A PHYSIOLOGICALLY BASED MODEL FOR TEMPORAL ENVELOPE ENCODING IN HUMAN PRIMARY AUDITORY CORTEX

Pierre Dugué[1,2], Régine Le Bouquin Jeannès[1,2], and Gérard Faucon[1,2]

[1]INSERM, U 642, Rennes, F-35000 France
[2]Université de Rennes 1, Rennes, F-35000 France
Laboratoire Traitement du Signal et de l'Image, Université de Rennes 1
Campus de Beaulieu, 35042 Rennes Cedex, France
phone: + (33) 2 23 23 69 19, fax: + (33) 2 23 23 69 17, email: pierre.dugue@.univ-rennes1.fr
web: www.ltsi.univ-rennes1.fr

ABSTRACT

Temporal envelope is an important sound feature involved in sound identification and speech processing. In this work, an auditory pathway model that reproduces the response to amplitude modulated white noise at different physiological levels is proposed. More particularly, this study focuses on primary auditory cortex and the validation of the model is based on temporal modulation transfer function.

Firstly, the model of the primary auditory cortex is tuned to reproduce background activity. Secondly, by varying coupling parameters between structures (medial geniculate body, thalamic reticular nucleus and primary auditory cortex), three classes of responses to amplitude modulated white noise observed in human primary auditory cortex are generated by the model. In spite of its simplicity, this model confirms the importance of interaction between thalamic and cortical structures in the temporal processing of sounds.

1. INTRODUCTION

Temporal envelope characterizes speech sounds. This component largely contributes to understand a vocal message [17]. Temporal envelope is coded in two ways by neurons of the auditory pathway. Some neurons discharge synchronously with envelope for a given modulation frequency range, others have their discharge rate that increases when the frequency of modulation is close to a specific value [19]. This last coding is a rate place one because information about amplitude modulation depends on firing rate and neuron's position. Here we focus on the processing of temporal information. That is why only synchronous coding will be studied.

An auditory pathway model from outer ear to primary auditory cortex is presented. This work is interesting because the model is based on physiology and includes each stage of temporal envelope processing. This model has to reproduce amplitude modulation processing. Its response to amplitude modulated white noise will be compared to human data using temporal modulation transfer function [13]. To our knowledge, there is no other existing model which simulates responses to these stimuli from outer ear to the primary auditory cortex.

2. AUDITORY PATHWAY MODEL

Modeling work involves simplifications. These are the result of a balance between main processing and available data. The model presented here is constrained by the following strong limitations. Firstly, only afferent pathway is modeled. Secondly, there is no difference between left and right pathways. Thirdly, there is no connection between left and right ears in the pathway involved in amplitude modulation processing. Binaural stimuli are used in most of the experimentations that address the question of auditory temporal processing. This method is supposed to avoid localization tasks but this also hide left and right temporal envelope processing specificity.

Based on Hewitt [7, 8] and Jansen [9] previous works, this model is composed of two parts. The first part is a four stage model dealing with the basilar membrane, the inner hair cells, the cochlear nucleus and the inferior colliculus. It is based on single unit responses. The second part of the model, based on neuronal population responses, consists in three interconnected neuronal populations representing: the medial geniculate body (MGB), the thalamic reticular nucleus (TRN) and the primary auditory cortex (PAC).

2.1 From outer ear to inferior colliculus

The stimulus is at first filtered by the outer and middle ear. Then it is analyzed by passing through the basilar membrane which can be considered as a filter bank. The base of the basilar membrane vibrates for high frequency components whereas the apex vibrates for low frequency components. This is the origin of the tonotopic organization of the auditory pathway. To model this membrane, the Dual Resonance Non Linear model of Lopez-Poveda [14] is used. It takes into account the compressive property of the membrane. In several points of the basilar membrane defined by their center frequency (CF), oscillations close to their CF induce a move on inner hair cells which realize the signal transduction.

