Computational Model of the Cochlear Implant User's Auditory System

João Francisco Felizardo[1], Alexandre Bernardino[2] and A. John van Opstal[3]

***Abstract*— A Cochlear Implant (CI) is the only option to treat patients with profound hearing loss, but requires the tuning of a large number of parameters. However, the necessary objective and reproducible patient-specific data are still missing. Building on recent research using reaction times to acoustic spectrotemporal modulations, we developed a realistic simulator of the auditory system, considering the CI, the electrode output, the auditory nerve population, and the ultimate decision of stimulus detection by the brain. The simulation results are in line with measured Reaction-Time tests to spectrotemporal modulations, while also showing some resemblance to actual patient data.**

***Clinical relevance*— The simulator forms the basis for an optimisation method to find optimal CI parameters for each patient, necessitating only a limited number of measurements.**

I. Introduction

According to the "World Report on Hearing" [20], people with Profound and Complete hearing loss (HL) (sensory-neural deafness, due to the nearly total loss of sensory hair cells in the cochlea, >80 dB HL) account for 0.4% of the world's population, approximately 30 million individuals. The only device currently available that can treat severe sensorineural HL is the Cochlear Implant (CI), which uses electric pulsatile stimulation of the auditory nerve through a discrete number of implanted electrode channels (about 10-20) to restore the patient's hearing.

A. Challenges

Effective use of a CI requires the optimal tuning of an extensive list of more than 50 device parameters ("fitting") to achieve improved hearing and speech understanding [3][8]. Among these parameters are the appropriate selection and mapping of frequency channels, adjustments due to implantation problems or different patient characteristics, or the input dynamic range for each channel, among others. This large array of parameters makes it difficult for manual optimisation, which is, however, still the case in current clinical practice. As described in [3][5][19], manual fitting procedures are based mainly on subjective tests, such as speech perception scores and questionnaires. These tests are prone to being affected by cognitive factors unrelated to the basic parameters of the user's performance. This often leads to noisy and irreproducible data, failing to reliably describe the actual auditory processing quality of the individual CI user. In the absence of more objective and reproducible perceptual data, clinicians typically resort to a 'one-size-fits-all' approach. This prevents a highly needed patient-specific approach to the fitting, which is necessary to address the significant variability in auditory perception in CI users [3].

[1]Instituto Superior Técnico, Institute for Systems and Robotics, Lisbon, Portugal j.francisco.felizardo@tecnico.ulisboa.pt

[2]Instituto Superior Técnico, Institute for Systems and Robotics, Lisbon, Portugal alex@isr.tecnico.ulisboa.pt

[3]Donders Centre for Neuroscience, Radboud University, Nijmegen, the Netherlands john.vanopstal@donders.ru.nl

B. Aim and Objectives

This work aims to develop a realistic model and system of the auditory process of individual CI users. We model each step of the hearing process, from the CI, to the electrode output, to the activity of the auditory nerve population, and ultimately to the decision by the brain to trigger a perceptual response. We use a simple reaction-time (or decision detection) test, which has been shown to yield reliable, reproducible, and sound-specific results, and to be sensitive to subtle acoustic manipulations.

II. Background

As described in [2], the cochlea is the inner ear structure that transforms the acoustic sound signal into activation of the Auditory Nerve (AN). The AN is activated by approximately 3500 cochlear inner hair cells, located on the $\sim$ 3.5 cm long basilar membrane (BM), which has tonotopic properties: each BM location, x_{BM}, is tuned to frequency, f, as a monotonic logarithmic function, $x_{BM}(f)$ with high frequencies at its base, $x_{BM} = 0$, and low frequencies at the apex, $x_{BM} = 3.5$ cm. The neurons of the AN follow the BM's tonotopy and can be roughly divided into three classes according to their firing rate behaviour to input strength: (i) low threshold, high spontaneous activity, small dynamic range, (ii) medium threshold, spontaneous activity and dynamic range, (iii) high threshold, low spontaneous activity, large dynamic range.

A. Cochlear Implant

Detailed in [21], the CI contains an external device with a microphone, sound processor, and an internal processor, surgically placed inside the head, and connected to an electrode array that is inserted into the cochlea. The CI utilises the BM tonotopy, as the different electrode contacts along the length of the BM provide direct frequency-specific electrical stimulation to the AN fibres. The external part of the CI receives the acoustic signal, $S(t)$, and performs a parametric transformation on the signal, leading to discretised stimulation pulse trains of the 10-20 CI electrode contacts. This transformation consists of using bandpass filters according to a Frequency Allocation Table (FAT), followed by an envelope extraction transformation and pulse-amplitude modulation at instantaneous envelope magnitudes within a specified dynamic range.

