A Functional Model of the Auditory Peripheral System

By Katuhiro Maki

A thesis submitted to
School of Information Science,
Japan Advanced Institute of Science and Technology
in partial fulfillment of the requirements
for the degree of
Master of Information Science
Graduate Program in Information Science

Written under the direction of
Associate Professor Masato Akagi

February 14, 1997

Abstract

This paper presents a functional model of the auditory peripheral system for obtaining input impulse trains to the central auditory system. To model the external ear, the middle ear, the basilar membrane (BM) and the outer hair cell (OHC), a dual analog model of the ascending path reported by Giguère et al. (1994) is adopted. In this paper, an inner hair cell (IHC) model is developed by extending Meddis's model (Meddis et al., 1986, 1988). This model can simulate nonlinear transduce functions of the IHC, which are depolarized and hyperpolarized peak responses as a function of the peak sound pressure level and the DC components of the receptor potential as a function of the stimulus level. An auditory nerve (AN) model is proposed using Hodgkin's cell membrane model (1952) to generate nerve impulse trains. These models are combined into a functional model of the auditory peripheral system. Output of the functional model is compared with physiological experimental data. The results show that the proposed model is in excellent agreement with the physiological data and that the model is effective in providing primary inputs to central auditory processing models. Additionally, using vowels as input data for the model, discharge patterns of all characteristic frequencies (CFs) can be obtained from the output of the model. These patterns show how the vowel features are represented in the auditory peripheral system.

Contents

Acknowledgment

There are many people who were involved in providing an environment in which this thesis and the ideas contained within could be realized.

I would like to extend my profound gratitude to my advisor, associate professor Masato Akagi, for infinite patience, kindness, encouragement, enthusiasm, and encyclopedic knowledge of speech signal processing and the speech perception. Despite being in high demand, he always made the time to advise and share his broadband wisdom in matters beyond just speech signal processing and the speech perception.

I also thanks to my thesis readers, professer Taizo Iijima and associate professor Kazunori Kotani and their insightful reading of my original proposal spurred me into conducting subjective tests, immeasurably improving the final work.

I would especially like to thanks Mr. Yuji Yonezawa, Unix sysadmin, who helped greatly and showed great patience and trust as I continually performed minor surgeries on our network. I would like to thanks Mr. A.C.R. Nandasena, who supported proofreading this thesis.

I would like to thank all members, past and present, of Acoustic Information Science laboratory and its heir for their part in marking a wouderful and supportive working community. Particular honors go to associate Koichi Iwaki, Mr. Yasuhiro Aoki, Mr. Tatsuya Kitamura, Mr. Takashi Ozeki, Mr. Masashi Unoki, Mr. Wataru Hashimoto, Mr. Koichi Kawamoto, Mr. Katsumi Hayashi, Mr. Tomoya Minakawa, Mr. Mitsuhiro Murumachi, Mr. Wahyu Widida.

My greatest thanks are to my parents, who made all this possible through their incredible support. I am proud to be able to dedicate this dissertation to them.

Chapter 1

Introduction

Most long-standing theories of speech perception and of the neural basis of language in general have emphasized the important role of neural structures within the cerebral cortex in normal language processing (Pont, 1991). However, as a result of recent investigations exploring the representation of speech and other complex sounds at the level of the auditory nerve, (Sachs and Young, 1979; Delgutte, 1980; Miller and Sachs, 1983; Sinex and Geisler, 1983; Carney and Geisler, 1986; Geisler, 1988; Sinex and MacDonald, 1988), it has become increasingly apparent that some sophisticated processing of the speech waveform occurs in the auditory periphery. Further processing occurs between the auditory nerve and cortex. If we are to understand fully the nature of speech perception - whether for reasons of scientific curiosity, to help in the prevention and treatment of various hearing disorders, or to assist in developing powerful speech recognizers - then the role of this processing must be investigated.

From the auditory nerve, information is transmitted to the cortex via the brain-stem nuclei. This includes the cochlear nucleus, the superior olivary complex, the inferior colliculus, and the thalamic auditory relay, the medical geniculate body (Hackney, 1987). Since all afferent fibers within the auditory nerve terminate in the cochlear nucleus, this nucleus seems a logical point to begin an investigation of neural processing beyond the auditory periphery.

The aim in this study was exploring the representation of speech in the cochlear nucleus by computer simulation. A model of the cochlear nucleus for studying its sophisticated functions must include procedures for dealing with action potentials and produces action potentials that can be compared with those from actual physiological experiments.

Action potentials must also be prepared as inputs to the model of the cochlear nucleus. A model of the auditory periphery within the auditory nerve must be prepared in other words. For this peripery model, (1) the action potentials must be a spike train, (2) output of the auditory peripheral model for the cochlear nucleus model must be in quantitative agreement with physiological data, and (3) the auditory peripheral model should consist of physiological agreed submodels, i.e., external ear, middle ear and inner ear. Therefore, this study marks the first stage in modeling the auditory peripheral system to satisfy above three requirements.

Previous auditory peripheral models over the last two decades have not included the

above three requirements. Although the peripheral models proposed by Carney, 1993, Jenison, 1991, Kates, 1991, Payton, 1988, and Schoonhoven, 1994, consisted of submodels, the models can not produce action potentials and the output of the model was not evaluated in detail. It was remarked that the models were not sufficiently modeling hair cell-primary fiber complex. Models of the hair cell-primary fiber complex was contributed for good evaluation. (Hewitte et al., 1991) The models of the hair cell-primary fiber complex proposed by Smith et al. 1980, Ross, 1996, Meddis et al., 1986 was in good agreement with physiological data but the model was a part of the peripheral model, that is, the models did not have external ear and middle ear. In addition to this, the models proposed by Smith et al. 1980, Meddis et al., 1986 did not produce action potentials. The other auditory models proposed by Ghitza, 1988, Senef, 1988, Deng et al., 1987, and Shamma, 1985 etc. have at least one of the drawbacks describes above.

This paper attempts to improve the drawbacks of these previous models. The models of the external ear, the middle ear, the basilar membrane (BM) and the outer hair cell (OHC) used in this work are based on the model reported by Giguère and Woodland (1994). The IHC model of this paper is obtained by extending Meddis's (1986, 1988) model, which was a good model of the hair cell-primary fiber complex, by taking into account the saturating nonlinearity of the DC component also. The auditory nerve (AN) model is proposed to generate a spike train and to model the spike discharge pattern directly These models are connected into a functional model of the auditory peripheral system. Response patterns of the auditory peripheral model are evaluated in detail in Table 1.

Chapter 2

Description of the model

A schematic description of the model is shown in Fig.1.

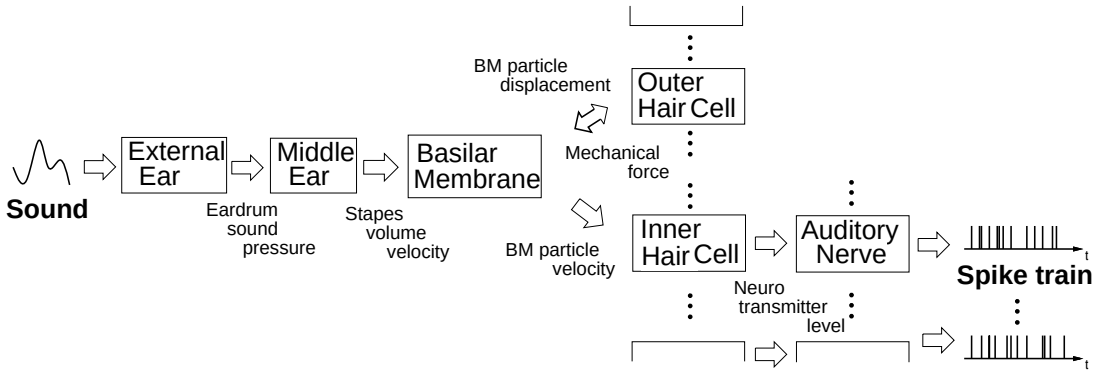


Figure 1. Schematic description of the auditory peripheral model.

The model consists of the external ear, the middle ear, the basilar membrane (BM), the outer hair cell (OHC), the inner hair cell (IHC) and the auditory nerve (AN) model.

2.1 External ear, middle ear, basilar membrane and outer hair cell model

The models of the external ear, middle ear, BM and OHC used in this study are based on the work of Giguère and Woodland (1994). They consist of a concatenation of electrical circuit submodels, and were implemented by applying the techniques of mesh analysis, Laplace transformation and bilinear transformation.

2.1.1 Model descriptions

The models of the external ear, middle ear, BM and OHC represented by electrical circuit are shown Figs.2, 3 and 4, respectively (Giguère and Woodland, 1994).

External ear model

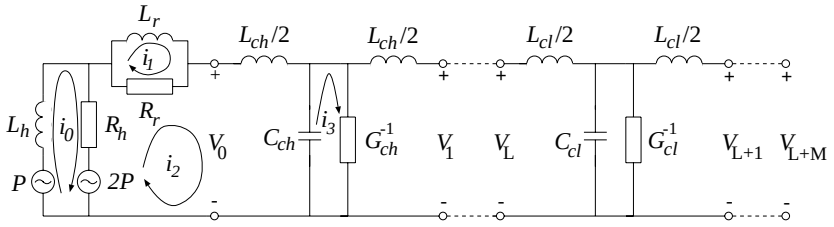


Figure.2. Electroacoustic network of the outer ear. Only the first segment (index 1) of the concha and the first segment (index L+1) of the auditory canal transmission lines are shown in fell.

From Giguère et al., it assumes that the principal obstacle (upper torso and head) confronting the incident sound wave can be represented by a solid sphere with effective radius a_s, while the ear opening can be represented by a small orifice on the surface of the sphere. The radius a_{ch} of the orifice corresponds to the effective radial size of the concha cavity at the base of the pinna. The pinna flange is not itself modeled. In Fig. 2, two voltage sources of amplitudes $P(t)$ and $2P(t)$ drive the external ear network in-phase, where $P(t)$ is analogous to the pressure of the incident free-field sound wave. These sources, together with elements L_h and R_h, model the sound diffraction associated with the principal obstacle. The latter elements are given by :

$$L_h = \frac{0.5\rho_a}{\pi a_s}, \quad R_h = \frac{\rho_a c}{\pi a_s^2}, \tag{2.1}$$

where ρ_a is the air density and c is the sound velocity. The parallel elements L_r and R_r form the equivalent circuit for the acoustic radiation impedance of the ear opening, i.e., the load seen by a hypothetical massless piston located at the concha entrance and radiating energy into the surrounding medium. They are given by :

$$L_r = \frac{0.7 rho_a}{\pi a_{ch}}, \quad R_r = \frac{\rho_a c}{\pi a_{ch}^2}. \tag{2.2}$$

The voltage $V_0(t)$ is analogous to the sound pressure at the entrance to the concha.

The concha cavity is an approximately cylindrical acoustic resonator of radius a_{ch} and length l_{ch} providing a broad pressure gain around its first normal mode of vibration at 4300 Hz. For frequencies up to the second normal mode ($\simeq$ 7100 Hz), Giguère et al. assumed that the concha can be represented by an $L-$segment uniform transmission line as shown in Fig. 2. By use of electroacoustic analogies for cylindrical tubes, the network

elements L_{ch} and C_{ch} characterizing each $T-$junction are equivalent to the acoustic mass and compliance of each discretized segment:

$$L_{ch} = \frac{\rho_a}{\pi a_{ch}^2}\Delta x, \quad C_{ch} = \frac{\pi a_{ch}^2}{\rho_a c^2}\Delta x, \tag{2.3}$$

where $\Delta x = l_{ch}/L$ is the segment length. Energy losses are more difficult to account for because they originate from multiple mechanisms (e.g., viscous friction, thermal conduction and vibrations at the walls) and are in general frequency dependent. To a first approximation, valid for small loss conditions, the damping mechanisms can be adequately modeled by lumping all effects into a single constant shunt conductance G_{ch} in each segment. From Flanagan (1972), an expression for this network element can be derived as

$$G_{ch} = \frac{2\alpha_{ch}}{Z_{ch}}\Delta x, \tag{2.4}$$

where α_{ch} is the effective attenuation constant of the propagating waves per unit length and $Z_{ch} = \sqrt{L_{ch}/C_{ch}}$ is the characteristic impedance of the line. The auditory canal is an acoustic waveguide of irregular shape whose length governs the primary resonance frequency of the complete outer ear at around 2600 Hz. For frequencies up to about 8000Hz, the canal geometry can be approximated by a straight cylindical tube of radius a_{cl} and length l_{cl}. This can be modeled as an $M-$segment uniform transmission line connected in series with the concha as shown in Fig. 2. In analogy to Eqs. (2.3)-(2.4), the network elements in each $T-$junction are given here by

$$L_{cl} = \frac{\rho_a}{\pi a_{cl}^2}\Delta x, \quad C_{cl} = \frac{\pi a_{cl}^2}{\rho_a c^2}\Delta x, \quad G_{cl} = \frac{2\alpha_{cl}}{Z_{cl}}\Delta x, \tag{2.5}$$

where $\Delta x = l_{cl}/M$ is the segment length, α_{cl} is the effective attenuation constant of the auditory canal per unit length, and $Z_{cl} = \sqrt{L_{cl}/C_{cl}}$ is the characteristic impedance of the line. The voltage $V_{L+M}(t)$ at the end of the canal is analogous to the eardrum sound pressure.

Middle ear model

Giguère et al. modified the terminal branch of the network of Lutman and Martinso as to allow the concha to be explicitly represented, as detailed in Giguère and Woodland (1992). The main modification was to replace elements R_c and L_c by an ideal transformer 1:r representing the effective acoustic transformer ratio between the eardrum and the oval window. This allows the cochlear network to be directly connected to the middle ear. It was also necessary to add a resistor R_{al} to account for acoustic resistance of the annular ligaments at the oval window. Bsed on the experimental work of Lynch et al. (1982) on cats, R_{al} was taken as about $\frac{1}{6}$ of the real part of the cochlear input impedance at midfrequencies. The terminal branch should also normally include a series inductor to account for the acoustic mass of the stapes. Its contribution, however, is very small compared to the imaginary part of the cochlear input impedance and was neglected. The

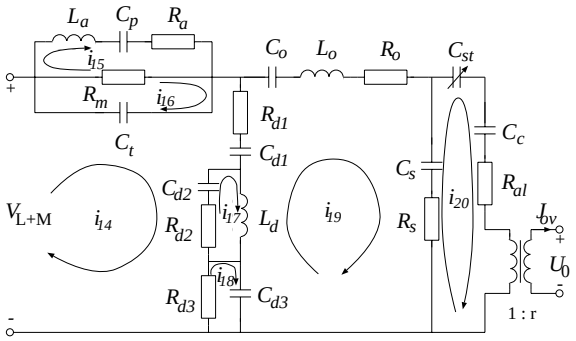


Figure.3. Final electroacoustic net work of the middle ear [adapted from Lutman and Martin (1979)]. It is connected to the left to the outer ear net work of Fig.2 and to the right to the cochlear net work of Fig.4.

current J_{ov} is analogous to the volume velocity of the stapes footplate. All other net work elements in Fig.3 are identical to those in Lutman and Martin (1979). In particular, C_c represents are combined acoustic compliance of the round window membrane and of the annular ligaments at the oval window. The time variant capacitor C_{st} models the variable acoustic compliance of the stapes suspension in response to stapedial muscle contractions.

