

Spectral Ripples in Normal and Electric Hearing Models

Martens, Savine S.M.; Briaire, Jeroen J.; Frijns, Johan H.M.

DOI

[10.3390/technologies13110505](https://doi.org/10.3390/technologies13110505)

Publication date

2025

Document Version

Final published version

Published in

Technologies

Citation (APA)

Martens, S. S. M., Briaire, J. J., & Frijns, J. H. M. (2025). Spectral Ripples in Normal and Electric Hearing Models. *Technologies*, 13(11), Article 505. <https://doi.org/10.3390/technologies13110505>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

Article

Spectral Ripples in Normal and Electric Hearing Models

Savine S. M. Martens ¹, Jeroen J. Briaire ¹ and Johan H. M. Frijns ^{1,2,3,*}

¹ Department of Otorhinolaryngology, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands; j.j.briaire@lumc.nl (J.J.B.)

² Leiden Institute for Brain and Cognition, Leiden University, P.O. Box 9600, 2300 RC Leiden, The Netherlands

³ Department of Bioelectronics, EEMCS, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands

* Correspondence: j.h.m.frijns@lumc.nl

Abstract

Devising a psychophysical test to assess spectral resolution has not been easy. Two tests that have been used previously are the spectral ripple test and the spectral-temporally modulated ripple test (SMRT). Over time, questions have been raised about the validity of these tests. We introduce a new computational electric hearing (with a cochlear implant, CI) model that can simulate how sound is transferred through a speech processor and is received by the cochlear nerve fibers. With this electric hearing model and a normal hearing model, we investigated whether the known limitations of these tests can be detected. For the spectral ripple test, we could show the limitations in the output of the CI, the information conveyed to the cochlear nerve, to estimate the threshold, and show the benefit of current steering. In addition, we reproduced the aliasing effect with normal hearing in the SMRT, as well as the reduced ripple resolution in CI users. Our computational modeling framework can serve as a first-step assessment of the validity of new psychophysical tests. Moreover, it could be used to test new speech coding strategies.

Keywords: neural modeling; cochlear implants; psychophysics; spectral ripple; auditory perception; spectral resolution; speech coding strategies



Academic Editor: Luc de Witte

Received: 22 July 2025

Revised: 30 October 2025

Accepted: 1 November 2025

Published: 5 November 2025

Citation: Martens, S.S.M.; Briaire, J.J.; Frijns, J.H.M. Spectral Ripples in Normal and Electric Hearing Models. *Technologies* **2025**, *13*, 505. <https://doi.org/10.3390/technologies13110505>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cochlear implants (CIs) are advanced neurostimulators that significantly improve speech perception for the deaf across all age groups [1,2]. Despite these benefits, CI users often struggle with understanding speech in noisy environments [3,4]. In addition, though CI users appreciate music, the pitch and timbre are not conveyed properly [5]. These challenges arise from the complexity of auditory perception, which involves multiple interrelated qualities.

CIs utilize the tonotopic organization of the cochlea, in which high frequencies are processed at the base and low frequencies at the apex. However, the insertion depth of the electrode array is limited, resulting in lower frequencies being presented at suboptimal locations, typically within two cochlear turns of the round window. After implantation, the auditory system is assumed to adapt its tonotopic organization to align with the frequency allocation of the CI [6,7].

Psychophysical tests have been developed to assess the auditory capabilities of CI users, such as loudness and pitch discrimination or forward masking effects. However, the process of designing these tests is complex, and it is crucial to identify and address confounding factors, preferably before involving participants. Cochlear modeling is a

valuable tool for this purpose, allowing preliminary evaluations of psychophysical tests before conducting clinical trials.

Advances in computational power have enabled the creation of highly detailed models that simulate auditory responses at the fiber level for the entire cochlea. These models provide insights into what listeners perceive and can reveal potential limitations or unintended cues in psychophysical tests. In this paper, we evaluate two psychophysical tests found to contain unintentional cues during clinical evaluations. We propose a modeling-based approach to enhance the reliability of these tests before their clinical use, reducing the need for patient involvement and improving resource efficiency.

1.1. Spectral Ripple Test

Assessing spectral resolution through psychophysical testing has proven to be challenging. Early approaches often utilized spectral ripple paradigms, which involve modulating a signal's spectral content [8,9]. Won et al. [10], developed the spectral ripple discrimination test, in which rippled stimuli with sinusoidal envelopes on a logarithmic frequency scale (100 Hz to 5000 Hz) are generated at varying ripple densities, measured in ripples per octave (RPO). In this test, participants use a three-alternative forced-choice (3AFC) format to distinguish between a standard reference stimulus and an inverted (phase-reversed) stimulus. However, its validity for CI users has been questioned. Studies suggest that ripple peaks align with CI filter spacing as the ripple density increases, which confounds the results because the filter spacing does not adequately sample the spectral peaks and valleys at high ripple densities [11–13]. Anderson et al. [11] analyzed this limitation using the Nyquist–Shannon sampling theorem, which states that the filter spacing must be less than half the ripple rate to reconstruct the signal accurately. Similarly, Gifford et al. [13] noted that performance above 2 RPO is constrained by the frequency assignment of CIs and hypothesized that CI users rely on both spectral and temporal processing. Anderson et al. [12] also observed non-monotonic behavior in spectral modulation transfer functions, suggesting that some participants unintentionally shift between perceptual cues. Winn and O'Brien [14] proposed that CI limitations and ripple distortions above 2 RPO exacerbate these issues. Interestingly, Drennan et al. [15] demonstrated that the current steering speech coding strategy (Fidelity 120 by Advanced Bionics) improves performance, achieving 3.42 RPO compared to 2.31 RPO with the HiResolution (HiRes) strategy.

1.2. Spectral-Temporally Modulated Ripple Test (SMRT)

To address the challenges noted above, Aronoff and Landsberger [16] introduced the spectral-temporally modulated ripple test (SMRT), incorporating temporal modulation to minimize confounding factors, such as loudness cues, spectral centroids, and changes to spectral edges [17]. SMRT stimuli are created using a non-harmonic tone complex with 202 pure-tone components with 33.333 carriers per octave (CPO) from 100 Hz to 6400 Hz [16]. Using a 3AFC format, participants are presented with two reference stimuli at 20 RPO and a target stimulus at a different ripple density. Performance is measured as the highest ripple density the participant can distinguish from the 20 RPO reference. Resnick et al. [18] found that performance scores were affected by the carrier density (33.333 CPO) used in the original test. Normal hearing participants performed better with a carrier density of 100 CPO but exhibited non-monotonic behavior with 33.333 CPO, including a drop in performance at 11 RPO and improvement at 16 RPO due to spectral aliasing caused by the Nyquist limit [18]. Narne et al. [19] hypothesized that amplitude fluctuations due to cochlear filtering around or above 6400 Hz may act as an unintentional cue, explaining why the SMRT yields higher thresholds (7.1 RPO) in normal hearing listen-

ers compared to a spectral ripple test (5.5 RPO). However, this spectral ripple test differed from the spectral ripple test by Won et al. in several characteristics, such as peak-to-valley ratio, a different frequency range, and filtering. For CI users, Lawler et al. [20] reported an average threshold of 3.1 RPO in the SMRT. However, both tests have shown correlations with speech understanding and speech recognition in noise, highlighting their importance despite their limitations [10,20,21].

1.3. Modeling Approach

This paper evaluates whether models of normal and electric hearing can replicate the unintentional cues and outcomes observed in the spectral ripple test and SMRT, such as aliasing, the benefits of current steering, and reduced ripple resolution in electric hearing. If successful, this would validate the use of model-based analyses to assess psychophysical tests without patient involvement. For normal hearing, we used the model from Bruce et al. [22], which simulates auditory signal transmission via sensory hair cells to auditory nerve fibers. This phenomenological model can simulate spiking patterns at characteristic frequencies in response to auditory inputs. For electric hearing, we introduce a new advanced modeling pipeline developed at our department that combines biophysical and phenomenological models [23–29]. This pipeline includes a speech processor to simulate spiking patterns in response to auditory inputs for a cochlear implant with an electrode array. More details on the models are provided in the Materials and Methods section. These models provide insights into what is perceived in normal hearing and what is conveyed by speech processors in electric hearing, enabling a deeper understanding of psychophysical test outcomes and paving the way for more reliable test designs.

2. Materials and Methods

A schematic overview of both models and their components is shown in Figure 1. The electric hearing pipeline consists of a model for a speech processor and a cochlear model. The speech processor translates a sound wave into an electrodogram, which contains the pulse amplitudes per electrode and is in steps of the pulse width. The implanted cochlea model, which consists of three separate models, receives the electrodogram as input and can calculate the spike times per fiber in response to this input, as will be described in Section 2.1. In comparison, the normal hearing model processes the sound wave directly without preprocessing. It is an elaborate and well-validated model with intricate processing of each step of normal hearing, and the steps within the model are described in Section 2.2. An overview of the settings of both models is provided in Table 1.

Table 1. Model settings of the electric hearing model and the normal hearing model. Both models had the same frequency range (250–8054 Hz), the same number of fibers (2416 fibers), the same number of trials (5 trials), and the same window size for moving average filtering (33, see Section 2.4.1). SPL: sound pressure level. FS: full scale. RMS: root mean square.

Parameter	Electric Hearing	Normal Hearing
Number of fibers per frequency	1	50
Loudness	65 dB SPL (FS 111.6 dB)	65 dB RMS SPL
Spontaneous rate [spikes/s]	10	0.1 (n = 10), 4 (n = 10), 70 (n = 30)

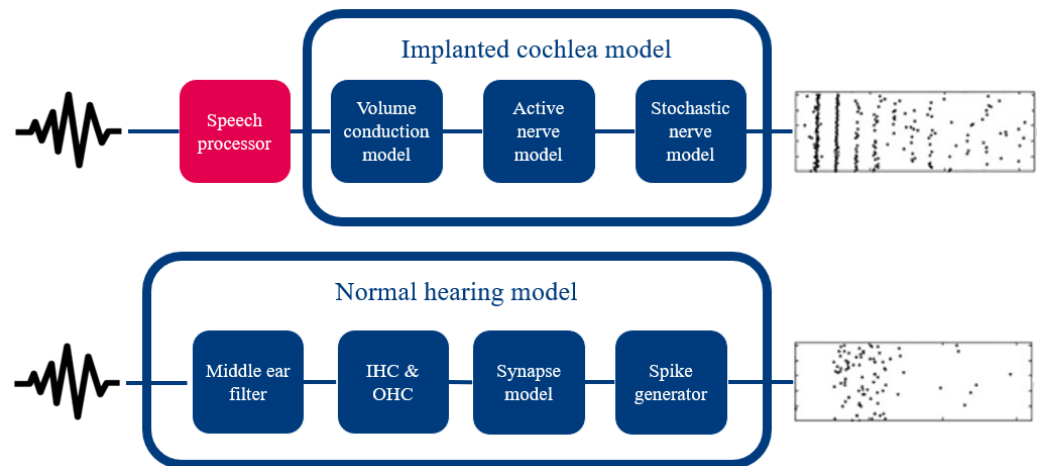


Figure 1. Hearing models. Top: Components of the electric hearing pipeline. The pipeline receives a sound wave as input and provides a spike pattern as output for each fiber. In blue are the components of the cochlear part of the pipeline. The speech processor provides these components with an electrogram. Bottom: The normal hearing model of Bruce et al. [22], which also receives a sound wave as input and provides a spike pattern as output for each characteristic frequency. The spiking per fiber (in vertical direction) is shown over time (in horizontal direction) as dots. IHC: inner hair cell. OHC: outer hair cell.