The inner hair cell model reproduces some of its main behaviors: signal rectification and adaptation to auditory nerve fibers [18]. The auditory nerve transmits the signal to cochlear nucleus cells. The auditory nerve signal is the realization of a geometrical law of the firing probability corresponding to the number of neurotransmitters in the synaptic cleft. The realization of the law is the number of fibers that fire.

In the cochlear nucleus, a group of neurons is connected to the same inner hair cell and consequently has the same CF. This group is divided in subgroups, each one is characterized by the modulation frequency for which it produces the most synchronized response. This modulation frequency is called the BMF (best modulation frequency). Neurons of the cochlear nucleus project in the inferior colliculus where they
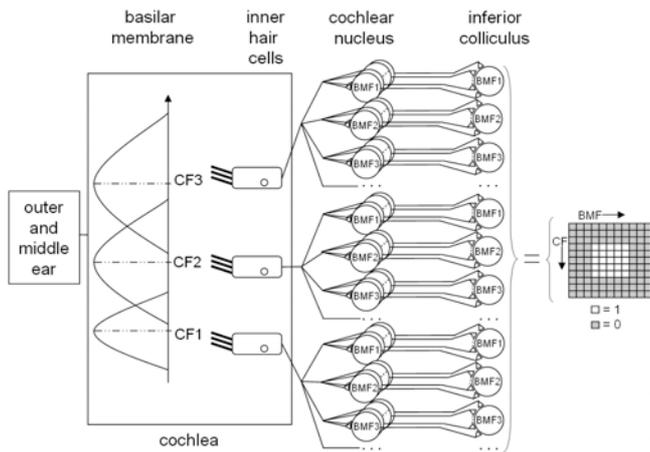
Figure 1: Structure of the model from outer ear to inferior colliculus. This first part reproduces single neuron response to amplitude modulated stimuli.

excite neurons that detect coincidence (Guérin *et al.* [5] studied different connectivity patterns). In this study, a subgroup of cochlear nucleus neurons of same CF and same BMF innervates an inferior colliculus neuron. To model neurons of the cochlear nucleus and those of the inferior colliculus, a McGregor model is used [15]. Figure 1 illustrates the first part of the model (four stages). The inferior colliculus output signals form a matrix of signals indexed by their CFs and BMFs. This part of the model is an updated version of Hewitt's model [8] which was validated with amplitude modulated stimuli. This version also reproduces responses to amplitude modulated stimuli close to physiological data in terms of firing rate and synchronization with the envelope.

2.2 From medial geniculate body to primary auditory cortex

The second part of the model is fed by inferior colliculus outputs to obtain the signal recorded in the PAC. Inferior colliculus outputs are integrated over CFs and BMFs to form a neural population response. Here we sum signals with CFs in the range $[500, 5000]$ Hz and BMFs (at 20dB) in $[60, 300]$ Hz. This integrated signal corresponds to an average pulse density of inferior colliculus afferent in MGB.

MGB receives excitatory afferent from inferior colliculus and PAC [16]. It also receives inhibitory afferent from TRN. Bushy cells are the main neurons of the MGB, they excite PAC and TRN neurons. These are excitatory neurons having AMPA, GABAa and GABAb receptors [1].

TRN receives excitatory afferent from MGB and PAC [16] and produces inhibition in MGB via GABAa and GABAb.