Basilar membrane (BM) model

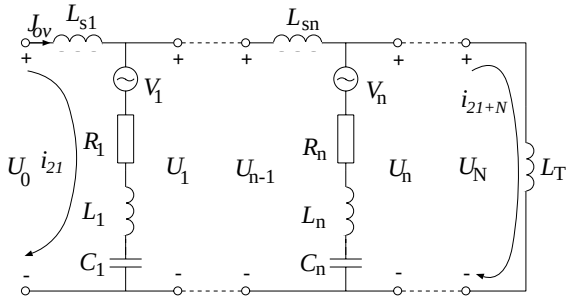


Figure. 4 electroacoustic analogue of the cochlea. Only the first and nth segments are shown in full together with the basal and apical boundary conditions.

The BM is spatially discretized into N segments of length Δx. The position, or place of a given segment indexed n is measured from the base of the cochlea. The voltage $U_n(t)$ is analogous to the pressure difference between the scala vestibuli and the scala tympani. The shunt current $I_n(t)$ represents the transversal volume velocity of the corresponding BM segment. The characteristic resonance frequency f_n of the basilar membrane is decreasing from the base to the apex. The cochlear mapping function of Greenwood (1990) is used to establish a formal correspondence. It can be expressed as

$$x_n = l_{bm} - \frac{1}{0.6 \ \mathrm{cm}^{-1}} log\left(\frac{f_n}{165.4 \ \mathrm{Hz}} + 1\right), \tag{2.6}$$

where l_{bm} is the total length of the BM. To reduce computational load in some applications, the BM is not discretized over its entire length, but only over a portion of interest (Giguère et al, 1994). The end points x_1 and x_N of this portion correspond to the maximum f_1 and minimum f_N desired auditory filter characteristic frequencies. It follows from Eq. (2.6) that the segment length is

$$\Delta x = \frac{(1/0.6\text{cm}^{-1}) log[(f_1 + 165.4\text{Hz})/(f_N + 165.4\text{Hz})]}{N-1}. \tag{2.7}$$

The transmission line network elements are derived from electroacoustic analogies and from the assumption that the natural frequency of the shunt second-order resonant circuit in each segment is equal to the characteristic frequency of the BM at that place. Inductor L_{sn} represents the acoustic mass of the scala vestibuli and tympani fluids:

$$L_{sn} = \frac{2\rho_w \Delta x}{A(x_n)}, \tag{2.8}$$

where ρ_w is the fluid density, and A(x) is the mean cross-sectional area of the scalae as a function of BM place. Inductor L_n, capacitor C_n, and resister L_{sn} represents the acoustic mass, compliance, and resistance components of the BM point impedance:

$$L_n = \frac{M_n}{b(x_n)\Delta x}, \quad C_n = \frac{1}{4pi^2 f_n^2 L_n}, \quad R_n = Q^{-1}\sqrt{\frac{L_n}{C_n}}, \tag{2.9}$$

where M_n is the transversal mass per area of BM, $b(x)$ is the width of the BM as a function of place, and Q_n is the quality factor of the shunt resonant circuit. The apical end of the line is terminated by an inductance L_T representing the acoustic mass of the cochlear fluid from the last BM segment to the helicotrema. From Eq. (2.8), Giguère et al. find

$$L_T = \int_{x_N+\Delta x}^{L} \frac{2\rho_w}{A(x)} dx. \tag{2.10}$$

The helicotrema is itself modeled as a short circuit. The cochlear network of Fig. 4 is explicitly connected to the middle ear network of Fig. 3.

Outer hair cell (OHC) model

From Giguère et al., this mechanism is assumed to comprise two main stages: (1) A nonlinear frequency-independent transduction of the transversal displacement of the organ of Corti into OHC receptor currents, and (2) an OHC force applied to the organ of Corti which depends on the receptor currents. The motion of the organ of Corti and the OHC receptor currents are thus bounded by a feedback loop. This feedback is primarily effective for frequencies near the characteristic frequency at a given BM place and its net effect is to reduce damping of the organ of Corti. In Fig4, the net pressure developed by the fast motile mechanism of the OHCs is represented in each segment by a voltage source saturating at high amplitudes:

$$V_n^{\text{ohc}} = GR_n \left(\frac{d_{1/2}}{d_{1/2} + |d(t)|} \right) I_n(t), \tag{2.11}$$

where $0 < G \leq 1$ is a gain factor, $d_n(t)$ is the BM particle displacement, and $d_{1/2}$ is a constant equal to the BM displacement at the half-saturation point of the nonlinearity. The proportionaly of $V_n^{\rm ohc}(t)$ to BM volume velocity $I_n(t)$ in Eq. (2.11) ensures that the effect of the OHCs is to reduce damping.

Using Eq. (2.11), the voltage $U_n(t)$ across BM segment n in Fig. 4 is

$$U_n(t) = R_n \left[1 - G\left(\frac{d_{1/2}}{d1/2 + |d(t)|}\right)\right] I_n(t) + L_n \frac{dI_n(t)}{dt} + \frac{1}{C_n}\int_{-\inf}^{t} I_n(t)dt\,, \quad (2.12)$$

$$= R_n^{\rm ohc}(t) I_n(t) + L_n \frac{dI_n(t)}{dt} + \frac{1}{C_n}\int_{-\inf}^{t} I_n(t)dt\,. \quad (2.13)$$

Thus, the series combination of voltage sourse $V_n^{ohc}(t)$ and resistor R_n can be equivalently represented by a time-variant resistor $R_n^{ohc}(t)$ function of the BM motion.

The BM particle displacement $d_n(t)$ and velocity $i_n(t)$ are the output variables of the cochlear network. From electroacoustic relation, they are given by

$$i_n(t) = \frac{I_n(t)}{b(x_n)\Delta x}, \quad (2.14)$$

$$d_n(t) = \frac{C_n V_{cn}(t)}{b(x_n)\Delta x}, \quad (2.15)$$

where $b(x_n)\Delta x$ is the BM segment area, and $V_{cn}(t)$ is the voltage drop across C_n.

2.1.2 Implementation

The electrical circuit submodels as shown in Figs. 2, 3, and 4 were implemented by applying the techniques of mesh analysis, Laplace transformation and bilinear transformation. All of parameter values of Fig. 2, 3 and 4 are identical to those of Giguère and Woodland (1994). But the segment number N of the basilar membrane model can set variably in this study.

First, we applied the techniques of mesh analysis to electrical circuit as shown in Figs. 2, 3 and 4. Resulting formula was given by Eq. (2.16). For father details of mesh analysis, see Desore (1969).

$$\boldsymbol{Z}_m(D)\boldsymbol{i} = \boldsymbol{e}_s i\,, \quad (2.16)$$

$$Z_{\rm m}({\rm D}) = \begin{bmatrix} L_h D+R_h & 0 & -R_h & 0 & \cdots & & \cdots & 0 \\ 0 & L_r D+R_r & -R_r & 0 & \cdots & & \cdots & 0 \\ -R_h & -R_r & \frac{L_{ch}}{2}+R_h+R_r+\frac{1}{C_{ch}D} & 0 & \cdots & & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & & & \cdots & 0 & L_{sN}+L_{(N-1)}D+R_{(N-1)}+\frac{1}{C_{(N-1)}D}+ & -L_N D-R_N-\frac{1}{C_N D} \\ & & & & & & L_N D+R_N+\frac{1}{C_N D} & \\ 0 & \cdots & & & \cdots & 0 & -L_N D-R_N-\frac{1}{C_N D} & L_T+L_N D+R_N+\frac{1}{C_N D} \end{bmatrix}, \quad (2.17)$$

$$\boldsymbol{i} = \begin{bmatrix} i_0 \\ i_1 \\ i_2 \\ i_3 \\ i_4 \\ \vdots \\ i_{(2l+N-3)} \\ i_{(2l+N-2)} \\ i_{(2l+N-1)} \\ i_{(2l+N)} \end{bmatrix}, \tag{2.18}$$

$$\boldsymbol{e}_s = \begin{bmatrix} P-2P \\ 0 \\ 2P \\ 0 \\ 0 \\ \vdots \\ 0 \\ V_1^{ohc} \\ V_2^{ohc} - V_1^{ohc} \\ V_3^{ohc} - V_2^{ohc} \\ \vdots \\ V_{N-1}^{ohc} - V_{N-2}^{ohc} \\ V_N^{ohc} - V_{N-1}^{ohc} \\ -V_N^{ohc} \end{bmatrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \cdots \text{line } 2l \\ \\ \\ \\ \\ \\ \cdots \text{line } (2l+N) \end{matrix} = \begin{matrix} \boldsymbol{V}_{s1} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{bmatrix} P-2P \\ 0 \\ 2P \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{matrix} \boldsymbol{V}_{s2} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ V_1^{ohc} \\ V_2^{ohc} - V_1^{ohc} \\ V_3^{ohc} - V_2^{ohc} \\ \vdots \\ V_{N-1}^{ohc} - V_{N-2}^{ohc} \\ V_N^{ohc} - V_{N-1}^{ohc} \\ -V_N^{ohc} \end{bmatrix}, \tag{2.19}$$

where, $\boldsymbol{Z}_m(D)$ is a mesh impedance matrix, $\boldsymbol{i}$ is a mesh current, and $\boldsymbol{e}_s$ is a mesh voltage source vector. The symbol "D^{-1}" is used to refer to a definite integral "$\int_0^t dt'$".

Equation (2.16) can be written in the form

$$(\boldsymbol{A}_2 D + \boldsymbol{A}_1 + \boldsymbol{A}_0 \frac{1}{D})\boldsymbol{i} = \boldsymbol{V}_{s1} + \boldsymbol{V}_{s2}. \tag{2.20}$$

Secondly, we apply the technique of Laplace transformation to each matrix equation shown in Eq. (2.16). Equation (2.21) is obtained from Eq. (2.20) by applying the technique of Laplace transformation and by substituting $\frac{1}{S}$ for D, where $s = 0$.

$$(\boldsymbol{A}_2 S + \boldsymbol{A}_1 + \boldsymbol{A}_0 \frac{1}{S})\boldsymbol{I}(s) = \boldsymbol{V}_{s1} + \boldsymbol{V}_{s2} \tag{2.21}$$

Thirdly, we apply the technique of Z transformation to Eq. (2.21) by substituting Eq. (2.22) for Eq. (2.23) in the bilinear transformation of Eq. (2.22) into Eq. (2.23).

$$S = \frac{2}{T}\frac{1 - Z^{-1}}{1 + Z^{-1}} = \alpha \frac{1 - Z^{-1}}{1 + Z^{-1}} \tag{2.22}$$

$$[(\alpha^2 \boldsymbol{A}_2 - \alpha \boldsymbol{A}_1 + \boldsymbol{A}_0)Z^{-2} + (-2\alpha^2 \boldsymbol{A}_2 + 2\boldsymbol{A}_0)Z^{-1} + (\alpha^2 \boldsymbol{A}_2 + \alpha \boldsymbol{A}_1 + \boldsymbol{A}_0)]\boldsymbol{I}(s) = \alpha(1 - Z^{-2})V_{1s} + \alpha(1 - Z^{-2})V_{2s} \tag{2.23}$$

If we define $\beta_2, \beta_1, \beta_0$ in the form

$$\beta_2 = -\alpha^2 \boldsymbol{A}_2 + \alpha \boldsymbol{A}_1 - \boldsymbol{A}_0, \tag{2.24}$$

$$\beta_1 = 2\alpha^2 \boldsymbol{A}_2 - 2\boldsymbol{A}_0, \tag{2.25}$$

$$\beta_0 = \alpha^2 \boldsymbol{A}_2 + \alpha \boldsymbol{A}_1 + \boldsymbol{A}_0, \tag{2.26}$$

then, Eq.(2.23) can be written as

$$\boldsymbol{I}(s) = \beta_0^{-1}\{(\beta_1 Z^{-1} + \beta_2^{-2})\boldsymbol{I}(s) + \alpha(1 - Z^{-2})V_{1s} + \alpha(1 - Z^{-2})V_{2s}\}. \quad (2.27)$$

Forth, we applied the technique of reverse Z transformation to Eq(2.27). The Equation (2.27) becomes

$$\boldsymbol{i}(n) = \beta_0^{-1}(\beta_1 \boldsymbol{i}(n-1) + \beta_2 \boldsymbol{i}(n-2) + \alpha(V_{1s}(n) - V_{1s}(n-2)) + \alpha(V_{2s}(n) - V_{2s}(n-2)) \quad (2.28)$$

From Eq(2.28), we can obtain a discreet-time representation of (2.16).

As the other form, Eq(2.29) is obtained from Eq(2.20) by substituting for $\boldsymbol{y} = \int_0^t \boldsymbol{i}\,dt'$ and by applying the technique of Laplace transformation

$$(\boldsymbol{A}_2 S^2 + A_1 S + A_0)\boldsymbol{Y} = \boldsymbol{V}_{s1} + \boldsymbol{V}_{s2} \quad (2.29)$$

The Eq(2.30) is obtained from Eq(2.29) by applying the techniques of bilinear transformation

$$\boldsymbol{Y}(s) = \beta_0^{-1}\{(\beta_1 Z^{-1} + \beta_2^{-2})\boldsymbol{Y}(s) + \alpha(1 - Z^{-1})^2 V_{1s} + \alpha(1 - Z^{-1})^2 V_{2s}\}. \quad (2.30)$$

Then, Eq(2.31) is obtained from Eq(2.30) by applying the techniques of Inverse Z transformation

$$\boldsymbol{y}(n) = \beta_0^{-1}(\beta_1 \boldsymbol{y}(n-1) + \beta_2 \boldsymbol{y}(n-2) + V_{1s}(n)2\,V_{2s}(n-1)\,V_{2s}(n-2) \quad (2.31)$$

where, β_0, β_1, β_2 is identical to Eqs.(2.24), (2.25), and (2.26).

The shunt current $I_n(t)$ and $V_{cn}(t)$ (for discreet representation, $I_n(t)$ is $I_n(k)$ and $V_{cn}(t)$ is $V_{cn}(k)$, $k = 1, 2, 3\cdots$.) , the voltage drop across C_n, is easily calculated using Eqs.(2.28) and (2.31).