2.1. Electric Hearing

In this paper, we propose a new extensive electric hearing pipeline that models how sound is processed by the fibers of the cochlear nerve of a CI user. The pipeline here models a speech coding strategy resembling the Fidelity 120 strategy of Advanced Bionics, which is described in Section 2.1.1. The implanted cochlea model uses the output of the speech processor to model the fiber responses to the input and is presented in Section 2.1.2. Lastly, an interpretation step in the form of a new frequency allocation is added to the output of the fibers to evaluate the perceived spectral content, which is explained in Section 2.1.3.

2.1.1. Speech Processor

The speech processor is taken from the open-source Advanced Bionics Generic-Python-Toolbox and generates pulse trains in response to sound waves. This is code for the spectral resolution (SpecRes) strategy (Figure 2), a research version of the commercial HiRes with Fidelity 120 (F120) strategy [29]. The input level was rescaled to 65 dB sound pressure level (SPL) with the full scale set to 111.6 dB. The range of the CI analysis bands is 306–8054 Hz.

The sound is preprocessed with a pre-emphasis filter, to attenuate lower frequencies and thus balance the frequency spectrum, and an automatic gain control, to maintain a suitable signal amplitude at its output, despite variation of the signal amplitude at the input. The filter bank consists of a fast Fourier transform (FFT), which reflects the frequency content per time frame or window, and is divided into 15 frequency bands. The output of the filter bank is used to assess the envelope of the input and to estimate the frequency per frequency band. The envelope information is used by a noise reduction algorithm, which is comparable to the ClearVoice algorithm of Advanced Bionics, in which the speech energy and noise energy are estimated, and a correction in the form of gains is applied. In parallel, the frequency estimate information is used to determine the carrier frequency and the frequency location within the band (for current steering, explained in the following paragraph). Lastly, these two parallel lines of information are used to map frequency content to the T- and M-levels to produce an electrogram.

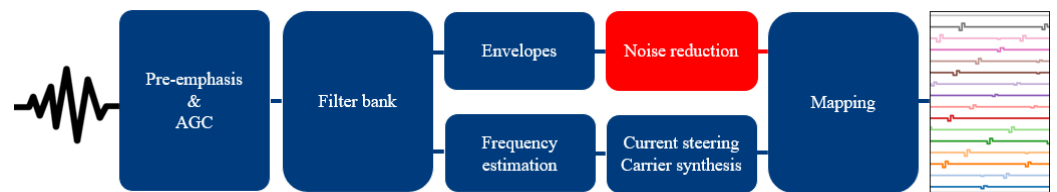


Figure 2. Overview of the SpecRes speech coding strategy. Sound waves are used as an input to create an electrogram. Red indicates the step from which the information was used to determine the electric frequency content per frequency band. The electrogram is fed to the cochlear model. AGC: automatic gain control.

This speech coding strategy uses current steering, a method to enhance the frequency resolution in cochlear implants, where two neighboring electrodes are stimulated simultaneously to elicit neural activation between these two physical contacts. Sixteen electrodes are used in this speech coding strategy, resulting in 15 pairs of electrodes called analysis bands or channels (see Table 2 for the frequency assignment). Thus, each analysis band has two assigned electrodes: the lower corner frequency is assigned to the lower electrode of the analysis band, and the upper electrode is assigned to the upper corner frequency or edge of the analysis band. By varying the weight (α) of the current in the electrodes in steps of 0.125, the peak location can be shifted. This results in nine possible sets of weight combinations and peak locations per frequency band per time frame. With the weight set to $\alpha_1 = 0.5$ and $\alpha_2 = 0.5$ for electrodes 1 and 2, respectively (i.e., equal amplitudes), the activation peak resides in the middle of electrodes 1 and 2. For example, a signal of 442 Hz (see Table 2) is assigned to electrode 2 (with $\alpha_2 = 1$ and $\alpha_3 = 0$), and a 578 Hz signal is assigned to electrode 3 ($\alpha_3 = 1$ and $\alpha_4 = 0$). For a more detailed explanation on current steering and the speech coding strategy, we refer to [29].

By setting the weights for all 15 channels to 0.5, we mimicked a different speech coding strategy without current steering, meaning all 15 peaks were placed in the middle of the electrodes. This is similar to the HiResolution strategy used by de Jong et al. [30], dubbed the HiRes FFT strategy. However, they used 16 channels, with the peaks situated directly at the electrodes by setting the α values to 0.

The frequency content per channel, meaning the energy estimated by the speech coding strategy per analysis band, was used to assess how much information can be conveyed by the speech coding strategy, as was done by Winn and O'Brien [14]. This assessment used the frequency content after the addition of the noise reduction gains to the envelope calculation per band (see red block in Figure 2). The T-levels and M-levels were set to those of the implanted cochlea model (described in Section 2.1.2).

Table 2. The 15 analysis bands of the Fidelity 120 speech coding strategy are divided over 16 electrodes. For each analysis band, the lower and upper frequency edges (corner frequencies) are listed in Hz. The upper edge of each band is the lower edge of the next band.

Analysis Band	Lower Edge	Upper Edge
1	306	442
2	442	578
3	578	646
4	646	782
5	782	918
6	918	1054
7	1054	1257
8	1257	1529
9	1529	1801
10	1801	2141
11	2141	2549
12	2549	3025
13	3025	3568
14	3568	4248
15	4248	8054

2.1.2. The Implanted Cochlea Model

The implanted cochlea model consists of a volume conduction model, an active nerve fiber (ANF) model, and a stochastic nerve model (see the top of Figure 1 for an overview). The 3D volume conduction model calculates the potential distribution along the modeled nerve fibers in the implanted cochlea per electrode contact [25,26]. The potential distribution is then used by the ANF model to compute a deterministic threshold for each electrode pair [23,27]. The cochlear geometry used in the ANF model is based on μ CT imaging data from human temporal bones, and realistic neural trajectories are defined according to histological data. Next, the phenomenological adaptive stochastic (PHAST+) model takes the deterministic threshold and adds stochasticity and temporal effects (refractoriness, accommodation, and adaptation) to simulate the resulting spike train per auditory nerve fiber in response to the pulse train [24,28]. An overview of the included neural parameters is listed in Table A1 in the Appendix A. The model can process the response of 3200 fibers distributed between 2.56 mm and 33.2 mm from the round window. However, for our purposes, we limited it to the fibers stimulated in proximity to the implanted electrode array, resulting in 2416 fibers. The ANF model modeled the thresholds, T-levels, and M-levels of the HiFocus MS electrode array (Advanced Bionics) in a mid-scalar position. The insertion angle of the most basal electrode contact was 27 degrees from the round window, and for the most apical electrode contact, 420 degrees from the round window, modeled following the mean angular insertion depth in van der Jagt et al. [31]. The T-levels and M-levels were estimated based on 0.5 mm and 3.5 mm excitation along the basilar membrane, respectively, per electrode. A known shortcoming of the ANF model is that the thresholds are approximately an order of magnitude too high compared to physiological values. To avoid affecting the speech coding strategy with this discrepancy, the T-levels and M-levels used by the speech coding strategy were divided by 3 to be in the same ballpark as physiological levels, and the amplitudes of the pulses of the electrodiagram were multiplied by 3 to use them as an input for the implanted cochlea model. We used five trials per stimulus to get an averaged response, meaning the neural response was calculated for a single stimulus five times and averaged over these five runs.

2.1.3. Fiber Frequency Allocation

Generally, the Greenwood frequencies are used based on the basilar membrane position. However, the consensus is that cochlear implant (CI) users adapt over time to the (frequency-to-place) mapping of the implant [7,32–34]. Here, we have assumed that neural adaptation to the new CI-based frequency distribution has occurred, and we have accordingly created a new, so-called ‘learnt’, fiber frequency allocation (see Figure 3A,B). Our frequency allocation approach is similar to the approach of de Nobel et al. [35]. Although this adaptation is presumed to occur higher up than the auditory nerve, we allocated these new frequencies to the fibers for interpretation of the neural processing, which greatly simplified the comparison between the two models. For each electrode, the fiber with the lowest threshold (see Figure 3) was assigned the corner frequency of that electrode (see Section 2.1.1). The fibers between these highest responding fibers were assigned frequencies log-spaced between the edges of the frequency bands (see Figure 3D). A different approach was applied to the first and last electrodes. The fibers situated just outside of the array, with a threshold up to M-level of these contacts, were included. For the first analysis band and fiber section, the fibers were assigned frequencies logarithmically spaced between 250 Hz and 442 Hz. A lower frequency edge was chosen to include frequencies to which the speech coding strategy might respond due to spectral leakage, due to the imperfect resolution of the FFT filter banks. This frequency was chosen heuristically by providing the speech coding strategy with low-frequency (below 306 Hz) pure sine stimuli and assessing when

the electric output increased. For the last analysis band and fiber section, the fibers were assigned frequencies logarithmically spaced between 4248 Hz and 8054 Hz.

