany and Bruce (2006) model by shifting the CF of the so-called control path filter by 1.2 mm on the BM. Without sound knowledge of how this mechanism works in humans, the default is retained here. However, a human frequency-place mapping for the BM is needed and has been updated from the original to the human fit from Greenwood (1990):

\[
f_c = A (10^{0.1f_c} - k)
\]

where $x$ is the distance on the BM apex in mm, and the constants are: $A = 165.4$, $a = 0.06$ and $k = 1$.

Figure 2: Schematic diagram of the auditory-periphery model reprinted from Zilany and Bruce (2007, 2006).

**Figure 2:** Schematic diagram of the auditory-periphery model reprinted from Zilany and Bruce (2007, 2006).

**Figure 3:** Frequency response of the human middle ear implemented in the AN model.

The original cat middle ear transfer function has been replaced by a human middle ear. This was based on the linear circuit model of Pascal et al. (1998) of human cadavers. The model magnitude response function is shown in figure 3.

It has been argued that humans have significantly sharper BM mechanical tuning than cats and other experimental animals (Shera et al., 2002). To incorporate this, the model equivalent rectangular bandwidth quality factor, $Q_{\text{ERB}}$, for cochlear tuning was defined to be:

\[
Q_{\text{ERB}} = 12.7 \left( \frac{f_c}{1000} \right)^{0.3}
\]

where $f_c$ is the center frequency of the BM filter. This function was taken from (Shera et al., 2002) and is applicable to humans at frequencies at and above 1 kHz. The choice of $Q_{\text{ERB}}$ will be further discussed later.

\[
Q_{\text{ERB}} = \frac{Q_{10} - 0.2085}{0.505}
\]

Also shown in figure 4b are the $Q_{\text{ERB}}$ from Shera et al. (2002) used to set the BM tuning in the model.

As described above and shown in figures 4a and 4b, the AN model tuning properties are determined by the frequency dependent $Q_{\text{ERB}}$ in equation 7. However, an additional delay function exists in the primary C1 filter path of the AN model. This acts as a so-called signal-front delay (see Ruggero and Trehub, 2007). This has been altered in the present model, to ensure that the model produces overall delays (signal front and travelling wave group delays) similar to the estimated BM delay reported in Shera et al. (2002).
function was determined, the envelope was extracted (via low-pass filtered hilbert envelope), and the latency of the peak of the enveloped recorded. The following logarithmic function was then fitted to the difference between the model output latencies and those reported in Shera et al. (2002):

$$\tau_{SF} = 10^{-3/3} \cdot \min \{0, -10.09 \cdot \log_{10}(f_c) + 29.23\} \quad (11)$$

By using this additional delay, it is hypothesised that physiologically plausible BM latencies can be approximated in the model. This is vital as it is well known that cochlear processing and delay has a strong influence on recorded brainstem evoked potentials (Dau, 2003; Dau et al., 2000; Wegner and Dau, 2002).

The unitary response

The combined unitary response in equation 6 describes the transformation of the output of the auditory nerve to the potential measured at electrodes placed on the scalp. The unitary response, like in Dau (2003), was obtained by deconvolving an experimentally recorded click ABR with the summed neural activity pattern for the click, generated by the AN model. The deconvolution is an ill-posed mathematical problem and has an infinite number of solutions. A stable and probable solution was found by using Tikhonov regularization (Tikhonov, 1963). The calculations were carried out in Matlab using a toolbox provided by Hansen (1998).

Tone-burst simulation

Auditory evoked potentials have been used historically to obtain indirect estimates of cochlear delay in humans. Tone-burst evoked ABRs have been studied extensively in the literature as a means of estimating BM delay (Gorga et al., 1988; Harte et al., 2009; Neely et al., 1988). Thus, this was a logical choice of basic stimuli to test if the AN model in the present study adequately modelled cochlear delay. In order to test if the BM delay introduced within the present model is reasonable, a simulation was run using hanning windowed tone bursts as stimuli, with CFs and durations given in Table 1. Levels of 40 to 100 dB pe SPL were used, in 10 dB steps.

The choice of stimuli was inspired by the experiments from Norton and Neely (1987) and Serbetçioğlu and Parker (1999). The tone-burst durations represent a trade-off between having an equal number of cycles for all frequencies and a relative narrow spread in their spectrum. The organisation of frequency along the cochlear partition is roughly logarithmic and tone bursts with a fixed number of cycles result in uniform energy splatter in log-frequency. The stimulus rise time is responsible for the simultaneous neural activation leading to the brainstem responses (Suzuki and Horiiuchi, 1981) and to obtain a detectable ABR response. A sharp stimulus onset (i.e., a short rise time) produces a large amount of synchronised neural activity, but also decreases the frequency specificity of the stimulus. Rise times for frequencies of 2 kHz and above include approximately 5 cycles and therefore ranged from 2.5 to 1.25 ms. Below 2 kHz it was felt that the reduced energy spread, by keeping a fixed number of cycles, would make it almost impossible to record a wave-V response. Therefore, a compromise was struck, similar to Gorga et al. (1988), between the need for rapid stimulus onsets and reduced energy spread in the choice of rise time. The number of cycles in the rise time were reduced to 3.25 at 1.5 kHz and approximately 2.5 for 1.0 kHz.