The BM particle displacement $d_n(t)$ and velocity $i_n(t)$ (for discreet representation, $d_n(t)$ is $d_n(k)$ and $i_n(t)$ is $i_n(k)$, $k = 1, 2, 3\cdots$.) which are the output variables of the cochlear network are calculated by Eqs.(2.14) and (2.15), respectively .

As the initial condition, $V_{1s}(0)$ and $V_{2s}(0)$ can not calculate from discreet-time representation Eqs.(2.28) and (2.31). Therefore, Eqs.(2.28) and (2.31) are rewritten as

$$\begin{aligned} \boldsymbol{i}(n) &= \beta^{-1}(\beta_1 \boldsymbol{i}(n-1) + \beta_2 \boldsymbol{i}(n-2) + \alpha(V_{1s}(n) - V_{1s}(n-2)) + \alpha(V_{2s}(n-1) - V_{2s}(n-3)), \\ \boldsymbol{y}(n) &= \beta^{-1}(\beta_1 \boldsymbol{y}(n-1) + \beta_2 \boldsymbol{y}(n-2) + V_{1s}(n-1)2\,V_{2s}(n-2)\,V_{2s}(n-3). \end{aligned} \quad (2.32)$$

To minimize calculation error, $i_n(k)$ and $d_n(k)$ are calculated from both $\boldsymbol{i}(n)$, $\boldsymbol{y}(n)$ as shown in Eq(2.32).

Figure 2.1 shows the response of the cochlear network to the vowel /i/. In the top row, time waveforms for the vowel /i/ are shown.

In Figure 2.1, the BM filtering is level dependent. The tuning curves are sharp at low levels and broad at high levels. Consequently, the response of cochlear network has the characteristics of a wideband analysis in the high-energy regions and those of a narrow band analysis in the low-energy regions. The phone boundaries are preserved, and there is compression and spectral sharpening of speech features.

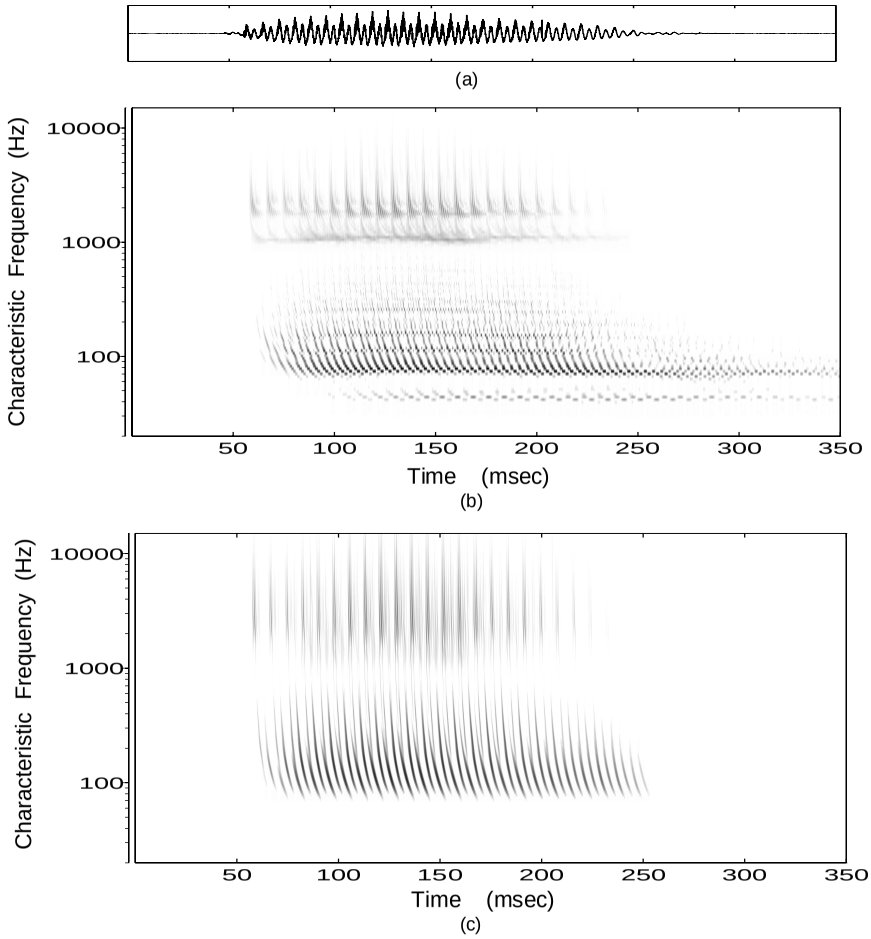


Figure.2.1. The response of cochlear network. (a) The input time waveform for the vowel /i/. (b), (c) The BM velocity response of the cochlear network to the vowel /i/. Input sound pressure level: (b) 60dB, (c) 100dB.

2.2 Inner hair cell (IHC) model

The IHC model proposed by Meddis (1986, 1988) is extended by modifying the membrane permeability as a function of amplitude of the BM velocity which is the output of the cochlear network.

2.2.1 Meddis's IHC model

The IHC model proposed by Meddis (1986、1988) can be fully understood in terms of the production, movements, and dissipation of transmitter substance in the region of the hair cell-auditory-nerve fiber synapse. The IHC model proposed by Meddis (1986、1988)

describes below.

$$k(t) = \begin{cases} g\frac{s(t)+A}{s(t)+A+B} & ,(s(t)+A) > 0 \\ 0 & ,(s(t)+A) \leqq 0, \end{cases} \tag{2.33}$$

where A and B are constant parameters of the model and $s(t)$ is the instantaneous amplitude of the signal.

$$\frac{dq}{dt} = y(m - q(t)) + xw(t) - k(t)q(t), \tag{2.34}$$

$$\frac{dc}{dt} = k(t)q(t) - lc(t) - rc(t), \tag{2.35}$$

$$\frac{dw}{dt} = rc(t) - xw(t). \tag{2.36}$$

An amount $q(t)$ of transmitter exists inside the cell wall near the synapse. A fraction $k(t)q(t)dt$ of this transmitter is released, between time t, and time $t + dt$ across the membrane into the cleft. A permeability factor $k(t)$ is a nonlinear function of the instantaneous amplitude of the signal after mechanical effects have been taken into account.

A fraction $lc(t)dt$ of the amount $c(t)$ of transmitter in the cleft is subject to chemical destruction or loss through diffusion. Another fraction $rc(t)dt$ is taken back up into the cell. The rest remains in the cleft to stimulate the postsynaptic membrane. It is assumed for the sake of simplicity, that spike occurrence in the auditory nerve is linearly probabilistically related to the residue of transmitter substance in the cleft. Accordingly, the quantity $c(t)$ is to be identified with the "excitation function" of Gaumond et al. (1982, 1983) or Gray's (1967) "recovered probability," i.e, the probability of spike emission disregarding refractory effects.

Transmitter taken back into the cleft is not immediately available for release again but is delayed in a reprocessing store. A fraction $xw(t)dt$ of the amount of transmitter $w(t)$ in this store is continuously transferred to the free transmitter pool. The transmitter originates in a manufacturing base or "factory" that replenishes the free transmitter pool at a rate $y[m - q(t)]$, where m is the (approximate) maximum amount of transmitter tube found in the pool. In the unquantified version of the model, m is set to unity and all transmitter amounts are construed as fractions of the total possible amount.

2.2.2 Improvement to the Meddis's IHC Model

The membrane permeability, $k(t)$, in Eq. (2.33) is a simple approximation of the depolarizing and hyperpolarizing peak responses as a function of peak sound pressure for the guinea pig (Russel and Sellik, 1978) and for the turtle (Crawford and Fettiplace, 1981). But, the peak responses described here are not thought to represent the nonlinear transducer function of the IHC (Dallos et al., 1989). In addition, Meddis's model does not take into account the properties of saturating nonlinearity of the DC component. Therefore, the firing rate, which is a function of sound level based on the output of the model, is quantitatively different from that of the physiological data (see Hewitte et al., 1991)

To cope with this problem, the input BM velocity level which directly influences the DC component of the internal cell voltage is transformed before inputting it to the IHC model. The BM velocity level is transformed using Eq. (2.37), which interpolates the DC component of the receptor potential as a function of the sound level (Zagaeski, 1994). We do not imply that the functions are rated for the property of DC component of the receptor potential. They are merely chosen to conveniently modifying the BM velocity level.

$$V_n(L) = \frac{RPmax}{1 + \exp^{(L_{n0} - L_n)/S}(1 + \exp^{(L_{n0} - L_n)/S})} \tag{2.37}$$

$$\hat{k}_n(t) = \begin{cases} g\frac{V \cdot i_n(t)/i_n^{max} + A}{V \cdot i_n(t)/i_n^{max} + A + B} & , (V \cdot i_n(t)/i_n^{max} + A) > 0 \\ 0 & , (V \cdot i_n(t)/i_n^{max} + A) \leq 0 \end{cases} \tag{2.38}$$

Here, $\hat{k}_n(t)$ is the modified membrane permeability at the BM model segment n. The parameters $RPmax$, L_{n0}, and S respectively specify the maximum BM velocity, the offset from the BM velocity level axis and the slope of the function. The parameter L, which is the output of BM model, is BM velocity level represented by logarithmic $i_n(t)$ is the BM velocity at the BM model segment n and i_n^{max} is the peak BM velocity at the segment n.

The functions $\frac{dq}{dt}$, $\frac{dc}{dt}$, and $\frac{dw}{dt}$ of extended IHC model was identical to those of Meddis's IHC model.

The final output of the extended IHC model was the amount of transmitter level in the cleft c_t.

2.3 Auditory nerve model

Although many AN models have been proposed, in almost all cases output is either the mean firing rate or the synchrony rate, or some other similar rate, and consequently these models are not appropriate for the central auditory system input. Therefore, we propose an AN model in which the outputs of the model are a train of spikes. In this model, the change of the time course of postsynaptic conductance caused by the neurotransmitter effect is modeled by Eqs. (2.39), (2.3), and (2.41).

$$G(t - \lambda_j) = C_A A_j \; t \; e^{-t/(C_T \tau_j)} / \tau_j \quad for \;\; t \geq \lambda_j \tag{2.39}$$

$$\lambda_j = t_j + d_j + N(0, f(A_j, \tau_j)) \tag{2.40}$$

$$f(A_j, \tau_j) = C_\lambda \tau_j / {A_j}^n \tag{2.41}$$

Here, G is postsynaptic conductance, τ_j is the time period between the adjacent local minimam of the neurotransmitter output level c_t, A_j is the peak neurotransmitter level between τ_j, C_T is a weighting factor related to τ_j, C_A is the amplitude constant, t_j is the starting time instance of each local minimum, d_j is a time delay, The time delay, d_j, set $\frac{1}{2}\tau_j$, and $N(0, f(A_j, \tau_j))$ is a normal distribution with a zero mean and a variance

$f(A_j, \tau_j)$, where $f(A_j, \tau_j)$ is the modeled discharge probability at time $t_j + d_j$ relating A_j and τ_j. The resulting conductance at time t is given by Eq. (2.42).

$$g(t) = \int_0^t G(t-\lambda)\, d\lambda \tag{2.42}$$

The cell membrane potential V is computed by Eq. (2.43) based on the cell membrane model (Hodgkin et al., 1952) as

$$\frac{dV(t)}{dt} = \frac{1}{c_m}[(V_L - V(t))g_L + (E_{syn} - V(t))g(t)] \tag{2.43}$$

where c_m is membrane capacitance, V_L is leakage resting potential, g_L is leakage conductance, and E_{syn} is the sum of ionic equilibrium potentials of sodium ion and potassium ion. The output of the AN model consists of an all-or-none action potential S with unit amplitude given by Eq. (2.44). Action potentials are generated when the membrane potential V of the cell crosses a threshold $T(t)$ when the cell has not fired for the refractory period t_r. The thresholds for the AN model have a random distribution to generate the neuronal-like outputs. $T(t)$ returns successive random numbers in the range from a to b.

$$S(t) = \begin{cases} 1 & V(t) \geq T(t), \text{ and } S(\lambda) = 0 \text{ for } \lambda \in [t - t_r, t] \\ 0 & \text{otherwise,} \end{cases} \tag{2.44}$$

Figure 2.2 shows response patterns of the AN model fiber to a 350ms tone burst for stimulus level 43dB. The CF of the model fiber is 1kHz, the spontaneous rate is 7 spikes/sec and threshold is 0dB.

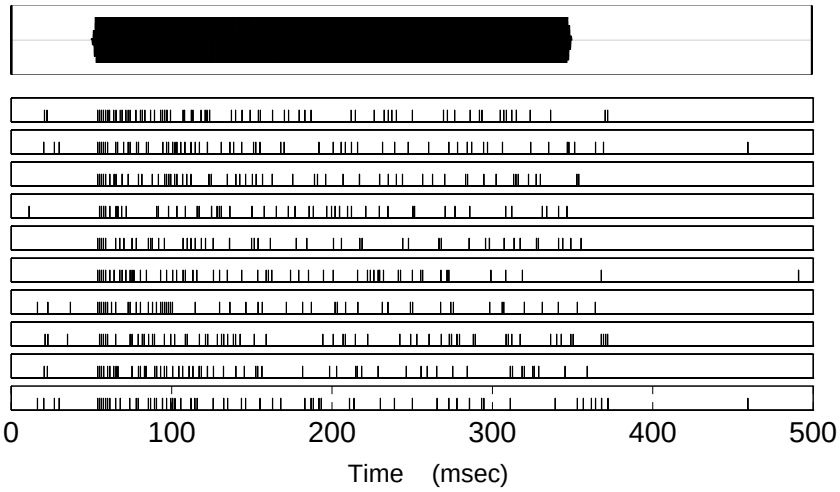


Figure.2.2. Example model output. Simulations form top to bottom receive as input a 1 kHz tone burst. In the top of the row, time waveforms of the stimulus is shown. Others are the AN model output obtained by different simulation. The output of the peripheral model is AN spikes as shown in this figure.

In Figure. 2.2, it is seems that spike trains evoked by identical stimuli are highly variable, because of the stochastic nature of action potential generation as modeled by Eq. (). For further details of the firing property, see section 3. 1.

Chapter 3

Evaluation

Response patterns of the auditory peripheral model were evaluated in detail as shown in Table 1. The threshold intensity of the AN model is normalized uniformly at relative intensity 0dB and sound levels in the simulated results indicate relative intensity. The model was tested based on evaluation methods suggested by Ross (1996) and Hewitt, et al. (1991). All of the tests are described in the following subsections.

Table 1. The list of evaluation to the eight previous IHC model and preset peripheral model. Each column of table is based on Ross (1996) and Hewitt, et al. (1991).