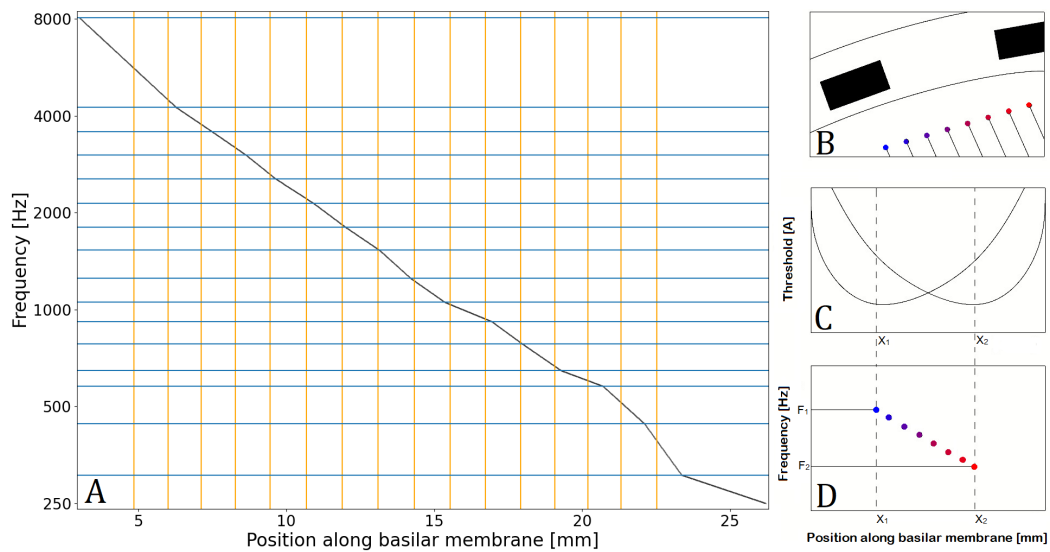


Figure 3. (A) Learnt fiber frequency allocation. Black indicates the fiber location and corresponding learnt fiber frequency. Orange lines indicate the electrode contact's position. Blue lines indicate the electrode contact's frequency. The position is determined from base to apex. (B) Schematic view of electrode contacts, Electrode 1 and Electrode 2 (black rectangles), and the nearest fibers (dots). (C) The threshold curves for each fiber along the basilar membrane for two electrode contacts. (D) Frequency assignment at the fiber level. The blue fiber with the lowest threshold for Electrode 1 is assigned the corner frequency F_1 of Electrode 1, situated at position X_1 along the basilar membrane. The red fiber is assigned the corner frequency F_2 of Electrode 2 and is situated at X_2 . The fibers between X_1 and X_2 are assigned log-spaced frequencies between F_1 and F_2 .

2.2. Normal Hearing Model

The phenomenological model from Bruce et al. [22] was used to simulate the response of single fibers in normal hearing. This model captures the transmission of auditory signals from the ear canal, via the middle ear and basilar membrane, to the sensory hair cells in the inner ear, to the auditory nerve fibers, simulating spiking patterns in response to specific auditory inputs across chosen characteristic frequencies. The model includes cochlear filtering, which replicates the frequency-selective response of the basilar membrane using a bank of bandpass filters that mirror the cochlea's tonotopic organization. The next stage involves inner hair cell processing, where basilar membrane motion is transformed into receptor potentials to reflect the dynamic range of human hearing. Synaptic transmission is modeled to simulate neurotransmitter release from the inner hair cells to auditory nerve fibers, including short-term adaptation dynamics that influence nerve responses. The auditory nerve fiber response is then captured, simulating the spiking behavior of fibers with stochastic elements to reflect the natural variability in spike timing. The model represents different types of nerve fibers with low, medium, and high spontaneous rates to account for the diversity in auditory nerve responses. The output per characteristic frequency was created using the model's default settings: 10 low-spontaneous, 10 medium-spontaneous, and 30 high-spontaneous fibers. The spikes of these fibers were summed and averaged over five trials. The default setting of variable fractional Gaussian noise was used. The basilar membrane tuning for humans was based on Shera et al. [36]. The approximate implementation of the power-law functions in the synapse model was chosen. The input sound wave of the normal hearing model was set to 65 dB root mean square (RMS) SPL.

The characteristic frequencies of the fibers were set to the same 2416 frequencies as used in the electric hearing pipeline, ranging from 250 Hz to 8054 Hz.

2.3. Sounds

2.3.1. Spectral Ripple Test

The first test used the ripple sound files of Won et al. [10]. In this test, the listener has to distinguish between two noisy stimuli with the same ripple density; the ripples in the frequency content are shifted, creating standard and two inverted stimuli (i1 and i2, see Figure 4). The inverted stimuli i1 and i2 have the same pattern in the power spectral density but have slightly different randomly included frequencies. The stimulus duration was 0.5 s, and the stimuli were ramped with 150 ms rise/fall times. Frequencies from 100 Hz up to 5 kHz are included. The test becomes more difficult with increasing ripple density as the peaks of the ripples fall closer together. Thus, this test makes use of the ability to discriminate between frequencies along the basilar membrane. The original ripple test contained 14 different densities, which differed by ratios of 1.414 (0.125, 0.176, 0.250, 0.354, 0.500, 0.707, 1.000, 1.414, 2.000, 2.828, 4.000, 5.657, 8.000, and 11.314 RPO). The peak-to-valley ratio of these stimuli was 30 dB, and they were filtered with a long-term speech-shaped filter [37]. For each ripple density, 30 stimuli with different starting phases in the spectral domain are possible. All figures are created with the first phase (indicated with ‘_1’ in the sound filename) and comparing s and i1 (not to overcrowd the figures), except for the spectral ripple threshold determination (see Section 2.5).

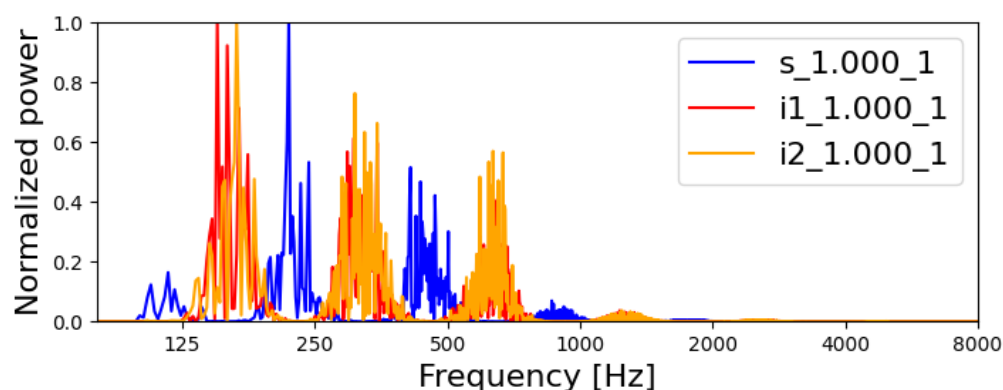


Figure 4. Normalized power spectral density of the standard (‘s_1.000_1’) and an inverted spectral ripples (‘i1_1.000_1’ and ‘i2_1.000_1’) with 1.000 RPO and the first phase, following Won et al. [10].

2.3.2. SMRT

For the second test, the SMRT stimuli (see Figure 5) described in Aronoff and Landsberger [16] were recreated in Matlab (2023a) with a sampling rate of 44.1 kHz. The amplitude modulations A applied to amplitude of the pure tone carrier P to derive stimulus S with

$$S(t) = \sum_{i=1}^C P_i(t) \cdot 10^{\frac{A(i,t)}{20}}, \quad (1)$$

where t is time, i is the index of the pure tone, with the number of carriers C at 201 for the lower carrier density (33.333 CPO) and 602 for the higher carrier density (100 CPO). The amplitude modulation A is calculated as follows:

$$A(i, t) = |D \cdot \sin(0.03\pi Qi + R\pi t)| - D, \quad (2)$$

with a ripple depth D of 20 dB, a ripple repetition rate R (the number of repetitions of the ripple per second) was set to 5 Hz, the ripple density Q expressed in ripples-per-

octave (RPO). The stimuli were generated using a complex of tones with 201 or 602 equal-amplitude pure-tone frequency components with 33.333 CPO or 100 CPO (indicated by red dots on the left-hand axis of Figure 5) of an octave from 100 Hz to 6500 Hz for the lower and higher carrier density, respectively. A higher density would mean a greater number of red dots. The stimuli had a duration of 0.5 s, and linear ramps of 0.1 s were added to the beginning and end of the sound.

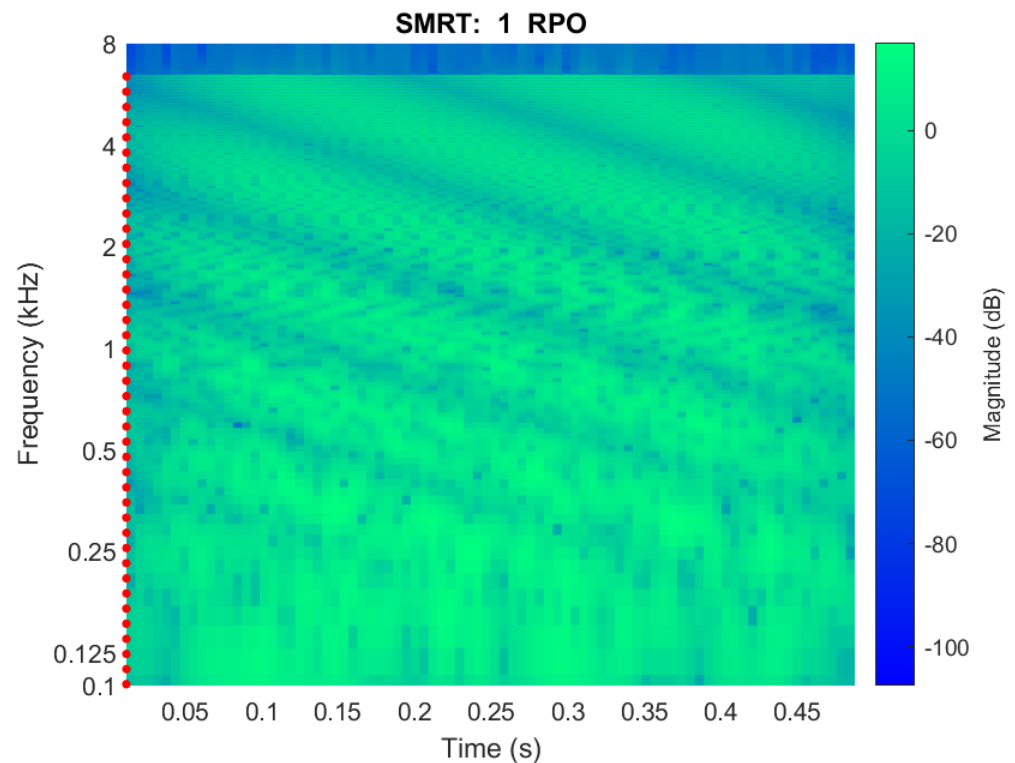


Figure 5. Spectrogram of the SMRT stimulus with 1 RPO (carrier density of 33.333 CPO). The red dots indicate the included pure-tone frequency components. Every fifth frequency is plotted to avoid overcrowding the y-axis.