Auditory cortex is organized in layers and columns. Cytoarchitectonic studies show that the cortex can be divided in six layers, from layer I at the surface of the cortex to layer VI at the other end. Moreover, there are neurons whose ramifications run perpendicularly to the surface of the cortex. A group of about thousand neurons (200 to 500 μm wide) forms a cortical column. Their functional role is not known but neurons of a cortical column respond similarly to the same stimulus. In our model, a single cortical column of the PAC is considered. The model of cortical column
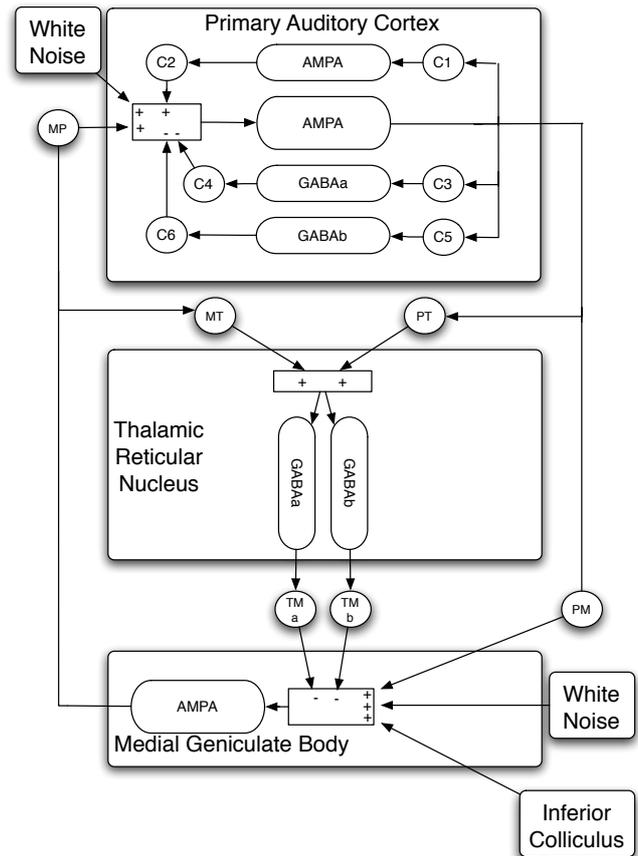

Figure 2: Structure of the model from medial geniculate body to the primary auditory cortex. This second part of the model reproduces neuronal population responses to amplitude modulated white noise. Neural populations are boxes defined by pyramidal receptors they are connected to (AMPA, GABAa and GABAb). These populations are modeled by a linear filter and a sigmoidal function. C1,..., C6 are coupling coefficients between neural populations in the primary auditory cortex. MP, PM, MT, PT, TMa and TMb are coupling coefficients between medial geniculate body, thalamic reticular nucleus and primary auditory cortex.

is simplified and takes into account only two layers: layer IV that receives MGB input and layer V that sends cortico-thalamic feedback (directly and via TRN). PAC receives excitatory afferent from MGB. This cortical column model is composed of a principal neuronal population of pyramidal neurons and three subtypes of interneurons: excitatory interneurons, fast and slow inhibitory interneurons (that act respectively on GABAa and GABAb pyramidal receptors).

Figure 2 gives the structure of the second part of the model. It represents interconnections between neuronal populations (AMPA, GABAa and GABAb boxes). Based on Jansen's work [9], a population model consists in two stages: the first one transforms the pulse density into post-synaptic potential via a causal linear filter. The second one transforms the post-synaptic potential into pulse density via a sigmoidal function. The causal linear filter impulse response is:
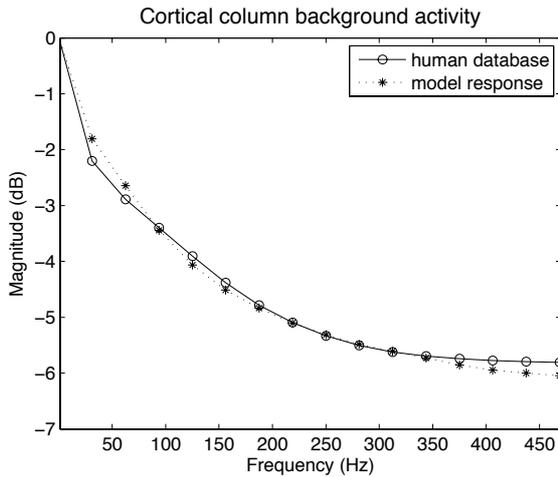
$$h(t) = A\,a\,t\,e^{-at} \quad t \geq 0, \tag{1}$$