Rapid and short-term adaptation
(1) Larger dynamic range for onset than for steady state
(2) Adaptation as a sum of (at least) two exponentials
(3) Exact fit to data sets for individual fibers
Properties of synchronized activity
Interval histograms for driven activity
Period histograms without clipping at high levels
Synchronization Index versus intensity
Sync. coefficient versus frequency of stimulation
Hazard function for driven activity
(1) For Low CF fiber
(2) fiber with weak phase locking
Recovery from adaptation
(1) as single exponential
(2) with smaller exponent for onset than steady state
Recovery of spontaneous activity
(1) as single exponential with realistic time constant
(2) with "dead" period
Response changes to intensity
Additivity for increments of stimulation
Response to decrement versus delay
(1) decreasing function for 1-ms window
(2) flat function for 10-ms window
Other evaluation
Output of the real action potentials
Model consist of concatenation of sub models
Making of the PSTH from real action potentials

3.1 Rapids and short-termadaptation

Adaptationisthevariationinresponsewhich occursduringa constant stimulus condition. The adaptationofthefiringrateofan auditorynerve fiberinresponsetolongtone bursts appearsto consistof severaldecayingcomponents,atleasttwo of which occur within the firstfewhundred milliseconds(Harriset al.,1979; Smith et al.,1975; Westermanet al., 1984). Estimatesof the time constant of short-termadaptationare40-50ms incat and guineapig,and 15 ms inchinchillar.apidadaptationappearstobe an orderof magnitude faster(Harriset al.,1979; Smith et al.,1975). Smith and co-workers(Smith et al.,1975, 1985; Smith, 1977, 1979) attempted to distinguishbetween the propertiesof short-term and rapidadaptationby obtainingrate-itensity functionsusinghistogramwindows of variousduration.Histogramwindows of 10ms or more emphasize the effectsof short-term adaptation,and produce onsetrate-itensity functionswhich have the same shape as the steady-staterate-itensity functionss,uggesting that the time courseand relativeamount of short-termadaptationareindependent of intensity. Histogramwindows of 2ms or less emphasize the rapidlyadapting component of the onset response. Corresponding onset rate-itensity functionsincreasemore at highintensitiesthan do the steady-staterate-intensity functions,suggestingthat relativelymore of the rapidlyadapting component existsat highintensities.

Westermanand Smith (1984) reportedmore directapproach tostudyingthe components of adaptation. They developed a least-squaressurve-fittingprocedure to fit the sum of two exponentialsplusa constant to experimentaldata. The fittingtechnique was appliedtothe responses of singleauditorynerve fibersof Mongolianger bil produced by constant-intensity tone burststimuli.By use of the fittingprocedure,they attempted to determinethe propertiesof both the rapidand short-termadaptationprocessesquantitatively Here, the fittingtechnique was appliedtothe responses of AN model and then fittingresultscompared tothat of theirphysiologicadata.

3.1.1 Methods

From Westermanet al.,(1984),the stimuliconsistedof 300-ms tone bursts(2.5-msrise time) at characteristifrequency(CF) and the post-stimulustime (PST) histogramwas generatedby summing 384 repetitionsof a tone burststimuluswith a constant intensity, usinga 1-ms bin width. For the simulation,the same stimuluswas used inthisstudy and PST histogramswere made inthesame way. The parameter valuesof the model are determinedthat the absoluteintensity fromthe model isagreedwith relativeintensity fromexperimentaldata reportedby Westermanand Smith (1984).

3.1.2 Data analysis

Westermanand Smith (1984) characterizedthe adaptationfunctionas the sum of two exponentialdecay functionsplusa constant, as shown by

$$A(t) = A_R\ \mathrm{e}^{-t/\tau_R} + A_{ST}\ \mathrm{e}^{-t/\tau_{ST}} + A_{SS}, \tag{3.1}$$

where A_R and A_{ST} respectively indicate the magnitude of the rapid (*Rapid*) and short-term (*Short*) adaptation components, and τ_R and τ_{ST} are respectively the decay time constants of the components, and A_{ss} is the steady-state response.

The bins of the histogram represent a set of discrete samples $[a_i, i = 1, \cdots N]$ of the response rate $A(t)$. The fitting was done by a least-squares procedure which involves minimizing the sum of squared errors given by

$$\begin{aligned} & S(A_R, \tau_R, A_{ST}, \tau_{ST}, A_{SS}) \\ & \quad = \sum [a_i - A_R \, \mathrm{e}^{-t/\tau_R} - A_{ST} \, \mathrm{e}^{-t/\tau_T} - A_{SS}]^2, \end{aligned} \tag{3.2}$$

with respect to the five parameters A_R,τ_R,A_{ST},τ_{ST} and A_{SS}. The least-squares procedure in this study is realized by Marquardt method. For both the simulation and the experiment, the spontaneous rate (SR) was counted during the 40 ms of silence preceding each tone burst. Several constants calculated for the following way, is the same way as Westerman et al (1984). The rapid (A_R), short-term (A_{ST}), and steady-state (A_{SS}) components were derived by fitting Eq. (3.1) to the histograms. The driven-steady-state rate is the A_{SS} minus the SR. The *Onset* is the sum of A_R, A_{ST} and A_{SS}.

3.1.3 Results

Figure 5 shows a set of response histograms from the AN of a gerbil at CF 1170 Hz on the right and histograms from the model fiber at CF 1170 Hz on the left.

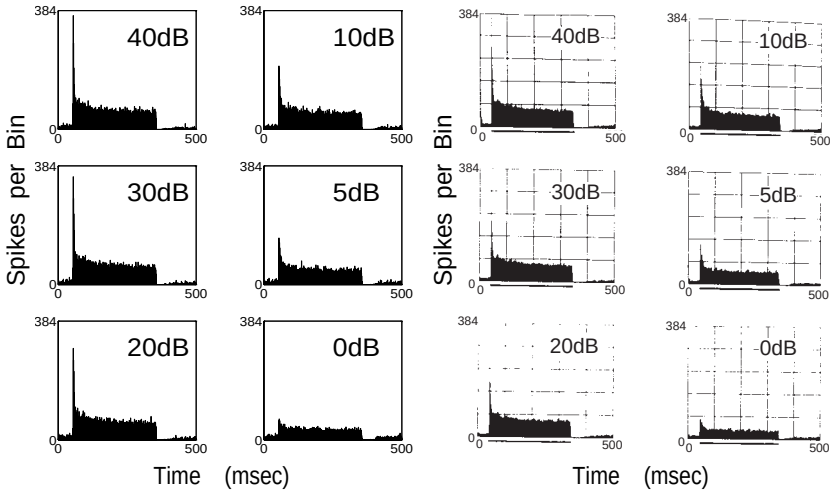


Figure 5. PST histograms at different sound levels. Left column: data from the model. Right column: physiological data from Westerman et al. (1984).

Figure 6 shows the magnitudes of the response components derived by fitting the characteristic response Eq. (3.2), and the response components for fibers at CFs of 900 Hz, 1170 Hz, 4940 Hz and 6100 Hz.

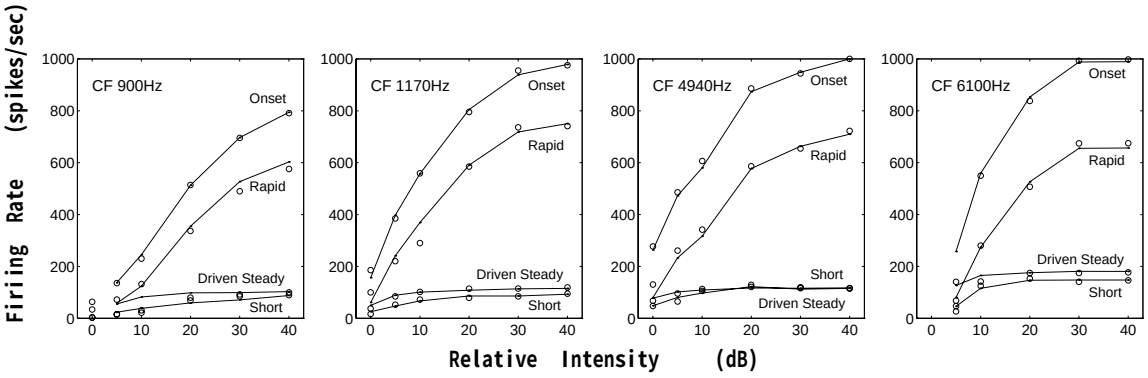


Figure 6. Fitted magnitude components for four AN fibers at CF's 900 Hz, 1170 Hz, 4940 Hz, and 6100 Hz as reported by Westerman (1984) (individual open circles) together with data generated by the model (marked points connected by straight lines).

In Fig. 6, the onset, rapid, short and driven-steady components for individual fibers agree with those for the physiological data quantitatively and the onset components (Onset) have larger dynamic ranges than the steady-state component (Driven Steady).

Figure 7 shows fitted decay time constants for fibers at CF's of 900 Hz, 1170 Hz, 4940 Hz and 6100 Hz.

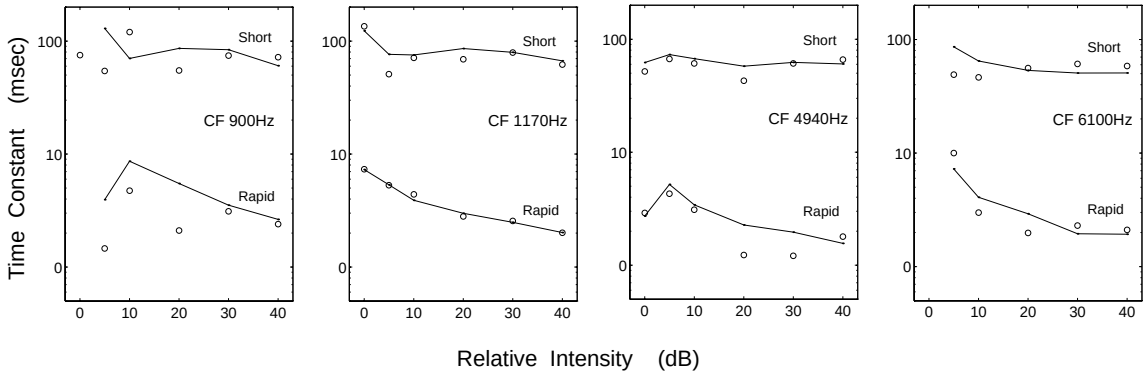


Figure 7. Fitted decay time constants for four AN's at CF's 900 Hz, 1170 Hz, 4940 Hz, and 6100 Hz (individual open circles) together with data generated by the model (marked points connected by straight lines). These data correspond to the component magnitudes shown in Figure. 6.

In Fig. 7, results of the fit of short and rapid components with physiological data are somewhat distributed compared to those of the amplitude components. The scale of the time constant is logarithmic in Fig. 7, and physiological data are also distributed compared to those of the amplitude component. We can thus conclude that the short and rapid components for individual fibers are in agreement with those of the physiological data. This is the first report that PST histogram generated by summing several hundred repetitions from the model are agree with physiological data quantitatively

3.1.4 Discussion

In Figures 6 and 7, the individual model fibers exactly fit the physiological data sets described above. This exact fitting of the individual model fibers turned out to be caused by modifying the membrane permeability of the IHC model as shown in subsection 2.2. The rate-intensity curve derived from Meddis's IHC model was not quantitatively agree with the physiological data. (See Meddis, 1988, Fig. 2). The parameter value of the model in order to fit the physiological data with various CF was not known. The model parameter values using in this section were given in Table 2.

Table 2. List of parameter values in the IHC and AN model section.

name	CF 900 Hz	CF 1170 Hz	CF 4940 Hz	CF 6100 Hz
L_0	220.0	135.0	370.0	220.0
s	30.0	17.5	22.5	16.5
$RP_{\max}$	17.0	12.0	15.0	8.0
A	2.0	2.0	2.0	2.0
B	300	300	370	400
g	6400	6400	8000	8000
y	5.05	5.05	5.05	5.05
l	2800	2800	2800	2800
r	6580	6580	6580	6580
x	59.3	52.0	69.0	78.5
m	1.0	1.0	1.0	1.0
C_A	500000	500000	500000	500000
C_T	0.047	0.047	0.047	0.047
C_λ	7.4×10^{-6}	5.7×10^{-6}	1.34×10^{-6}	1.09×10^{-6}
n	2.0	2.0	2.0	2.0
c_m	0.001	0.001	0.001	0.001
V_L	0.0	0.0	0.0	0.0
g_L	0.1	0.1	0.1	0.1
E_{syn}	40.0	40.0	40.0	40.0
a	2.720	1.850	0.400	0.300
b	2.603	1.764	0.391	0.292
t_r	0.8ms	0.8ms	0.8ms	0.8ms

In Table. 2, the IHC model parameters A, y, g, l, r, m and the AN model paprameters C_T, C_A, n, c_m, V_L, g_L, E_{syn}, and t_r were constant with irrespective of CF. Only nine parameters were changed CF to CF. The IHC-AN model proposed by Rose, 1996 also exact fit to data set for individual fibers, but 23 parameter have to be changed CF to CF. Therefore, Rose's model was difficult to fit with physiological data. Contrary the peripheral model in this study can be fitted to physiological data easily.

The slope of "Onset", "Rapid", "Short", and "Driven Steady" in Fig. 7 were more sensitive to the parameter L_0, s, $RP_{\max}$ than other parameters. These parameters were

appended to the IHC model proposed by Meddis (1986, 1988) to modify the membrane permiability of the model as shown in section 2.2.2. Onset firing rate mainly depended on the parameter value a and steady-state firing rate depended on b. Through fitting the parameters with physiological data, it was found that the rapid time constants are sensitive to the IHC model parameter x which is the weighing factor related to recovering transmitter level. Therefore, the rapid time constant can be fitted to physiological data by modifying only one parameter x.

In this simulations, spontaneous rate (SR) was set constant, but SR can be controlled by changing the IHC model parameter A or C_A independently in silencient interval.

3.2 Properties of synchronized activity

In response to low-frequency sinusoidal stimuli, discharges in fibers of the auditory nerve do not occur randomly in time. Because action potentials are elicited by unidirectional movements of the basilar membrane (Davis et al., 1950; Goldstein, 1968; Brugge et al., 1969), discharges occur within a well-defined time window relative to a single cycle of the sinusoid. This occurrence of action potentials at preferred times has been termed synchronization or phase locking of the discharge (Rose et al., 1967). Synchronization or phase-locking occurs in auditory neurons of all vertebrate classes, squirrel monkey, cat, chinchilla, guinea-pig, birds, crocodile, turtle, frog and fish. Synchronization or phase-locking in response to low-frequencies can be observed in all auditory nerve fibers irrespective of their best frequency.