2.4. Visualization

2.4.1. Spectrum

For the spectral ripple sounds [10], a spectrum view was used to indicate the spectral content of the sound. Three types of spectrum views were used: acoustic, electric, and neural. The latter portrays the amount of neural activity, the other spectra the power in a frequency bin over the entire stimulus duration. All spectrum views were min–max-normalized for better comparison.

The *acoustic* spectrum shows the normalized power spectral density (Figures 4 and 6A–D). In contrast to Winn and O’Brien’s method, the actual sound was used to calculate the electric spectrum and neural spectrum, not an ideal spectrum [14]. As a result, the ripples varied in height due to the sounds being filtered with a long-term speech-shaped filter [10,37]. For better visualization of the basilar membrane response, the same pre-emphasis filter in the speech processor was applied to the acoustic spectra.

The *electric* spectrum (Figure 6E–H) shows the normalized output (the perceived power per analysis band) of the speech coding strategy after the noise reduction step, before mapping, to evaluate the frequency content per frequency band (see red block in Figure 2).

For the *neural* spectrum, the number of spikes generated by the neural models over the stimulus in a frequency bin of the neural models was used. For all of the following neural spectrum figures in the Results (Figures 7 and 8), the y-axis indicates the normalized

response over all fiber responses, and the x-axis indicates the mapped frequency of the fiber. For the normal hearing output, this is the characteristic frequency. Individual fiber responses varied due to stochastic properties and were rather large in number, resulting in a somewhat chaotic spectrum. Therefore, in addition, the frequency response of the individual fibers was also spatially low-pass filtered to show the general response along the basilar membrane with a symmetric moving average filter.

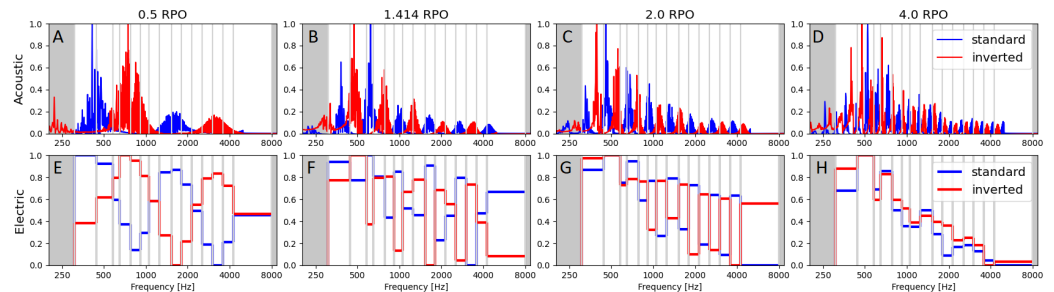


Figure 6. Acoustic and electric spectra. (A–D) The normalized power spectral density of sounds with different ripple densities: (A) 0.5 RPO, (B) 1.414 RPO, (C) 2.0 RPO, and (D) 4.0 RPO. (E–H) The normalized content per frequency band of the speech coding strategy for different ripple densities: (E) 0.5 RPO, (F) 1.414 RPO, (G) 2.0 RPO, and (H) 4.0 RPO. The blue line indicates the standard ripple, and the red line the inverted ripple (i1). The shaded areas indicate frequencies outside of the frequency bands. The vertical lines correspond to the frequency band edges of the speech coding strategy.

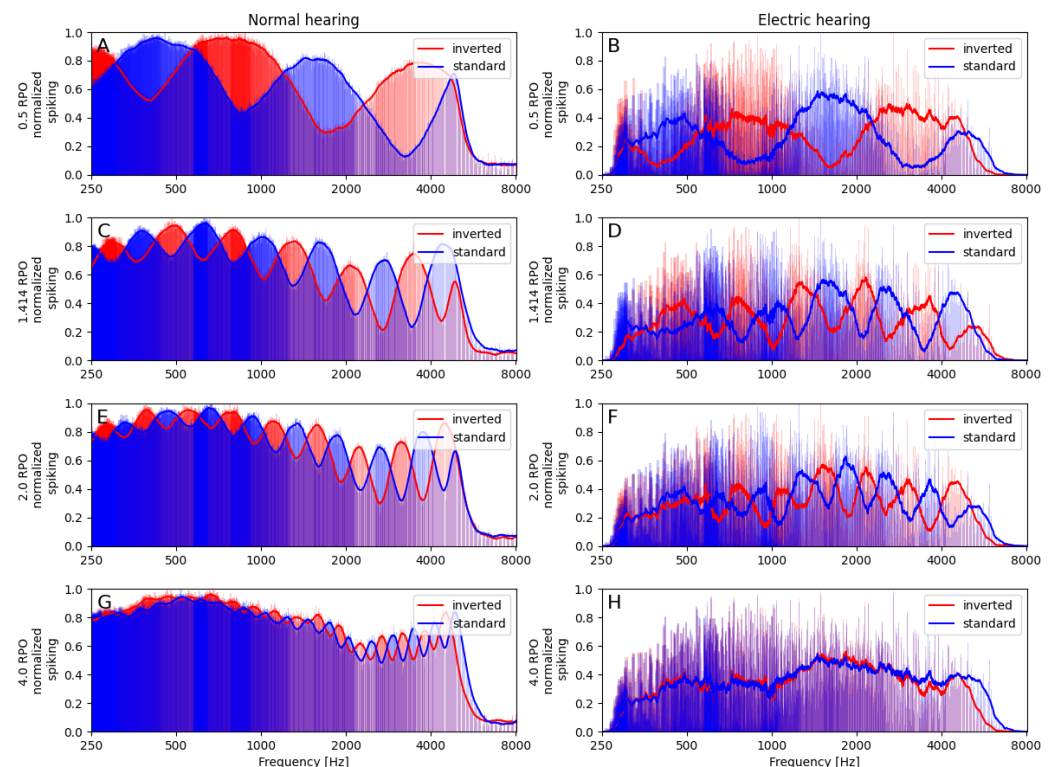


Figure 7. Spectra of neural activation in response to standard (blue) and inverted (i1 in red) spectral ripples of various ripple densities. The bars indicate the spikes per characteristic frequency. The thick lines are the low-pass filtered frequency responses. Left: Normalized spiking spectra for normal hearing with (A) 0.5 RPO ($d' = 32.77$), (C) 1.414 RPO ($d' = 21.86$), (E) 2.0 RPO ($d' = 16.39$), and (G) 4.0 RPO ($d' = 6.61$). Right: Normalized spiking spectra for electric hearing with (B) 0.5 RPO ($d' = 4.50$), (D) 1.414 RPO ($d' = 3.64$), (F) 2.0 RPO ($d' = 2.61$), and (H) 4.0 RPO ($d' = 0.98$). The fiber frequencies are based on the assumed frequencies after neural plasticity.

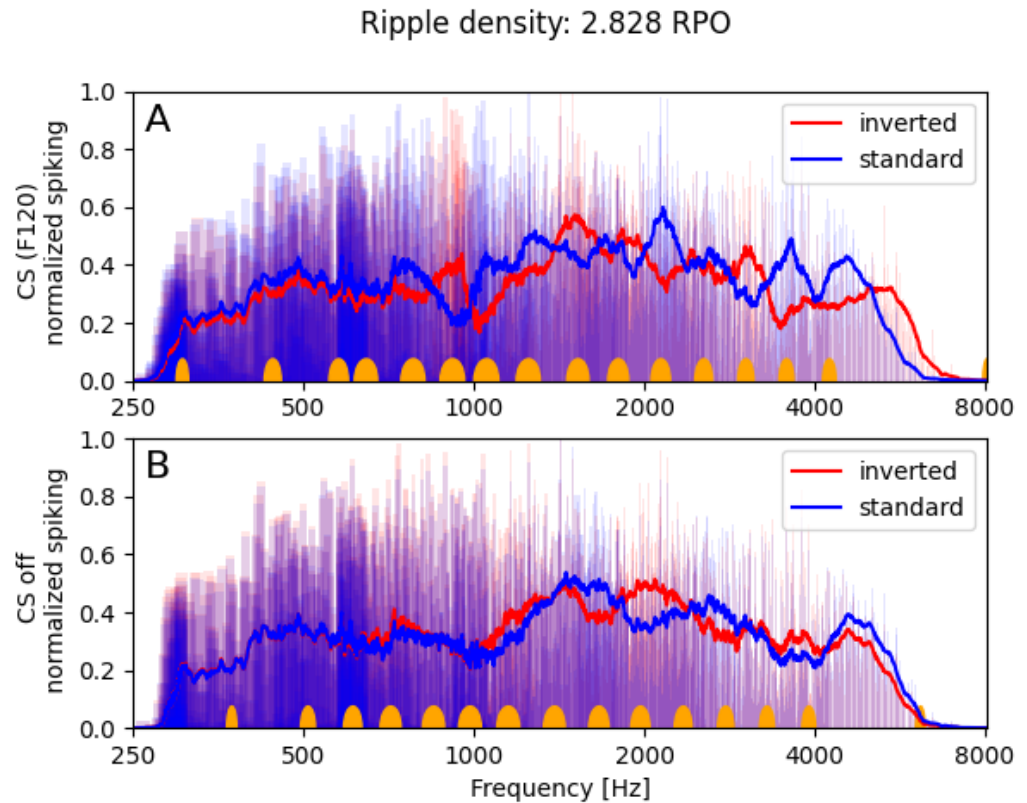


Figure 8. Comparison of the effect of current steering. The fiber frequencies are the assumed frequencies after neural plasticity. **(A)** Neural activation with the F120 speech coding strategy. The halved orange ellipses indicate the 16 electrode positions ($d' = 1.58$). **(B)** Neural activation when current steering (CS) is turned off ($d' = 0.76$). This is done by setting the weights to 0.5, resulting in 15 peak positions as indicated in orange. The blue line indicates the standard ripple, and the red line the inverted ripple (i1).

2.4.2. Neurogram

To assess the presence of aliasing in the temporal domain, neurograms were chosen as the method for visualization (Figures 10 and 11). A neurogram is made by stacking the post-stimulus time histograms (PSTHs) of all fiber responses. The PSTHs were created with a bin size of 5 ms. Neurograms portray the neural response of the auditory nerve in terms of the neural spiking intensity of multiple auditory nerve fibers as a function of time and location on the basilar membrane. As auditory nerve fibers have a topological order in terms of characteristic frequency, neurograms look similar to spectrograms. The colors of the bins in the neurograms indicate the spike rate per frequency and time bin. Brighter colors reflect a higher spike rate.