Figure 3: Comparison between human PAC background activity (circle) (extracted from [4]) and background activity given by the model (asterisk).
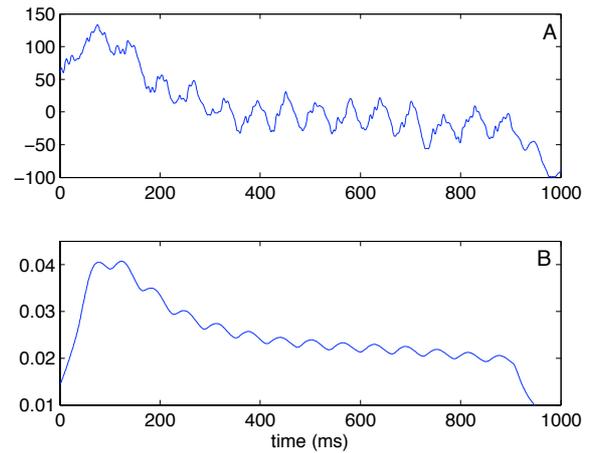

Figure 4: Auditory evoked potential in response to a 75 dB amplitude modulated white noise at 16 Hz (A) in human PAC (50 repetitions) [4] and (B) reproduced by the model (20 repetitions).

A is the amplitude constant and *a* is the time constant. The non linearity introduced in the model by the sigmoidal function is defined by:

$$y(t) = \frac{5}{1 + e^{r(v_0 - x(t))}} \qquad (2)$$

where $x(t)$ and $y(t)$ are respectively the input and the output of the non linearity, $r = 0.56$ is the compression ratio and $v_0 = 6$ the maximum output level. Table 1 lists the values corresponding to each kind of receptor. These values have been set to reproduce excitatory and inhibitory post-synaptic potentials ([3] for AMPA and GABAa, [11] for GABAb receptors). A uniformly distributed white noise is added to MGB and PAC inputs. It represents non stimulus specific afferent and simulates a pulse density in the range $[120, 320]$ spikes per second. There are two kinds of coupling coefficients between neural populations. Coupling coefficients C1, C2, C3, C4, C5 and C6 correspond to neuronal connectivity between populations in a cortical column. They are defined in section 4.1. MP (MGB to PAC), MT (MGB to TRN), PT, PM, TMa and TMb are also coupling coefficients between the three structures. Three sets of coupling coefficients are proposed to reproduce different PAC temporal behaviors (see section 4.3).

3. DATABASE

To be validated, the model must reproduce different behaviors of human response to amplitude modulated noise. The database used to compute validation criteria is presented here and these ones will be presented in sections 4.1 and 4.3.

	AMPA	GABAa	GABAb
A (mV)	5.5	1.34	11
a (s^{-1})	77.4	57	11

Table 1: Amplitude and time constant for each kind of receptor.

Twenty epileptic patients suffering from drug-resistant partial epilepsy participate in this study and are implanted with chronic SEEG (stereoelectroencephalographic) electrodes in various cortical structures (left or right hemispheres) for epileptic studies. Physiology of their auditory cortex has been studied previously and considered as normal. The patients are informed about the research protocol during SEEG monitoring and gave their fully informed consent for participating in this study. Stimuli are 1 s amplitude modulated white noise. They are shaped by rising and falling 25 ms cosine ramps to avoid auditory response to sudden sound rise. Sounds are presented binaurally via headphones to the listener by series of 50 to 100 stimuli of two randomly alternated modulation frequencies (4/32 Hz, 8/64 Hz, 16/128 Hz) with a 100 % modulation depth.

4. RESULTS

4.1 Background activity

The first step of this study was to tune the cortical column model by optimizing coupling parameters to reproduce background activity recorded in humans. More precisely, the averaged spectrum of signals recorded in PAC when the patient is not stimulated was computed. The background activity is modeled with a seven order AR model [4] (according to bayesian and Akaike information criteria). Cortical column tuning was achieved by minimizing the mean squared error between the AR model and the spectrum of a 10 s cortical column output signal. Figure 3 illustrates the result for the parameters given in table 2. The model we proposed reproduces with a slight error the AR spectrum.