There are many ways of documenting neural discharge synchronization. Two of most commonly employed analyses are based on period and interspike interval (ISI) histograms. The principal difference between period and ISI histograms is that ISI histograms provide additional information about the tendency of successive discharges to skip stimulus periods. That is, if a spike occurs at one point in the stimulus period and the next spike occurs at the same point in the next stimulus period, then the time interval between the spikes will be equal to one period of the stimulus. On the other hand, if the neuron fails to discharge on the next stimulus period and discharges on the one immediately following, then the time interval will be equal to two stimulus periods. In either case, the neuron is sending precise temporal encoded information about the stimulus period and, by implication, about the stimulus frequency as well (Javel et al., 1988).

Therefore period and ISI histograms were generated from the model output and then the histograms from the model are compared with that of physiological data. Synchronization index versus intensity and synchronization coefficient versus frequency of stimulation were also compared with physiological data in a similar way.

3.2.1 Period histograms without clipping at high levels

A period histogram is a plot of the accumulated numbers of occurrences of discharges as a function of time within the stimulus period. To compile a period histogram, the stimulus period is divided into discrete bins, a pointer cycle through the histogram bins as the stimulus waveform is presented acoustically, and the occurrence of each spike causes the appropriate bin to be incremented by one. The histogram pointer resets to the beginning with each repetition or period of the stimulus. Any tendency of the discharges to be synchronized will ultimately show up as peaks in the period histograms.

Examples of period histogram from the model are shown in Fig. 8. The left side shows data from the model and the right side shows physiological data reported by Javel et al. (1988), obtained as the intensity of a tone burst stimulus increasing.

The stimulus using simulation and experiment was a 120 ms, 833.3 Hz sinusoid and the bin size of each histograms were 0.0384 ms.

In Fig. 8, the shapes of the period histograms based on the output of the model are very similar to those of the physiological data. The synchronization has a certain threshold,

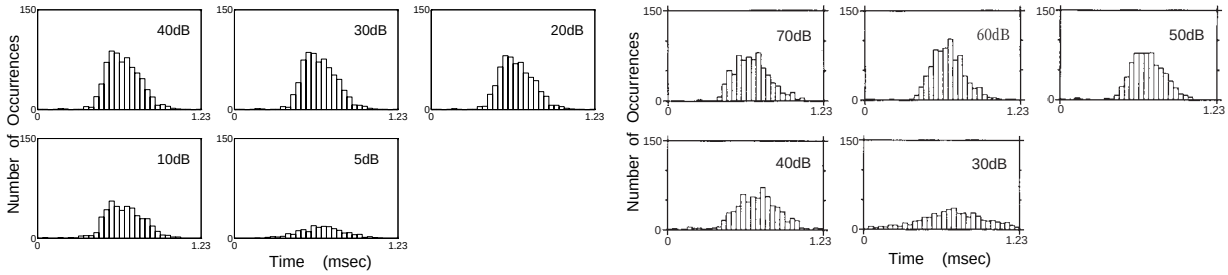


Figure 8. Period histograms of the stimulus at different intensities. Left column: data from the model. Right column: physiological data from Javel (1988)

below which discharges are unsynchronized. The degree of discharge synchronization increases with increasing stimulus intensity. There is an intensity at which synchronization saturates and, further increments in intensity do not result in further synchronization. The function $f(A_j, \tau_j)$ of Eq. (2.41) models data at this type of saturation described above.

3.2.2 Interval histograms for driven activity

An ISI histogram is a plot of the accumulated number of occurrences of given time intervals between successive spikes. As was the case for period histograms, synchronization of discharges will produce peaks in the ISI histogram with the set of peaks possibly repeating at integer multiples of the stimulus period.

Figure 9 shows a set of ISI histograms of the stimulus at different intensities for fibers at CF 1111 Hz. The left side shows data from the model and the right side shows physiological data reported by Rose et al. (1967). The duration of the stimulus tone is 200-ms and it is presented only 111 Hz. Each histogram shows responses obtained over 50 trials using 1 ms bin width.

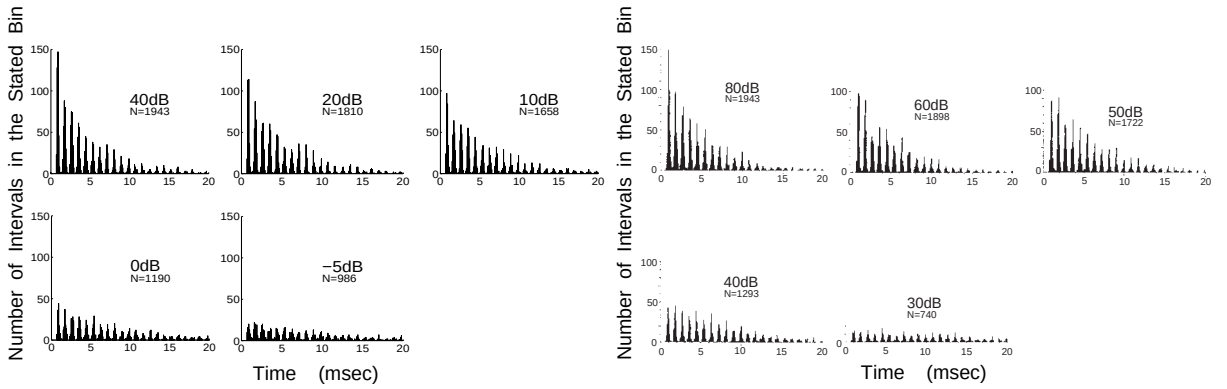


Figure 9. Interspike interval histograms for different stimulus intensities. Left column: data from the model. Right column: physiological data from Rose et al. (1967).

In Figure 9, the shapes of the ISI histograms are not the output of the model are very similar to those of the physiological data. Each histogram is simply a set of local distributions located at multiples of the 0.9-ms period of the stimulus waveform. The period of split tinged distributions into the ISI histogram can be controlled by the AN model parameters C_λ, C_T, and c_m. C_λ is most effective in the period of local distributions

the ISI histograms. As C_λ of the AN model parameter was decrease, the period of local distributions was decreased.

The length of the modal ISI, or "center of gravity" of the ISI histogram, tends to decrease as more intense stimuli induce increases in discharge rate, that is, as more intervals of shorter length arise.

The modality of the ISI histograms can be controlled by the AN model parameters C_A and $T(t)$.

3.2.3 Synchronization Index versus intensity

The relationship between average rates and synchrony has been studied extensively (Johnson, 1980). To obtain a measure of the amplitude of the synchronized response, we adopted synchronization index reported by Johnson (1980) (for calculation method of synchronization index, see Johnson, 1980). The synchronization index obtained from period histograms for the model fiber at CF 1000Hz are plotted in Fig.10 (a) and the avarage rate are plotted in Fig.10 (b) obtained the same data. The stimulus frequency is 1kHz.

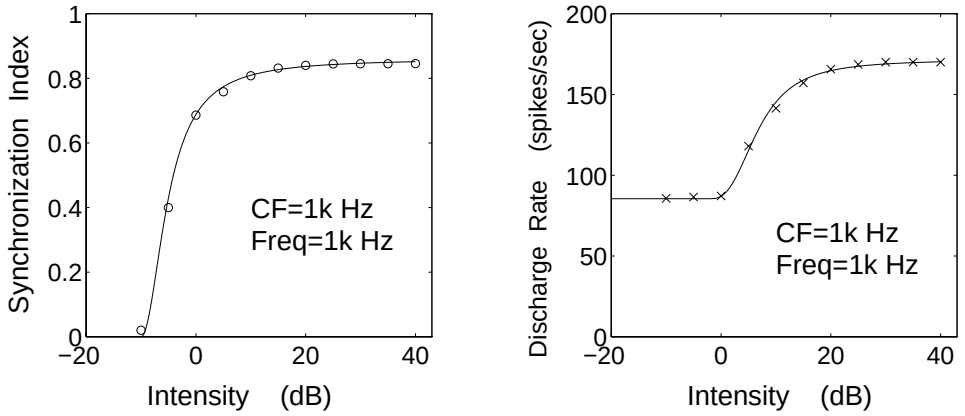


Figure.10. (a) Synchronization index for the model fiber at CF 1kHz as a function of stimulus level (b) Average rate obtained the same data for (a).

The normalization shown in Fig.11 was produced by taking the results of the fitting procedure and manipulating them so that each curve in Fig.10 spanned the range 0-1. From Figure 11, it can be seen that the shapes of normalized curve for synchronization and for average rates have threshold. Although the shape of the curves are the same, synchronization has a lower threshold than average rate in Fig.11. In term of this, data from the model is corresponded to physiological data from Javel (1988).

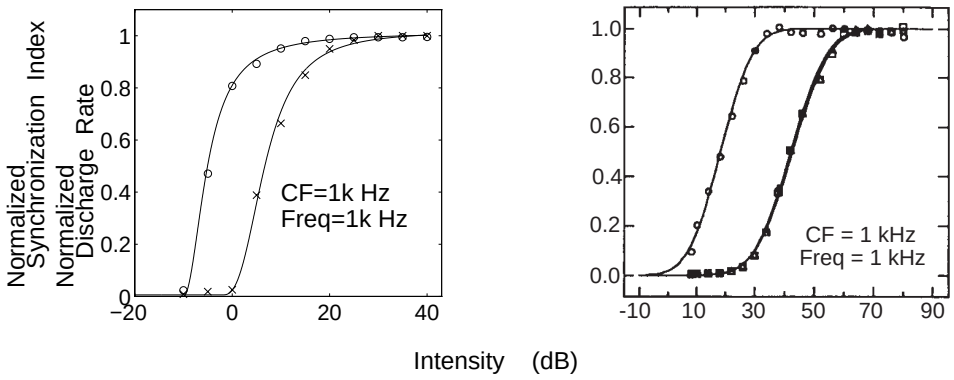


Figure.11. Normalized synchronization index (cross or square) and average rate (open circle). Left column: data from the model. Right column: physiological data from javel (1988).

3.2.4 Synchronization coefficient versus frequency of stimulation

Period histograms (PHs) show that AN-fiber responses phase-lock to the positive half-cycle of low frequency tones. At high signal frequencies, however, the period histogram shows no relationship to the signal's phase characteristics. Rose et al. (1967) quantified the loss of synchrony with the synchronization coefficient (the density of the most populous half of the period histogram divided by total density). To measure model fibers's synchronization coefficients, stimuli of 1000, 2000, 3000, 4000, and 5000 Hz sinusoids were used. The results for the model fiber at CF 4000Hz are presented in Fig.12 which is also illustrates physiological data reported by Rose et al, (1967).

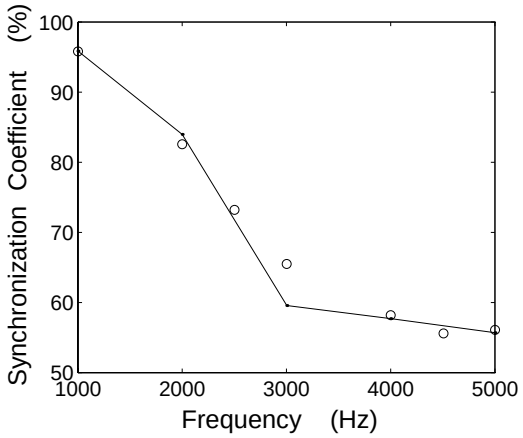


Figure. 12. Synchronization coefficients for the model (solid line) compared to Rose's (1967) physiological data (open circle) for the fiber at CF 4000Hz.

In Figure 12, the model simulates the fall off of synchronization with increasing frequency accurately compared to the physiological data reported by Rose (1967).

The synchronization coefficient can be controlled the AN model parameter C_T, c_m, and C_λ. The synchronization coefficient is increased with increasing the parameter C_T and with decreasing the parameter c_m, and C_λ. The parameter C_λ is directly and easely controled the value of synchronization coefficient compared to C_T and c_m.

3.2.5 Discussion

In this section, properties of synchronized activity of the model responses were tested by using various methods. As we have seen, the degree of phase-locking or synchronization depends on the AN model parameter C_λ, C_A etc. The Eq. (2.3) was modeled to satisfy the properties of synchronized activity. Although the proposed model was a functional model, the properties of discharge patterns were known from the form of Eq. (2.3). The main reason is that the properties of synchronized activity from the model output were good agree with those from physiological data.

The properties of synchronized activity mainly depend on the AN model parameter. Therefore, only the AN model parameters are given in Table. 3.

Table 3. List of parameter values in the AN model section.

name	Period H	ISI H	Syn. Index	Syn. Coefficient
C_A	500000	500000	500000	500000
C_T	0.047	0.047	0.047	0.047
C_λ	4.2×10^{-6}	6.0×10^{-6}	6.7×10^{-6}	1.6×10^{-6}
n	2.0	2.0	2.0	2.0
c_m	0.001	0.001	0.001	0.001
V_L	0.0	0.0	0.0	0.0
g_L	0.1	0.1	0.1	0.1
E_{syn}	40.0	40.0	40.0	40.0
a	1.85	1.83	1.85	0.40
b	1.764	1.82	1.764	0.39
t_r	0.8ms	0.8ms	0.8ms	0.8ms

In Table. 3, the parameter values were determined that the output of the model was agreed with physiological data by referring to Table. 2.

As shown in Table. 3, the values of parameters C_A, C_T, n, c_m, V_L, g_L, E_{syn} and t_r were constant irrespective of simulation paradigms.

3.3 Hazard function for driven activity

When a neuron generates an action potential, the neuron enters a state of absolute refractoriness for a brief interval, the cell is incapable of generating second action potential. Then gradually the excitability of the cell returns to its normal level. During this period

of relative refractoriness,the probability of firing is repressed due to the lingering effect of the spike on the neuron.

Gaumond (1980) found that the recovery of firing probability following a spike in an auditory nerve fiber was determined by the product of two functions, one depends only on the temporal variation of the stimulus, and another depends only on the interval since the preceding spike. Gaumond termed the former the "excitation" function and the latter the "recovery" function (see also Lutkenhoner et al., 1980). Gaumond et al. (1983) proposed a Markov chain model of the recovery process, which can be used to estimate the underlying excitation functions for a given spike train, provided that the recovery function is known (or can be estimated). Westerman (1985) also proposed the method to derive the recovered excitation function. We adopted the method proposed by Westerman (1985) to derive the function. The method utilized in the present study is described below. The recovery functions are derived from the hazard function by normalizing it to a value of unity at long intervals. In Westweman (1985), the steady-state driven response to tone bursts was used to derive the recovery function. An interval histogram was constructed from all the interspike intervals occurring during the last 200 ms of the standard 300 ms tonal stimulus.

From Westweman (1985), the hazard function is derived from the interval histograms in the following way: The estimated value of the hazard function at each bin in the interval histogram is the contents of the bin, divided by the sum of the contents of that bin plus the bins of all longer intervals. The hazard function is the probability that the neuron will fire at a given time (number in a particular bin) given that the neuron has not fired since the discharge at time zero (sum of contents of all intervals equal or longer than this interval). The quotient is an estimate of the hazard function. The hazard function is then normalized to yield the recovery function.