2.5. Spectral-Ripple Threshold

We modeled the neural spectrum for multiple starting phases per ripple density in addition to multiple trials per stimulus. This provided us with enough trials to calculate d' for the standard and the inverted stimulus at each ripple density. We calculated d' following the method of [38,39], which assumes unequal variances, using

$$d' = \frac{|\mu_s - \mu_i|}{\sqrt{\frac{1}{2}(\sigma_s^2 + \sigma_i^2)}}, \quad (3)$$

where μ_s is the mean of the neural spectrum for the standard stimulus, μ_i is the mean for the inverted stimulus, σ_s is the standard deviation of the neural spectrum in response

to the standard ripple, and σ_i is the standard deviation of the inverted ripple. For the inverted ripple, two versions (i1 and i2) are available per phase, providing three options for a 3-alternative forced-choice test. The average d' of both inverted stimuli (d'_{si1} and d'_{si2}) was used to simulate the proportion of correct P_c to plot the psychometric curve for a 3AFC test:

$$P_c(d') = \int_{-\infty}^{\infty} \Phi(x)^2 \varphi(x - d') dx, \quad (4)$$

where $\Phi(x)$ is the cumulative distribution function of a standard normal distribution and $\varphi(x)$ is the probability density function of a standard normal distribution [40–43]. We estimate the spectral-ripple threshold T in this 3AFC test:

$$T = \gamma + \frac{1 - \gamma}{2}, \quad (5)$$

with the guessing rate $\gamma = \frac{1}{3}$, resulting in a threshold at a proportion of correct $P_c = 0.667$ or 66.67% correct [43]. This assessment was performed on (moving average filtered) neural spectra.

3. Results

3.1. Spectral Ripple Test

3.1.1. Acoustic and Electric Spectra

Winn and O'Brien [14] looked at the acoustic and electric content for the spectral ripples from Won et al. [10]: the frequency spectrum of the sound and the frequency content that is transferred through the speech coding strategy. Their results were recreated with our model in Figure 6. In the top row, the acoustic content of four different spectral ripple stimuli, with standard and inverted versions, is shown in order of increasing ripple density from left to right: 0.5 RPO, 1.1414 RPO, 2.0 RPO, and 4.0 RPO. In the acoustic spectrum (Figure 6A–D), the standard (blue) and inverted (red) peaks are easily distinguished up to 4.0 RPO. From the grey vertical lines, indicating the frequency band edges of the speech processor, it is clear that for electric hearing, the standard and inverted peaks are often in the same analysis band at 4.0 RPO. The bottom row shows the corresponding electric spectrum (i.e., the spectral information conveyed by the speech coding strategy per analysis band). Similarly, the spectral content of the standard (blue) and inverted (red) stimuli is shown in order of increasing ripple density. Spectral content below 340 Hz was not transmitted through the speech coding strategy, as the lowest frequency band ranged from 306 Hz to 442 Hz. Thus, no content was shown in the electric spectrum. At low ripple density (0.5 RPO, Figure 6E), the speech coding strategy could accurately follow the ripples in the spectral domain, as a wavy pattern was visible in the spectrum. With greater ripple density (1.414 and 2.0 RPO, Figure 6F,G), the CI processor could still follow the ripple as the peaks alternated between the edges of the frequency band. At 4.0 RPO (Figure 6H), the frequency content was less discernible between the standard and inverted ripples due to multiple peaks being within the same analysis band, resulting in similar content for the standard and the inverted stimulus. Moreover, a descending ripple or wave pattern in the electric spectrum of (Figure 6H), similar to aliasing, was seen in the inverted content and more clearly in the standard content with peaks situated at the 442–578 Hz, 646–782 Hz, 1257–1529 Hz, 2549–3025 Hz. In other words, the ripple density of the ripple pattern in the electric spectrum in Figure 6H was lower than the ripple density in the acoustic spectrum (Figure 6D). However, some bands differed between the standard and inverted content, such as between 306–442 Hz and 1801–2141 Hz, indicating that some CI users may be able to use this loudness cue to distinguish the stimuli.

3.1.2. Neural Spectra

Figure 7 shows the neural activation for normal hearing and the neural activation with the F120 speech coding strategy for the standard and inverted stimuli in ascending ripple density per row. The average d' values over all phases are listed in the caption per ripple density. Similar to the ripples in the normalized power spectral density in Figure 7A–D, the neural spectrum shows ripples along the basilar membrane. For normal hearing, the peaks of the standard and inverted stimuli were discernible at every depicted ripple density. With increasing ripple density, the peaks became slimmer but were still specific to the location on the basilar membrane, though the valleys became shallower with increasing ripple density due to the peaks being closer to each other. This alludes to increased difficulty in discrimination. For electric hearing, the discernibility between the peaks of the standard and inverted stimuli rapidly decreased with increasing ripple density. At 0.5 RPO (Figure 7B), the peaks were situated at different locations on the basilar membrane. At 2.0 RPO (Figure 7F), this also appeared to be the case, but the peaks were less distinctly located than what was observed at the lower ripple density. At 4.0 RPO (Figure 7H), the peaks of the standard and inverted ripples coincided at the same location of the basilar membrane for electric hearing, which was not the case for normal hearing. In electric hearing, the peaks in the highest frequency band (4–8 kHz) did not coincide, reflecting that the upper standard peak is higher than the upper inverted peak (Figure 6D). In the normal hearing response, a clear decrease in neural activation occurred after 5 kHz, the upper limit of the frequency content of the spectral ripple; this was not the case for electric hearing. For normal hearing, the activation was relatively strong over all frequencies up to 5 kHz. For electric hearing, the drop in activation is not as distinct as with normal hearing. The spatial profile of activation (derived by spatially filtering the individual fiber responses), shown as thick lines in the spectra for normal hearing, consistently concurred with the peaks of the bins. In contrast, for electric hearing, the thick line fell lower than the peaks of the individual fibers, a lower height than for normal hearing. This indicates that the fiber response rates are more variable for electric hearing.

Figure 8 shows the neural activation with the F120 speech coding strategy (Figure 8A) and without current steering (Figure 8B) for the standard and inverted stimuli with 2.828 RPO. The average d' values over all phases are listed in the caption for both strategies. The summed squared difference of the individual fiber responses between the inverted and standard stimuli was larger with current steering (25.5) than without current steering (6.2). For the filtered response (thick line), the summed squared difference was 21.6 with current steering and 5.4 without current steering. Apparently, current steering offers a benefit in discernibility; for example, the peaks within the ranges of 782–1054 Hz (channels 5 and 6), 1257–1801 Hz (channels 8 and 9), 2549–3568 Hz (channels 12 and 13), and 4248–8054 Hz (channel 15) are separable.

3.1.3. Spectral Ripple Test Threshold

The filtered neural spectra were used to determine the spectral ripple threshold by simulating the psychometric curves. The psychometric curves, shown in Figure 9, have been fitted with a logistic function on filtered spectra (the thick lines in Figures 7 and 8) of normal hearing ($R^2 = 0.999$), electric hearing ($R^2 = 0.995$), and electric hearing without current steering ($R^2 = 0.995$). The results per phase and inverted stimulus are shown as smaller dots. Around the threshold, there is a greater spread in performance over the phases. Per ripple density, two large dots are plotted from each average d' of 30 phases; however, they often concur. For normal hearing, the threshold is not reached within the tested sample of ripple densities. The interpolated spectral ripple threshold from the fitted

psychometric curve with F120 and current steering ($T = 3.7$ RPO) is higher than for F120 without current steering ($T = 2.2$ RPO).

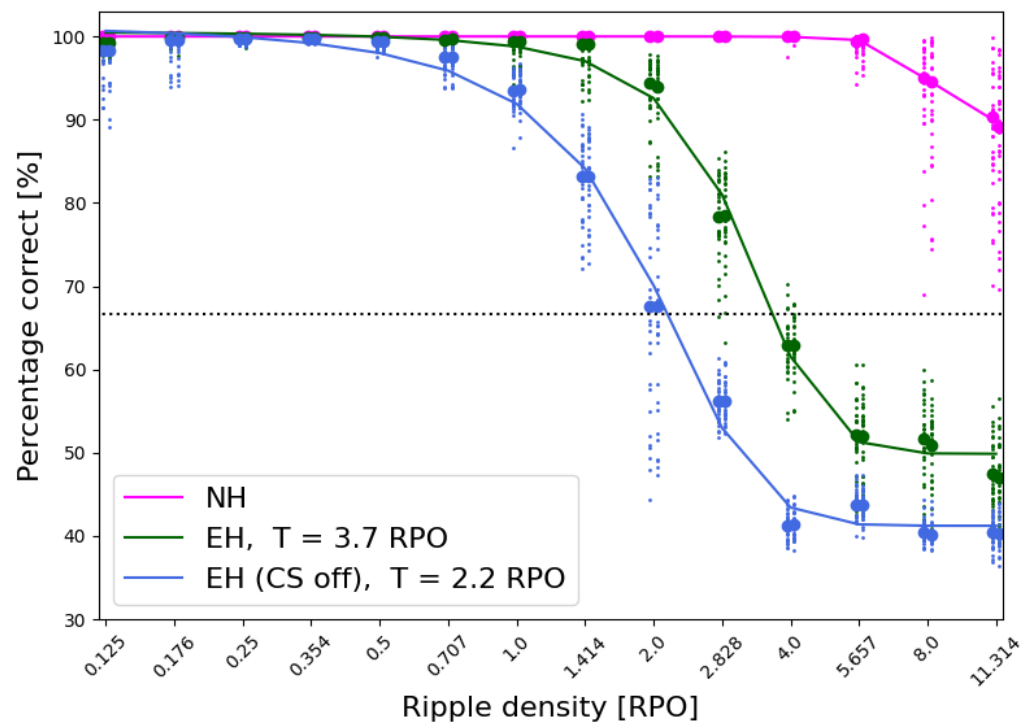


Figure 9. Simulation of psychometric curves with spectral ripples using filtered neural spectra of normal hearing (NH), electric hearing (EH) with current steering, and electric hearing with current steering (EH CS off) turned off (by setting the weights to 0.5, resulting in 15 peak positions). The smaller dots indicate the performance per phase and inverted stimulus (i1 and i2). The dots are shifted slightly to separate the distribution of the two inverted stimuli. The larger dot is the average performance per inverted stimulus for that ripple density. The average of these two is used to fit the psychometric curve. For each psychometric curve, the threshold (T) was interpolated at 66.67% (indicated with the dotted line) when possible.

