Auditory-nerve Model including Efferent Dynamic Gain Control with Inputs from Cochlear Nucleus and Inferior Colliculus

Afagh Farhadi,[1] Skyler G. Jennings,[2] Elizabeth A. Strickland,[3] and Laurel H. Carney[4,a]

[1] *Departments of Electrical and Computer Engineering, University of Rochester, Rochester, New York 14642, USA*

[2] *Department of Communication Sciences and Disorders, University of Utah, Salt Lake City, Utah 84112, USA*

[3] *Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, Indiana 47907, USA*

[4] *Departments of Biomedical Engineering and Neuroscience, University of Rochester, Rochester, New York 14642, USA*

The focus of most existing auditory models is on the afferent system. The auditory efferent system contains descending projections from several levels of the auditory pathway, from the auditory cortex to the brainstem, that control gain in the cochlea. We developed a model with a time-varying, gain-control signal from the efferent system that includes sub-cortical ascending and descending neural pathways. The medial olivocochlear (MOC) efferent stage of the model receives excitatory projections from both fluctuation-sensitive neurons in the inferior colliculus (IC) and wide-dynamic-range neurons in the cochlear nucleus (CN). The response of the model MOC stage controlled cochlear gain dynamically. Changes in the rates of IC neurons in awake rabbit to long-duration amplitude-modulated (AM) noise were employed to adjust the parameters of the proposed model. In response to AM stimuli, physiological response rates of most IC neurons with band-enhanced (BE) modulation transfer functions (MTFs) increased over a time course consistent with the dynamics of the MOC efferent feedback. The time constant of the MOC model that best matched the IC physiology was compared to available descriptions of the MOC. Responses of the proposed

[a] Laurel.Carney@Rochester.edu

subcortical model to AM noise simulate the trend of increasing rate over time, while the model without the efferent system did not show this trend.

I. INTRODUCTION

Efferent feedback from the central auditory nervous system to the cochlea has been hypothesized to improve auditory perception in noise, including improvements in speech intelligibility in background noise (de Boer & Thornton, 2008; Ghitza et al., 2007; Giraud et al., 1997; Winslow & Sachs, 1988). Understanding speech in background noise is a challenging problem for all listeners, especially those with hearing loss, but the underlying mechanisms, even for normal hearing listeners, are yet to be revealed (Holmes & Griffiths, 2019). An accurate model for subcortical auditory pathways, including the efferent feedback system, would guide investigations of the mechanisms of hearing perception in background noise.

The dynamics of the efferent system support its potential role in improving speech perception. Specifically, the latency and time constant of the MOC efferent system (Backus & Guinan, 2006; Roverud & Strickland, 2010; Warren III & Liberman, 1989) are comparable to typical phonemic durations (Jacewicz et al., 2007; Umeda, 1977). Therefore, the MOC efferent system would presumably affect auditory neural responses both within the time course of phonemes and also in running speech, for which the perception of a phoneme would be influenced by the previous phoneme (co-articulation) (C. Chambers et al., 2017; Recasens, 1984, 2018). Despite the potential beneficial role of this feedback system, it is still poorly understood (Jennings, 2021).

A detailed, accurate, and comprehensive model capable of simulating auditory-nerve (AN) responses to arbitrary stimuli is a powerful tool for testing hypotheses, leading to insights on mechanisms in hearing (Meddis et al., 2010). In addition, cochlear implants, hearing aids, and automatic speech recognition devices could also benefit from improved auditory models (Dávila-Chacón et al., 2018; Saeedi et al., 2014; Zouhir & Ouni, 2016). Recent studies of speech enhancement have focused on improving the acoustic waveform and testing the improvement using neural-network-based models for the auditory cortex (Fu et al., 2017; K. Tan & Wang, 2018);

however, this method overlooks the role of the physiological nonlinearities in the auditory periphery (Drakopoulos et al., 2022; Zaar & Carney, 2022). Speech recognition systems that include a physiologically realistic model of the auditory periphery have been shown to be more robust in noisy environments than the model that does not include the physiological properties of the auditory system (Stern et al., 2012; Stern & Morgan, 2012).

Several models for AN responses exist that include detailed representations of the periphery, including nonlinearities such as cochlear compression, inner-hair-cell (IHC) transduction saturation, and AN rate saturation. However, most of these models do not include the efferent pathway(Rahman et al., 2019; Verhulst et al., 2012; Zilany et al., 2014; Zweig, 2015; Lopez-Poveda et al., 2016, 2017).), which is at least as large as the ascending pathway in terms of numbers of anatomical projections (Schofield, 2011). The efferent system consists of several feedback projections, spanning all levels of the auditory pathways from the auditory cortex to the cochlea (Guinan, 2006; Warr, 1992). The olivocochlear efferents, the projections from the brainstem to the cochlea, include two systems (Guinan, 2006): lateral olivocochlear (LOC) and medial olivocochlear (MOC) efferents. The MOC efferent system, the focus of this study, controls cochlear gain, and has been hypothesized to improve hearing in noisy backgrounds (Christopher Kirk & Smith, 2003; S. Kim et al., 2006; Mertes et al., 2018, 2019). The MOC efferents receive several convergent inputs; the model presented here will focus on subcortical pathways, specifically inputs from the IC and cochlear nucleus (CN). MOC neurons modulation transfer functions (MTFs) (Gummer et al., 1988a; Huffman & Henson Jr, 1990; Schofield, 2011) from these inputs. MOC efferents also receive ascending projections from several types of cells in the CN that have wide dynamic ranges, including the small cell cap and T-stellate cells (Guinan Jr, 2018; Romero & Trussell, 2021; Ye et al., 2000) (Fig. 1). MOC neurons have tuning curves that are similar in shape, but slightly wider, than AN fibers (M. C. Brown, 2016; Guinan Jr, 2018; Wilson et al., 1991). The MOC neurons project to the

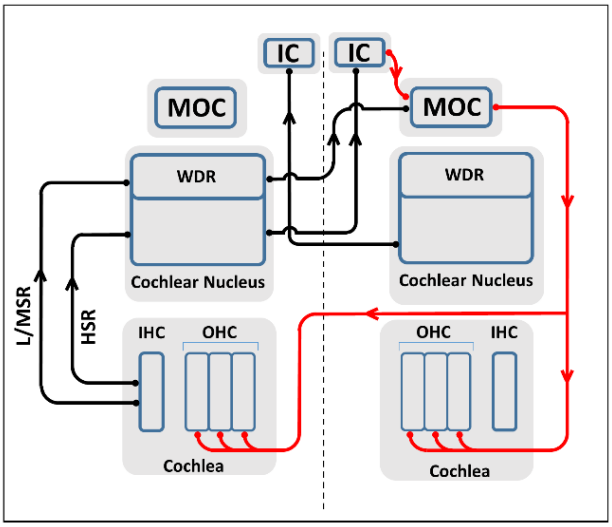


FIG. 1 Schematic illustration of the peripheral afferent (black arrows) and MOC efferent (red arrows) pathways. Here for a simple presentation we have shown the afferent pathways on the left side and the efferent pathways on the right side, but anatomically both pathways exist on both sides of the brainstem.

outer hair cells (OHCs), and this projection is tonotopic along the cochlea (M. C. Brown, 2016). MOC efferent activity suppresses cochlear gain, often described in terms of two time frames (Guinan Jr, 2018). First, the fast effect of the MOC efferents has rise and fall times of approximately 100 ms for most mammals, including humans (Backus & Guinan, 2006; Roverud & Strickland, 2010; Warren III & Liberman, 1989). Second, the slow effect of the MOC efferents occurs over a time course of tens of seconds, and this effect is believed to be relatively small in humans (Guinan, 2018).

Despite several studies of the physiology of the efferent system (Guinan, 2006; Gummer et al., 1988; Liberman & Guinan, 1998; Schofield, 2011), many aspects are not well understood. Studying the efferent system is challenging due to the complex projections within this pathway (Schofield, 2011). Additionally, there exist limited experimental methods for understanding the efferent system. Most of the information available is based on a limited number of recordings from single neurons located deep in the brainstem in anesthetized animals (Guinan, 2018); however, anesthesia

suppresses MOC responses (Aedo et al., 2015; A. R. Chambers et al., 2012; Guitton et al., 2004). Noninvasive methods used to study MOC effects on cochlear gain is to measure either otoacoustic emissions (OAEs) (Guinan, 2006) or cochlear microphonic (Jamos et al., 2020, 2021) with and without MOC activation by electrical stimulation or contralateral broadband noise, respectively. This method, however, does not provide information about higher-level projections in the pathway, such as from the IC or auditory cortex to the MOC (Delano et al, 2007). Furthermore, the middle-ear reflex also affects OAEs (Sun, 2008). The proposed model of the MOC efferent system may help in designing future experiments for studying efferent function, for example using recordings from structures that are accessible in awake animals, such as the midbrain (Kuwada et al., 1987).

Several previous AN models include efferent signals driven by the periphery (N. R. Clark et al., 2012; Giguere & Woodland, 1994; Kwan et al., 2019; Smalt et al., 2014; Yasin et al., 2020). In some of these previous implementations, cochlear gain remains constant over time, omitting the dynamics of MOC efferents (G. J. Brown et al., 2010; Ferry & Meddis, 2007; Giguere & Woodland, 1994; Jennings et al., 2011). Smalt et al. (2014) proposed a model based on Zilany et al. (2014) but with an efferent system modeled by a time-varying gain-control feedback system.However, in this model, a local gain-control system was implemented based on the output of model OHCs, which is inconsistent with the overall structure of the MOC efferent system (Smalt et al., 2014; Schofield, 2011). None of these models included the large descending input to the MOC efferent system from the midbrain (Schofield, 2011). The focus of this study is to include both peripheral and midbrain inputs to a dynamic MOC efferent system in a model for AN responses.