3.3.1 Results

An example of the hazard function for the model fiber shown in the left side of Fig. 13 and for a single gerbil auditory nerve fibers reported by Westerman (1985) shown in the right side of Fig. 13.

The hazard function shown in Fig. 13 was derived from the responses of a model and actual fiber with a CF of 4300 Hz. The instantaneous firing probability was zero immediately following a spike (absolute refractory period). This period lasted 0.8 ms in the model fiber and 1.2 ms in the actual fiber from Fig. 13. This refractory period can be controled by the AN model parameter t_r. The parameter t_r set 0.8ms in the left side of Fig. 13.

Figure 13 shows that the firing probability then increased rapidly over the next millisecond, and more slowly over the next tens of milliseconds, to an asymptotic value. Westerman (1985) described that this function can be converted to a recovery function by normalizing all data points to the asymptonic value, right side of Fig. 13 about 0.06. The hazard function from the model also converted to a recovery function divided by 0.05. This hazard function is very similar in shape and time course to the archetypal recovery function shown by Westerman (1985, right side of Fig. 13). Westerman (1985) was

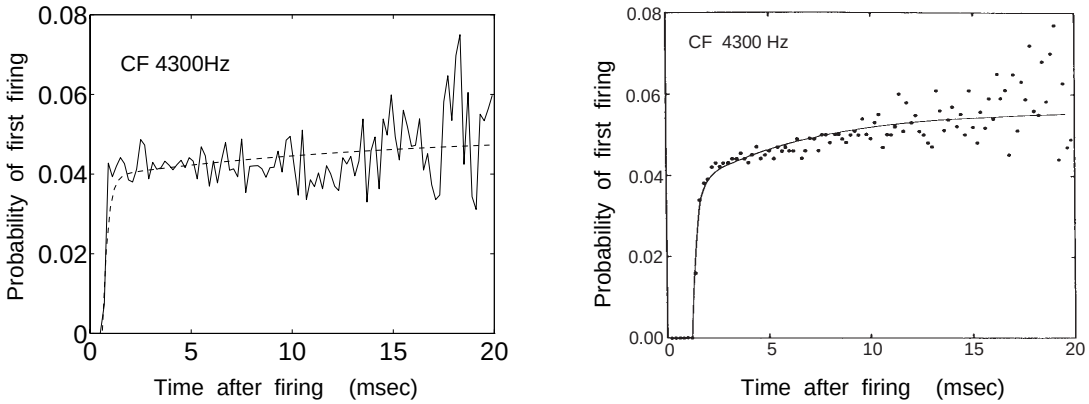


Figure 13. Hazard function of the model and actual fiber at CF 4300 Hz. Left column: data from the model. Right column: physiological data from Westerman (1985). The total number of spikes from the model fiber was 13778, and that form the actual fiber was 52065. Bin width of each hazard function was 200 μ sec. Stimulus intensity was 43 dB above AV threshold for both simulation and experiment. The curve (broken line in left side and straight line in right side) is the least-squares fit of a two-component exponential function to the data (see text).

approximated the recovery function using an analytical expression which consisted of two exponential components which started after a brief delay. Westerman denoted that the fast component had a time constant of 0.5 ms and contributed 70-75 % of the total probability recovery and the slow component had a time constant of 20 ms. From Westerman (1985), the analytical form of the recovery function was

$$p(t) = \begin{cases} 0 & \tau < \tau_0 \\ 1 - P_F e^{-(\tau-\tau_0)/\tau_F} - P_S e^{-(\tau-\tau_0)/\tau_S} & \tau > \tau_0. \end{cases} \qquad (3.3)$$

The fitted result of the model shows in the left side of Fig. 13 (broken line). The fast time constant of the model was 0.2 ms and contributed 85.5% of the total probability recovery and the slow time constant was 14 ms. The fitted curve of the left and right sides of Fig. 13 demonstrated the degree to which the analytical form described the actual recovery. Westerman also indicated that the parameters P_F, τ_F, P_S, τ_S varied somewhat from fiber to fiber. Therefore, the parameter values from the model were very appropriate for physiological data.

Westweman (1985) denoted that the timing of nerve discharge is controlled by the recovery from refractoriness and by stimulus phase, however, phase locking is strong only with low frequency stimulations, so that the data of Fig. 13 do not display this phenomenon. Therefore, Westweman generates the hazard function for low frequency units as seen in the right side of Fig. 14. Similarly, we generate the hazard function for low CF model fiber as shown in the left side of Fig. 14.

These data were collected from the model and actual fiber with a CF of 842 Hz. The left side of Fig. 14 is very similar in content to one shown as Westerman (1985, right side of Fig. 14), and demonstrates that the recovery of firing probability following a spike reflects the temporal variation of stimulus for low frequency stimulation.

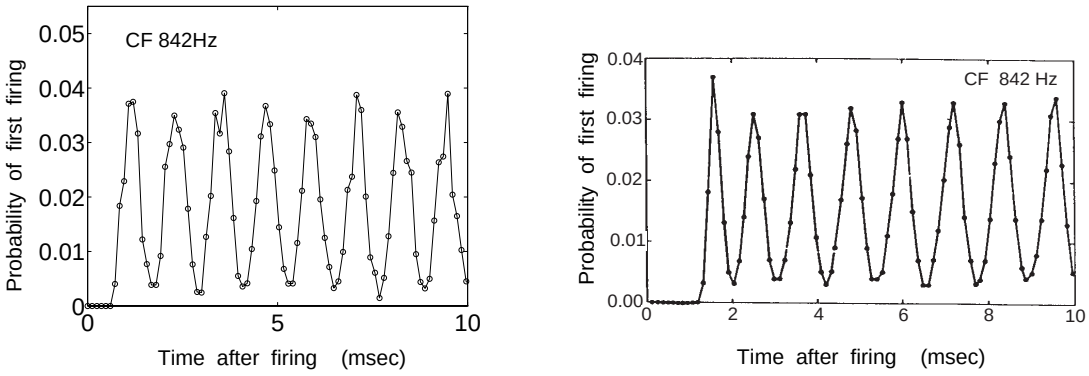


Figure 14. Hazard function for low-CF fiber (842 Hz). Left column: data from the model. Right column: physiological data from Westerman (1985). Interspike intervals occurring within the last 200 msec of the responses to 300 msec constant intensity tone burst stimuli, were used in the both of calculations. The total number of spikes from model was 13778, and from actual fiber was 147056 Analysis bin width for each function was 200 μ sec. Stimulus intensity was 43 dB above AV threshold for both simulation and experiment.

It follows from what has been said that the hazard function for fiber with weak phase locking and for low CF fiber from the model is quite agreement with experimental data.

3.3.2 Discussion

The only difference between data from the model and physiological data in Figs. 13 and 14 was the value of refractory period. But the refractory period can be controlled by the AN model parameter t_r as described above. Therefore, the proposed model passed the test for the hazard function for fiber with weak phase locking and for low CF fiber.

The parameter values used in this section (Figs. 13 and 14) were used the values near the CFs as shown in Table 3. But only the parameter C_λ was modified to 6.4×10^{-6} in Fig. 14.

3.4 Recovery from adaptation

To examine the recovery process, data were collected from single AN fibers of the model using a forward masking paradigm. Masking, as originally applied to auditory sensation, was the effect of one sound being smothered or concealed by another sound, thereby being "rendered powerless to affect the ear". Different terminologies have been used to differentiate the "one sound" from "another sound". We will use the term "probe" to apply to the signal to be detected, and will use "masker" to refer to the sound introduced to interfere with the detection of the probe. Masking has also been used to describe the process by which the loudness of the probe is reduced in the presence of masker. Physiological masking has acquired a similar definition as a shift in the threshold or a reduction in the magnitude of the response evoked by the probe stimulus caused by the introduction of the masking stimulus. As applied specifically to the responses of individual auditory neurons, masking is manifest as a reduction in probe-evoked firing rate or as a shift in the fine structure of the temporal response pattern from one phase-locked to the probe to one phase-locked to the masker. (Harris and Dallos, 1979)

The work of Smith (1977) and Harris and Dallos (1979) investigated the recovery from adaptation of single AN fibers in gerbil and chinchilla. The faster component of recovery was studied in detail by Westerman (1985). In this study, we refer the work of Westerman (1985).

3.4.1 Methods

The work of Westerman (1985) used the paradigm shown in Fig 15.

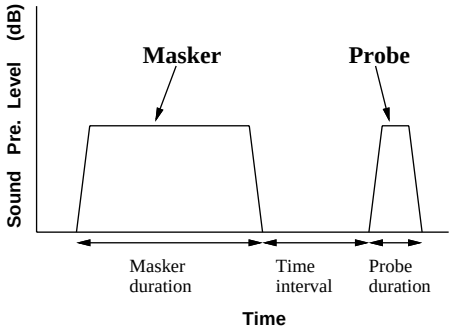


Figure 15. Forward masking paradigm used by Westerman (1985) to study recovery of response after adaptation. Masker duration 300ms, probe duration 30ms, masker and probe at 43dB above AV threshold. Δt varied between 0 and 200ms.

The masker stimulus were 300 ms constant-intensity tone bursts with frequency at fiber CF and intensity at 43 dB above AV threshold (dB above AV threshold, Westerman, 1985). The probe stimulus were 30 ms tone pips (2.5 ms rise/fall time) presented at various interval following the masking stimulus. 3.84 s stimulus epoch were recorded each

condition. Bin width of PST histograms is set at 1ms. In this study, the stimulus intensity and frequency used for the simulation are the same described above.

3.4.2 Results

An example of the response from the model obtained from the recovery paradigm is shown in left side of the Fig.16 and from Westerman is shown in the right side of Fig.16.

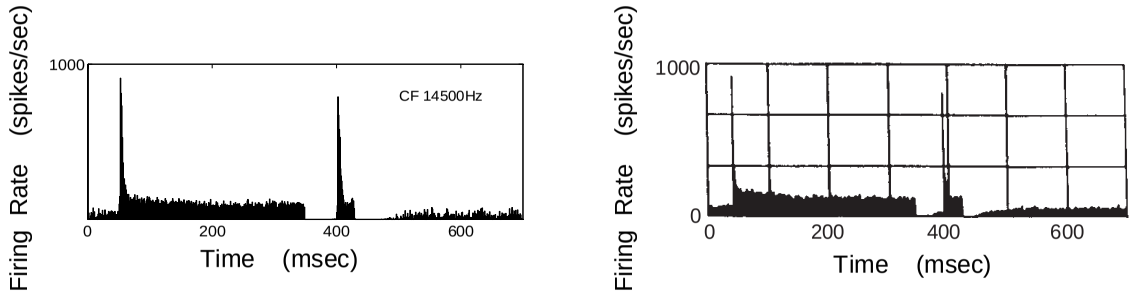


Figure.16. Recovery of response following adaptation. Left column: data from the model for fiber at CF 14500 Hz. Right column: physiological data from Westerman et al. (1984). Fiber E2947, CF 14600Hz. Each figure shows the normalized PST histograms recorded in response to a 300ms constant intensity probe tone, and a 30ms masking tone presented at various intervals after the offset of the probe tone. The silent interval was 25 ms. Intensity of adapting and test tones was 43 dB above AV threshold for both simulation and experiment. 384 stimulus repetitions were recorded at left and right side of the Fig.16. Bin width in Fig.16 was 1 ms.

Right side of the Figure 16 shows that the onset response generated by the probe tone was affected by the masking tone. This phenomenon was also observed in the data from model as shown in left side of the Fig.16.

Westerman (1985) characterized the recovery process following way: the corresponding measures were calculated for the (fully recovered) response to the adaptor, and the decrement of the test tone response, relative to the recovered response, determined. It was assumed that the recovery of each measure could be described as a single exponential process using an equation of the form

$$A_m(\tau) = A_r(1 - e^{-\frac{\tau}{\tau_\rho}}), \quad \tau > 0, \tag{3.4}$$

where A_m is the masked response at time τ after adaptor offset, A_r is the recovered response, and τ_ρ is the time constant of recovery of the measure. Eq(3.4) was applied to the recovery of both the onset and short-term measures. A_r and A_m were estimated from the spike counts, and a plot of the decrement $(A_r - A_m(\tau))$ on semi-log coordinates yielded estimates of τ_ρ for the rapid and short-term components. The maximal onset 1ms count of the first 10 ms, and the count of the last 20ms of the tone pip, were taken as indicators of the rapid and short-term components of the masking response for both simulation and experiment.

An example of the physiological data derived from the forward masking paradigm is shown in the right side of Fig.17 (Westerman, 1985).

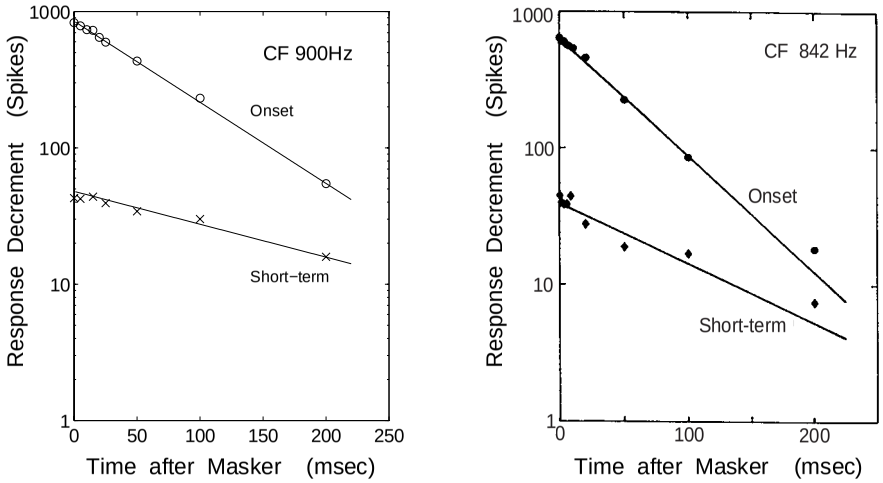


Figure.17. Recovery of response components. Left column: data from the model for fiber at CF 900Hz. Right column: physiological data from Westerman(1988), fiber E27B3, CF 842Hz. The figures show the difference between the response to the initial portion of the 30ms masking tone and the response to a 30ms test tone of equal intensity, presented at various intervals after the offset of the masking tone. The onset count was the maximal 1ms count within the first 10 ms of the stimulus; the short-term counts were the number of spikes occurring between 10 and 30ms after stimulus onset. Stimulus intensity was 43 dB above AV threshold for both simulation and experiment.