3.2. SMRT

For the four SMRT spectrograms in Figure 1 from Resnick et al. [18] neurograms were simulated with the same ripple density (4.0 and 16.0 RPO) and both carrier densities (33.333 CPO and 100 CPO). The neurograms for normal hearing are shown in Figure 10. At 4.0 RPO, the ripples were visible at both carrier densities (Figure 10A,B). However, at the carrier density of 33.333 CPO with 16.0 RPO (Figure 10C), ripples were visible in the opposite direction (bottom left to top right), similar to Resnick's spectrogram for this situation. With the carrier density of 100 CPO, the ripples could no longer be discerned at 16.0 RPO (Figure 10D). Although the SMRT stimuli were created with frequencies up to 6500 Hz, some neural activation was seen above this frequency (6500 Hz to 8000 Hz) in synchrony with the activation at 6500 Hz. The darker areas at the beginning and end of the stimuli are due to the ramps. When focusing on a slimmer frequency band, such as from 3500 Hz to 4500 Hz (similar to looking at a critical band), these ripples were visible as amplitude modulations in the neural activation. In some regions, vertical lines can be discerned. These lines are also present in the acoustic stimulus (see Figure 5).

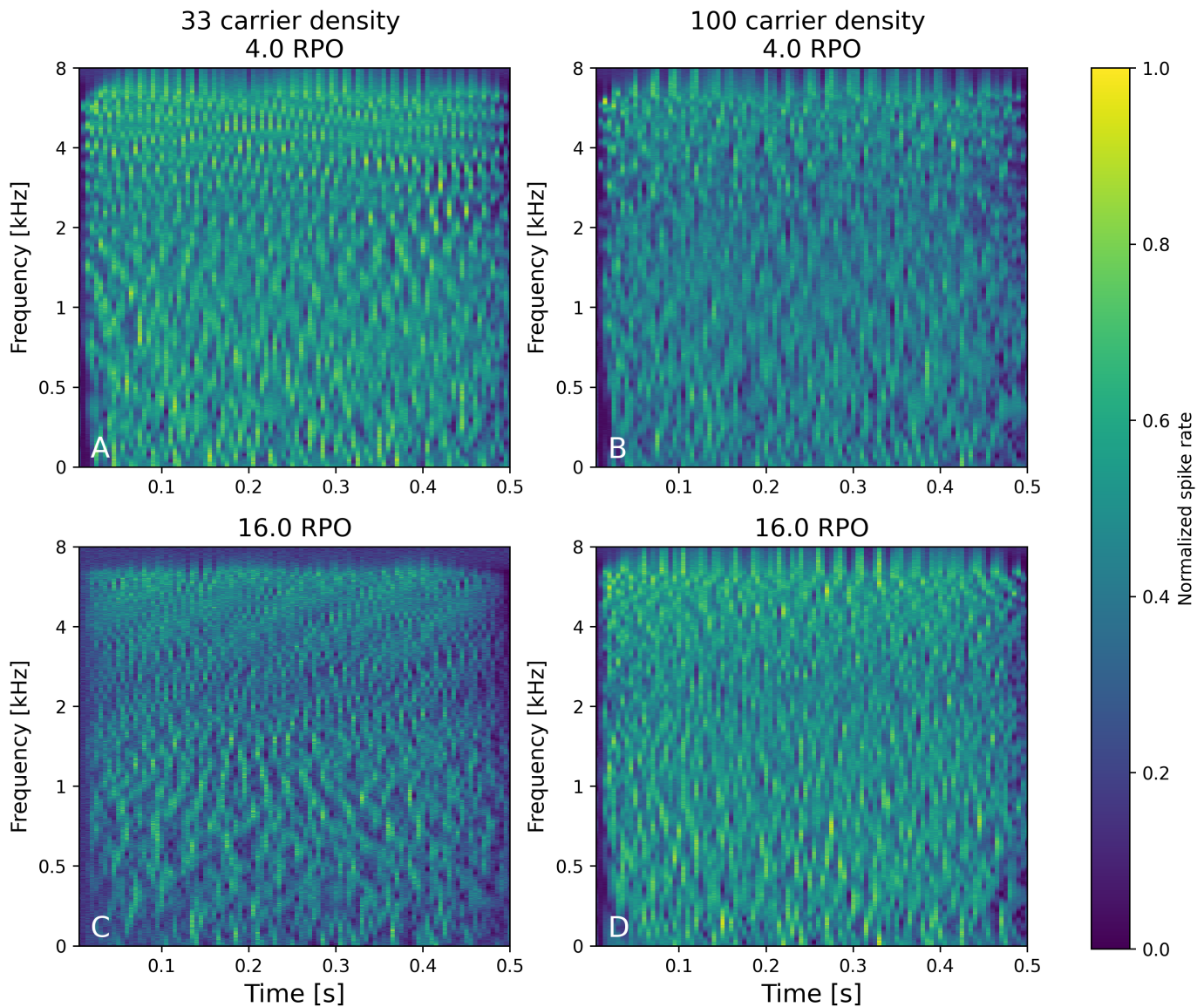


Figure 10. Neurograms of 4.0 (A,B) and 16.0 (C,D) RPO with two different carrier densities ((A,C) 33.333 CPO, (B,D) 100 CPO) in normal hearing. Brighter colors indicate greater spiking.

The neurograms created with the electric hearing pipeline had a lower ripple density in order to be in the perceptible range of CI users. Figure 11 shows the neurograms with 1.0 RPO (Figure 11A), 2.0 RPO (Figure 11B), and 3.0 RPO (Figure 11C). At 1.0 RPO and 2.0 RPO, the ripples were visible and had a smooth course due to current steering. At 3.0 RPO, the ripples were difficult to discern. They were vaguely visible between 3 kHz and 4 kHz. Considering that 6500 Hz falls within the highest frequency band of channel 15 (4248–8054 Hz), the current will be allocated to electrode 16, inducing neural excitation in these high frequencies.

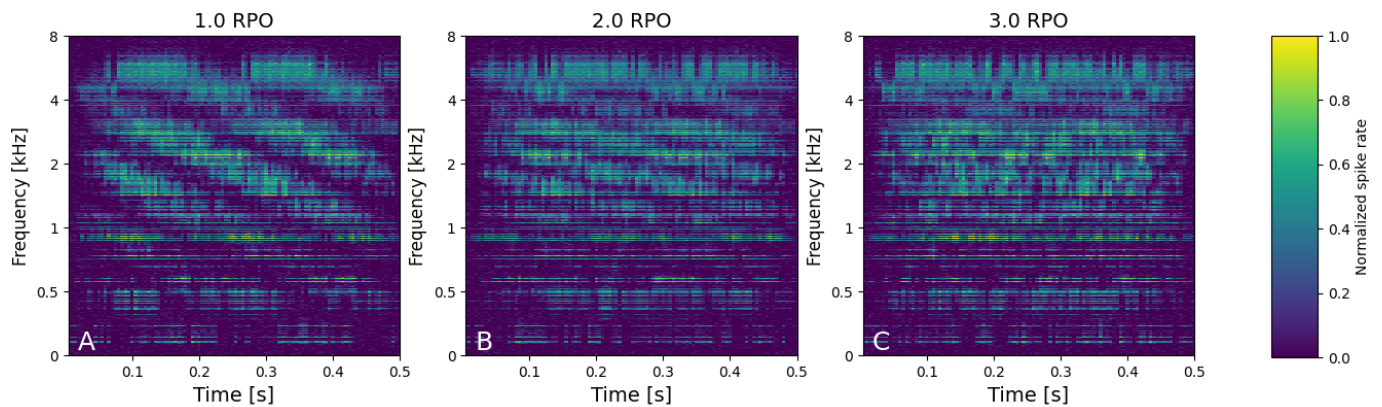


Figure 11. Neurograms with electric hearing of: (A) 1.0 RPO, (B) 2.0 RPO, and (C) 3.0 RPO. The fiber frequencies on the y-axis are based on the assumed new frequencies after neural plasticity. The stimuli with a carrier density of 100 CPO were used.

4. Discussion

Here, we have studied how the stimuli of the spectral ripple test and SMRT are processed with models of normal hearing and hearing with a CI. The models used here have demonstrated the anomalies in the clinical results of these tests.

Based on the information conveyed through the speech processor (Figure 6), the original spectral ripple test by Won et al. [10] is limited in the transmission of peaks and valleys for higher ripple densities. This aligns with the findings and hypotheses of earlier studies [11–13]. Moreover, unlike in the acoustic case, for CIs, spectral aliasing is visible at 4.0 RPO due to the limited spectral resolution of the filter bands, in the form of a slower wave pattern of the electric content. In other words, the CI was unable to follow the higher ripple density of the spectral ripples, resulting in the presentation of a ripple of lower ripple density. This result corroborates the theory postulated by Winn and O'Brien [14] that the speech processor cannot follow above 2.0 RPO, which also turns out to be the visible limit in our results. Compared to Winn and O'Brien [14], we used a more realistic method to portray the electric content; they used an idealized approach with ideal filters as square windows on an ideal ripple spectrum, whereas we used the true stimuli and the research version of a speech processor. Moreover, they simulated the pattern of excitation with a 2 dB/mm roll-off, whereas we included an implanted cochlea model with 3D volume conduction. This limitation of 2 RPO is also seen in the neural activation in Figure 7, where the peaks are no longer distinguishable at 4.0 RPO. However, the aliasing is not visible in the neural activation. The spectrum results are congruent with the performances of CI users and normal hearing subjects. Won et al. [10] found an average ripple threshold of 1.73 RPO in CI users, with a range of 0.6 to 4.87 RPO. The majority of CI users in this study used the ACE (Advanced Combination Encoder) strategy and the remainder used SPEAK (spectral-peak speech coding strategy), CIS (Continuous Interleaved Sampling), or HiRes [10]. In a similar paradigm, Henry et al. [44] found a spectral ripple detection average of 1.77 RPO for CI users and 4.84 RPO for normal hearing participants. However, these stimuli were also slightly different from the stimuli of Won et al. [10] in terms of spectral shape (sinusoidal shape on a linear amplitude axis versus on a logarithmic axis). In another similar paradigm, Anderson et al. [11] found an average ripple discrimination threshold of 1.68 RPO (range of 0.41 to 4.27 RPO), but with cutoff frequencies of the noise pass band at 350 and 5600 Hz and a stimulus duration of 400 ms. Drennan et al. [45] reported a higher average ripple discrimination threshold of 2.60 RPO with a range of 0.63 to 7.03 RPO over their CI users. In this study, a more diverse selection of strategies was included: ACE, CIS, SPEAK, HiRes, HiRes with F120, SAS (Simultaneous Analogue

Stimulation), FSP (Fine Structure Processing), and CCIS (conditioned CIS). The electric hearing pipeline has also shown the benefit of the F120 strategy in the case of spectral ripples, as the peaks at 2.828 RPO are better discriminable than with current steering turned off (Figure 8). Similarly, this improved discrimination was visible in our psychometric curves derived from the filtered neural spectra (Figure 9). Drennan et al. [15] had a similar finding when they compared the performance of the F120 strategy and the HiRes strategy in the spectral ripple test. The HiRes strategy employs real-time filtering with 16 logarithmically spaced bandpass filters and electrodes, whereas the F120 strategy uses an FFT filter bank and current steering over 15 frequency bands between 16 electrodes. They found a mean threshold of 3.42 RPO for the F120 strategy and 2.31 RPO for the HiRes strategy. Without any optimization or additional parameters, our results are in the same ballpark, with 3.7 RPO for F120 and 2.2 RPO for F120 without current steering. However, the strategy without current steering is not the same strategy as ours, as it does not make use of FFT and includes 16 electrodes. Our strategy without current steering is based on the HiRes FFT strategy used by de Jong et al. [30] and setting the weights to 0.5 with 15 frequency bands, resulting in 15 unique peak positions. The channel-specific temporal envelopes can be more faithfully represented in the original HiRes strategy, but a spectrum is unlikely to give us information on this difference. Nonetheless, de Jong et al. [30] found no difference in SMRT performance between the original HiRes strategy with a bandpass filter bank and their HiRes FFT strategy.