B. Reaction-Time test

The main problem with the currently used subjective tests is that cognitive factors unrelated to the user's basic auditory perception affect their results. Reaction times (RTs), on the other hand, are a reliable output measure of the perceptual (sensory-to-motor) system and have been demonstrated to be sensitive to subtle auditory manipulations. For example, RTs to a 5 ms sound burst are systematically longer than for a 3 ms sound burst [17]. In addition, RTs for sounds presented monaurally are longer than for binaurally presented sounds [13]. Furthermore, reaction times are stable over time, and measuring them does not require extensive learning or advanced cognitive skills (children, and even monkeys [18] can do the task). By measuring the RTs to a variety of spectral-temporal (ST) rippled stimuli, the listener's spectral-temporal sensitivity profile can be measured [4][6][18][13], which quantifies the patient's basic bottom-up auditory system capacities. ST ripples are composed of sinusoidal modulations of a broadband noise carrier in amplitude over time and/or frequency and become manifest as an ST envelope modulation of duration D and bandwidth Δf [4][6]. The envelope is given by the following parameterisation:

$$R(\omega,\Omega) = 1 + \Delta M \cdot \cos(2\pi(\Omega \cdot x \pm \omega \cdot t)), \qquad (1)$$

where ω (in Hz) is the temporal velocity, Ω (in cycles/octave) is the spectral density, $\Delta M \in [0,1]$ is the envelope's amplitude modulation depth, x (octaves above the lowest frequency) is the spectral position and t (s) is time after ripple onset. The Linear Approach to Threshold with Ergodic Rate (LATER) model [9] provides a simple (only four parameters) neuro-computational mechanism to explain RT distributions to sensory stimuli. The model considers a Gaussian stochastic neural process of an accumulating stimulus evidence variable, $r \in N[r_0, \sigma_0]$, that triggers a reaction as soon as it reaches a (fixed) decision threshold. The model's output can be conveniently represented in so-called reciprobit plots, in which the cumulative probability (on probit scale) of the inverse of -RT (negative promptness) is predicted to follow a straight line, as illustrated in Fig. 1. The resulting data, associated with different stimulus conditions, allow for several inferences: changes in the mean neural evidence rate, r_0, move the reciprobit lines laterally (the more to the right, the harder to detect the stimulus), while changes in the neural noise, described by σ_0 alter the slopes, with the higher σ_0, the shallower the slope.

For these reasons, a reaction-time task for CI users could be instrumental to measure their spectrotemporal sensitivity and to create a patient-specific model that can learn to improve their CI fitting. The first promising results of RT measurements in CI patients are now becoming available in the literature [12].

C. Auditory Models in the Literature

There exist several realistic auditory simulators, as collected in the Auditory Modelling Toolbox [14]. However, these models do not refer to the characteristics associated with CI users, and do not readily facilitate the scaling and practical parameterisation needed for this work. Other models have studied the CI and its effect on the auditory system. Some simulate and study a particular part of the auditory pipeline, but miss the relation to the percept, or to the fitting problem [8]. The model of [22] obtains estimates of parameters related to neuronal health and current spread, with comparison to patient data. Others focus on using Machine Learning (ML) to predict speech understanding test scores, based on patient characteristics and test results, without a physiological model of the auditory process [15]. Prediction of fitting parameters and an ML-aided fitting procedure is used in the FOX algorithm (Fitting to Outcome eXpert [16]). FOX-assisted fitting has achieved comparable results to manual fitting [11], which shows the potential of automatised, machine-aided procedures in a clinical setting. However, results are still limited, possibly related to the use of subjective, cognitive tests as an evaluation and predictive measure in FOX. Our model connects auditory modelling, CI fitting and objective perception measures (RT tests).

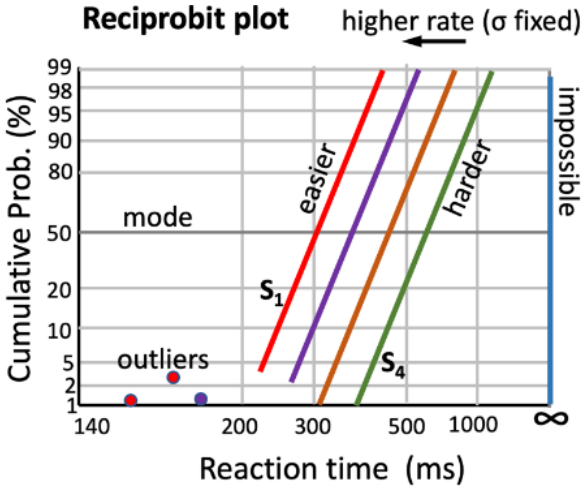

Fig. 1: Cumulative response probabilities on probit scale for response promptness (as -1/RT). Slopes relate to σ_0, and the mode (at 50%) relates to r_0. Outliers are easily identified, and at 0 promptness, a reciprobit line (blue) is shown as an example for non-responses. Image from [3].

III. Auditory Simulator

The basic outline of the auditory simulator is shown in Figure 2. The proposed simulator comprises four key modules: CI processing, electrode interaction, neuronal response, and perceptual decision-making. Below, we describe each module's function. Further implementation details can be found in [7].

A. CI Module

We follow, with some adaptations, the CI processing strategy of Advanced Bionics Corporation [1].

1) Filterbank: The spectral power of the input sound is distributed across the different frequency channels according to a Frequency Allocation Table (FAT). The electrode position is initialised in this block, where the electrodes are set with an equal distance from each other, and their equivalent characteristic frequency is obtained using the inverse of the

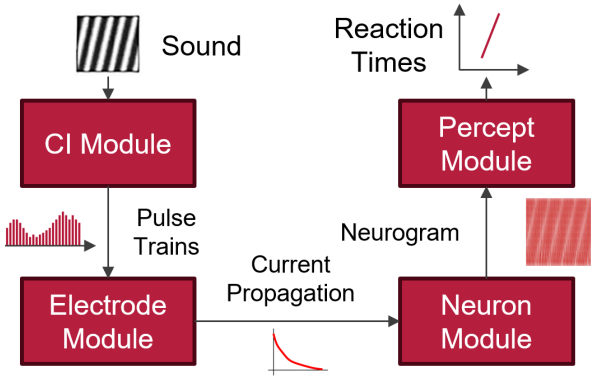

Fig. 2: The four processing modules that are incorporated in the Auditory Simulator. Input is a spectrotemporal rippled sound; the model output is a reaction time to the stimulus.