In this work, we propose a dynamic, sample-by-sample based, MOC efferent gain-control system with feedback loops from CN neurons and from IC neurons with band-enhanced (BE) MTFs. The ascending pathway from the CN is characterized by a wide dynamic range, driven by low-spontaneous-rate (LSR) AN fibers. This input represents both the small-cell-cap and T-stellate

cell projections to the MOC, as the average rate dynamic ranges of both types of neurons are similar to those of low- or medium-spontaneous-rate AN fibers (Blackburn & Sachs, 1990; Romero & Trussell, 2021). It is often assumed that AN average-rates profiles encode the spectrum of complex sounds, thus previous models for the MOC efferent system have AN rate as the primary control input. However, the majority of AN fibers, the HSR fibers, have average rates that are saturated at conversational speech levels (55–65 dB SPL, Olsen, 1998). Therefore, level-dependent neural discharge rates of HSR AN fibers do not include enough information to explain perception of complex sounds (Bharadwaj et al., 2014; Carney, 2018; Liberman, 1978). On the other hand, temporal fluctuations of HSR AN responses, referred to as neural fluctuations, have robust information at conversational speech levels, potentially encoding peaks and valleys in the stimulus spectrum because the amplitudes of fluctuations are lower for AN fibers tuned near peak frequencies and higher for other frequency channels (Carney et al., 2015; Carney, 2018). If AN response fluctuations play a key role in encoding complex sounds, then these fluctuations should also play a key role in the gain control system. The descending projection to the MOC neuronsfrom IC neurons, which are sensitive to the amplitudes of low-frequency fluctuations of their inputs (Carney, 2018; P X Joris et al., 2004; Krishna & Semple, 2000; Nelson & Carney, 2007), provides this information to the MOC gain control system. Re-examination of IC responses to amplitude-modulated (AM) noise revealed rates that vary over a time course that is consistent with the dynamics of MOC efferent gain control (Farhadi et al., 2021).

Here, we refine and test an AN model with both fluctuation- and wide-dynamic-range-driven gain control (Farhadi et al., 2021). First, we will review the modifications required to change the structure of previous AN models from processing the entire stimulus duration, stage-to-stage, to a model using sample-by-sample processing, as required to implement a dynamic feedback system. Then, we will introduce the novel control system used to implement the MOC efferent pathway.

The next step was to adjust the parameters of the MOC efferent model based on a physiological dataset from a previous study of IC responses to AM noise (Carney et al., 2014). Finally, we show that the model with efferents outperforms the model without efferents in simulating the dynamics of physiological responses in the IC to AM stimuli. This finding, based on the simulation result with and without activation of the efferent system, supports the hypothesis that the efferent system is a potential mechanism underlying the increase in rate trend observed in IC cell responses to AM noise.

II. METHOD

We adapted existing models to simulate responses of AN fibers and IC cells in order to implement the proposed model with MOC efferent gain control. First, we briefly introduce the foundational AN and IC models. Then, we describe the modifications that were applied, including changes to the structure of the existing models and new extensions added to simulate the MOC efferent system. Finally, we introduce the methods used to estimate the MOC efferent-system model parameters.

A. Existing models

The Zilany et al. (2009, 2014) AN model was expanded here. In this model, the stimulus first passes through a middle-ear filter. The output of the middle ear is the input to the cochlea (Fig. 2). Cochlear tuning is modeled using two filters, a relatively sharp chirping filter in the signal path tuned at the characteristic frequency (CF) and a wider bandpass filter in the feedforward control path tuned slightly higher than CF (Zhang et al., 2001; Tan & Carney, 2003; Zilany & Bruce, 2006). The output of the control-path filter and outer-hair-cell (OHC) model adjusts the time constant, and thus the gain and bandwidth, of the chirping filter. The constant value of the C_{OHC} cochlear

132 gainparameter in this model can be varied from 1 (normal OHC function) to 0 (completely impaired
133 OHC function) to simulate hearing loss related to OHC impairment. The inner hair cell (IHC) is
134 modeled with a nonlinear function followed by a low-pass filter. Hearing loss related to decreased
135 IHC sensitivity is implemented by varying the constant value of the C_{IHC} parameter between 1
136 (normal IHC function) and 0 (completely impaired IHC function).
137 We used the IC model from Mao et al. (2013), a band-pass filter centered at the best modulation
138 frequency (BMF) of the IC cell. The output of the AN model is the input to the IC model; this
139 structure is a simplified version of an IC model that includes the cochlear nucleus (CN) stage
140 between the AN and IC to simulate modulation tuning (Nelson and Carney, 2004).

B. Proposed model with MOC efferents

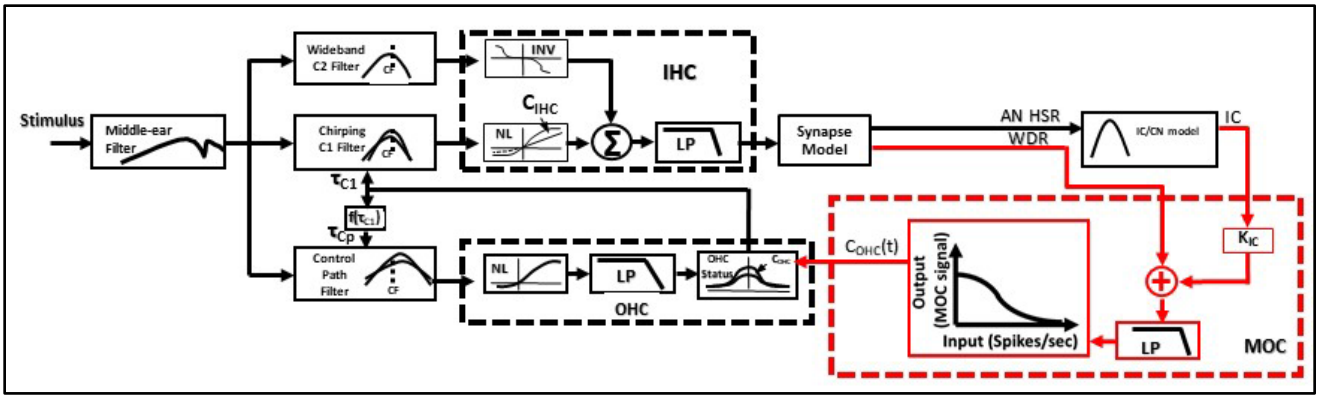


FIG. 2. Schematic illustrating overall structure of the new model components (red) with respect to the Zilany et al. (2014) AN model and Mao et al. (2013) IC/CN model (black). After (Zilany et al., 2009) and (Farhadi et al., 2021). The new model components includes the two feedback signals from IC and WDR cells to the MOC unit (dashed red box), where the two input signals are scaled and summed before going through a low-pass filter. The low-pass filtered signal goes to the rational input-output MOC function, which maps the input signal to a cochlear-gain scaler between 0 and 1, for that time sample.

One of the challenges of including the efferent system in the computational AN model is to reconfigure the implementation of the model to work sample by sample. In the previous model, since there was no feedback from IC to the cochlear, the AN and IC response were calculated one stage at a time for the whole duration of the stimulus. In the sample-by-sample configuration, for each time sample, a new value of the cochlear gain, controlled by the time-varying parameter, $C_{OHC}(t)$, must be computed. Each value of the gain is influenced by the output of a new stage, the model MOC (Fig. 2, dashed red box). The inputs to the MOC model are the previous time samples of the IC-model rate-function (time-varying rate, in spikes/sec) and CN wide-dynamic-range (WDR) response provided by the AN-model rate-function output. For simplicity, we used the AN LSR model response, which represents all inputs to the MOC with wide dynamic ranges, including the T-

152 stellate cells in the anteroventral cochlear nucleus (Oertel et al., 2011; Romero & Trussell, 2021) and
153 cells in the small cell cap of the CN that receive exclusive inputs from LSR and MSR AN fibers
154 (Ryugo, 2008). In the sample-by-sample implementation, for accurate performance of the filters in
155 the AN model, the previous states of all the filters are preserved and passed to the filter input to
156 process the next sample of the stimulus.

157 The Zilany et al. (2014) AN AN model was re-configured to compute responses, one sample at a
158 time, by storing the states of each filter stage from one sample to the next. Then, the gain-control
159 feedback signal, based on the current samples of the neural signals at the input to the MOC stage,
160 modulated the output of the control-path filter to calculate the gain for processing the next sample.
161 The model with this new configuration was tested to ensure that it matched the response of the
162 original model when the efferent stage was deactivated.

163 The two inputs to the proposed MOC model (Fig. 2, red) are the descending IC feedback signal
164 and the ascending WDR signal. The IC signal carries information related to the amplitude of AN
165 fluctuations at a given CF and works as a positive feedback loop to improve contrast in the the IC
166 rate profile across frequency channels. The rationale behind this statement is as follows: 1) AN
167 frequency channels tuned near local peaks in the stimulus spectrum have responses that are
168 "captured" by the spectral peak, and thus have low-amplitude fluctuations (Carney, 2018; Young and
169 Sachs, 1979). Consequently, for these frequency channels, model IC cells that are excited by
170 fluctuations (BE IC cells) have lower response rates. Activity of MOC neurons reduces cochlear gain
171 (Guinan, 2018); therefore, reduced input to the MOC in these channels results in a smaller gain
172 reduction, resulting in channels with relatively higher cochlear gains. The higher gains will tend to
173 drive IHCs closer to saturation over time, resulting in further reduction of the amplitude
174 fluctuations in AN HSR responses (i.e., positive feedback). 2) For frequency channels tuned away
175 from spectral peaks, AN fluctuations are stronger and the BE IC responses are therefore stronger.

The increased input to the MOC results in a reduction in cochlear gain, reducing IHC saturation and further increasing fluctuation amplitudes in the AN response. Overall, the IC provides a positive feedback that would enhance the fluctuation contrast across AN channels and thus the rate contrast across IC channels. This feedback potentially enhances the neural representation of the spectrum at the level of the IC.