ABR wave V is the wave with the largest amplitude and hence the most easily detectable. In the simulation, the ABRs for the tone burst stimuli were generated and the wave V latency calculated and plotted against Neely et al. (1988)’s empirically determined model of latency derived from tone burst stimulation:

$$\tau_{wave V} = a + b\c^i \left(\frac{f_c}{1000}\right)^{-d} \quad (12)$$

where $i$ is the tone-burst intensity (divided by 100), $f_c$ is the tone burst center frequency in Hertz, and $a = 5$ ms, $b = 12.9$ ms, $c = 5.0$ and $d = 0.413$ were fitted constants to Neely et al. (1988)’s data.
Experimental methods

A total of four normal hearing test subjects (four female) participated in the experimental part of this study, and were aged between 22-26 years. The experiments were conducted in an electrically and acoustically shielded audiometric booth (IEC 268-13). The basic stimulus used in this experiment was a 5 sample duration impulse played at 44.1 kHz. Five sets of stimuli conditions were presented at a constant inter-epoch rate of \( \approx 8 \) Hz (i.e. a duration of 125 ms). The first stimulus set was a single impulse to evoke a standard ABR, used to empirically determine the unitary response functions. The remaining sets were trains of impulses with a within-train rate of 40, 80, 190 and 250 Hz. A total of 4000 averages were made per stimulus type and repeated twice (three times for the single impulse condition) to ensure repeatability of results. The stimuli were all presented at a level of 80 dB pe SPL, to ensure reasonable SNR and test subject comfort.

The stimuli were generated in MATLAB and A/D conversion made through an RME ADI-8 Pro 24-bit sound card. The levels were set via a TDT PA5 programmable attenuator. The stimuli were presented to the left ear of the test subject via an ER-2 insert earphone. EEG activity was recorded differentially between the vertex and ipsilateral mastoid, with the ground electrode placed on the forehead. Silver/silver chloride electrodes were used, and an inter-electrode impedance was maintained below 5 kOmega. EEG activity was recorded on a SynAmps2 amplifier at a sampling rate of 10000 Hz, and band-pass filtered between 0.05 and 2000 Hz. After recording, the EEG-data were epoched and filtered again from 100 to 1500 Hz using a 200 tap FIR filter with zero phase delay. The epochs were averaged using an iterative weighted-averaging algorithm (Riedel et al., 2001).

RESULTS

Auditory brainstem response and unitary response

![Figure 5: Recorded (dotted line) and simulated (solid line) auditory brainstem response to single transient stimuli.](image)

Single transient evoked potentials were averaged across 12,000 epochs (all 3 runs) for subject ML and are shown by the dotted curve in figure 5. The recorded ABR shows the typical pattern with clear waves I, III, and V at latencies that are consistent with the literature. The wave V peak is the largest occurring at \( \approx 6.5 \) ms. Figure 6 shows the calculated unitary response functions. The remaining sets were trains of impulses with a within-train rate of 40, 80, 190 and 250 Hz. A total of 4000 averages were made per stimulus type and repeated twice (three times for the single impulse condition) to ensure repeatability of results. The stimuli were all presented at a level of 80 dB pe SPL, to ensure reasonable SNR and test subject comfort.

In the present study, linear super-position is assumed above the level of the AN, and thus the calculated unitary response function given in figure 6 was used for any input stimulus at any level.

Tone-burst simulation

![Figure 7: Simulated (solid curves) and modelled (dashed curves, based on equation 12) ABR wave V latencies as a function of tone-burst center frequency and level.](image)

Figure 7 shows the wave V latencies for the ABR model simulations to tone-burst stimuli, with center frequencies from 1 to 8 kHz and excitation levels 40 to 100 dB pe SPL in 10 dB steps. Also shown are dotted lines representing the empirically fitted latency model of Neely et al. (1988) given in eqn. 12. Both the simulated ABR and modelled latencies show exponentially decreasing delays as a function of frequency. At the lowest levels of excitation, the simulated ABR latencies have a slope similar to that seen in Neely et al. (1988)’s modelled latencies. This is logical as the AN model tuning and delay was based on Shera et al. (2002)’s stimulus frequency otoacoustic emission delay estimates, made at 40 dB SPL. Further, as excitation
levels increase the simulated ABR rate of change of latency with frequency decreases. The overall spread of simulated ABR latencies with level is reasonable at lower frequencies (1-2 kHz), but seems compressed at higher frequencies relative to Neely et al. (1988)’s results.