tonotopic logarithmic mapping function, $x_{BM}(f)$, yielding:

$$f = f_{max}\left(\frac{1}{2}\right)^{x_{BM}/\lambda} \tag{2}$$

where the maximum frequency, $f_{max} \approx 8.5$ kHz, and $\lambda = 0.4$ cm. In our implementation, the FAT contains the center frequencies corresponding to the 10 electrode positions and a Gaussian bandwidth (with $\sigma = 10/44$ octaves for the standard setting) to weigh the spectral power. A third-order band-pass filter was created for each frequency channel, allowing full coverage of the relevant frequency domain. Then, a Hilbert Transform is applied to the filtered signals at the output of the filterbank, which then pass through a Sound Pressure Level (SPL) transformation, yielding the instantaneous SPL-scaled channel levels, L (in dB).

2) Current mapping: The mapping of the channel levels (in dB) to current intensity (in current units, CU) on each channel is provided by

$$I_e = \frac{(M-T)}{IDR}(L - S + IDR + G) + T. \tag{3}$$

Here, I_e represents the current to be sent to the electrodes in CU (1 CU = 10 μA) and it is clipped to the levels 0 and M. L is the filterbank output for each channel, M and T are the Maximum and Threshold levels of current stimulation (in CU), respectively, IDR represents the input dynamic range (in dB), G is an added logarithmic gain (in dB), and S is the saturation level (in dB).

3) Pulse Generation: Subsequently, the current values obtained at each time step are used for pulse-train generation. Pulses have a biphasic shape, an anodic (positive) phase immediately followed by a cathodic (negative) phase, a pulse width of $50-80\mu$s, and a pulse rate between 1000 and 1300 Hz. Interleaved stimulation is applied to minimise electric-field interactions. Electrodes are stimulated in a predefined order where the distance between consecutively stimulated electrodes is maximised. Tables I and II provide the CI parameters for the default simulation.

B. Electrode Module

This module describes the spatial spread and attenuation of the current generated by each electrode, as it propagates

TABLE I: Frequency Allocation Table used for a 10-electrode CI.

Electrode	1	2	3	4	5	6	7	8	9	10
Center Frequency (Hz)	442	603	824	1126	1538	2101	2870	3920	5355	7316
Bandwidth	10/44	10/44	10/44	10/44	10/44	10/44	10/44	10/44	10/44	10/44
Frequency Interval (Hz)	377 - 518	515 - 706	703 - 965	961 - 1319	1313 - 1801	1794 - 2460	2451 - 3360	3348 - 4589	4574 - 6269	6249 - 8565
Stimulation order	1	4	7	10	2	5	8	3	6	9

TABLE II: Typical parameter values used in the simulations.

Pulse Rate (Hz)	Pulse width (μs)	M (CU)	T (CU)	IDR (dB)	Gain (dB)	Sat (dB)
1000	80	100	10	60	12	90

to activate the AN fibres. We simplified the cochlea as a stretched linear tube with the electrode array inside, where the neurons and electrodes are placed as shown in Fig. 3.

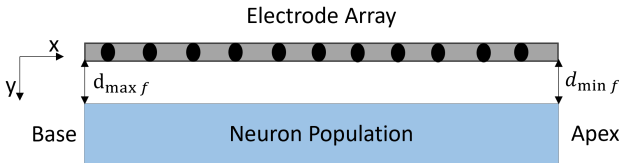

Fig. 3: Schematic of the Electrode Module.

A neuron's position (x_i, y_i) is specified relative to the electrode array, where the neuron closest to the maximum frequency (at x_{BM} = 0 mm) has position coordinates $x_1 = 0, y_1 = d_{maxf}$ mm, and the neuron at the minimum frequency, $x_{N_e} = x_{max}, y_{N_e} = d_{minf}$ mm. Neurons are homogeneously distributed along x within the neuron population, and logarithmically associated with frequency through $x_{BM}(f)$.

While current travels through the medium inside and outside the cochlea, its intensity decays with distance. The current-spread function is assumed to have an exponential decay $e^{-d_j/\lambda}$, with d the distance from electrode contact j and λ (in mm) the spatial current attenuation constant, as in [8]. Thus, the total current from the electrode array, indexed by j, at neuron i is described by

$$I_i(t) = \sum_{j=1}^{N_e} S_{e_j} \cdot I_{e_j}(t) \cdot e^{\frac{-d_{i,j}}{\lambda}} \tag{4}$$

where $S_{e_j} \in [0,1]$ is a scaling parameter that represents the integrity of electrode j, I_{e_j} is the current from electrode contact j, and $d_{i,j}$ is the distance between electrode j and neuron i. The attenuation coefficient of $\lambda = 0.7238$ mm was chosen to provide -12 dB/mm decay and a realistic current spread in the model.