The other input to the MOC stage is from the WDR cells in the CN, which carry the stimulus sound-level information. This level-dependent input to the MOC acts as a negative feedback, keeping the operating point of the IHC transducer function near the knee point, so that local maxima in the stimulus spectrum saturate the IHCs. This signal ensures sufficient baseline fluctuation contrast across frequency channels, so that the IC feedback signal can then enhance this contrast. Without the WDR signal, regardless of the peaks and valleys in the stimulus spectrum, if the stimulus level were too low or high, all AN channels would have fluctuating or flat responses, respectively. Both extremes would degrade the ultimate neural representation of the spectrum, reducing the contrast in fluctuation amplitudes across IC frequency channels.

The excitatory input to the MOC stage is the sum of the IC and the WDR rate functions. A constant scalar, K_{IC}, is applied to the IC rate signal to approximately balance the amplitudes of the two MOC inputs (Table 1). The MOC system was modeled with a low-pass filter representing slow dynamics of the MOC efferent system, followed by a decaying rational function (Eqn. 1) that maps the MOC inputs to the time-varying cochlear gain parameter, COHC. The rational function was chosen to be second order to ensure asymptotic behavior for both small and large input signals.

$$C_{OHC}(t) = \frac{1}{1 + (B \times {M_{in(t)}}^2)} \qquad (Eqn. 1)$$

In this equation, $M_{in(t)}$ is the sum of the projections from the CN and IC to the MOC unit that is low-pass filtered. B is a constant that adjusts the shape of MOC input-output function (Fig. 2). The optimization of the value of parameter B is discussed later (Section C of the methods). The MOC output, C_{OHC}(t), is a value between 0 and 1 which controls the gain of the chirping filter. Increasing the input to the MOC decreases the output of the MOC and thus cochlear gain. Figure 3 shows the intermediate responses at each level of the composite model to a tone-in-noise stimulus. For the first 200 ms the stimulus is noise only (Fig. 3, black); because of the fluctuation in AN HSR rates (Fig. 3, blue), the BE IC response is robust (Fig. 3, purple), thus driving the MOC model and decreasing the cochlear gain (Fig. 3, green). Between 200 ms and 500 ms, a tone at the model CF (5 kHz) is added to the noise stimulus, resulting in saturation of the IHC (not shown) and flattening the neural fluctuation amplitudes in the AN HSR response. As a result of the reduced AN HSR neural fluctuations, the BE IC rate response decreases. The reduced IC rate results in a weaker MOC input and thus an increase in the cochlear gain signal (Fig. 3, green). Note that the mean rate of the AN HSR fiber is not changed appreciably by addition of the tone due to rate saturation; however, the model WDR rate function is increased during presentation of the tone. After the tone

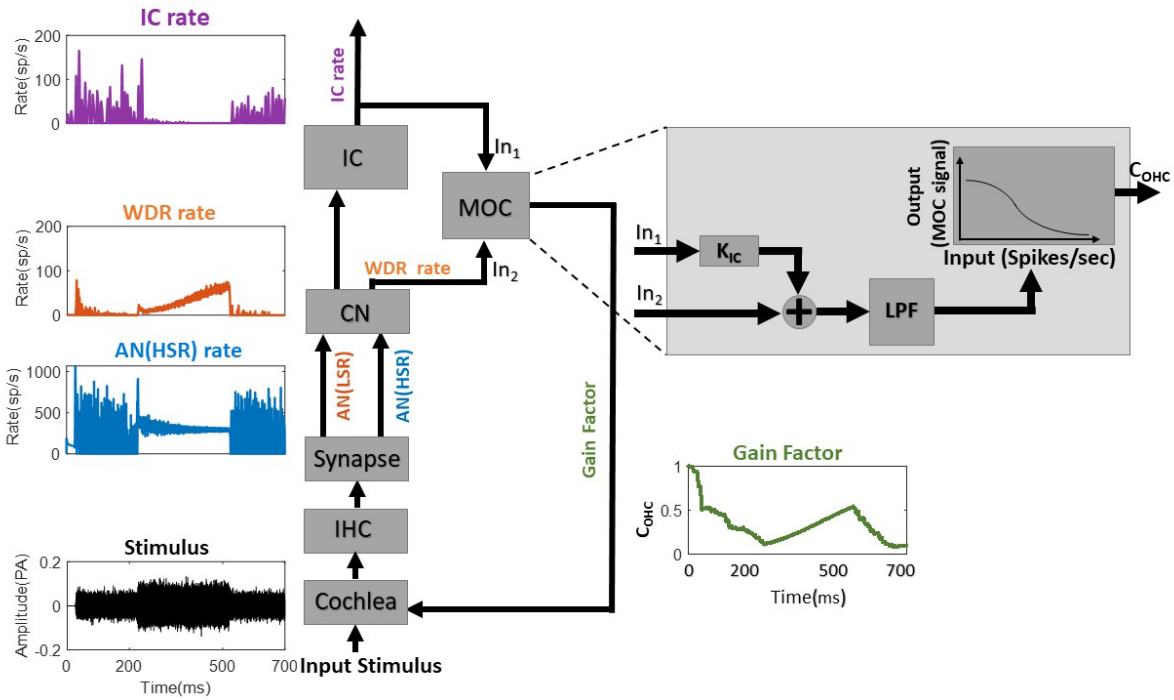


FIG. 3. The MOC efferent model diagram with intermediate responses at different levels of auditory periphery in response to tone-in-noise stimulus. The stimulus (black) is a wide band noise and after 200 ms a tone is added to the noise for a duration of 300 ms. The response of the models at AN (HSR) stage (blue) shows that during the noise-only stimulus (first 200 ms) the fluctuation increases over time as the gain decreases (green), and as a result the IC (purple) rate increases. After the addition of the tone (200 -500 ms) the fluctuations in the AN HSR responses are reduced in amplitude, resulting in a drop in IC response rate. As the In1 input to the MOC block (on the right side of this figure) decreases, the gain starts to recover (green). After the offset of the tone (500-700 ms) the gain decreases again, and the AN fluctuations and IC rate increase.

ends, the IHC is no longer saturated, AN fluctuations reappear, the IC response increases. The increase in the IC response that excites the MOC reduces the cochlear gain. The changes in the WDR response rate (Fig.3, red) also influence cochlear gain; as described above, the WDR signal maintains the IHC operating point such that the IC response has an effect on C_{OHC}.

218 **C. Adjusting the parameters of the MOC efferent model**

219 Parameters of the MOC efferent model include those of the MOC input/output function, B
220 (Eqn.1, Fig.3) the scaler constant for IC rate (K_{IC}), and the low pass filter (LPF) time constant (cut-
221 off frequency). We used an existing physiological dataset to adjust these parameters (Carney et al.,
222 2014).

223 *1. Stimuli*

224 The stimulus for both the physiological dataset and the model simulation during the parameter
225 fits was a sinusoidally amplitude-modulated (SAM) wideband noise (100 Hz to 20 kHz) and
226 modulation depths that varied from -30 to 0 dB (fully modulated) in 5 dB steps, and an
227 unmodulated wideband noise with the same bandwidth as the SAM noise. We analyzed the model
228 and physiological responses to stimuli at three spectrum levels, -4, 16, 36 dB re 20μPa (RMS: 36, 56,
229 76 dB SPL overall level). The stimuli in response to 36 dB SPL were used for adjusting
230 modelparameters because at this level as the trend of increasing rate over time was strongest. The
231 duration of the stimuli in this dataset was 500 ms, with 50-ms raised-cosine ramps.

232 *2. Physiological method and data analysis*

233 The physiological dataset used in this work included previously published extracellular, tetrode
234 recordings of responses in the central nucleus of the IC in three awake rabbits. Details of the
235 physiological methods can be found in Carney et al. (2014). Briefly, IC cells were classified based on
236 their modulation transfer functions (MTFs, the average rate in response to AM stimuli

237 as a function of stimulus modulation frequency) into band-enhanced (BE), band-suppressed (BS),
238 hybrid, flat and unusual types (Kim et al., 2015, 2020) (Fig. 4). The focus of this study was on BE
239 cells; this IC MTF type was hypothesized in to project to the MOC based on BE MTFs reported for
240 MOC neurons (Gummer et al., 1988b). IC BE cells are excited by low frequency (5-250 Hz)
241 fluctuations (Kim et al., 2020). MTF shape was classified similar to Kim et al. (2020), using a
242 criterion ratio of 1.2 peak MTF rate to reference rate which yielded an equally distributed class of
243 MTFs (Fig. 4). MTFs in response to 36-dB SPL spectrum level wideband noise were used to identify
244 a cell's best modulation frequency (BMF), the frequency of the peak rate in the MTF. In the dataset
245 used here, the modulation frequency was matched approximately to the cell's BMF; datasets with
246 more than an octave difference between stimulus modulation frequency and the cell's BMF were
247 excluded in these analyses.