Click-train ABR

Figure 8 shows the recorded (dot-dashed curve) and simulated (solid curve) ABR to a single click and click-train stimuli with within-train rates of 40, 80, 190 and 250 Hz for one illustrative subject. The noise floor for the recorded ABR is shown by the vertical bar near 0 ms on each trace. The vertical line to the right of the single click ABR indicates the scale on the figure.

As the within-train rate increases the smaller waves that make up the click ABR (waves I, II, III and IV ) become more difficult to distinguish and only the wave V seems to be visible. As the within-train rate increases, the peak amplitudes of the wave V decrease for rates higher than 80 Hz. The first peaks are typically the largest, and these then decrease as rates increase. The modelled ABR seems to accurately predict the recorded ABR at moderate within-train rates of 40 Hz. Wave V amplitude seems unchanged within trains and latencies seem well modelled. As the within-train rate increases, the modelled ABR amplitude seems to decrease faster than the recorded ABR. In addition, the timing of the peaks of modelled ABR are faster for higher rates than for the recorded potentials at the same rate. For the highest rate stimuli, the simulated ABR wave V peaks drop in magnitude seemingly exponentially for successive stimuli. The recorded ABR on the other hand tends to have a sharp initial drop in magnitude and does not demonstrate such an exponential decrease. Similar trends were observed for all of the subjects tested, though the magnitudes and timing of the responses demonstrated some subject-dependent variability.

DISCUSSION

Frequency-dependent delay

The intrinsic relationship between frequency and travel time in the cochlea is fairly well represented by the AN and the ABR model. Gorga et al. (1988), in the original study on tone-burst evoked ABR wave V latency, did not specify the earphones they used to present the stimuli nor the coupler used to calibrate them. Therefore there is some ambiguity as to the exact levels used by Neely et al. (1988) to model these, and reproduced here in equation 12. With that in mind, one could not expect an exact fit of the present simulated ABR wave V latencies with those modelled by equation 12. The range of latencies across level and frequency, should be covered however. As mentioned earlier, the simulated ABR latencies at higher frequencies seem compressed relative to those seen within the literature. This could be an indication that the level-dependent bandwidth is not well implemented in the AN model. This could not be humanised at present, due to a lack of sensible data to fit the model to. At present the CF of the control path filter is shifted approximately 1.2 mm on the BM. This will be further investigated in future studies.

At low excitation levels, the simulated ABR wave V latencies accurately reproduces the latencies across frequency seen in the literature. The frequency dependent delay in the AN model used here arose due to the cochlear tuning, $Q_{ERB}$, incorporated. This was given in equation 7 and the additional delay in equation 11. There is some contention in the literature about accurate estimates of $Q_{ERB}$ in humans (Bentsen et al., submitted). In the present study, $Q_{ERB}$ estimates from Shera et al. (2002) were used. These $Q_{ERB}$ values were obtained by averaging objective (based on stimulus frequency otoacoustic emission group delay) and behavioural (forward masking) estimates. In these $Q_{ERB}$ values, as seen in figure 4b, the auditory filters are very sharp.
and become effectively sharper as frequency increases. Alternative estimates of $Q_{ERB}$ suggest much broader tuning, and a near frequency independence. These estimates come from objective stimulus frequency otocoustic emission isoo-suppression tuning curves (Keefe et al., 2008), and behavioural simultaneous masking (Glasberg and Moore, 1990).

Ruggero and Temchin (2007) offered an alternative novel estimate of in vivo cochlear delay in humans, using post-mortem delay estimates with post-mortem effects compensated for via comparison with experimental animal data. Bentsen et al. (submitted) showed that Ruggero and Temchin (2007)’s cochlear delay estimates led to $Q_{ERB}$ estimates similar to those obtained with simultaneous masking and stimulus frequency otocoustic emission isoo-suppression tuning curves. If $Q_{ERB}$ were much smaller than those used in the present model (where Ruggero and Temchin (2007)’s were approximately 2.5 times shorter than Shera et al. (2002)), then the latency estimates of the modelled wave V’s seen in figure 7 would be much shorter. Thus a greater degree of disparity would be seen between the modelled and historically reported latencies. This provides some indirect evidence to support Shera et al. (2002)’s estimates of $Q_{ERB}$.