C. Neuron Module

This module calculates the spiking activity of each neuron in the auditory nerve population in response to the pulsatile inputs of (4). To calculate the instantaneous spiking activity of a neuron, we implemented an adapted version of the original membrane equations of the Izhikevich model [10]. The discretized neural membrane potential ($v[mV]$, at 10 μs time steps) is given by

$$\begin{cases} v_{t+1} = v_t + \frac{1}{100}(0.04v_t^2 + 4.1v_t + 108 - u_t + v_{I_t} + I_{bg_t}) \\ u_{t+1} = u_t + \frac{1}{100}\,a\,(bv_t - u_t). \end{cases} \tag{5}$$

along with after-spike reset

$$\begin{aligned} &\text{if } v_{t+1} \geq 30\ mV: \\ &\text{then } v_{t+1} = c \text{ and } u_{t+1} = u_{t+1} + d. \end{aligned} \tag{6}$$

In these equations, u is a dimensionless membrane recovery variable and index t is discrete time in 10 μs. a, b, c, and d are neuron parameters that can be tuned to simulate different firing behaviours. Further, I_{bg} (background current) and v_I (current stimulation) are the external inputs to the neurons, described by $I_{bg} = c_{bg}|\mathcal{N}(0,1)|$ and $v_{I_t} = R(I_t)$, with c_{bg} a scaling factor for the Gaussian ($\mathcal{N}[0,1]$) background, and $R(I)$ a nonlinear resistance function that transforms the neuron's current input into membrane voltage (mV) and realistic firing rates. By exploring the parameters of the Izhikevich model, along with different resistance functions (Table III), we generated the three neuronal classes of the auditory nerve [2] with their distinct firing behaviours as a function of current strength I, shown in Fig. 4. The simulated neuron population base characteristics are shown in Table IV.

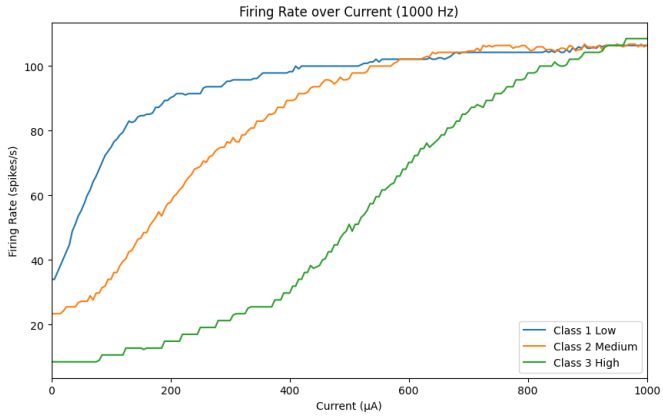

Fig. 4: Firing rate as a function of current strength, with cathodic-anodic pulses (duration 50μs), with the parameters of Table III.

TABLE III: Neuron Class parameters and resistance types.

	a	b	c	d	R(I) (mV)	Threshold
Class 1	0.1	-0.1	-60	6	$1500\frac{2}{\pi}\arctan(\frac{I}{0.05\times10^{-3}})$	Low
Class 2	0.025	-0.1	-60	2	$3300(\frac{1}{1+exp(I/0.135\times10^{-3})}-0.5)$	Medium
Class 3	0.02	-0.1	-60	10	$4800(\frac{1}{1+exp(I/0.2\times10^{-3})}-0.5)$	High

TABLE IV: Default neuronal population characteristics.

Groups	Neurons per Class	f_{min} (Hz)	f_{max} (Hz)	d_{min} (mm)	d_{max} (mm)
3	1001	350	9000	1.0	1.0

Neuronal health: CI users will have different conditions that affect their auditory nerve activity, some of them possibly related to the aetiology of the HL, and others linked to the CI implantation. To provide flexibility of the neuronal parameters, different neuronal health conditions are implemented: (i) varying distance between electrode array and auditory nerve, which impacts the current spread, and consequently the neuronal activity; (ii) varying number of neurons per class; (iii) extra neuronal noise N, by adding $N[i] = n_f\alpha[i]M$ to the membrane potential calculations in (5), where i represents each neuron; n_f (relative noise factor) is a scalar between 0 and 1; α is a vector with a length corresponding to the number of neurons, where each element of the vector is drawn at each time step from a normal distribution; M is the maximum noise gain, set experimentally at 500 mV; (iv) damaged neuronal groups, where the neurons in a certain frequency band provide no spiking activity.

D. Percept Module

The final module of the simulator calculates a dynamic decision signal from the neurogram (collective activation map of all neurons over time), which will lead to the reaction time. The onset of the acoustic spectrotemporal ripple triggers the response of the perceptual module. This response is quantified as an instantaneous monotonically increasing decision signal. A straightforward method to make this association does not yet exist in the literature. The ripple stimuli comprise an initial non-modulated broadband noise (of 1 s), followed by a 1 s envelope modulation over time and/or frequency of that same noise as in (1).

1) The decision signal: The decision signal was created from a continuous comparison of the statistical properties of the neurogram along the sliding window concerning a base window, with the static noise response at the start of the stimulus. The number of spikes in the base window (of duration Δt_B ms) is calculated at t_{start} for N_b neuron bands (each containing n_b neurons), with $N_b = N_{tot}/n_b$ and $N_{tot} = 3003$ the total number of neurons in the population. The sliding window (of duration Δt_S) advances until the end of the neurogram with increments Δt_i. This results in an $N_b \times 1$ vector $\mathbf{B}$ for the spike counts in the base window, and an $N_b \times N_S$ matrix for the number of spikes in the neuron bands for $N_S = 2/\Delta t_i$ iterations of the sliding window, $\mathbf{S}(t)$.