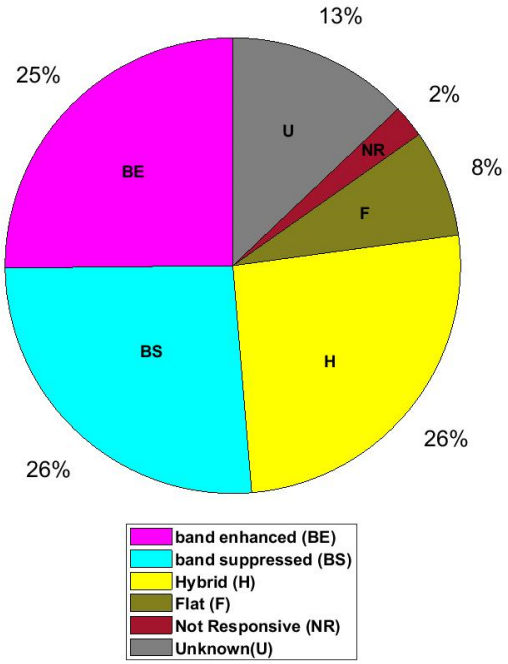


FIG. 4. Distribution of MTF classes of the IC neurons in the physiological dataset (Carney et al., 2014) using the MTF-categorization method of Kim et al. (2020) (D. O. Kim et al., 2020)

248 To analyze the physiological responses to AM noise, a Gaussian filter was used to smooth
249 the spike trains to avoid aliasing introduced by PSTH bin-widths (Lehky, 2010). The standard
250 deviation of the Gaussian filter was inversely proportional to each cell's BMF; the number of points
251 in the filter was 6 times the standard deviation of the Gaussian window. Responses were averaged
252 over 50 repetitions, excluding 200 ms at the response onset.
253 Trends in rate over the course of the response were characterized using a simple linear
254 regression. The sign of the linear model's slope indicated whether the response increased or
255 decreased over time. IC BE responses with rates that increased over time were used to find the time
256 constant of the LPF in the MOC efferent model by maximizing the correlation between the IC BE
257 spiking rate and the model simulations in response to a stimulus with a modulation depth of -10 dB.
258 -10 dB modulation depth was chose here because the majority of the IC neurons had stronger
259 increases in rate over time at this modulation depth as compared to higher modulation depths.
260 Constrained minimization (fmincon function in MATLAB) was used to simultaneously fir the time
261 constant τ_{MOC}, the rational function parameter, B (Eqn. 1), and the IC constant scaler K_{IC}. A 300-ms
262 duration analysis window beginning 200 ms after the stimulus onset was applied to the simulation
263 results and to the physiological response. Properties of the model such as CF, BMF, and modulation
264 frequency were matched to each cell in the physiological dataset. Based on the parameter value fits
265 for each neuron, a distribution of the parameters was established, and the means of the distributions
266 were chosen for the final model parameter values. The sensitivity of the model to the parameter
267 values was studied by comparing the average correlation between the physiological and the model
268 response when each model parameter was either set to the value fit to each neuron or to the mean
269 parameter value for the population.

III. RESULTS

Section A below reviews our hypothesis regarding the MOC efferent model response to AM noise stimuli, analysis of the physiological dataset in response to AM noise, and implementation of the MOC efferent model with its parameters (K_{IC},B, τ_{MOC}). Section B describes how these parameters were optimized and compares the responses of models with and without the MOC efferent control system to physiological data. In section C we use the new model with MOC efferent control to simulate responses to a few different stimuli and compare these responses with physiological data from the literature.

A. The hypothesis and physiological dataset

We hypothesized that MOC efferent gain control would result in an increase in the average rate of IC BE neurons over time in response to an AM noise stimulus. The rationale for this argument is that the low-frequency fluctuations in the envelope of AN fiber responses to modulated noise stimuli excite IC neurons with BE MTFs. BE IC neurons were then assumed to excite MOC cells, and MOC activity decreases cochlear gain. Reduction in cochlear gain effectively decreases the slope of the IHC transduction nonlinearity and therefore increases the effective modulation depth in the output of the IHC (Fig. 5), ultimately increasing the BE IC response. Thus, the MOC efferent control represents positive feedback on IC rates in response to AM noise; IC rates would increase until a saturation point was reached.

We re-analyzed an existing physiological dataset (Carney et al., 2014), which includes recordings from IC cells in the midbrain of awake rabbits. Consistent with our hypothesis, the IC rates increase over time in most IC neurons with BE MTFs. Figure 6 presents six examples of BE IC neuron responses for which the trend of increasing rate over time seems to be very noticeable and also these examples were chosen as they have different BMFs and CFs; the left panel in each example is the rate MTF, and the right panel is the IC response to AM noise modulated at a frequency near the neurons BMF, with a range of modulation depths. The gray time range is the time window after the

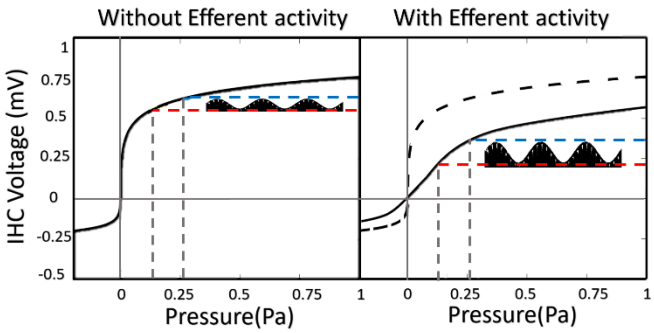


FIG. 5. Effect of MOC efferent activity on IHC input-output function. Left panel: The effective modulation depth in IHC output voltage is limited due to the saturation. Right panel: with the same input pressure, effective modulation depth is increased due to the MOC efferent activity and decrease in gain.

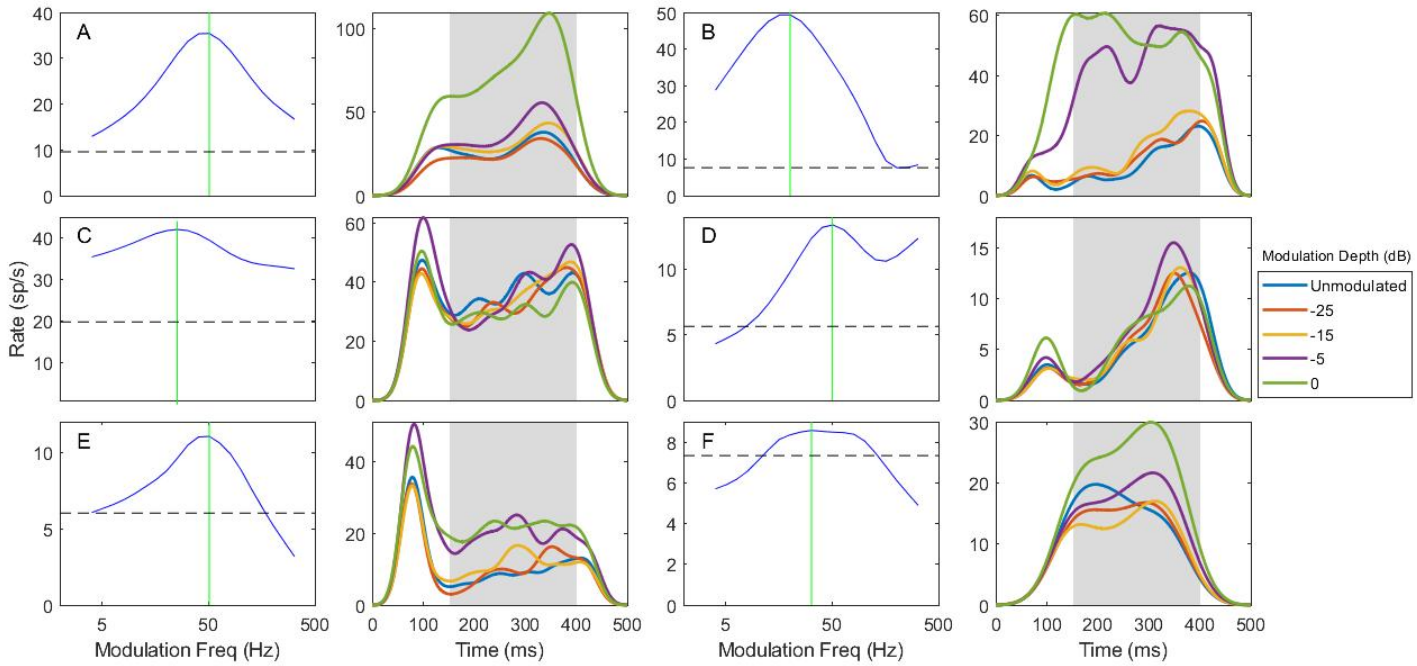
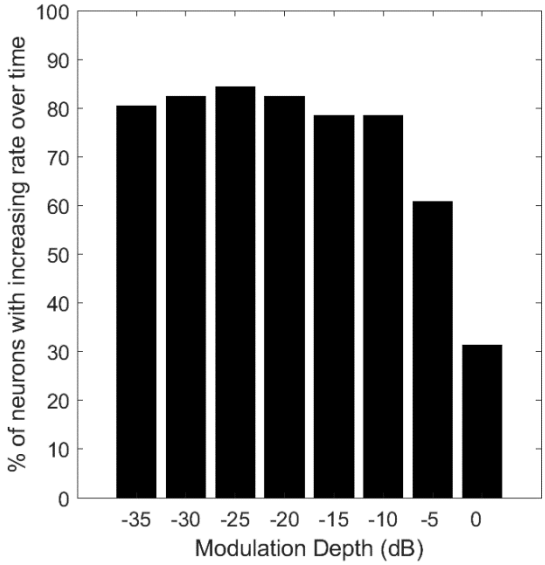


FIG. 6. A-F) Six examples of IC BE neuron responses to AM noise stimulus. The left-hand panel in each example is the MTF for AM noise stimuli; the vertical green lines indicate BMF. The right-hand panels show smoothed PSTHs that illustrate the time course of the rate during responses to AM noise stimuli with a modulation frequency near each neuron's BMF and a range of modulation depths (see legend).

onset response over which we analyzed the increase in rate. All of the examples in figure 6 had increasing rate over time during the analysis window, especially in response to lower modulation depths. In some neurons (e.g., Fig. 6B) the rate was nearly constant over time in response to fully modulated noise, but the IC rate was high for this modulation depth from the beginning of the analysis window. In the framework of our hypothesis, this observation suggests that because of the strong modulation in the stimulus, the fluctuation in AN PSTHs would be large enough to maximize the IC rate, such that a change in cochlear gain as a result of MOC efferent activity could not further increase the IC rate.