C1	C2	C3	C4	C5	C6
70	20	5	10	100	20

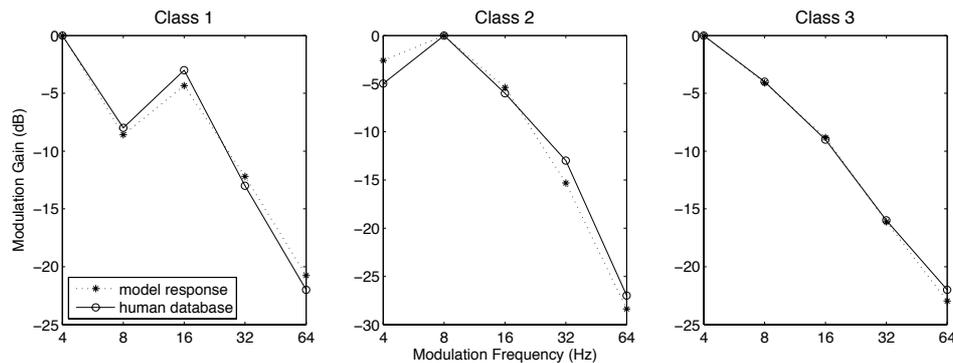Table 2: Cortical column coupling parameters

Figure 5: The three most representative temporal modulation transfer functions extracted from human PAC records [4] (circles) are reproduced by the model (asterisk).

4.2 Temporal response

An epoch is the response to a stimulus presented once and recorded via an electrode. For a given stimulus, an average of epochs forms the auditory evoked potential (AEP). The average process reduces additive noise and enhances responses which are phase locked with the stimulus. Figure 4.A shows an AEP measured in the PAC corresponding to a 16 Hz amplitude modulated white noise [4]. This AEP is composed of three parts: the transient response, the oscillatory response and the response to the stimulus ending. Oscillation frequency corresponds to the stimulus modulation frequency. This highlights the important amount of neurons that code amplitude modulation in a temporal way (only for low modulation frequencies). Figure 4.B shows the AEP given by the model in response to the same stimulus. AEP produced by the model is also composed of transient and oscillatory parts. The transient response of the model rises slowly compared to that of the database. As observed on human AEPs, oscillation frequency corresponds to the stimulus amplitude modulation one. AEPs oscillatory parts are evaluated in section 4.3.

4.3 Temporal modulation transfer function

Responses to amplitude modulated stimuli are characterized by temporal modulation transfer functions (TMTFs). These curves show the evolution of the degree of the response synchronization (temporal coding) with the modulation frequency. These curves are defined for a given stimulus level because the auditory pathway is a non linear system. The amplitude of the highest spectral peak in an interval around the modulation frequency is used to evaluate the synchronization of the response. The more synchronized the response and the stimulus, the greatest the amplitude of the oscillation in the AEP. Several methods exist to evaluate this amplitude. Gourévitch [4] compared the most common ones. Here we used the one he advocates.

Let $X^{(j)}(n)$ be the epoch recorded during the jth stimulation and n the current time sample. AEP is defined by: $X(n) = \frac{1}{J}\sum_{j=1}^{J} X^{(j)}(n)$ with J the number of recorded epochs corresponding to stimuli with the same modulation frequency. $S_{X^{(j)}}(f) = \frac{1}{N}\sum_{n=1}^{N} X^{(j)}(n) e^{-2i\pi\frac{f}{f_e}n}$ is the discrete Fourier transform of $X^{(j)}(n)$, with N the number of samples in the oscillatory part of the response and f_e the sampling frequency. Power spectrum density of the oscillation is eval-