In the right side of Fig.17, the values for the decrement in the onset response appear to be well-fitted by a single straight line, implying that the recovery reflected a exponential process. The estimated time constant of the process is 49ms. The recovery of the short-term measure was also fitted by a single process, with a time constant of 68ms. The time constants of recovery of rapid components are smaller than that of short-term components. Westerman(1985) repoted that the average for the time constants of recovery of rapid and short-term components is 48ms (s.d.=25ms) and 169ms (s.d.=79ms), respectively from estimation on 12 fibers.

The data from the model fiber at CF 900Hz (is the near CF at right side of the Fig. 17) are shown in left side of the Fig. 17. The recovery of the onset and short-term measure was fitted by a single process and the time constants of recovery of rapid components are smaller than that of short-term components similar to the physiological data. The estimated onset time constant from the model is 74ms and the short-term time constant is 180ms.

It seems reasonable to suppose that the value of the time constants from the model is realistic compared to that from experimental results.

In the model and the actual fibers displaying moderate or high spontaneous level of discharge, the spontaneous activity rate was retransiently suppressed following the offset of the adapting tone (see e.g. Fig. 5). Westerman(1985) described that the recovery of spontaneous firing also appeared to follow a single exponential function, the measurable recovery however, commencing at some delay after stimulus offset. A curve of this type

$$A_{sp}(t) = \begin{cases} 0 & t < \tau_0 \\ A(1 - e^{-(\tau - \tau_0)/\tau_{sp}}) & t \geq \tau_0 \end{cases} \quad (3.5)$$

was fitted to the PST records of spontaneous activities following the masker offset. Here, τ_0 represents the delay, and τ_{SP} is the time constant of the recovery of spontaneous activity.

The spontaneous activity of the actual fiber (right side of Fig.17) were recorded with a time constant of 38 ms. The spontaneous activities of the model fiber also follow a single exponential function. Time constant of the recovery of spontaneous activity from the model (left side of Fig.17) was 35 ms. The time constant from model, 35 ms, was realistic value compared to that from the actual fiber.

3.4.3 Discussion

In Fig 17, the recovery of the onset and short-term measures from the model yielded more straight than that from actual fiber. Through simulation, it is clear that the IHC parameter A affected with recovery process. When the value of the parameter A was decrease, decremental responses in logarithmic scale yielded more straight. The parameter value of A used in this section was 0.3. Other parameters were almost identical as shown in Table.2.

Meddis, (1988) reported that the IHC model shows equivalent time constants for both recovery processes. The time constants of recovery of rapid components from the proposed model, in contrast, are smaller than that of short-term components similar to the physiological data. This is because that the AN model has threshold $T(t)$ as shown in Eq.(2.44)

3.5 Response changes to intensity

Auditory-nerve fibers exhibit a characteristic decay or adaptation in firing rate in response to a tone of constant sound intensity (see chapter 3.1). The decay appears to be comprised of several exponential components, including rapid adaptation with a time constant of several milliseconds, and short-term adaptation with a time constant of several tens of milliseconds. Adaptation produces an emphasis on the response to change in intensity and may play an important role in the encoding of dynamic stimuli such as speech (Smith et al., 1985).

Previous experimental studies have led to the conclusion that short-term adaptation is basically additive in nature and not the result of a sensitivity or gain change. For example, when an increment in sound intensity is added to an ongoing background sound, the change in firing rate produced by the increment is constant and independent of the amount of prior adaptation (Smith and Zwislocki, 1975; Smith, 1979). Similar additivity was found for responses to decrement in intensity and for test tones applied after adapting tones (Smith, 1977; Abbas, 1979).

Smith and Zwislocki (1975) studies involved measuring firing rates over relatively large time intervals of 10ms or more in order to emphasize the properities of short-term adaptation and reduce the contribution of rapid adaptation. Subsequently, it was discovered that the shapes of rate-intensity functions depend to some extent on the size of time interval used to compute firing rates. In particular, smaller time intervals produce functions with larger operating ranges, so that the rapidly adapting component appears to have a larger operating range than the short-term component (Smith and Brachman, 1980; Westerman and Smith, 1984).

Smith (1985) study an additional extension of the previous findings to smaller time intervals in order to further studies the properities of rapid adaptation. Specifically, dynamic responses to sinusoidal and step changes in intensity were obtained as a function of the time delay between the change in intensity and the onset of an ongoing background sound.

In this study, simulated results from the model are mainly compared to the experimental results reported by Smith et al. (1985).

3.5.1 Methods

The stimulus paradigms used by Smith and his colleagues are shown in Figure 18. The stimulus was a 60ms tone burst with an intensity 13 dB above the AV threshold. A 6 dB decrement or increment occurred at time Δt. The time delay in Δt is changed between 0ms to 40ms. The stimulus used in this study is same as shown Figure 18. The stimulus frequency with a decrement in intensity is 3500Hz and with an increment is 4000Hz.

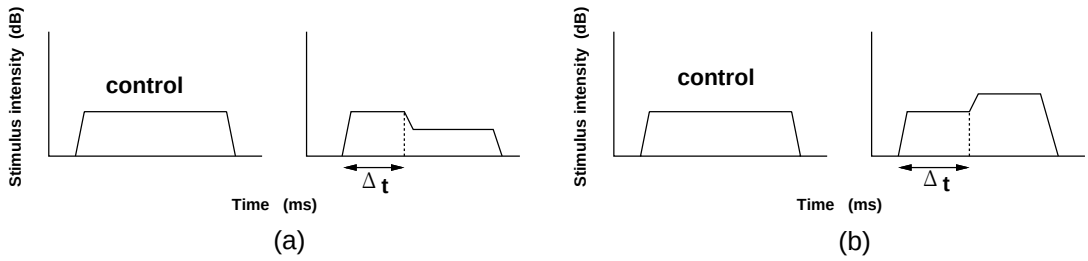


Figure 18. Schematic representation of the paradigms. (a) Decrements of 6dB applied at time delays(Δt) after the onset of a pedestal. (b) Increments of 6dB applied to a pedestal.

3.5.2 Results

Responses to decrements in intensity

An example of PST histogram evoked by decrements in intensity at the shown in the left-hand column of Fig. 19(a). The PST histograms obtained with 3000 stimulus repetitions are smoothed using gms bin width. The CF of the model fiber is 3500Hz. The time delay at the decrement segment is aligning with the onset of the bottom of the

Fig.19 (a). The histograms on the right are the corresponding difference histograms. The difference histogram is obtained by subtracting the response to a tone with a decrement from the response to a constant-intensity tone. Consequently, decreases in firing rate were computed as positive. Figure 19(b) illustrates experimental data reported by Smith et al(1985).

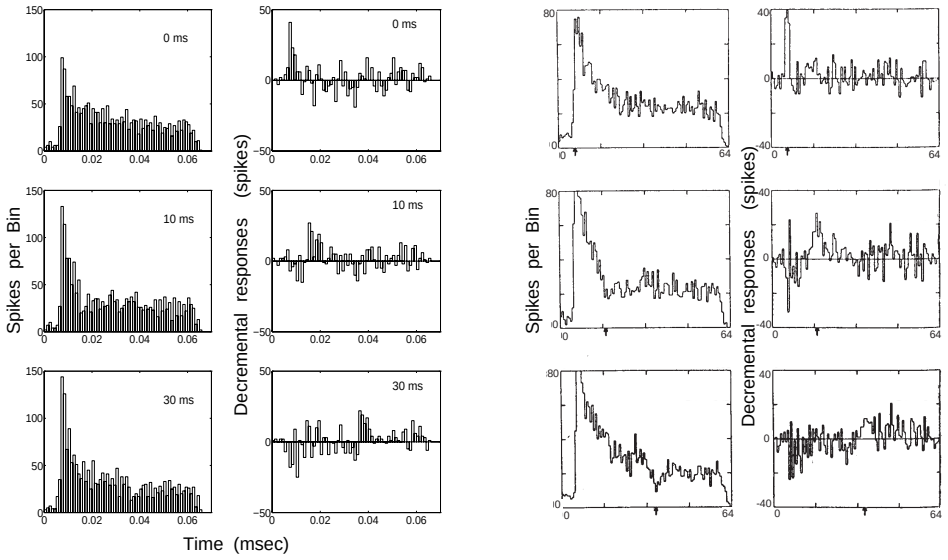


Figure.19. PST histograms of the response to a decrement in intensity. The left-hand column shows total responses and right-hand column the corresponding decremental responses obtained by subtraction as explained in the text. (a) Data from model fiber at CF 3.5kHz. Binwidth of histogram is 1ms. (b) Physiological data from fiber at CF 3.58kH. Bin width of histogram is 640 μ sec..

The effects of time delay are shown by another way in Fig 19.

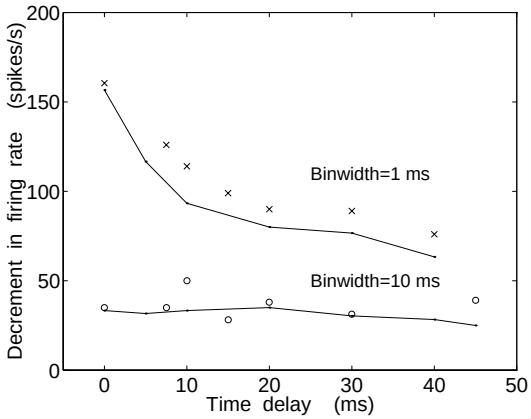
The onset decremental response versus time delay, measured with a small window of 1ms, is plotted in Fig 20 (solid line), is decrease with time. In contrast, the decremental response measured with a large window of 10ms, shown in Fig 20 (solid line), is approximately constant as a function of time delay, as was found data from Smith et al (1985) in shown in Fig 20 (open circle)

Consequently, the onset decremental response measured using both small and large windows from the model is consistent with physiological data reported by Smith (1985).

Response to increments in intensity

Incremental responses were obtained using paradigms identical to those used for decrements, but with a 6dB increase in intensity applied instead of a decrease. Typical incremental responses, measured using both small and large windows, are shown in Fig 21.

In contrast to the onset decremental response, neither measure of the incremental response decreases as a function of the time delay from background onset. Both the onset



Figuer.20. Decremen t in resp onse caused by 6 dB decremen t in lev el of stim ulation applied at various dela ys after onset. Solid line: data from the mo del as measured 1ms and 10ms bin width, Cross: empirical data from Smith et al., 1985. Bin width is 0.64ms, Open circles: empirical data from Smith et al., 1985. Bin width is 10.2ms

increment, measured with a small window, and the increments measured with a large window either remain constant or increase slightly as a result of time delay. Physiological data derived from Smith et al. (1985) and Winter et al. (1993) are shown Fig. 21 (open circle and cross, respectively)

In Figuer 21, both the onset increments, measured with a small window, and the increments measured with a large window is constant for both data from the mo del and physiological data repoted by Smith (1985). Consequently, the onset incremental response measured using both small and large windows from the mo del is consistent with physiological data.

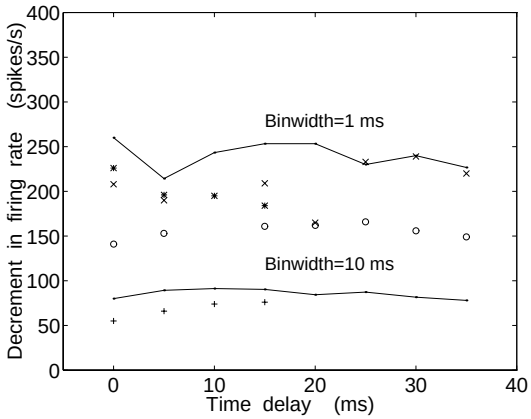


Figure 21. Increment in response produced by a 6 dB increment in level of stimulation applied at various delays after onset of stimulation. Solid line: data from the model as measured 1ms and 10ms bin width; Cross: epirical data from Smith et al., 1985. Bin width is 0.64ms. Open circle: epirical data from Smith et al., 1985. Bin width is 10.2ms. Asterisk: epirical data from Winter et al. 1998. Bin width is 0.64ms; Plus: epirical data from Winter et al. 1998. Bin width is 10.2ms.

3.6 Response to steady-state vowels

To investigate the speech processing of the auditory periphery, we compile the response of the vowel using the auditory peripheral model.

Using the vowel /i/ as input data for the model, the discharge patterns of all CFs can be obtained from the model's output. An example of discharge patterns is shown in Fig. 22. PST, ISI and period histograms are based on responses obtained over 100 trials. The bin size of the PST and ISI histograms is 1 ms and 0.1ms respectively and that of the period histograms is 0.2 ms. The period histograms were made by accumulating the number of occurrences of discharges in pitch period intervals (for /i/ in this work, 8.2 ms). ISI and period histograms were calculated using steady-state responses of the peripheral model. The dynamic range of the AN model is set uniformly at 45 dB and SR is set uniformly at 0 spikes/sec. Sound pressure level of the stimulus is 60dB.

The patterns show how the vowel features are represented in the auditory peripheral system.

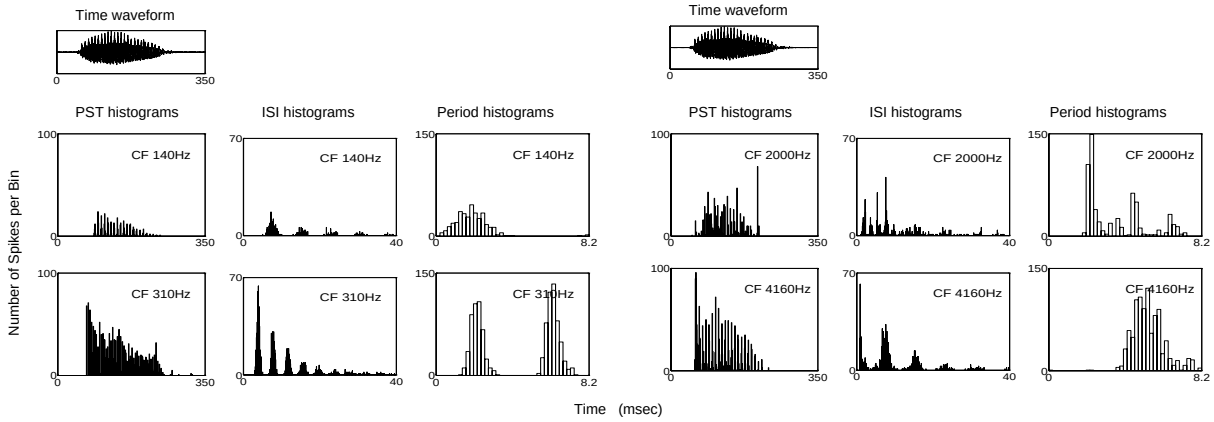


Figure 22. PST histograms, ISI histograms and "pitch-synchronous" period histograms of the responses of four different CFs fibers. The Top left and right above PST histograms are time waveforms for the vowel /i/.

Chapter 4

General discussion

The evaluated results are summarized in Table 4. Table 4 shows a comparison of the previous eight IHC models and the proposed peripheral model.