303 The trend of increasing rate over time occurs in approximately 80% of neurons with BE MTFs
304 in response to lower modulation depths (modulation depth less than -15 dB in Fig. 7). This
305 percentage decreases at high modulation depths (see Fig. 6). As described above, the decrease at
306 high modulation depths can be explained by a ceiling effect on the IC rate. In general, response rates
307 are expected to decrease over time as a result of rate adaptation. The increase or decrease of rates
308 over time was determined by the sign of the slope of a linear regression fit to these responses. The
309 result, that in fact 80% of the IC BE cells had rates that increased over the time course of the AM
310 noise stimuli, is consistent with the hypothesis that MOC efferent activity results in an increase in
311 these IC rates over time.



312 FIG. 7. Percentage IC BE neurons with increasing rate over the time course of wideband AM noise
313 stimuli.

314 **B. Adjusting the model parameter**

315 We used constrained minimization (fmincon function in MATLAB) to adjust the MOC model
316 parameters, including the time constant of the MOC system, τ_{MOC}, the bias parameter in the MOC
317 input/output function, B, and the IC constant scalar, K_{IC}. We searched for parameter values

within a reasonable range that maximized the correlation between the model response and the physiological data over the analysis window. Parameters were estimated for each neuron in the dataset. The distribution of the optimized parameter for neurons in this study was very broad (Fig. 8). The median of the distribution for each parameter was also tested as the final value for that parameter (K_{IC}=20, τ_{MOC}=236 ms, B=0.01) (Eqn.1). The estimated time constant of 236 ms for the LPF in the MOC's model, τ_{MOC}, agrees with values of the time constant that have been suggested for

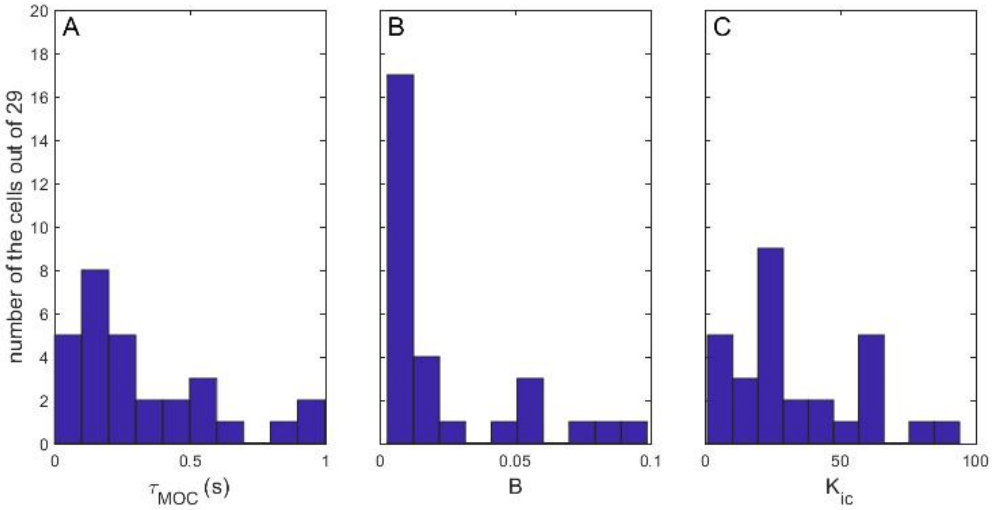


FIG. 8. Distribution of the fit values of, A) time constant (τ_{MOC}), B) the rational function parameter (B), and C) the IC scaler (K_{IC}) for the model for 29 neurons in this study for which the rates increased during the SAM stimuli, the MTF type was BE, and the modulating frequency of the stimulus matched the BMF of the cell..

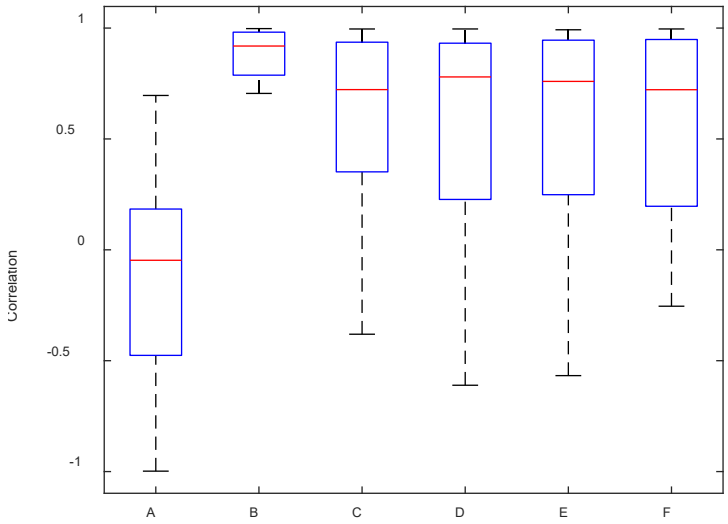


FIG. 9. Correlations between the physiology and models A) without efferent , B) with efferent with all parameters individually identified for each neuron, C) with efferent where parameters were fixed at the median of the distributions, D) with efferent where B and K_{IC} were fixed at median of the distributions and the time constant was individually identified for each neuron, E) model with efferent where time constant and K_{IC} were fixed at the median of the distributions and B was individually identified for each neuron, F) model with efferent where time constant and B were fixed at the median of the distributions and K_{IC} was individually identified for each neuron

the MOC efferent system based on other types of measurements (Backus & Guinan, 2006; Roverud & Strickland, 2010; Warren & Liberman, 1989).

The median correlation between the model and physiological responses during the analysis time window was calculated for models with and without the efferent feedback system (Fig. 9). The negative correlation of -0.15 between data and model without efferent response (Fig. 9-A) demonstrates that the model without efferents did not simulate the trend of increasing rate over time in the physiological dataset. Simulations using a MOC efferent model with parameters

identified for each neuron resulted in a correlation of 0.87 between data and the model response (Fig. 9-B). The correlation reduced to 0.72 after replacing the individual values with the median values of the parameter distributions. (Fig. 9-C). This result shows that the model was sensitive to the parameters but still simulated the increased rate in response to AM noise using fixed parameters. We further studied the effect of each parameter on the correlation independently, where either B, K_{IC} or τ_{MOC} was identified for each neuron while the two other were fixed based on the mean of the distribution of estimates for each parameter (Fig. 9-D, E, F respectively). These results showed that the model was equally sensitive to all these parameters.

Unlike the physiological data, the response of the model without efferent feedback did not increase over time during the analysis window. As expected from the correlation results and the

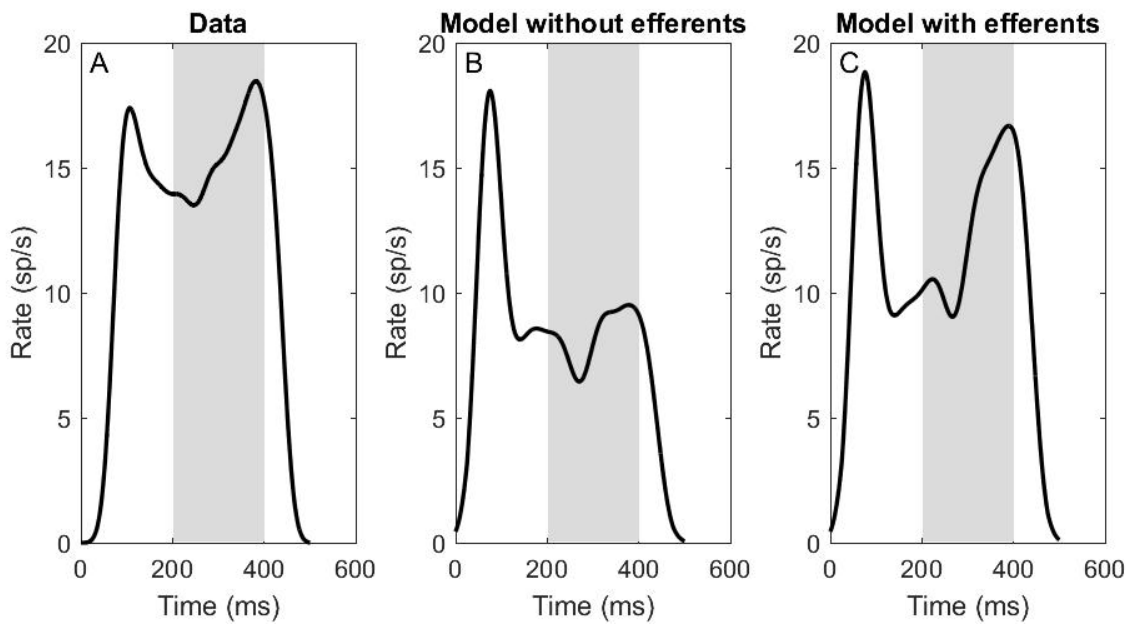


FIG. 10. A) Physiological data from a BE IC cell in awake rabbit, and IC model responses to AM noise with modulation depth of -10 dB B) without and C) with efferent activation. The CF and BMF of the model were matched to the example neuron (BMF is 80 Hz and CF is 3500Hz in this example neuron) and the Model is using the mean value for the model parameters (Fig. 9-C).

hypothesis, the model with MOC efferent gain control had an IC response that increased over time during the analysis window.Figure 10 shows responses to AM noise for a single representative IC cell and for models with MOC efferent using the mean value for the model parameters (Fig. 9-C) and without efferents (Fig. 9-A). The models had BMF and CF that matched the example neuron. The model with MOC efferent feedback simulates the increase in rate over time consistent with the hypothesis of this study that a model with MOC efferent system can simulate the response to AM noise more accurately than the model without efferents.

C. Validating the proposed model responses to tone stimuli

The model responses to tone stimulus is compared to the previous model and physiological data in the literature. The purpose of presenting this result is to test the new model's responses to some basic stimuli. However, it is important to note that the physiological data were recorded from ANs in anesthetized animals, in which the MOC efferent system was likely suppressed (Guitton et al., 2004). Thus, direct correspondence between the physiological responses and the responses of the model with MOC efferents was not expected for complex stimuli.However, for tone stimuli Rhode and Kettner (1987) showed that anesthesia does not affect post-stimulus time histograms (LSTHs) or rate-level functions of primary-like neurons in the cochlear nucleus (and one presumed AN fiber). Therefore, we validated the proposed model by comparing simulated tone responses to physiological examples.