uated computing epochs cross-spectra:
$\widehat{\gamma}_G(f) = \frac{1}{J(J-1)}\sum_{k=1}^{J}\sum_{j=1,j\neq k}^{J} S^*_{X^{(k)}}(f)S_{X^{(j)}}(f)$.
The asterisk is the conjugate complex operator. Then we look for f_0 in $[f_m - 0.5, f_m + 0.5]$ such as $\widehat{\gamma}_G(f_0) \geq \widehat{\gamma}_G(f), f \in [f_m - 0.5, f_m + 0.5]$ where f_m is the modulation frequency of the stimulus. $\widehat{A}(f_m) = \sqrt{\widehat{\gamma}_G(f_0)}$ is the evaluation of the oscillatory response amplitude. Amplitude is evaluated for each modulation frequency available in the database (see section 3). Temporal modulation transfer function is defined as follows: $TMTF(f_m) = 20log(\widehat{A}(f_m)/\widehat{A}_{max})$, where $\widehat{A}_{max}$ is the maximum of evaluated amplitudes over all modulation frequencies.

All TMTFs extracted from the database were grouped into classes according to their modulation frequency selectivity [4]. The model reproduces the three main classes obtained on database (see table 3). This is achieved by varying coupling parameters between structures in order to minimize the mean quadratic error between the two curves. Figure 5 illustrates this result. It appears that class 1 is characterized by a weak coupling between MGB and TRN whereas the transmission from PAC to TRN is strong. At the opposite, class 3 is obtained for strong connection between MGB and TRN and weak feedback. As class 1, class 2 requires strong PAC feedback and particularly via GABAa inhibition, because there is no GABAb inhibition.

	MT	MP	PT	PM	TMa	TMb
Class 1	3.6	1.8	5.9	3.4	1.2	1.6
Class 2	5.5	1.3	7.1	8.1	4.3	0
Class 3	7.2	2.5	1.3	3	1.8	1.1

Table 3: MGB, TRN and PAC coupling parameters (see figure 2).

5. DISCUSSION

As illustrated in figure 4, the simulated cortical response is smoother than the one of the database. This can be explained by the fact that SEEG electrodes record activity from other layers and from cortical columns of other cortical areas. This activity is not necessarily temporally correlated with the stimulus and a residual activity may be observed after the averaging process. The transient part of the response

is poorly reproduced but the main source of this transient wave seems to be in the secondary auditory cortex [6]. This wave is recorded in PAC because of its electromagnetic propagation.

The use of one cortical column may appear as an underestimation of the PAC complexity. MGB, PAC and TRN are tonotopically organized but amplitude modulated white noise stimulus is not frequency selective. One can suppose that the response to amplitude modulated white noise is similar across CF. Langner and colleagues [12] showed a specific modulation frequency selectivity perpendicular to the tonotopic axis but this result was not confirmed [10]. Stimuli spectra and lack of information about PAC organization lead to only one column model. However, results obtained with this simplified structure help us to understand the main connections in the amplitude modulation process. The three classes reproduced represent 65% of the TMTFs recorded in the PAC. To obtain these classes, coupling coefficients between the three structures were adjusted whereas the PAC cortical column coefficients (C1, C2, C3, C4, C5 and C6) were unchanged. In the database, some patients present several recording sites in PAC. Each recording site has a different TMTF. So, it is possible to find these classes in the PAC of one patient. This supports our approach because the connectivity pattern may be different for columns that have the same organization and which produce the same background activity.

To our knowledge, there is no characterization of AEP in response to amplitude modulated stimuli recorded in the TRN. Concerning MGB, single neuron responses to amplitude modulated stimuli were studied. Creutzfeld [2] observed that a neuron which receives MGB afferent can code temporally amplitude modulation for frequencies up to 20 Hz whereas its afferent neuron can follow modulation frequencies up to 200 Hz. This phenomenon is supposed to be a consequence of synaptic depression. Unfortunately, the model presented here does not take into account such a mechanism. That is why this behavior is not reproduced. This will lead us to include these mechanisms to refine the model.

