

Modeling neurons in the anteroventral cochlear nucleus for amplitude modulation (AM) processing: Application to speech sound

Ping TANG* and Jean ROUAT**

Université du Québec à Chicoutimi, Sciences Appliquées,
555 boul. de l'Université, CHICOUTIMI, QC, CANADA, G7H 2B1
(*INRS-télécommunication **Institut de Physiologie de Lausanne)
email: ptang@inrs-telecom.quebec.ca & jrout@uqac.quebec.ca

ABSTRACT

The aim of this work is to explore the representation of speech in the anteroventral cochlear nucleus by computational models based on neuroanatomical and neurophysiological data. A computational model of the anteroventral cochlear nucleus stellate cells with chop-S type response properties is described. Input that is excitatory and inhibitory to the model is in the form of simulated auditory-nerve spikes. The model is capable of generating a wide range of realistic chop-S unit responses. We investigate the speech information processing performed by our system that includes a peripheral model (cochlea, hair cells, auditory nerve spikes generation and chop-S units).

1. INTRODUCTION

The analysis and the recognition of speech spoken in noisy environment is a difficult and crucial task. Most of the speech recognizer alleviate the difficulties of these task by training on noisy data, assuming that the statistical properties of the noise will remain unchanged between the learning and the recognition phase. Other techniques assume that the speech source and the interference noise sources are geographically different. In summary, most of the effective techniques are useful for only specific conditions.

On the other hand, perceptive analysis and auditory models enhance the discriminant information from non stationary noise and are supposed to yield good performance in adverse conditions. However, their complexity and the difficulty of exploiting the dynamic output information with standard pattern recognition algorithms, restricts their integration in speech recognizer. Still, a good understanding of the auditory system, without an exact mimic of the perception, should help the scientist to design robust speech systems.

Two important questions can be addressed regarding the usefulness of perceptive approaches:

- (1) Which features robust to noise is extracted and/or enhanced by the auditory system?
- (2) How can these features be integrated in speech processing (or recognition) systems?

The present paper attempts to answer to the first question by exploiting the representation of speech in the anteroventral cochlear nucleus. A partial answer to the second question can be found in Rouat and Garcia [8].

In the mammalian auditory system, the acoustical information is first encoded into the auditory nerve by the cochlea. Then, the auditory nerve fibres terminate in the first important relay of the auditory system: the cochlear nucleus. It is comprised of various types of neurons that extract specific features from the signal representation observed at the level of the auditory nerve [9].

The cochlear nucleus complex is divided into three major divisions (anteroventral, posteroventral, dorsal) on the basis of differing cytoarchitecture. The three divisions play different functional roles in the perception of speech. We focus the present work on the study of the stellate cells with chop-S responses that are located in the anteroventral cochlear nucleus. The chop-S cells enhance the amplitude modulation (AM) characteristics of an input signal, a perceptually important feature of speech.

The chop-S seem to be organized in a systematic way according to their Best Modulation Frequency sensitivity [2].

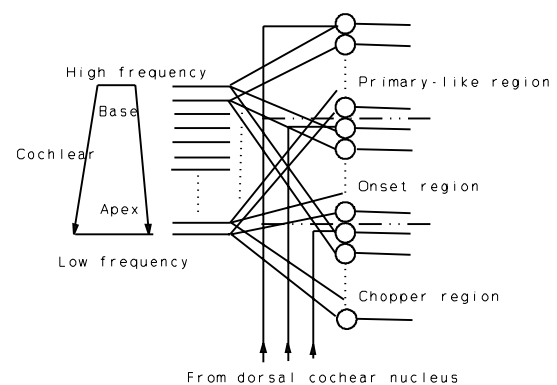


Figure 1: Array of cells with systematically varied AM sensitivity.

The acoustical information coming from primary auditory-nerve fibres is divided into parallel streams, each of which undergoes processing and encoding for transmission to higher auditory centres (Figure 1).

In section 2 we present the peripheral model, Section 3 details the model of the chop-S cell in the anteroventral cochlear nucleus. In section 4 we show that the model is capable of generating a wide range of realistic chop-S unit behaviour. We show also that our model can encode the amplitude modulation characteristics of an input signal.

2. THE PERIPHERAL MODEL

We first use the alternative spectral analysis module developed by Giguere and Woodland [3] as a cochlear model. It is a transmission line filterbank that takes into consideration the physiology of the cochlea and reproduces some of the non linear active processes occurring via the outer hair cells. The output signal of the cochlea is then fed into the inner haircell model as proposed by Meddis [6]. The spikes in the auditory nerve are randomly generated according to the work by Hewitt et al [4]. A flow diagram of the composite model used in our study is shown in figure 2.