Response to pure tones and recovery of spontaneous rate

Figure 11 shows data and simulated of responses to a 500-ms tone at CF followed by 500-ms silence. The purpose of this task was to study both the tone response and the AN model's recovery to spontaneous rate. Recovery of spontaneous activity depends on stimulus level and on the spontaneous rate of the fiber (Kiang, 1965). Here we compare model responses to physiological recordings from two HSR fibers with CFs of 1.82 and 10.34 kHz from Kiang (1965) (Fig. 11). The

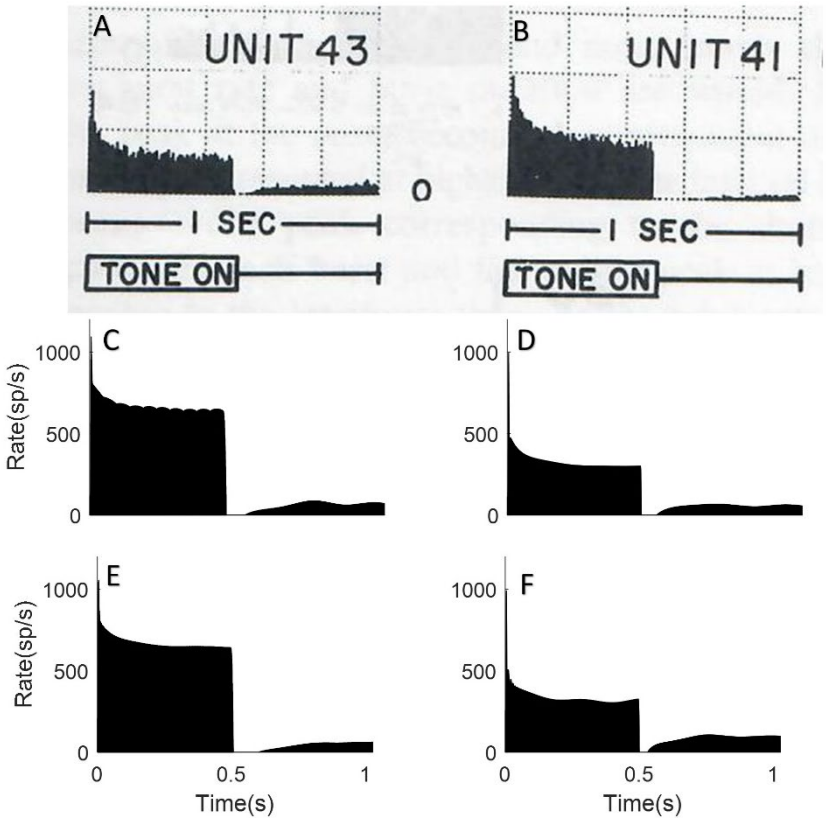


FIG. 11. AN and model responses to tones at CF: (A) Unit 43 : CF = 1.82 kHz, 55 dB SPL, SR = 60.6 s/sec, (B) Unit 41: CF = 10.34 kHz, 30 dB SPL, SR = 51.7 s/sec),C, D) Responses to the same tones for the model without efferents, and E, F) with MOC efferents.For all model simulations (C to F) the AN model type was HSR with CFs matching the AN fibers in the physiological study. The stimulus sound level was also matched to stimuli for the physiological data from Kiang (1965, permission requested.)

model response during the tone is comparable to the physiological examples, and the model shows the same initial pause and time course of recovery after the stimulus as in the physiological responses. Note that the response to a pure tone would primarily activate the AN/CN input to MOC efferent pathway, because the IC stage of the model would mainly respond at the onset and offset of the tone.

IV. DISCUSSION

In this work, we modified an existing model for the peripheral auditory system to simulate both the ascending pathway and the MOC efferent feedback with both AN/CN and IC inputs. The physiological data and simulation results are consistent with the hypothesis that the MOC efferent system influences neural fluctuations over time (Carney, 2018), as observed here in the increasing responses of BE IC cells over the time course of an AM noise. The simulation results support the hypothesis that a sub-cortical model with MOC efferent gain control better simulates IC responses to AM noise compared to a model without efferents.

One of the components of the proposed model is a LPF which was implemented to simulate the dynamics of the MOC efferent system. The time constant (or cut-off frequency) of this filter was adjusted based on the physiological dataset. The dynamics of the MOC efferent system based on the fit parameter agree with the reported dynamics in the literature for the MOC efferent fast effect. Backus and Guinan (Backus & Guinan Jr, 2006) used stimulus-frequency otoacoustic emission (SFOAE) measurements to estimate the human MOC reflex time constant. They reported an overall onset time constant of 277±62 ms and decay time constant of 159±54 ms. Warren & Liberman (1989) estimated the MOC response to have a time constant of 250 ms in barbiturate-anesthetized cats. Roverud and Strickland (2010) used an optimized psychophysical technique to estimate the

time course of cochlear gain reduction. They estimated an average exponential decay of 100 ms, which is faster than the time constant estimated by Backus and Guinan. Yasin et al. (2020) used a computational model of human hearing that included the MOC efferent feedback and evaluated the model's speech recognition for different MOC efferent time constants. Their results suggest the improvement in recognition of speech in AM noise is the highest for a MOC efferent time constant of 200 ms. In general, all of the time constants suggested in the literature for the MOC efferent system are in the same range as our results based on rabbit midbrain physiology. However, the MOC efferent system time constant could differ between rabbits and humans. Using the proposed MOC efferent model to simulate psychoacoustic experiments could better inform our understanding of the time constant for human listeners. It should also be noted that there is not a pure delay in this proposed model as we assumed that the time constant of the proposed model can simulate the delay that is suggested in James et al. (1987). Adding a pure delay to the model would require re- adjusting of the time constant of the low-pass filter in the proposed model.

Another important component of the proposed model MOC efferent system was the input-output function that maps the MOC input to a gain factor that scales the cochlear gain coefficient between 0 and 1. The second-order rational function used in this model was chosen based on its asymptotic behavior at low and high input levels. The shape of the proposed input-output function was also similar to the MOC efferent attenuation as a function of input sound level shown in Russell and Murugasu (1997). Also, the asymptotic trend of this function for higher values is similar to the asymptotic responses shown in Liberman (1988) for MOC efferent activity as a function of input sound level.

Other physiological and psychoacoustic studies can be used to further fine tune the model's parameters and structure. For example, forward masking studies with long-duration maskers would be expected to activate the MOC efferent system. At the level of IC, further animal studies of the

temporal properties of responses to complex sounds would contribute to this modeling effort, such as investigating the changes in synchrony of IC responses to the envelope of AM stimuli over time. Another direction for future work is to study the influence of efferent activity on the responses of other MTF types in the IC, such as BS or Hybrid MTFs. Here we focused on BE MTFs, as the increase in their rate over the time course of a stimulus cannot be confused with rate adaptation.

The peripheral model studied here is a monaural model, with "one-sided" brainstem and midbrain stages. A binaural model would be a powerful tool for studying the crossed and uncrossed effects of the efferents.

In general, recording from AN fibers requires an anesthetized animal preparation. Anesthesia suppresses MOC efferent activity (Guitton et al., 2004). As a result, there are no known references in the AN physiology literature for comparison to the AN model response with active MOC efferents. However, recording from more accessible levels of the auditory pathway, such as the IC in awake animals, could provide further information regarding the MOC efferent system. Recording from the CN in awake animals and comparing it with the recording from anesthetized animal or recording from AN fibers in decerebrate animals could also potentially be beneficial for studying the MOC efferent system (Kim et al., 1990a,1990b, Chang, et al., 1990).

Other phenomena phenomenon could explain the increase in IC response rates over time in response to AM noise, either in addition to or instead of efferent activity. For instance, adaptation of inhibition (Viemeister & Bacon, 1982) could also be a potential reason for the increase in rate we observed in the IC response to AM noise in this study. Adaptation of inhibition has be suggested as an explanation for auditory enhancement, in which a target signal embedded in a masker is preceded by the same masker but without the target frequency. Adaptation of inhibition results in an improvement in the threshold of detecting the tone (Byrne et al., 2011) However, neural mechanisms underlying adaptation of inhibition are not clear; the related enhancement effect and its

437 properties, such as time constant and level-dependence, could be further examined using the MOC
438 efferent model in the future.
439

440 ## V. ACKNOWLEDGMENTS

441 This work was supported by NIH-010813.
442

443 ## REFERENCES

444 Aedo, C., Tapia, E., Pavez, E., Elgueda, D., Delano, P. H., & Robles, L. (2015). Stronger efferent
445 suppression of cochlear neural potentials by contralateral acoustic stimulation in awake than
446 in anesthetized chinchilla. Frontiers in Systems Neuroscience, 9, 21.
447 Altoè, A., Pulkki, V., & Verhulst, S. (2014). Transmission line cochlear models: improved accuracy
448 and efficiency. The Journal of the Acoustical Society of America, 136(4), EL302–EL308.
449 Backus, B. C., & Guinan Jr, J. J. (2006). Time-course of the human medial olivocochlear reflex. The
450 Journal of the Acoustical Society of America, 119(5), 2889–2904.
451 Bharadwaj, H. M., Verhulst, S., Shaheen, L., Liberman, M. C., & Shinn-Cunningham, B. G. (2014).
452 Cochlear neuropathy and the coding of supra-threshold sound. Frontiers in Systems
453 Neuroscience, 8, 26.
454 Blackburn, C. C., & Sachs, M. B. (1990). The representations of the steady-state vowel sound/e/in
455 the discharge patterns of cat anteroventral cochlear nucleus neurons. Journal of
456 Neurophysiology, 63(5), 1191–1212.
457 Brown, G. J., Ferry, R. T., & Meddis, R. (2010). A computer model of auditory efferent suppression:
458 Implications for the recognition of speech in noise. The Journal of the Acoustical Society of
459 America, 127(2), 943–954. https://doi.org/10.1121/1.3273893