From this spike-count information, we constructed the instantaneous decision value by

$$d(t) = a_d|\Delta\mu(t)| + b_d|\Delta\sigma(t)| + c_d|\Delta\beta(t)|. \tag{7}$$

where a_d, b_d and c_d are weights, and $\Delta\mu(t)$, $\Delta\sigma(t)$ and $\Delta\beta(t)$ are differences in the statistics of the neurogram between the base window's vector and the sliding window's vector at bin t. $\Delta\mu(t)$ compares the mean firing rates in the windows and is given by

$$\Delta\mu(t) = \mu_B - \mu_S(t) = \frac{1}{N_b\Delta t_B}\sum_{i=1}^{N_b} b_i - \frac{1}{N_b\Delta t_S}\sum_{i=1}^{N_b} s_{i,t} \tag{8}$$

where b_i and $s_{i,t}$ are the spike counts in base window B and sliding window S at time t. It is most sensitive to purely temporal modulations, as they affect the whole window at the same time.

The second statistic compares the standard deviations within the neuron band firing rates, and is described by

$$\Delta\sigma(t) = \sigma\left(\frac{B}{\Delta t_B}\right) - \sigma\left(\frac{S(t)}{\Delta t_S}\right) \tag{9}$$

with $\sigma(\cdot)$ the standard deviation function. This compares the spread of neural activation within the two windows and should be more sensitive to spectrotemporal modulations.

The third statistic quantifies the relative band contributions:

$$\Delta\beta(t) = \sum_{i=1}^{N_b} (|\beta_{B_i} - \beta_{S_i}(t)|) \tag{10}$$

where β_{B_i} and $\beta_{S_i}(t)$ are the entries of contribution vectors β_B and $\beta_S(t)$, given by:

$$\begin{cases} \beta_B = \text{softmax}\left(\frac{100B}{\Delta t_B \mu_B}\right) \\ \beta_{S(t)} = \text{softmax}\left(\frac{100S(t)}{\Delta t_S \mu_S(t)}\right) \end{cases} \text{with } \text{softmax}(\mathbf{z})_i = \frac{e^{z_i}}{\sum^N e^{z_j}}. \tag{11}$$

The contribution values are constrained to $[0,1]$ due to the softmax operation. It also polarises the results by assigning extra importance to the higher values in the vector and less to the lower values, highlighting small differences in the contribution vectors. This is done with the multiplication by 100, which is equivalent to choosing a softmax temperature of 0.01, decreasing the entropy of the distribution. This statistic is meant to respond more to purely spectral modulations.

2) Percept Analysis: By calculating the ongoing decision signal during a neurogram sliding window, we can follow the contributions of the different factors to the total decision signal. An example is shown in Fig. 5, with the parameters for the decision function listed in Table V.

TABLE V: Default decision signal parameters.

Δt_B (ms)	Δt_S (ms)	t_{start} (ms)	Δt_i (ms)	N_b	a_d	b_d	c_d
100	50	100	10	30	1	1	30

Note that after ripple onset at t = 1 s, $\Delta\mu$ (green) and $\Delta\sigma$ (blue) both increase noticeably faster than for the non-modulated sound ($t < 1.0$ s). The total decision signal (red) serves as a direct measure for response promptness (=RT^{-1}) to reach the decision threshold. To calculate it, we set a fixed threshold ($d_\Theta = 1950$) for the decision signal value. The reaction time for this standard simulation is then calculated, obtaining a RT = 417 ms. At the end of the stimulus (2.0 s), the decision signal is $d_{2.0} \approx 3545$.

IV. Results

Some representative results of our simulations are presented in Figures 6 to 10. Further experiments, respective results, and discussion can be found in [7].

A. Module Sensitivity

Fig. 6 shows the neurogram for a purely temporal ripple and the associated decision signal. The neuronal response reaches a much higher decision value ($d_{2.0} \approx 9000$) than the ripple of Fig.5A ($d_{2.0} \approx 3545$), which demonstrates the sensitivity of the simulator for different spectrotemporal modulations. Due to the faster rising decision signal, the temporal ripple is easier to detect, leading to RT = 148 ms.