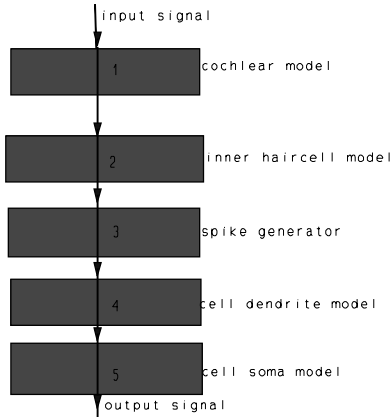


Figure 2: The composite model

As illustrated in figure 2, the composite model is comprised of five modules. The first is the cochlear model [3]. The second is the neural encoding module that is based on the inner haircell model of Meddis [6]. The third is the auditory nerve spike generator [4]. The fourth is the chop-S cell dendrite model and the last is the chop-S cell membrane model [1].

3. THE CHOPPER-S MODEL

We briefly give the formulation of the cell-soma and cell-dendrite model. Again, the chop-S cell model is developed according to the appropriate physiological data based on Banks and Sachs [1]. The model is similar in some respects to other models of chopper cells (Banks and Sachs 1991[1], Hewitt and al. 1992 [4]). It is based on a digital simulation of the Hodgkin and Huxley equations. The actual implementation of the model involves two components: the dendritic tree and the soma.

Figure 3 shows the electrical circuit representations for the soma and the two dendritic compartments. I_{Na} , I_K , I_L are the ion currents

that flow in response to the equivalent battery in series with conductance G_{Na} , G_K and G_L . C_m is the membrane capacitance and I_c is its charging or discharging current. V_m is an external voltage applied to the wire inside the axon and I is the resulting current. If the external source is an open circuit, V_m is the membrane potential and $I=0$.

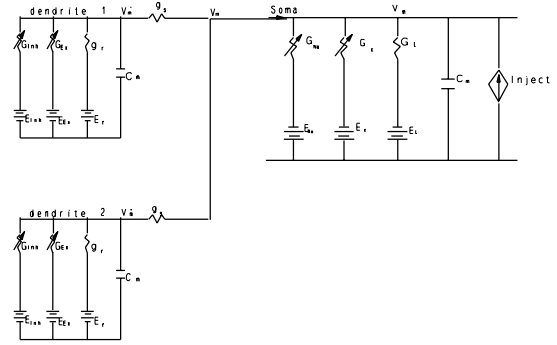


Figure 3: Model of cell dendrite and soma

The work by Hodgkin and Huxley can be partially summarized by the following six equations

$$I = I_c + I_{Na} + I_K + I_L \quad (1)$$

or

$$I = C_m \frac{dV_m}{dt} + G_{Na}(V - V_{Na}) + G_K(V - V_K) + G_L(V - V_L) \quad (2)$$

The equations and parameters describing the voltage-dependent processes are based largely on work by Banks and Sachs [1]. We also include a voltage-dependent scaling factor on the potassium activation time constant that allows greater flexibility in adjusting the firing frequency and the first spike latency in response to injected current. Three activity coefficients, m , h , n , are defined:

$$G_{Na} = 1.20m^3h, \quad G_K = 36n^4, \quad G_L = 0.3. \quad (3)$$

The activity coefficients are related to voltage and time as follow

$$\frac{dm}{dt} = \alpha_m(1-m) - \beta_m \cdot m \quad (4)$$

$$\frac{dn}{dt} = \alpha_n(1-n) - \beta_n \cdot n \quad (5)$$

$$\frac{dh}{dt} = \alpha_h(1-h) - \beta_h \cdot h \quad (6)$$

From this, the conductance values, the sodium, potassium, and chloride currents can be calculated as a function of time during an action potential.

4. SIMPLE STIMULUS RESPONSES

The chop-S model has been tested with different stimuli. We report responses to simple stimuli presented at the level of the

dendritic tree or of the soma. The results for a more complex sound are reported in section 5 where the cochlear and auditory nerve models are included.

An important stellate cell membrane characteristic is a linear steady-state response to small current injection on the soma.

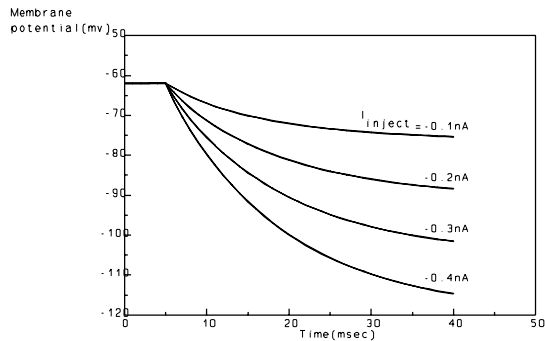


Figure 4: Membrane potential at the soma for a small current.