460 Brown, M. C. (2016). Recording and labeling at a site along the cochlea shows alignment of medial
461 olivocochlear and auditory nerve tonotopic mappings. Journal of Neurophysiology, 115(3),
462 1644–1653.
463 Bruce, I. C., Erfani, Y., & Zilany, M. S. A. (2018). A phenomenological model of the synapse
464 between the inner hair cell and auditory nerve: Implications of limited neurotransmitter
465 release sites. Hearing Research, 360, 40–54.
466 Byrne, A. J., Stellmack, M. A., & Viemeister, N. F. (2011). The enhancement effect: evidence for
467 adaptation of inhibition using a binaural centering task. The Journal of the Acoustical Society
468 of America, 129(4), 2088–2094.
469 Carney, L. H. (2018). Supra-threshold hearing and fluctuation profiles: implications for sensorineural
470 and hidden hearing loss. Journal of the Association for Research in Otolaryngology, 19(4),
471 331–352.
472 Carney, L. H., Zilany, M. S. A., Huang, N. J., Abrams, K. S., & Idrobo, F. (2014). Suboptimal use of
473 neural information in a mammalian auditory system. Journal of Neuroscience, 34(4), 1306–
474 1313.
475 Chambers, A. R., Hancock, K. E., Maison, S. F., Liberman, M. C., & Polley, D. B. (2012). Sound-
476 evoked olivocochlear activation in unanesthetized mice. Journal of the Association for
477 Research in Otolaryngology, 13(2), 209–217.
478 Chambers, C., Akram, S., Adam, V., Pelofi, C., Sahani, M., Shamma, S., & Pressnitzer, D. (2017).
479 Prior context in audition informs binding and shapes simple features. Nature
480 Communications, 8(1), 1–11.
481 Christopher Kirk, E., & Smith, D. W. (2003). Protection from acoustic trauma is not a primary
482 function of the medial olivocochlear efferent system. Journal of the Association for Research
483 in Otolaryngology, 4(4), 445–465.

Clark, L. F., Kiang, N.-S., & Thomas, E. C. (1965). Discharge Patterns of Single Fibers in the Cat's Auditory Nerve.

Clark, N. R., Brown, G. J., Jürgens, T., & Meddis, R. (2012). A frequency-selective feedback model of auditory efferent suppression and its implications for the recognition of speech in noise. The Journal of the Acoustical Society of America, 132(3), 1535–1541. https://doi.org/10.1121/1.4742745

Dávila-Chacón, J., Liu, J., & Wermter, S. (2018). Enhanced robot speech recognition using biomimetic binaural sound source localization. IEEE Transactions on Neural Networks and Learning Systems, 30(1), 138–150.

de Boer, J., & Thornton, A. R. D. (2008). Neural correlates of perceptual learning in the auditory brainstem: efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. Journal of Neuroscience, 28(19), 4929–4937.

Delano, P. H., Elgueda, D., Hamame, C. M., & Robles, L. (2007). Selective attention to visual stimuli reduces cochlear sensitivity in chinchillas. Journal of Neuroscience, 27(15), 4146-4153.

Diependaal, R. J., Duifhuis, H., Hoogstraten, H. W., & Viergever, M. A. (1987). Numerical methods for solving one-dimensional cochlear models in the time domain. The Journal of the Acoustical Society of America, 82(5), 1655–1666.

Drakopoulos, F., Vasilkov, V., Vecchi, A. O., Wartenberg, T., & Verhulst, S. (2022). Model-based hearing-enhancement strategies for cochlear synaptopathy pathologies. Hearing Research, 108569.

Farhadi, A., Jennings, S. G., Strickland, E. A., & Carney, L. H. (2021, June). A Closed-Loop Gain-Control feedback model for the medial efferent system of the descending auditory pathway. In ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 291-295). IEEE.

Ferry, R. T., & Meddis, R. (2007). A computer model of medial efferent suppression in the mammalian auditory system. The Journal of the Acoustical Society of America, 122(6), 3519–3526. https://doi.org/10.1121/1.2799914

Fu, S.-W., Tsao, Y., Lu, X., & Kawai, H. (2017). Raw waveform-based speech enhancement by fully convolutional networks. 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 6–12.

Ghitza, O., Messing, D., Delhorne, L., Braida, L., Bruckert, E., & Sondhi, M. M. (2007). Towards predicting consonant confusions of degraded speech. In Hearing–From Sensory Processing to Perception (pp. 541–550). Springer.

Giguere, C., & Woodland, P. C. (1994). A computational model of the auditory periphery for speech and hearing research. II. Descending paths. Journal of the Acoustical Society of America, 95(1), 343–349. https://doi.org/10.1121/1.408367

Giraud, A. L., Garnier, S., Micheyl, C., Lina, G., Chays, A., & Chéry-Croze, S. (1997). Auditory efferents involved in speech-in-noise intelligibility. Neuroreport, 8(7), 1779–1783.

Guinan Jr, J. J. (2006). Olivocochlear efferents: anatomy, physiology, function, and the measurement of efferent effects in humans. Ear and Hearing, 27(6), 589–607.

Guinan Jr, J. J. (2018). Olivocochlear efferents: Their action, effects, measurement and uses, and the impact of the new conception of cochlear mechanical responses. Hearing Research, 362, 38–47.

Guitton, M. J., Avan, P., Puel, J.-L., & Bonfils, P. (2004). Medial olivocochlear efferent activity in awake guinea pigs. Neuroreport, 15(9), 1379–1382.

Gummer, M., Yates, G. K., & Johnstone, B. M. (1988a). Modulation transfer function of efferent neurones in the guinea pig cochlea. Hearing Research, 36(1), 41–51. https://doi.org/10.1016/0378-5955(88)90136-0

532 Gummer, M., Yates, G. K., & Johnstone, B. M. (1988b). Modulation transfer function of efferent
533 neurones in the guinea pig cochlea. Hearing Research, 36(1), 41–51.
534 Heil, P., & Peterson, A. J. (2015). Basic response properties of auditory nerve fibers: a review. Cell
535 and Tissue Research, 361(1), 129–158.
536 Holmes, E., & Griffiths, T. D. (2019). 'Normal' hearing thresholds and fundamental auditory
537 grouping processes predict difficulties with speech-in-noise perception. Scientific Reports,
538 9(1), 1–11. https://doi.org/10.1038/s41598-019-53353-5
539 Huffman, R. F., & Henson Jr, O. W. (1990). The descending auditory pathway and acousticomotor
540 systems: connections with the inferior colliculus. Brain Research Reviews, 15(3), 295–323.
541 Jacewicz, E., Fox, R. A., & Salmons, J. (2007). Vowel duration in three American English dialects.
542 American Speech, 82(4), 367–385.
543 James, A. L., Harrison, R. V., Pienkowski, M., Dajani, H. R., & Mount, R. J. (2005). Dynamics of
544 real time DPOAE contralateral suppression in chinchillas and humans Dinámica de la
545 supresión contralateral de las DPOAE en tiempo real en chinchillas y
546 humanos. International journal of audiology, 44(2), 118-129.
547 Jamos, A. M., Kaf, W. A., Chertoff, M. E., & Ferraro, J. A. (2020). Human medial olivocochlear
548 reflex: contralateral activation effect on low and high frequency cochlear response. Hearing
549 Research, 389, 107925.
550 Jamos, A. M., Hosier, B., Davis, S., & Franklin, T. C. (2021). The role of the medial olivocochlear
551 reflex in acceptable noise level in adults. Journal of the American Academy of
552 Audiology, 32(03), 137-143.
553 Jennings, S. G. (2021). The role of the medial olivocochlear reflex in psychophysical masking and
554 intensity resolution in humans: a review. Journal of Neurophysiology, 125(6), 2279–2308.

Jennings, S. G., Heinz, M. G., & Strickland, E. A. (2011). Evaluating adaptation and olivocochlear efferent feedback as potential explanations of psychophysical overshoot. JARO - Journal of the Association for Research in Otolaryngology, 12(3), 345–360. https://doi.org/10.1007/s10162-011-0256-5

Joris, P X, Schreiner, C. E., & Rees, A. (2004). Neural processing of amplitude-modulated sounds. Physiological Reviews, 84(2), 541–577.

Joris, Philip X, & Yin, T. C. T. (1992). Responses to amplitude-modulated tones in the auditory nerve of the cat. The Journal of the Acoustical Society of America, 91(1), 215–232.

Kim, D. O., Carney, L., & Kuwada, S. (2020). Amplitude modulation transfer functions reveal opposing populations within both the inferior colliculus and medial geniculate body. Journal of Neurophysiology, 124(4), 1198–1215.

Kim, S., Frisina, R. D., & Frisina, D. R. (2006). Effects of age on speech understanding in normal hearing listeners: Relationship between the auditory efferent system and speech intelligibility in noise. Speech Communication, 48(7), 855–862.

Krishna, B. S., & Semple, M. N. (2000). Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. Journal of Neurophysiology, 84(1), 255–273.

Kuwada, S., Stanford, T. R., & Batra, R. (1987). Interaural phase-sensitive units in the inferior colliculus of the unanesthetized rabbit: effects of changing frequency. Journal of Neurophysiology, 57(5), 1338–1360.

Kwan, T. J. M., Zilany, M. S. A., Davies-Venn, E., & Abdul Wahab, A. K. (2019). Modeling the effects of medial olivocochlear efferent stimulation at the level of the inferior colliculus. Experimental Brain Research, 237(6), 1479–1491. https://doi.org/10.1007/s00221-019-05511-4

Lehky, S. R. (2010). Decoding poisson spike trains by gaussian filtering. Neural Computation, 22(5), 1245–1271.