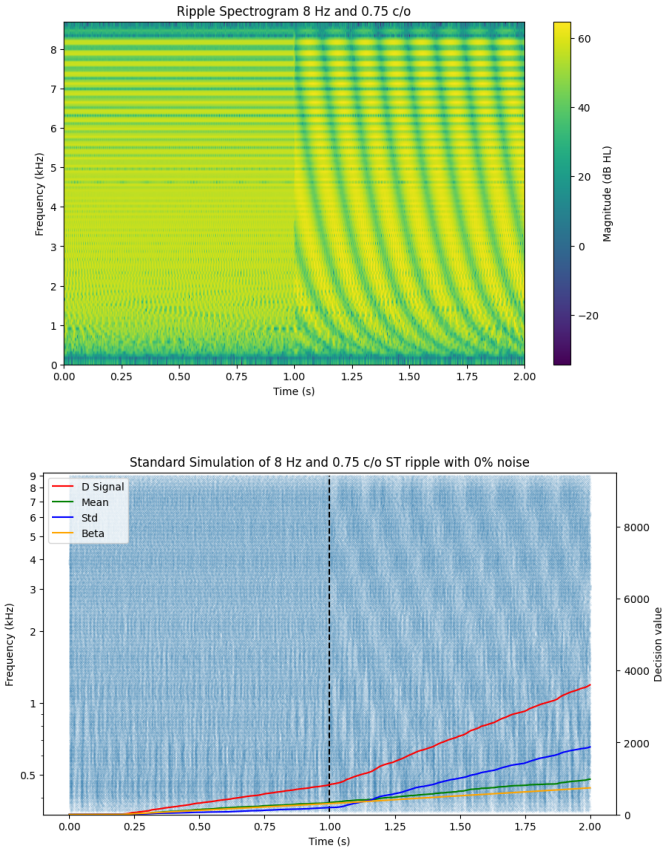

Fig. 5: Standard Reaction Test Simulation. (a) Spectrogram of the standard test ripple, with $\omega = 8$ Hz, $\Omega = 0.75$ c/o and $\Delta M = 0.8$. (b) Neurogram of the population (Table IV), and the associated decision signal (red) and its constituents (legend) for this standard test.

B. CI parameter changes

Our simulator will be instrumental to the training of automated models for CI tuning. All parameters of the CI can be modified and their effect on the reaction times analysed. For instance, Fig. 7 shows the result of reducing the IDR by 10 dB, from 60 dB to 50 dB, with respect to the results of Fig. 5. This change in the fitting led to a different neurogram that reached a final decision value $d_{2.0} = 4581$ and an RT = 296 ms for the same stimulus. This result nicely illustrates that differences in the CI fitting lead to sensible neural and perceptual outcomes.

C. RT Distributions

Because of the neural noise, an identical stimulus will generate different neural responses, which could lead to different reaction times. By repeating the same stimulus (here 20 times), we generated a distribution of reaction times to that stimulus, which was transformed into a reciprobit plot (as in Fig. 1). It is interesting to investigate whether the model produces different reciprobit plots for different ST ripples, and whether these resemble results from real subjects. Figure 8 shows the RT distributions on reciprobit scale (cf. Fig. 1) for nine different ST ripples: 3 purely spectral (0.25, 0.5, 0.75 c/o, in blue shading), 3 purely

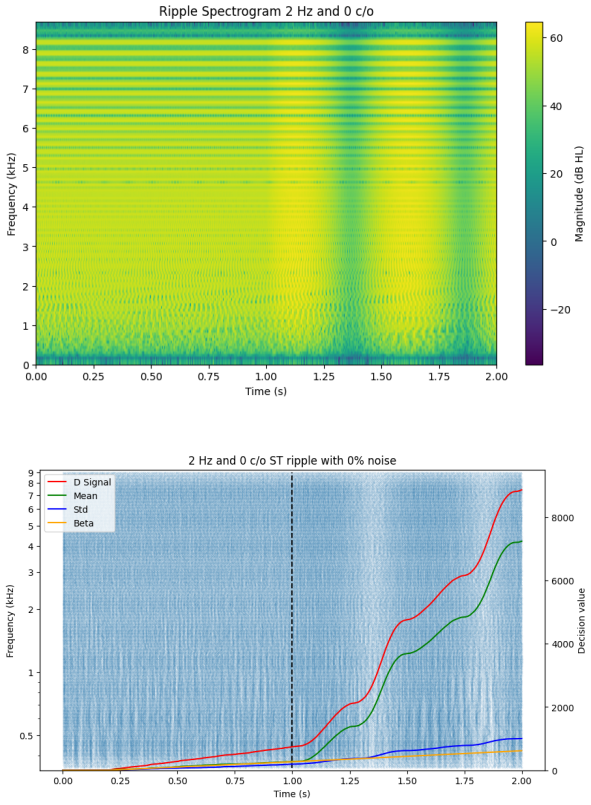

Fig. 6: Model's response to a temporal ripple. (a) Spectrogram of the ripple: $\omega = 2$ Hz, $\Omega = 0$ c/o. (b) Associated neurogram and decision signals.

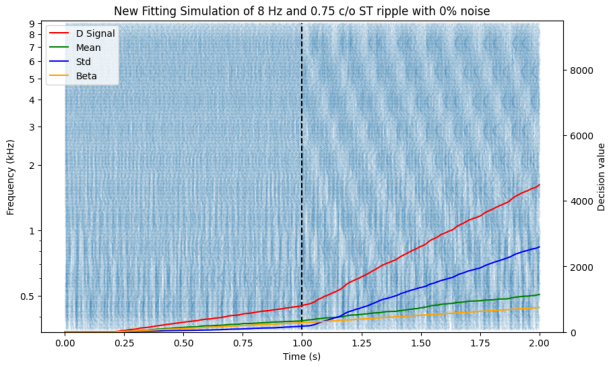

Fig. 7: Decision signals and neurogram for the ω =8 Hz and Ω =0.75 c/o ripple, for a different CI fitting (IDR = 50 dB; cf. Fig.5).

temporal (2, 8, 16 Hz, in red shading) and 3 spectrotemporal ripples (in green shading). All simulations had a 20% noise factor ($n_f = 0.2$; see Neuronal Health section). The RTs for each stimulus presentation were determined by setting a fixed threshold, $d_{RT,\Theta}$ = 1950, to the decision signal. We took the values of the decision signal at the onset and the end of the ripple, and linearly interpolated these points to determine the time to reach threshold, as RTs could exceed ripple duration.