6. ACKNOWLEDGMENTS

REFERENCES

[1] E. L. Bartlett and P. H. Smith, "Anatomic, intrinsic, and synaptic properties of dorsal and ventral division neurons in rat medial geniculate body," *J. Neurophysiol.*, vol. 81, pp. 1999–2016, 1999.

[2] O. Creutzfeld, F. C. Hellweg, and C. Schreiner, "Thalamocortical transformation of responses to complex auditory stimuli," *Exp. Brain Res.*, vol. 39, pp. 87–104, 1980.

[3] D. Golomb, X. J. Wang, and J. Rinzel, "Propagation of spindle waves in a thalamic slice model," *J. Neurophysiol.*, vol. 75, pp. 750–769, 1996.

[4] B. Gourévitch, *Etude de mécanismes corticaux auditifs par l'analyse de l'enveloppe temporelle*. Ph.D. Thesis, Université de Rennes 1, 2005.

[5] A. Guérin, R. Le Bouquin Jeannès, J. Bès, G. Faucon, and C. Lorenzi, "Evaluation of two computational models of amplitude modulation coding in the inferior colliculus," *Hear. Res.*, vol. 211, pp. 54–62, 2006.

[6] A. Gutschalk, R. D. Patterson, M. Scherg, S. Uppenkamp, and A. Rupp, "Temporal dynamics of pitch in human auditory cortex," *NeuroImage*, vol. 20, pp. 755–766, 2004.

[7] M. J. Hewitt and R. Meddis, "A computer model of a cochlear-nucleus stellate cell: response to amplitude modulated and pure-tone stimuli," *J. Acoust. Soc. Am.*, vol. 91, pp. 2096–2109, 1992.

[8] M. J. Hewitt and R. Meddis, "A computer model of amplitude modulation sensitivity of single units in the inferior colliculus," *J. Acoust. Soc. Am.*, vol. 95, pp. 2145–2159, 1994.

[9] B. H. Jansen and V. G. Rit, "Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns," *Biol. Cybern.*, vol. 73, pp. 357–366, 1995.

[10] P. X. Joris, C. E. Schreiner, and A. Rees, "Neural representation of amplitude-modulated sounds," *Physiol. Rev.*, vol. 84, pp. 541–577, 2004.

[11] U. Kim, M. V. Sanchez-Vives, and D. A. McCormick, "Functional dynamics of GABAergic inhibition in the thalamus," *Science*, vol. 278, pp. 130–134, Oct. 1997.

[12] G. Langner, M. Sams, P. Heil, and H. Schulze, "Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: evidence from magnetoencephalography," *J. Comp. Physiol. A.*, vol. 181, pp. 665–676, 1997.

[13] C. Liégeois-Chauvel, C. Lorenzi, A. Trébuchon, J. Régis, and P. Chauvel, "Temporal envelope processing in the human left and right auditory cortices," *Cereb. Cortex*, vol. 14, pp. 731–740, 2004.

[14] E. A. Lopez-Poveda, "An approximate transfer function for the dual-resonance nonlinear filter model of auditory frequency selectivity," *J. Acoust. Soc. Am.*, vol. 114, pp. 2112–2117, 2003.

[15] R. J. McGregor, *Neural and Brain Modeling*. San Diego: Academic, 1987.

[16] D. Oertel, R. R. Fay, and A. N. Popper, *Integrative functions in the mammalian auditory pathway*. New York: Springer-Verlag, 2002.

[17] S. Rosen, "Temporal information in speech: acoustic, auditory and linguistic aspects," *Philos. Trans. Soc. Lond. B. Biol. Sci.*, vol. 336, pp. 367–373, 1992.

[18] C. J. Sumner, E. A. Lopez-Poveda, L. P. O'Mard, and R. Meddis, "A revised model of the inner-hair cell and auditory nerve complex," *J. Acoust. Soc. Am.*, vol. 111, pp. 2178–2188, 2002.

[19] X. Wang, T. Lu, and L. Liang, "Cortical processing of temporal modulation," *Speech Communication*, vol. 41, pp. 107–121, 2003.