Liberman, M. C. (1978). Auditory-nerve response from cats raised in a low-noise chamber. The Journal of the Acoustical Society of America, 63(2), 442–455.

Liberman, M. C., & Guinan Jr, J. J. (1998). Feedback control of the auditory periphery: anti-masking effects of middle ear muscles vs. olivocochlear efferents. Journal of Communication Disorders, 31(6), 471–482.

Lopez-Poveda, E. A., Eustaquio-Martín, A., Stohl, J. S., Wolford, R. D., Schatzer, R., & Wilson, B. S. (2016). A binaural cochlear implant sound coding strategy inspired by the contralateral medial olivocochlear reflex. Ear and hearing, 37(3), e138.

Lopez-Poveda, E. A., Eustaquio-Martín, A., Stohl, J. S., Wolford, R. D., Schatzer, R., Gorospe, J. M., ... & Wilson, B. S. (2017). Intelligibility in speech maskers with a binaural cochlear implant sound coding strategy inspired by the contralateral medial olivocochlear reflex. Hearing Research, 348, 134-137.

Mao, J., Vosoughi, A., & Carney, L. H. (2013). Predictions of diotic tone-in-noise detection based on a nonlinear optimal combination of energy, envelope, and fine-structure cues. The Journal of the Acoustical Society of America, 134(1), 396–406.

Meddis, Ray, Lopez-Poveda, E. A., Fay, R. R., & Popper, A. N. (2010). Computational models of the auditory system. Springer.

Meddis, Raymond, Lecluyse, W., Clark, N. R., Jürgens, T., Tan, C. M., Panda, M. R., & Brown, G. J. (2013). A computer model of the auditory periphery and its application to the study of hearing. Basic Aspects of Hearing, 11–20.

601 Mertes, I. B., Johnson, K. M., & Dinger, Z. A. (2019). Olivocochlear efferent contributions to
602 speech-in-noise recognition across signal-to-noise ratios. The Journal of the Acoustical
603 Society of America, 145(3), 1529–1540.
604 Mertes, I. B., Wilbanks, E. C., & Leek, M. R. (2018). Olivocochlear efferent activity is associated
605 with the slope of the psychometric function of speech recognition in noise. Ear and Hearing,
606 39(3), 583.
607 Nelson, P. C., & Carney, L. H. (2007). Neural rate and timing cues for detection and discrimination
608 of amplitude-modulated tones in the awake rabbit inferior colliculus. Journal of
609 Neurophysiology, 97(1), 522–539.
610 Oertel, D., Wright, S., Cao, X.-J., Ferragamo, M., & Bal, R. (2011). The multiple functions of T
611 stellate/multipolar/chopper cells in the ventral cochlear nucleus. Hearing Research, 276(1–
612 2), 61–69.
613 Olsen, W. O. (1998). Average speech levels and spectra in various speaking/listening conditions.
614 Rahman, M., Willmore, B. D. B., King, A. J., & Harper, N. S. (2019). A dynamic network model of
615 temporal receptive fields in primary auditory cortex. PLoS Computational Biology, 15(5),
616 e1006618.
617 Rahman, M., Willmore, B. D. B., King, A. J., & Harper, N. S. (2020). Simple transformations capture
618 auditory input to cortex. Proceedings of the National Academy of Sciences, 117(45), 28442–
619 28451.
620 Recasens, D. (1984). Vowel-to-vowel coarticulation in Catalan VCV sequences. The Journal of the
621 Acoustical Society of America, 76(6), 1624–1635.
622 Recasens, D. (2018). Coarticulation. In Oxford Research Encyclopedia of Linguistics.
623 Romero, G. E., & Trussell, L. O. (2021). Distinct forms of synaptic plasticity during ascending vs
624 descending control of medial olivocochlear efferent neurons. Elife, 10, e66396.

Roverud, E., & Strickland, E. A. (2010). The time course of cochlear gain reduction measured using a more efficient psychophysical technique. The Journal of the Acoustical Society of America, 128(3), 1203–1214.

Ryugo, D. K. (2008). Projections of low spontaneous rate, high threshold auditory nerve fibers to the small cell cap of the cochlear nucleus in cats. Neuroscience, 154(1), 114–126.

Saeedi, N. E., Blamey, P. J., Burkitt, A. N., & Grayden, D. B. (2014). Application of a pitch perception model to investigate the effect of stimulation field spread on the pitch ranking abilities of cochlear implant recipients. Hearing Research, 316, 129–137.

Schofield, B. R. (2011). Central descending auditory pathways. In Auditory and vestibular efferents (pp. 261–290). Springer.

Smalt, C. J., Heinz, M. G., & Strickland, E. A. (2014). Modeling the time-varying and level-dependent effects of the medial olivocochlear reflex in auditory nerve responses. JARO - Journal of the Association for Research in Otolaryngology, 15(2), 159–173. https://doi.org/10.1007/s10162-013-0430-z

Stern, R. M., & Morgan, N. (2012). Hearing is believing: Biologically inspired methods for robust automatic speech recognition. IEEE Signal Processing Magazine, 29(6), 34–43.

Stern, R. M., Morgan, N., Virtanen, T., Raj, B., & Singh, R. (2012). Features Based on Auditory Physiology and Perception. Techniques for Noise Robustness in Automatic Speech Recognition, 193227.

Sumner, C. J., Lopez-Poveda, E. A., O'Mard, L. P., & Meddis, R. (2003). Adaptation in a revised inner-hair cell model. The Journal of the Acoustical Society of America, 113(2), 893–901.

Sun, X.-M. (2008). Contralateral suppression of distortion product otoacoustic emissions and the middle-ear muscle reflex in human ears. Hearing Research, 237(1–2), 66–75.

Tan, K., & Wang, D. (2018). A Convolutional Recurrent Neural Network for Real-Time Speech Enhancement. Interspeech, 2018, 3229–3233.

Tan, Q., & Carney, L. H. (2003). A phenomenological model for the responses of auditory-nerve fibers. II. Nonlinear tuning with a frequency glide. The Journal of the Acoustical Society of America, 114(4), 2007–2020.

Umeda, N. (1977). Consonant duration in american english. The Journal of the Acoustical Society of America, 61(3), 846–858.

Verhulst, S., Dau, T., & Shera, C. A. (2012). Nonlinear time-domain cochlear model for transient stimulation and human otoacoustic emission. The Journal of the Acoustical Society of America, 132(6), 3842–3848.

Viemeister, N. F., & Bacon, S. P. (1982). Forward masking by enhanced components in harmonic complexes. The Journal of the Acoustical Society of America, 71(6), 1502–1507.

Wang, K., & Shamma, S. A. (1995). Spectral shape analysis in the central auditory system. IEEE Transactions on Speech and Audio Processing, 3(5), 382–395.

Warr, W. B. (1992). Organization of olivocochlear efferent systems in mammals. The Mammalian Auditory Pathway: Neuroanatomy, 410–448.

Warren III, E. H., & Liberman, M. C. (1989). Effects of contralateral sound on auditory-nerve responses. I. Contributions of cochlear efferents. Hearing Research, 37(2), 89–104.

Westerman, L. A., & Smith, R. L. (1984). Rapid and short-term adaptation in auditory nerve responses. Hearing Research, 15(3), 249–260.

Westerman, L. A., & Smith, R. L. (1987). Conservation of adapting components in auditory-nerve responses. The Journal of the Acoustical Society of America, 81(3), 680–691.

Willmore, B. D. B., Schoppe, O., King, A. J., Schnupp, J. W. H., & Harper, N. S. (2016). Incorporating midbrain adaptation to mean sound level improves models of auditory cortical processing. Journal of Neuroscience, 36(2), 280–289.

Wilson, J. L., Henson, M. M., & Henson Jr, O. W. (1991). Course and distribution of efferent fibers in the cochlea of the mouse. Hearing Research, 55(1), 98–108.

Winslow, R. L., & Sachs, M. B. (1988). Single-tone intensity discrimination based on auditory-nerve rate responses in backgrounds of quiet, noise, and with stimulation of the crossed olivocochlear bundle. Hearing Research, 35(2–3), 165–189.

Yasin, I., Drga, V., Liu, F., Demosthenous, A., & Meddis, R. (2020). Optimizing Speech Recognition Using a Computational Model of Human Hearing: Effect of Noise Type and Efferent Time Constants. IEEE Access, 8, 56711–56719. https://doi.org/10.1109/ACCESS.2020.2981885

Ye, Y., Machado, D. G., & Kim, D. O. (2000). Projection of the marginal shell of the anteroventral cochlear nucleus to olivocochlear neurons in the cat. Journal of Comparative Neurology, 420(1), 127–138.

Young, E. D., & Sachs, M. B. (1979). Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. The Journal of the Acoustical Society of America, 66(5), 1381-1403.

Zaar, J., & Carney, L. H. (2022). Predicting speech intelligibility in hearing-impaired listeners using a physiologically inspired auditory model. Hearing Research, 108553.

Zilany, M. S. A., & Bruce, I. C. (2006). Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery. The Journal of the Acoustical Society of America, 120(3), 1446–1466.

Zilany, M. S. A., Bruce, I. C., & Carney, L. H. (2014). Updated parameters and expanded simulation options for a model of the auditory periphery. The Journal of the Acoustical Society of America, 135(1), 283–286.

Zilany, M. S. A., Bruce, I. C., Nelson, P. C., & Carney, L. H. (2009). A phenomenological model of the synapse between the inner hair cell and auditory nerve: long-term adaptation with power-law dynamics. The Journal of the Acoustical Society of America, 126(5), 2390–2412.

Zouhir, Y., & Ouni, K. (2016). Feature Extraction Method for Improving Speech Recognition in Noisy Environments. J. Comput. Sci., 12(2), 56–61.

Zweig, G. (2015). Linear cochlear mechanics. The Journal of the Acoustical Society of America, 138(2), 1102–1121.