D. Effect of Modulation Depth

Figure 9 displays the reciprobit RT distributions for three ripples at five different modulation depths, ΔM: a purely spectral (0.5 c/o, in blue), a purely temporal (8 Hz, in red),

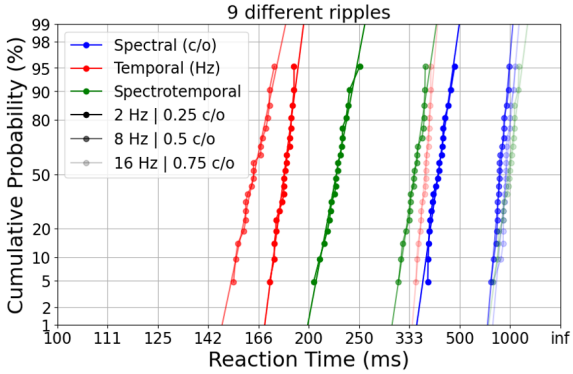

Fig. 8: RT distributions of 9 ripples on reciprobit scale. Abscissa: negative promptness: $P = -1/RT$, with fast reactions on the left (e.g., RT = 250 ms: P = -4.0; RT = ∞: P = 0). Ordinate: cumulative probability of promptness on probit scale. In this format, a Gaussian promptness distribution results in a straight line with positive slope. Light colours correspond to the 16 Hz and 0.75 c/o stimuli, which yield the slowest responses. Note that purely temporal ripples yield much faster reactions than pure spectral ripples.

and a spectrotemporal (4 Hz, 0.25 c/o, in green) ripple, at $\Delta M \in$ [100%, 80%, 60%, 40%, 20%] (dark to light). All simulations were performed with standard parameters, using the same criterion for the RT as above ($d_{RT,\Theta}$ = 1950), with noise factor, $n_f = 0.2$.

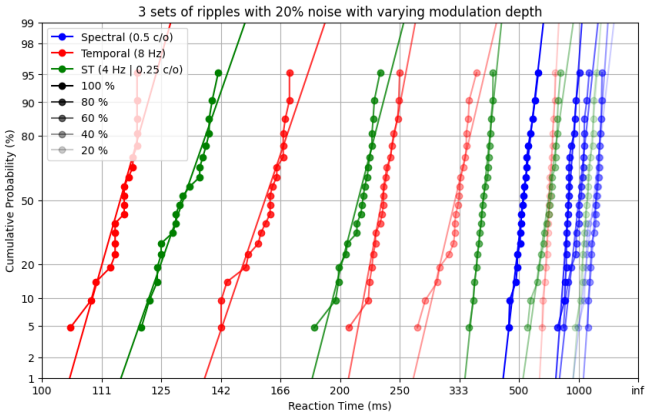

Fig. 9: Reaction Time distributions for 3 sets of ripples with changing modulation depth. Light colours equal lower modulation depths. The highest modulation depths yield the shortest RTs, and purely temporal ripples lead to much faster reactions than purely spectral ones.

E. Spectrotemporal Modulation Sensitivity

To assess the auditory performance of the model, we tested its response to a set of 72 different ripples, close to the range of modulations that are relevant for speech. By presenting different ripples to the model, it is possible to study how the auditory nerve neurons react to the CI stimulation strategy, and to construct a spectral-temporal modulation sensitivity plot [4][6]. Figure 10 shows the final values of the decision signal (at 2 s) in colour scale for the different ripples, averaged over 10 simulation runs per ripple, and standard model parameters. Note that the strongest sensitivity of the model (i.e., fastest reaction times) is obtained for low spectral densities ($\Omega < 0.75$ c/o) and for temporal modulations ($\omega \in [1-10]$ Hz).

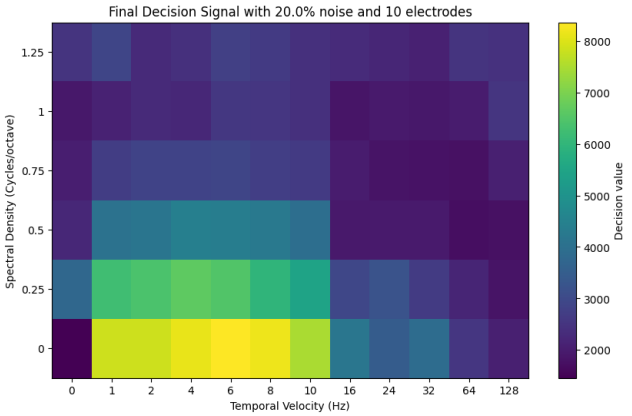

Fig. 10: Averaged result (10 runs) of the decision signal for a set of 72 different spectrotemporal ripples and a 20% noise factor. CI parameters from Tables I and II.

V. Discussion

In this work, a cochlear implant auditory simulator was built and tested with different patient characteristics. Table VI provides a summary of the current simulator's patient-specific properties and the tunable CI parameters. Our results (Figs. 8-10) show encouraging resemblance to what has been reported in the literature on actual patient data[4][13].

TABLE VI: Simulated CI and Patient Parameters.

Patient		Tuneable	
Electrode Module	Electrode position	**CI Module**	M,T-Levels
	Electrode number		IDR
	Electrode strength		Saturation
Neuron Module	Neuron position		Gain
	Neuron Classes		Stimulation Order
	Noise factor		FAT (# electrodes)
	Damaged nerves		Pulse width, rate
Percept Module	Decision Parameters		Pulse shape
	Decision Threshold		Sensitivity

In future work, the simulator can be used to create an ML algorithm that learns to adapt the parameters of Table VI to improve spectral-temporal performance of the modelled CI user (i.e., as an RT reduction to speech-relevant ripples). If the model outputs match the CI-user's RT distributions, the patient-specific parameters can be used to simulate (and predict) the auditory perception of the CI user to other sounds, including speech signals.