Figure 4 illustrates the membrane potential at the soma in response to a depolarizing and hyperpolarizing small current. The injected current is made of pulses of equal magnitude (-0.1, -0.2, -0.3, -0.4 nA) beginning at $t = 5$ ms and lasting $t=40$ ms. $t = 40$ ms is the time required by cell potential to reach its steady state value.

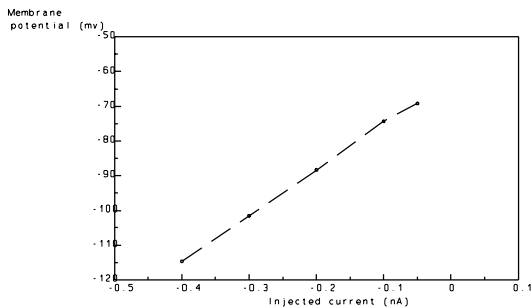


Figure 5: Steady-state membrane potential and injected current.

Figure 5 shows the steady state responses of the cell when small currents are injected with different intensity values. It is the potential of the membrane at $t=40$ ms obtained with the same stimuli as in figure 4. The results are similar to what is observed for stellate cells in vitro [7]. In fact, the responses are linearly distributed over a range of membrane potentials according to the work by Oertel [7].

A useful measure of regularity is the coefficient of variation CV, that is the ratio of the standard deviation of a random variable to its mean, in this case $CV = \sigma_R / \mu_R$. σ_R and μ_R are respectively the standard deviation and mean. Young et al.[10] characterize distinct chopper populations by using the CV measure. The chop-S units are regular choppers with CVs less than 0.3 that show only slight

variation over time. The CVs that we measured at post-stimulus onset are comprised between 0.0019 and 0.0031. It seems to be typical of chop-S neural data.

Chopper units are known to have an initial response that is highly regular, resulting in the characteristic multimodal post-stimulus time histogram (PSTH). Chopper PSTHs show series of regularly spaced peaks of discharge that become less distinct after about 50 ms.

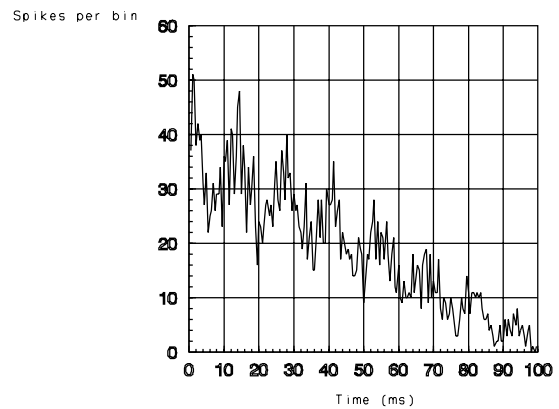


Figure 6: PSTH of responses to a 5kHz frequency tone with a 75Hz modulation rate.

On figure 6 we have reported the PSTH of the model in response to a 5-kHz tone that was amplitude-modulated over a rate of 75 Hz. The duration of the stimulus is 100ms with an amplitude of 100 units.

The chop-S cells in the anteroventral cochlear nucleus enhance the amplitude modulation with respect to their input signals [10]. We evaluate the modulation transfer function of our chop-S model. The spikes of the auditory nerve are generated from the hair-cell firing probability according to the work by Hewitt et al. [4].

The modulation transfer functions MTF are obtained with a 5kHz carrier tone that is amplitude modulated over a range of frequencies from 10Hz to 1000Hz. A MTF is illustrated in figure 7 and has been obtained by using a 100 units maximum amplitude signal. The MTF has a band-pass response that is a qualitative feature of chop-S units.

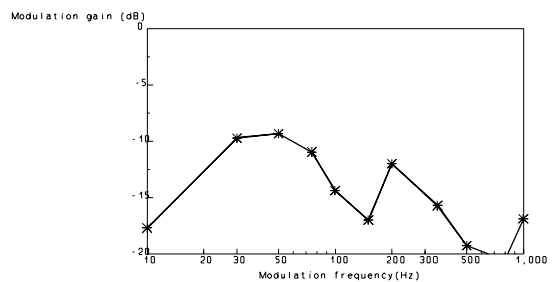


Figure 7: Model modulation transfer function of a chop-S unit.