These tests and models have given insights into the differences in the spectral resolution of normal hearing and electric hearing. The spectral resolution, in terms of ripple visibility along the spectrum or ripple visibility in the neurograms, is discernible up to higher ripple densities in normal hearing. Especially above 4000 Hz, greater discriminative power was discerned for normal hearing. This analysis band in the F120 strategy spans the largest frequency range and thus provides the least frequency resolution with current steering. In the neural spectrum for the spectral ripples, a difference in smoothness is similarly evident. The ripples in the electric hearing spectrum are jagged compared with those in the normal hearing spectrum. This is also visible in the filtered response, which varied in distribution. The threshold profiles created for this array are of a similar choppy nature [25]. As postulated by Narne et al. [19], amplitude modulation cues at the outputs of auditory filtering centered at the upper edge of or above the frequency range may be the reason why the ripple threshold is higher for SMRT than the spectral ripple test they used. In our neurograms, amplitude modulations above the frequency limit of SMRT were also visible (Figure 10).

The models have provided an interesting visualization of the perception of SMRT stimuli. In the normal hearing neurograms, the ripples were visible at both carrier densities for 4.0 RPO (Figure 10A,B). Furthermore, the aliasing artefact was visible at 16.0 RPO with a carrier density of 33.333 CPO (Figure 10C), which is consistent with the spectrogram of Resnick et al. [18]. For electric hearing, the ripple discrimination in neurograms, in terms of ripple visibility, was reduced compared to normal hearing (Figure 11). The ripples were hardly visible at 3.0 RPO, which is similar to what has been found among CI users in our department [46], whereas with normal hearing, these ripples were still clearly visible at 4.0 RPO. We can infer from these neurograms that the current steering strategy is capable of a more precise distribution of the peak in this range with 1.0 RPO (e.g., between 0.2 s and 0.3 s), but not with greater ripple density. At 3.0 RPO, the ripples are barely visible, except for the range between, for instance, 3 kHz and 4 kHz. In this regard, another cue that has been overlooked in the SMRT is that the CI participant may selectively use the difference in amplitude modulation within one channel, instead of the full spectrum, to identify the odd one out [47]. To address this cue, a new spectral resolution test was developed by Archer-

Boyd et al. [47], the Spectro-Temporal Ripple for Investigating Processor Effectiveness (STRIPES). In this test, listeners must distinguish between complexes of sine downward and upward sweeps. The difficulty is increased by increasing the spectro-temporal density of the stimulus. In addition, noise bookends were added to avoid onset and offset cues. However, analysis of this psychophysical test is beyond the scope of this paper but will be assessed in future research.

Several limitations of this study should be addressed. Both neural models used in this paper have only been validated on single-fiber data, whereas the output here is extrapolated to the entire cochlea. However, the first two components of the implanted cochlea model (Figure 1) are based on the entire cochlea. In the configuration of the normal hearing model used here, the frequency resolution was relatively high, and we might have overestimated the resolution of the filters that reflect the basilar membrane response. Similarly, we may have overestimated the extent to which a CI user learns to adapt to the new frequency allocation of the array, or we may have misaligned the fibers to the frequency allocation. Moreover, we assumed that the information we can gather visually is similarly processed by or distinguishable to the ear and auditory cortex. The extraction of processed information would be more objective with the addition of an interpretation model, such as the one described by Hamacher [48], but the discrimination between ripple densities of the SMRT was not that straightforward because it did not yield a monotonic decrease in the model performance. In other words, a higher ripple density did not translate to a trend of more difficult discrimination, as was possible with the spectral ripple test shown here in the psychometric curves (Figure 9). Similarly, with the d' implementation we were not able to find such a pattern for the modeled SMRT neurograms. Again, this implies that for such intricate signals, a more elaborate interpretation model is necessary to replicate cortical processing. Alternatively, a neural network approach could produce better results. Nonetheless, the models here corroborate clinical results, giving the impression that they work well.

The pipeline presented in this paper, with normal and electric hearing, offers a tool for assessing other psychophysical tests, such as STRIPES, or for future psychophysical tests before assessment in participants. With this pipeline, unintentional cues present in the stimuli can be filtered out in advance. In addition, this pipeline could be a tool for assessing future speech coding strategies. The current degree of temporal and spectral resolution of the information conveyed by speech coding strategies leaves room for improvement. The pipeline with both normal hearing and electric hearing allows for an easy first assessment of a new speech coding strategy 'in silico'. Another purpose for this pipeline could be to evaluate the characteristics of the CI or the speech coding strategy, such as the FFT assignment or filter banks to electrodes. Using artificial intelligence, a speech coding strategy could be optimized based on the difference between the outputs of the two models. If the neural activation of electric hearing is as similar to that of normal hearing as possible, we hypothesize that the sound experienced by CI users will be improved and more similar to that of normal hearing.

Author Contributions: Conceptualization, S.S.M.M.; methodology, S.S.M.M.; software, S.S.M.M.; validation, S.S.M.M.; formal analysis, S.S.M.M.; investigation, S.S.M.M.; resources, S.S.M.M.; data curation, S.S.M.M.; writing—original draft preparation, S.S.M.M.; writing—review and editing, J.J.B. and J.H.M.F.; visualization, S.S.M.M.; supervision, J.J.B. and J.H.M.F.; project administration, S.S.M.M.; funding acquisition, J.J.B. and J.H.M.F. All authors have read and agreed to the published version of the manuscript.

Funding: The collaboration project TEMPORAL is co-funded by PPP Allowance awarded by Health~Holland, Top Sector Life Sciences & Health, to stimulate public-private partnerships (grant number LSHM20101).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The PHAST+ model is available at <https://github.com/jacobdenobel/PHAST> (accessed on 21 July 2025), The code of the used implementation of the HiRes Fidelity 120 strategy is available at <https://github.com/jacobdenobel/abt> (accessed on 21 July 2025). Data and code for this specific evaluation of the electric hearing model will be made available on request. The normal hearing model of Bruce et al. [22] can be found at <https://doi.org/10.1016/j.heares.2017.12.016>. The slightly adjusted code for this model, with our settings, will be made available upon request as well.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

3AFC	3-alternative forced choice
ACE	Advanced Combination Encoder
CPO	Carriers-per-octave
CI	Cochlear implant
CIS	Continuous interleaved sampling
F120	Fidelity 120
FFT	Fast Fourier transform
FS	Full scale
HiRes	High Resolution
RMS	Root mean square
RPO	Ripples-per-octave
SMRT	Spectral-temporally modulated ripple test
STRIPES	Spectro-temporal ripple for investigating processor effectiveness
T	Threshold

Appendix A. Electric Hearing: Neural Parameters

Table A1. The included neural parameters for the electric hearing pipeline. The parameters are listed with their abbreviation or variable as listed in the reference. SD: standard deviation.

Parameter	Value (SD)	Reference
Absolute refractory period (ARP)	0.0004 s (0.0001 s)	[24]
Relative refractory period (RRP)	0.0008 s (0.0005 s)	[24]
Relative spread (RS)	0.06 (0.04)	[24]
Accommodation rate (θ_a)	0.014	[28]
Adaptation rate (θ_s)	19.996	[28]
Accommodation amplitude (ω_a)	0.072	[28]
Adaptation amplitude (ω_s)	7.142	[28]
Fiber spacing	9.6 μm (0.46 μm)	[25]
Spontaneous rate (SR)	10 Hz	

References

1. Budenz, C.L.; Cosetti, M.K.; Coelho, D.H.; Birenbaum, B.; Babb, J.; Waltzman, S.B.; Roehm, P.C. The effects of cochlear implantation on speech perception in older adults. *J. Am. Geriatr. Soc.* **2011**, *59*, 446–453. [CrossRef]
2. Calmels, M.N.; Saliba, I.; Wanna, G.; Cochard, N.; Fillaux, J.; Deguine, O.; Fraysse, B. Speech perception and speech intelligibility in children after cochlear implantation. *Int. J. Pediatr. Otorhinolaryngol.* **2004**, *68*, 347–351. [CrossRef]
3. Caldwell, A.; Nittrouer, S. Speech perception in noise by children with cochlear implants. *J. Speech Lang. Hear. Res.* **2013**, *56*, 13–30. [CrossRef]