Furthermore, the model can be extended to also include additional CI-user experiments that have been reported in the literature (e.g., electrode cross-impedances, RTs to changes in direct electrode stimulation patterns), which are conducted to gain objective insights into the capacities of the CI user's auditory system, by bypassing the sound processor[3]. These experiments could be useful to adapt the model's patient characteristics to fit the performance of an actual CI user.

In this pilot work, the full scope of possibilities regarding patient characteristics and/or effects of CI parameters has not yet been explored. Moreover, several improvements can be made to the simulator, e.g. at the level of detail of the AN neurons, or to the different components of the decision function. Due to the model's flexibility, it can eventually be combined with more refined auditory-system models (e.g., from [14]) that could potentially lead to further improvements of the simulator regarding its biological realism.

Acknowledgements

Supported by NWO - TTW grant 204141 'Otocontrol-2.0' to AJVO, JFF; and by the Portuguese Recovery and Resilience Plan (RRP), through project number 62, Center for Responsible AI, LARSyS FCT funding (DOI: 10.54499/LA/P/0083/2020, 10.54499/UIDP/50009/2020, 10.54499/UIDB/50009/2020) and FCT HAVATAR project (10.54499/PTDC/EEI-ROB/1155/2020), to AB, JFF.

References

[1] Advanced Bionics. Intensity Coding, 2018.
[2] AJ Van Opstal. The Auditory System and Human Sound-Localization Behavior. Elsevier, Academic Press, pages 113–169, 2016.
[3] AJ Van Opstal and E Noordanus. Towards personalized and optimized fitting of cochlear implants. Frontiers in Neuroscience, 17, 2023.
[4] T Chi, Y Gao, MC Guyton, P Ru, and S Shamma. Spectro-temporal modulation transfer functions and speech intelligibility. The Journal of the Acoustical Society of America, 106(5), 1999.
[5] DB Pisoni, WG Kronenberger, MS Harris and AC Moberly. Three challenges for future research on cochlear implants. World J Otorh Head-Neck Surgery, 2018.
[6] TM Elliott and FE Theunissen. The modulation transfer function for speech intelligibility. PLoS computational biology, 5(3), 2009.
[7] JF Felizardo. Computational auditory process simulator for optimising the parameters of cochlear implants. Master's thesis, IST Lisbon, 2024.
[8] H Yang, JH Won, I Choi, J Woo. A computational study to model the effect of electrode-to-auditory nerve fiber distance on spectral resolution in cochlear implant. PLOS One, 2020.
[9] I Noorani, RHS Carpenter. The LATER model of reaction time and decision. Neuroscience and Biobehavioral Reviews, 64:229–251, 2016.
[10] EM Izhikevich. Simple model of spiking neurons. IEEE Transactions on neural networks, 14(6):1569–1572, 2003.
[11] J Wathour, PJ Govaerts, L Derue, S Vanderbemden, H Huaux, et al. Prospective Comparison Between Manual and Computer-Assisted (FOX) Cochlear Implant Fitting in Newly Implanted Patients. Ear and Hearing, 44, 2023.
[12] LC Veugen. Bimodal Stimulation - towards Binaural Integration. PhD Thesis, Radboud University - Donders series nr. 287. ISBN 978-94-6284-121-5, 2017.
[13] LCE Veugen, AJ van Opstal, MM van Wanrooij. Reaction Time Sensitivity to Spectrotemporal Modulations of Sound. Trends in Hearing, 2022.
[14] P Majdak, C Hollomey, and R Baumgartner. Amt 1.x: A toolbox for reproducible research in auditory modeling. Acta Acust., 6:19, 2022.
[15] MG Crowson, P Dixon, R Mahmood, JW Lee, D Shipp, et al. Predicting Postoperative Cochlear Implant Performance Using Supervised Machine Learning. Otology and Neurotology, 41(8), 2020.
[16] PJ Govaerts, B Vaerenberg, G De Ceulaer, K Daemers, C De Beukelaer, K Schauwers. Development of a software tool using deterministic logic for the optimization of cochlear implant processor programming. Otology and Neurotology, 31(6), 2010.
[17] PM Hofman and AJ Van Opstal. Spectro-temporal factors in two-dimensional human sound localization. J Acoust Soc America, 103(5), 1998.
[18] RF van der Willigen, H Versnel, AJ van Opstal. Spectral-temporal processing of naturalistic sounds in monkeys and humans. Journal of Neurophysiology, 2024.
[19] RP Carlyon and T Goehring. Cochlear Implant Research and Development in the Twenty-first Century: A Critical Update. JARO, 2021.
[20] The World Heath Organization. World Report on Hearing, 2021.
[21] B Vaerenberg. Programming Cochlear Implants for Auditory Performance. PhD thesis, The University of Antwerp, Belgium, 2014.
[22] X Xia, DX Gao, T Brochier, and DB Grayden. Estimating user-specific current spread and neural health parameters in a model of hearing with cochlear implants. In IEEE Engineering in Medicine and Biology Society (EMBC), pages 1–4. IEEE, 2024.