An alternative source of error lies with the unitary response function. In the present ABR model, the only frequency dependent delay is due to the BM filtering in the AN model. It is implicitly assumed that linear super-position holds at higher stages in the model, with the frequency- and level-independent unitary response function. If the unitary response function were to be strongly frequency- or level-dependent, then the wave V latencies simulated in figure 7 would be significantly altered. However, there is good physiological evidence to suggest this is not the case. Wave-V latency is often considered to be composed of the sum of a synaptic delay, $\tau_{synaptic}$, a neural delay, $\tau_{neural}$, as well as the cochlear delay $\tau_{BM}$ (Neely et al., 1988). The synaptic delay is the time between the inner hair-cells activity and the auditory-nerve fibers firing. It is typically around 1 ms (Burkard and Secor, 2002; Kiang, 1975; Kim and Molnar, 1979; Möller and Jannetta, 1983) and frequency- and level-independent (Don et al., 1998). The neural conduction time (neural delay) is the time between the auditory-nerve activity and the place generating the ABR wave. Synaptic delay and cochlear delay are both included in the AN model. However, the neural conduction time is not, and is implicitly in the unitary response function. There is no historical neurophysiological evidence to suggest that the neural conduction time is frequency-dependent (Don and Eggermont, 1978; Don and Kwong, 1980; Eggermont and Don, 1980). However, it would still be prudent to investigate both the frequency and level dependence of the unitary response function in future studies.

Click-train ABR and neural adaptation

The simulated ABR were successful at modeling the recorded ABR for within-train rates of 40 Hz, as seen in figure 8. At these relatively slow rates, little or no neural adaptation was expected. Figure 9a shows the output of the summed AN model in the present study, for the 40 Hz within-train rate stimuli. The model output clearly reverts to baseline (50 spikes/s, AN spontaneous rate) after each click, and the peak of the response for each new stimulus click within the train does not decrease significantly. Thus the stimuli do not interfere with each other within the AN model.

As the within-train rate increases, the ABR wave V tends to dominate the response due to the convolution of smaller peaks and the reduction in amplitude of the spikes in the summed AN model output, as seen in figure 9. For the higher-rate stimuli the summed AN model output never returns to baseline, and the peak magnitudes reduce. The model does not return to baseline due to the ringing of the filters in the AN model. The reduction in the peak spike rates is linked with adaptation and appears to follow an exponential decrease with each new click. Zilany and Bruce (2006)’s rate adaptation at the synapse between IHC and AN fibers was a purely exponential model, albeit with multiple short and long time constants. Zilany et al. (2009) have suggested a new rate adaptation model incorporating both exponential and power-law dynamics. Incorporating this model revision into the present model might help to improve the under-predicted wave V amplitudes at high rates. This will be investigated in future versions of the ABR model.

Figure 9: Summed auditory nerve model output for within-click train rates of (a) 40 Hz, (b) 80 Hz, (c) 190 Hz and (d) 250 Hz.

Outlook

It was stated in the introduction that the role of neural adaptation in AEP recording is important to understand, due to the clinical use of high-rate stimuli. In addition to this, there is a trend in AEP studies to use steady-state signals, where neural adaptation will play an even greater role. Auditory steady state responses (ASSR) are typically responses to carrier signals with amplitude modulation (AM) imposed on them at different rates. Such ASSRs give excellent frequency specificity as the response will mainly contain energy at the AM from a narrow band of AN fibers at the carrier frequency (John and Picton, 2000). This is obviously an advantage clinically to test auditory function at specific frequencies. Invasive animal studies and magnetoencephalographic (MEG) source analysis studies in humans have shown that the ASSR is generated in different brain regions, depending on the modulation frequency of the stimulus (Kuwada et al., 2002; Schoonhoven et al., 2003). For low rates of AM, around 40 Hz, a number of studies have demonstrated that the ASSR can be predicted from the convolution of single middle-latency and brainstem transient responses with a click train with the appropriate repetition rate (Bohórquez and Özdamar, 2008; Galambos et al.; 1981; Gutschalk et al.; 1999; Hari et al.; 1989; Picton et al.; 1987; Plourde et al.; 1991). This is further supported by the finding in the present study, that little or no interaction occurs in the AN model for the different clicks in the 40 Hz click train, as seen in figure 9a.

For modulation rates above 80 Hz, ASSRs are typically argued to be generated by neurons in the brainstem that both respond to transient stimuli and are locked to the envelopes of AM tones (John and Picton, 2000; Kuwada et al., 2002; Sininger and Cone-
Wesson, 2002). The different within-train rates were chosen in the present study to span the AM rates investigated in the literature. The present study has the potential to help understand the brainstem contribution to ASSRs. This is an advantage as sources due to the brainstem are hard to investigate using classic dipole source modeling (Scherg, 1990), due to the brainstem sources depth and small signal strength.

ACKNOWLEDGEMENTS

The authors thank Ian Bruce and co-workers for supplying the first version of the humanised AN model, as well as answering our many questions. The authors also thank Sara Møllenbach and Stine Claessen for their help with collecting the click-train AEP data.

REFERENCES