4. Fetterman, B.L.; Domico, E.H. Speech recognition in background noise of cochlear implant patients. *Otolaryngol.—Head Neck Surg.* **2002**, *126*, 257–263. [[CrossRef](#)]
5. Riley, P.E.; Ruhl, D.S.; Camacho, M.; Tolisano, A.M. Music appreciation after cochlear implantation in adult patients: A systematic review. *Otolaryngol. Head Neck Surg.* **2018**, *158*, 1002–1010. [[CrossRef](#)] [[PubMed](#)]
6. Irvine, D.R. Plasticity in the auditory system. *Hear. Res.* **2018**, *362*, 61–73. [[CrossRef](#)] [[PubMed](#)]
7. Reiss, L.A.; Turner, C.W.; Erenberg, S.R.; Gantz, B.J. Changes in pitch with a cochlear implant over time. *J. Assoc. Res. Otolaryngol.* **2007**, *8*, 241–257. [[CrossRef](#)] [[PubMed](#)]
8. Bernstein, L.R.; Green, D.M. Detection of changes in spectral shape: Uniform vs. non-uniform background spectra. *Hear. Res.* **1988**, *34*, 157–165. [[CrossRef](#)]
9. Henry, B.A.; Turner, C.W. The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. *J. Acoust. Soc. Am.* **2003**, *113*, 2861–2873. [[CrossRef](#)]
10. Won, J.H.; Drennan, W.R.; Rubinstein, J.T. Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. *J. Assoc. Res. Otolaryngol.* **2007**, *8*, 384–392. [[CrossRef](#)]
11. Anderson, E.S.; Nelson, D.A.; Kreft, H.; Nelson, P.B.; Oxenham, A.J. Comparing spatial tuning curves, spectral ripple resolution, and speech perception in cochlear implant users. *J. Acoust. Soc. Am.* **2011**, *130*, 364–375. [[CrossRef](#)] [[PubMed](#)]
12. Anderson, E.S.; Oxenham, A.J.; Nelson, P.B.; Nelson, D.A. Assessing the role of spectral and intensity cues in spectral ripple detection and discrimination in cochlear-implant users. *J. Acoust. Soc. Am.* **2012**, *132*, 3925–3934. [[CrossRef](#)] [[PubMed](#)]
13. Gifford, R.H.; Noble, J.H.; Camarata, S.M.; Sunderhaus, L.W.; Dwyer, R.T.; Dawant, B.M.; Dietrich, M.S.; Labadie, R.F. The relationship between spectral modulation detection and speech recognition: Adult versus pediatric cochlear implant recipients. *Trends Hear.* **2018**, *22*, 2331216518771176. [[CrossRef](#)]
14. Winn, M.B.; O'Brien, G. Distortion of spectral ripples through cochlear implants has major implications for interpreting performance scores. *Ear Hear.* **2022**, *43*, 764–772. [[CrossRef](#)] [[PubMed](#)]
15. Drennan, W.R.; Won, J.H.; Nie, K.; Jameyson, E.; Rubinstein, J.T. Sensitivity of psychophysical measures to signal processor modifications in cochlear implant users. *Hear. Res.* **2010**, *262*, 1–8. [[CrossRef](#)]
16. Aronoff, J.M.; Landsberger, D.M. The development of a modified spectral ripple test. *J. Acoust. Soc. Am.* **2013**, *134*, EL217–EL222. [[CrossRef](#)]
17. Azadpour, M.; McKay, C.M. A psychophysical method for measuring spatial resolution in cochlear implants. *J. Assoc. Res. Otolaryngol.* **2012**, *13*, 145–157. [[CrossRef](#)]
18. Resnick, J.M.; Horn, D.L.; Noble, A.R.; Rubinstein, J.T. Spectral aliasing in an acoustic spectral ripple discrimination task. *J. Acoust. Soc. Am.* **2020**, *147*, 1054–1058. [[CrossRef](#)]
19. Narne, V.K.; Sharma, M.; Van Dun, B.; Bansal, S.; Prabhu, L.; Moore, B.C. Effects of spectral smearing on performance of the spectral ripple and spectro-temporal ripple tests. *J. Acoust. Soc. Am.* **2016**, *140*, 4298–4306. [[CrossRef](#)]
20. Lawler, M.; Yu, J.; Aronoff, J.M. Comparison of the spectral-temporally modulated ripple test with the Arizona Biomedical Institute Sentence Test in cochlear implant users. *Ear Hear.* **2017**, *38*, 760–766. [[CrossRef](#)]
21. Jeddi, Z.; Lotfi, Y.; Moossavi, A.; Bakhshi, E.; Hashemi, S.B. Correlation between auditory spectral resolution and speech perception in children with cochlear implants. *Iran. J. Med. Sci.* **2019**, *44*, 382.
22. Bruce, I.C.; Erfani, Y.; Zilany, M.S. A phenomenological model of the synapse between the inner hair cell and auditory nerve: Implications of limited neurotransmitter release sites. *Hear. Res.* **2018**, *360*, 40–54. [[CrossRef](#)]
23. Kalkman, R.K.; Briaire, J.J.; Frijns, J.H. Current focussing in cochlear implants: An analysis of neural recruitment in a computational model. *Hear. Res.* **2015**, *322*, 89–98. [[CrossRef](#)] [[PubMed](#)]
24. van Gendt, M.J.; Briaire, J.J.; Kalkman, R.K.; Frijns, J.H. A fast, stochastic, and adaptive model of auditory nerve responses to cochlear implant stimulation. *Hear. Res.* **2016**, *341*, 130–143. [[CrossRef](#)]
25. Kalkman, R.K.; Briaire, J.J.; Dekker, D.M.; Frijns, J.H. The relation between polarity sensitivity and neural degeneration in a computational model of cochlear implant stimulation. *Hear. Res.* **2022**, *415*, 108413. [[CrossRef](#)] [[PubMed](#)]
26. Briaire, J.J.; Frijns, J.H. 3D mesh generation to solve the electrical volume conduction problem in the implanted inner ear. *Simul. Pract. Theory* **2000**, *8*, 57–73. [[CrossRef](#)]
27. Frijns, J.H.; Briaire, J.J.; Schoonhoven, R. Integrated use of volume conduction and neural models to simulate the response to cochlear implants. *Simul. Pract. Theory* **2000**, *8*, 75–97. [[CrossRef](#)]
28. de Nobel, J.; Martens, S.S.; Briaire, J.J.; Bäck, T.H.; Kononova, A.V.; Frijns, J.H. Biophysics-inspired spike rate adaptation for computationally efficient phenomenological nerve modeling. *Hear. Res.* **2024**, *447*, 109011. [[CrossRef](#)]
29. Nogueira, W.; Litvak, L.; Edler, B.; Ostermann, J.; Büchner, A. Signal processing strategies for cochlear implants using current steering. *EURASIP J. Adv. Signal Process.* **2009**, *2009*, 531213. [[CrossRef](#)]
30. de Jong, M.A.; Briaire, J.J.; Frijns, J.H. Take-Home Trial Comparing Fast Fourier Transformation-Based and Filter Bank-Based Cochlear Implant Speech Coding Strategies. *BioMed Res. Int.* **2017**, *2017*, 7915042. [[CrossRef](#)]

31. van der Jagt, M.A.; Briaire, J.J.; Verbist, B.M.; Frijns, J.H. Comparison of the HiFocus Mid-Scala and HiFocus 1J electrode array: angular insertion depths and speech perception outcomes. *Audiol. Neurotol.* **2017**, *21*, 316–325. [[CrossRef](#)]
32. Aronoff, J.M.; Staisloff, H.E.; Kirchner, A.; Lee, D.H.; Stelmach, J. Pitch matching adapts even for bilateral cochlear implant users with relatively small initial pitch differences across the ears. *J. Assoc. Res. Otolaryngol.* **2019**, *20*, 595–603. [[CrossRef](#)] [[PubMed](#)]
33. Reiss, L.A.; Turner, C.W.; Karsten, S.A.; Gantz, B.J. Plasticity in human pitch perception induced by tonotopically mismatched electro-acoustic stimulation. *Neuroscience* **2014**, *256*, 43–52. [[CrossRef](#)]
34. Carlyon, R.P.; Macherey, O.; Frijns, J.H.; Axon, P.R.; Kalkman, R.K.; Boyle, P.; Baguley, D.M.; Briggs, J.; Deeks, J.M.; Briaire, J.J.; et al. Pitch comparisons between electrical stimulation of a cochlear implant and acoustic stimuli presented to a normal-hearing contralateral ear. *J. Assoc. Res. Otolaryngol.* **2010**, *11*, 625–640. [[CrossRef](#)]
35. de Nobel, J.; Briaire, J.J.; Baeck, T.H.; Kononova, A.V.; Frijns, J.H. From Spikes to Speech: NeuroVoc—A Biologically Plausible Vocoder Framework for Auditory Perception and Cochlear Implant Simulation. *arXiv* **2025**, arXiv:2506.03959.
36. Shera, C.A.; Guinan, J.J., Jr.; Oxenham, A.J. Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 3318–3323. [[CrossRef](#)]
37. Byrne, D.; Dillon, H.; Tran, K.; Arlinger, S.; Wilbraham, K.; Cox, R.; Hagerman, B.; Hetu, R.; Kei, J.; Lui, C.; et al. An international comparison of long-term average speech spectra. *J. Acoust. Soc. Am.* **1994**, *96*, 2108–2120. [[CrossRef](#)]
38. Simpson, A.J.; Fitter, M.J. What is the best index of detectability? *Psychol. Bull.* **1973**, *80*, 481. [[CrossRef](#)]
39. Das, A.; Geisler, W.S. Methods to integrate multinormals and compute classification measures. *arXiv* **2020**, arXiv:2012.14331.
40. Macmillan, N.A.; Creelman, C.D. *Detection Theory: A User's Guide*; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 2005.
41. DeCarlo, L.T. On a signal detection approach to m-alternative forced choice with bias, with maximum likelihood and Bayesian approaches to estimation. *J. Math. Psychol.* **2012**, *56*, 196–207. [[CrossRef](#)]
42. Green, D.M.; Swets, J.A. *Signal Detection Theory and Psychophysics*; Wiley New York: New York, NY, USA, 1966; Volume 1.
43. Kingdom, A.A.; Prins, N. *Psychophysics: A Practical Introduction*; Academic Press: Cambridge, MA, USA, 2016.
44. Henry, B.A.; Turner, C.W.; Behrens, A. Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners. *J. Acoust. Soc. Am.* **2005**, *118*, 1111–1121. [[CrossRef](#)] [[PubMed](#)]
45. Drennan, W.R.; Anderson, E.S.; Won, J.H.; Rubinstein, J.T. Validation of a clinical assessment of spectral-ripple resolution for cochlear implant users. *Ear Hear.* **2014**, *35*, e92–e98. [[CrossRef](#)] [[PubMed](#)]
46. Knipscheer, B.; Briaire, J.J.; Frijns, J.H. Characterization of a Psychophysical Test Battery for the Evaluation of Novel Speech Coding Strategies in Cochlear Implants. In Proceedings of the Abstracts of the Conference on Implantable Auditory Prosthesis. Conference on Implantable Auditory Prosthesis, Tahoe, CA, USA, 9–14 July 2023; p. 194.
47. Archer-Boyd, A.W.; Southwell, R.V.; Deeks, J.M.; Turner, R.E.; Carlyon, R.P. Development and validation of a spectro-temporal processing test for cochlear-implant listeners. *J. Acoust. Soc. Am.* **2018**, *144*, 2983–2997. [[CrossRef](#)] [[PubMed](#)]
48. Hamacher, V. Signalverarbeitungsmodelle des Elektrisch Stimulierten Gehörs. Ph.D. Thesis, RWTH Aachen, Aachen, Germany, 2004.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.