Table 4. The evaluated results and comparison of the eight previous IHC models and the present peripheral model. Each column of tables is based on Ross (1996) and Hewitt et al. (1991). The codes used are the same as those used by Ross. The codes for entries are as follows: "+," test passed; "(+)," test partly passed; "No," test failed; "(No)," test partially failed; "Unr," test passed only with unrealistic parameter value (s); "NA," not applicable owing to nature of model; "?," not tested.

	Scröder and Hall	Allen	Oono and Sujuka	Cooke	Schwid and Geisler	Meddis	Smith and Brachman (1980)	Ross (1996)	Meki
Rapid and short-term adaptation									
(1) Larger dynamic range for onset than for steady state	Unr.	?	?	+	+	+	+	+	+
(2) Adaptation as a sum of (at least) two exponentials	No	No	+	+	+	+	+	+	+
(3) Exact fit to data sets for individual fibers	?	?	?	?	?	?	?	+	+
Properties of synchronized activity									
Interval histograms for driven activity	?	NA	NA	NA	NA	+	NA	+	+
Period histograms without clipping at high levels	+	?	?	?	+	Too flat	+	(+)	+
Synchronization Index versus intensity	?	?	?	?	?	+	?	+	+
Sync. coefficient versus frequency of stimulation	No	+	No	No	No	+	+	+	+
Recovery from adaptation									
(1) as single exponential	No	No	No	No	No	+	No	+	+
(2) with smaller exponent for onset than steady state	NA	NA	NA	NA	NA	No	NA	+	+
Recovery of spontaneous activity									
(1) as single exponential with realistic time constant	Unr.	Unr.	Unr.	Unr.	Unr.	+	Unr.	+	+
(2) with "dead" period	No	No	No	No	No	No	No	+	+
Response changes to intensity									
Additivity for increments of stimulation	No	No	No	+	+	No	+	+	+
Response to decrement versus delay									
(1) decreasing function for 1-ms window	?	?	?	?	?	+	(+)	(+)	+
(2) flat function for 10-ms window	?	?	?	?	?	No	+	No	+
Hazard function for driven activity									
(1) For Low-CF fiber	?	?	?	?	?	?	?	?	+
(2) fiber with weak phase locking	?	NA	NA	NA	NA	?	NA	+	+
Other evaluation									
Output of the real action potentials	No	No	No	No	No	No	No	+	+
Model consist of concatenation of sub models	No	No	No	No	No	No	No	No	+
Making of the PSTH from real action potentials	NA	NA	NA	NA	NA	NA	NA	No	+

The proposed model passed all tests as shown in Table 4.

The IHC-AN model proposed by Ross successes to a certain extent. But the model was not connected with external ear and middle ear. On the contrary, the proposed model consists of concatenation of sub models, i.e. external ear, middle ear, inner ear. the proposed model was a peripheral model. Therefore, the proposed model can produce the response of vowels or the response of more complex sounds that can not ignore the role of external ear and middle ear.

As mentioned in Introduction, other peripheral models consist of concatenation of sub models have been proposed (Carney , 1993; Jenison, 1991; Kates, 1991; Payton, 1988, and Schoonhoven, 1994), but output of the models was not evaluated in detail as shown in Table. 4. Therefore, the proposed peripheral model was the most effective in providing primary input to the central auditory model.

The number of parameters of proposed model was lessthan that of other models e.g. Ross, (1996). From Tables 2. and 3, a half of parameters of IHC and AN models were constant through various evaluations. Accordingly, the parameter values can fit to agree with physiological data easily Additionally, the parameter dependence on evaluation was clear to a certain extent as shown in above section.

Chapter 5

Conclusion

We present a functional model of the auditory peripheral system that can generate a train of spikes of all CFs. The output of the model were compared with various physiological data.

For the test of rapid and short-term adaptation, the short and rapid components for individual model fibers are in quantitatively agreement with those of the physiological data to modify the membrane permeability of the IHC model. Clearly, through fitting the parameters with physiological data, the rapid time constants are sensitive to the IHC model parameter x which is the weighting factor related to recovering transmitter level. Therefore, the rapid time constant can be fitted to physiological data by modifying only the parameter x. For the test of period histograms without clipping at high levels, the aspects of the period histograms based on the output of the model are very similar to those of the physiological data. The function $f(A_j, \tau_j)$ is most effective in modeling these properties. For the test of interval histograms for driven activity, the shapes of the ISI histograms based on the output of the model are very similar to those of the physiological data. The period of splitting local distributions in ISI histograms can be controlled by the AN model parameters C_λ, C_T, and c_m. C_λ is the most effective in the period of local distributions in ISI histograms. The modality of ISI histogram can be controlled by the AN model parameters C_A and $T(t)$. For the test of synchronization Index versus intensity, data from the model is matched to the physiological data with regard to threshold. For the test of synchronization coefficient versus frequency of stimulation, the model simulates the falloff of synchronization with increasing frequency accurately compared to the physiological data. For the test of recovery of onset and short-term process, the hazard function for fiber with weak phase locking and for low CF fiber from the model is quite agreement with emprical data. For the test of recovery from adaptation, the value of the estimated time constants from the model is realistic compared to that from experimental results. For the test of recovery of spontaneous activity, the spontaneous activities of the model fiber follow a single exponential function and value of time constant was matche to the physiological data. For the test of reponse changes to intensity, both the onset increments, measured with a small window, and the increments measured with a large window is constant for both data from the model and physiological data.

As described above, the model was in excellent agreement with physiological data

compared to previous models over the last two decades as shown in Table 4. Additionally the number of parameters of the proposed model was less than other models that is in agreement with physiological data to a certain extent. Accordingly the parameter values can be set to agree with physiological data easily.

To begin an investigation of neural processing beyond the auditory periphery the model is effective in providing primary inputs to central auditory processing models.

Bibliography

[1] Carney, L. H. (**1993**). "A model for the response of low-frequency auditory-nerve fibers in cat," J. Acoust. Soc. Am. **93**, 401-417.

[2] Carney, L. H., and Geisler, C. D. (**1982**). "A temporal analysis of auditory-nerve fiber responses to spoken stop consonant-vowel syllables," J. Acoust. Soc. Am. **79**, 1896-1914.

[3] Dallos, P. and Cheatham, M. A. (**1989**). "Nonlinearities in cochlear receptor potentials and their origins," J. Acoust. Soc. Am. **86**, 179-1796.

[4] Dallos, P. and Cheatham, M. A. (**1990**). "Effects of electrical polarization on inner hair cell receptor potentials," J. Acoust. Soc. Am. **87**, 1636-1647.

[5] Delgutte, B. (**1980**). "Representation of speech-like sounds in the discharge patterns of auditory-nerve fibers," J. Acoust. Soc. Am. **68**, 843-857.

[6] Deng, L., and Geisler, C. D. (**1987**). "A composite auditory model for processing speech sounds," J. Acoust. Soc. Am. **82**, 2001-2012.

[7] Desoer, C. A., and Kuh, E. S. (**1969**). "Basic Circuit Theory" (McGraw-Hill, New York).

[8] Ghitza, O. (**1988**). "Temporal non-place information in the auditory-nerve firing patterns as a front-end for speech recognition in a noisy environment," J. Phonet. **16**, 109-123.

[9] Geisler, C. D. (**1988**). "Representation of speech sounds in the auditory nerve," J. Phon. **16**, 19-35.

[10] Giguere, C., and Woodland, P. C. (**1994**). "A computational model of the auditory periphery for speech and hearing research. I. Ascending path," J. Acoust. Soc. Am. **95**(1), 331-342.

[11] Greenwood, D. (**1990**). "A cochlear frequency-position function for several species-29 years later," J. Acoust. Soc. Am. **87**, 2592-2605.

[12] Gaumond, P. R., Kim, D. O., and Molnar, C. E. (**1983**). "Response of cochlear nerve fibers to brief acoustic stimuli: Role of discharge-history effects," J. Acoust. Soc. Am. **74**, 1392-1398.

[13] Harris, D. M., and Dallos, P. (**1979**). "Forward masking of auditory nerve fiber response," J. Neurophysiol. **42**, 1083-1107.

[14] Hackney, C. M. (**1987**). "Anatomical features of the auditory pathway from cochlea to cortex," Br. Med. Bull. **43**, 780-801.

[15] Hewwitt, M. and Meddis, R. (**1991**). "An evaluation of eight computer models of mammalian inner hair-cell function," J. Acoust. Soc. Am. **90**, 904-917.

[16] Javel, E., McGee, J. A, Horst, J. W., and Farley, G. R. (**1988**). "Temporal mechanisms in auditory stimulus coding," *in Auditory Function: Neurological bases of Hearing*, edited by G.E.Edelman, W.E.Gall, and W.M.Cowan(Wiley, New York), Chap. 17.

[17] Jenison, R. L., Greenberg, S., Kluender, K. R., and Rhode, W. S. (**1991**). "A composite model of the auditory periphery for the processing of speech based on the filter response functions of single auditory-nerve fibers," J. Acoust. Soc. Am. **90**, 773-786.

[18] Johnson, D. H. (**1980**). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," J. Acoust. Soc. Am. **68**, 1115-1122.

[19] Kates, J. M. (**1991**). "A time-Domain Digital Cochlear Model," IEEE Trans. Signal Proc. **39**, 2573-2592.

[20] Lutman, M. E., and Martin, A.M. (**1979**). "Development of an electroacoustic analogue model of the middle ear and acoustic reflex," J. Sound. Vib. **64**, 133-157.

[21] Miller, M. I., and Sachs, M. B. (**1983**). "Representation of stop consonants in the discharge patterns of auditory-nerve fibers," J. Acoust. Soc. Am. **74**, 502-517.

[22] Maki, K. and Akagi, M. (**1996**). "A Real-Time Functional Model of the Auditory Peripheral System," Trans. Tech. Com. Psycho. Physio. Acoust. H-96-74.

[23] Maki, K. and Akagi, M. (**1997**). "A Functional Model of the Auditory Peripheral System," Proc. International Symposium on ASVA.

[24] Meddis, R. (**1986**). "Simulation of Mechanical to Neural Transduction in the Auditory Recepter," J. Acoust. Soc. Am. **79**, 702-711.

[25] Meddis, R. (**1988**). "Simulation of Auditory-Neural Transduction: Further Studies," J. Acoust. Soc. Am. **83**, 1506-1063.

[26] Meddis, R., and Hewitt, M. J. (**1991**). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery I: Pitch identification," J. Acoust. Soc. Am. **89**, 2866-2882.

[27] Payton, K. L. (**1988**). "Vowel processing by a model of the auditory periphery : A comparison to eighth-nerve responses," J. Acoust. Soc. Am. **83**, 145-162.

[28] Pont, M. J. and Damper, R. I. (**1991**). "A computational model of afferent neural activity from the cochlea to the dorsal acoustic stria," J. Acoust. Soc. Am. **89**, 1213-1228.

[29] Rose, J. E., Brugge, J. F., Anderson, D. J. and Hind, J. E. (**1967**). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey," J. Neurophysiol. **30**, 769-793.

[30] Ross, S. (**1996**). "A functional model of the hair cell - primary fiber complex," J. Acoust. Soc. Am. **99**(1), 222-228.

[31] Sachs, M. B. and Young, E. D. (**1979**). "Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate," J. Acoust. Soc. Am. **66**, 470-479.

[32] Seneff, S. (**1988**). "A joint synchrony/mean-rate model of auditory speech processing," J. Phonet. **16**, 55-76.

[33] Shamma, S. (**1985**). "Speech processing in the auditory system II. Lateral inhibition and the processing of speech-evoked activity in the auditory nerve," J. Acoust. Soc. Am. **78**, 1612-1621.

[34] Sinex, D. G., and Geisler, C. D. (**1983**). "Responses of auditory-nerve fibers to consonant vowel syllables," J. Acoust. Soc. Am. **73**, 602-615.

[35] Sinex, D. G. and McDonald, L. P. "Average discharge rate representation of voice onset time in the chinchilla auditory nerve," J. Acoust. Soc. Am. **83**, 1817-1827.

[36] Smith, R. L., and Zwislocki, J. J. (**1975**). "Short-term adaptation and incremental responses of single auditory-nerve fibers," Biol. Cybern. **17**, 169-182.

[37] Smith, R. L. "Short-term adaptation in single auditory nerve fibers: some poststimulatory effects," (**1977**). J. Neurophysiol. **40**, 1098-1112.

[38] Smith, R. L. (**1979**). "Adaptation, saturation, and physiological masking in single auditory-nerve fibers," J. Acoust. Soc. Am. **65**, 166-1098.

[39] Smith, R. L., and Brachman, M. L. (**1980**). "Operating range and maximum response of single auditory-nerve fibers," Brain Res. **184**, 499-505.

[40] Smith, R. L., Brachman, M. L., and Frisina, R. D. (**1985**). "Sensitivity of auditory-nerve fibers to changes in intensity: A dichotomy between decrements and increments," J. Acoust. Soc. Am. **78**, 1310-1316.

[41] Tateno, T. and Hirahara, T. (**1995**). "An auditory peripheral model based on physiological structure" Hearing Tech. Report, ASJ, H-95-54.

[42] Westerman, L. A., and Smith, R. L. (**1984**). "Rapid and Short Term Adaptation in Auditory-Nerve Responses," Hear. Res. **15**, 249-260.

[43] Westerman, L. A., and Smith, R. L. (**1987**). "Conservation of adapting components in auditory-nerve responses," J. Acoust. Soc. Am. **81**, 680-691.

[44] Westerman, L. A. (**1985**). "Adaptation and recovery of auditory nerve responses," Ph. D. Thesis, Syracuse University, Syracuse, N. Y.

[45] Wilson, M. A. and Bower, J. M. (**1988**). "The Simulation of Large-Scale Neural Networks," in *Methods in Neuronal Modeling* edited by Koch, C. and Segev, I. (The MIT press), Chap. 9.

[46] Zagaeski, M., Cody, A. R., Russell, J. and Mountain, D. C. (**1994**). "Transfer characteristic of the inner hair cell synapse: Steady-state analysis," J. Acoust. Soc. Am. **95**, 3430-3434.

Presented papers

牧　勝弘、　赤木　正人 (**1996**). “聴覚末梢系における実時間モデルの検討” [1]、日本音響学会聴覚研究会資料、H-96-74.

Maki, K. and Akagi, M (**1997**). “A Functional Model of the Auditory Peripheral System,” Proc. International Symposium on ASVA

牧　勝弘、　赤木　正人 (**1997**). “聴覚末梢系における機能モデルの検討 ”[2]、日本音響学会講演論文集、 1-8-11

[1] Maki, K. and Akagi, M. “A Real-time Functional Model of the Auditory Peripheral System”
[2] Maki, K. and Akagi, M. “A Functional Model of the Auditory Peripheral System”