5. COMPLEX STIMULUS RESPONSES (WITH COCHLEAR MODEL)

The complete composite model is used to analyze responses of chop-S units to natural speech.

On figure 9 the responses of 5 channels with different central frequencies have been reported along with the input signal (the short segment of a vowel in figure 8). The same chop-S unit has been used with each channel.

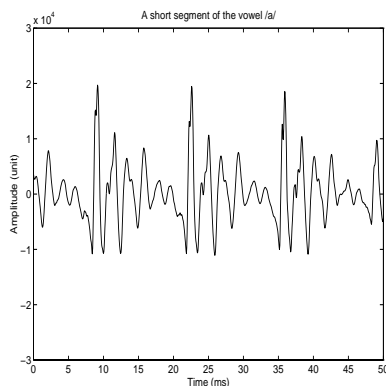


Figure 8: A short segment of the vowel /a/ (50ms).

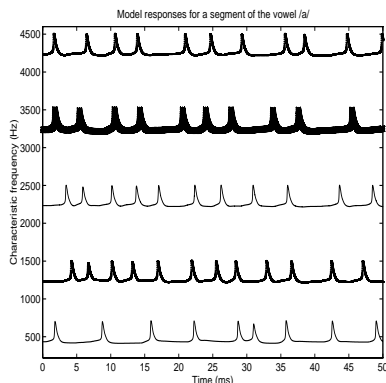


Figure 9: Responses of an array of Chop-S cell to a short segment of the vowel /a/.

6. CONCLUSIONS

We have presented and evaluated a computer model of a stellate cell with chop-S type responses in the anteroventral cochlear nucleus. The model can reproduce the main features of the chop-S cell responses. In order to better understand the contribution of these cells to the encoding of AM in speech, more experiments have to be done. Furthermore, to be more realistic, chop-S units with various best modulation frequencies should be used in combination with various channel central frequencies.

ACKNOWLEDGEMENTS

Many thanks to C. Lorenzi and F. Berthommier for stimulating discussions on chop-S. This work has been supported by the NSERC of Canada, by the Canadian Microelectronics Corporation and by the "fondation" from Université du Québec à Chicoutimi.

REFERENCES

- [1] Banks M. I. and Sachs M. B. (1991), "Regularity analysis in a compartmental model of chopper units in the anteroventral cochlear nucleus", *J. of Neurophysiology*, Vol. 65.No. 3, 606-629.
- [2] Evans E. F. (1991), "Auditory processing of complex sounds: an overview", *Proceedings of a Royal Society Discussion Meeting held on 4 and 5 December 1991, organized and edited by R. P. Carlyon, C. J. Darwin and I. J. Russell*, Oxford Press, 296-306.
- [3] Giguere, C. and Woodand, P. C. (1994), "A computational model of the auditory periphery for speech and hearing research", *J. acoust. Soc. Am.* 95, 331-342.
- [4] Michael J. Hewitt, Ray Meddis and Trevor M. Shackleton (1992), "A computer model of a cochlear-nucleus stellate cell: Responses to amplitude-modulated and pure-tone stimuli", *J. Acoust. Soc. Am.* 91 (4), Pt. 1, 2096-2109.
- [5] Kim D. O., Sirianni J. G., and Chang S. O. (1990), "Responses of DCN-PVCN neurons and auditory nerve fibres in unanesthetized decerebrate cats to AM and pure tones: Analysis with autocorrelation/power-spectrum", *Hear.Res.* 45, 95-113.
- [6] Meddis, R. (1988), "Simulation of auditory-neural transduction: Further studies", *J. acoust. soc. Am.* 83, 1056 -1063.
- [7] Oertel D., Wu S. H., and Hirsch J. A. (1988), "Electrical characteristics of cells and neural circuitry in the cochlear nucleus studied with intracellular recording from brain slices", In *Auditory Function*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan. New York:Wiley 313-336.
- [8] Rouat J. and Garcia M. (1996), "A prototype speech recogniser based on associative learning and nonlinear speech analysis", in *computational auditory scene analysis*, eds. H. Okuno and D. Rosenthal, Erlbaum, to appear.
- [9] Rouiller Eric M., "Organisations fonctionnelle des voies auditives", Editions INSERM 1992, EDITE PAR Romand, 50-65.
- [10] Young E. D., Robert J-M., and Shofner W. P. (1988), "Regularity and latency of units in the ventral cochlear nucleus: Implications for unit classification and generation of response properties", *J. Neurophysiol.* 60, 1-29.