A MODEL OF THE CENTRAL PROCESSING OF AUDITORY NERVE ACTIVITY

A LATERAL-INHIBITORY-NETWORK MODEL OF THE CENTRAL PROCESSING OF AUDITORY NERVE ACTIVITY

By

JENNIFER KO, B.A.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Applied Science

McMaster University

Master of Applied Science (2004)
(Electrical Engineering)

McMaster University
Hamilton, Ontario

TITLE: **A Lateral-Inhibitory-Network Model of the Central Processing of Auditory Nerve Activity**

AUTHOR: Jennifer Ko, B.A.Sc. (University of Toronto)

SUPERVISOR: Dr. Ian C. Bruce

NUMBER OF PAGES: xiii, 143

Abstract

Lateral inhibitory networks (LINs) of neurons are thought to be prominent in sensory systems and are known to enhance spatial edges and peaks in their input excitation patterns. It is postulated, based on experimental findings, that lateral inhibition contributes to central, sub-cortical, auditory processing. Previous computational LIN models of the central processing of auditory nerve activity were based on highly simplified, non-spiking models of neurons. A more biologically realistic LIN model of spiking neurons was thus developed to investigate the plausibility of such networks achieving contrast enhancement and speech feature extraction.

The model developed is a single-layer, uniform, recurrent LIN structure. Each neuron in the LIN is described by a leaky integrate-and-fire model with conductance-based synaptic input. Input spike instances were obtained from Bruce and colleagues' [2003] model of the auditory periphery for synthesized speech stimuli or from a Bernoulli approximation of a Poisson process to represent spontaneous activity from the auditory nerve.

The effect of neural and network parameters on contrast enhancement exhibited in the mean spike rates was measured. It was found that the spiking LIN is able to achieve contrast enhancement if the values of the neuronal parameters fall within a very specific and narrow range. Furthermore, the spatial edge in the input had to be high and steep. Compared to non-spiking neuron models, it is quite difficult for spiking neurons to achieve contrast enhancement. The spiking LIN was found to be capable of making formant frequencies more distinct in the average rate profiles of speech stimuli presented at high intensities. However, synchrony of the neural activity to the formant frequencies was largely lost. This architecture of spiking neurons is therefore unlikely to be how contrast enhancement and speech feature extraction is realized in the central auditory system.

Acknowledgments

Many people have contributed as teachers, friends and colleagues to my graduate education. Their guidance, love and support, made my M.A.Sc. candidature the most enjoyable two years I have spent in academia.

I would first like to thank Dr. Ian Bruce for his attentive supervision. His patience for my ignorance and my love of playing sports was remarkable. I am grateful to him for all that he has taught me, for editing all of my work (including this thesis), and for his friendship. Ian and his family made the seemingly bleak city of Hamilton a much warmer place for me.

My sincerest gratitude goes to the following people for their direct and indirect contributions to this work: Mr. Harjeet Bajaj for coding the earliest version of the model; Ms Silvia Pavuk, Mr. Gregory Hasiuk, and Dr. Melissa Dominguez for doing the unthinkable — proof-reading this thesis; Mr. Barry Rawn for acquiring journal articles for me and for competing against me in the Thesis Race; Mr. Marc Sibson, Mr. Chris Lambacher, and Mr. Karl Martin for their invaluable assistance in typesetting this document in LaTeX; Mr. Kamran Mustafa for being a great lab-mate and for serving as Computer Techie; and all my friends at McMaster for their care, for helping me keep a well balanced student-life and for getting me through graduate courses - the legend of the Mad Cows and Team Stochastic will never be forgotten.

Most importantly, I am forever indebted to my parents, Dick and Lena Ko, who have been so willing to support my endeavours as a Professional Student. To them, I dedicate this thesis, for their unfailing love and support.

This research was funded by the Canadian Institutes of Health Research (New Emerging Teams grant 54023).

Contents

List of Figures

List of Abbreviations

Abbreviation	Term
AAF	Anterior Auditory Field
ABR	Auditory Brain Stem Response
AI	Primary Auditory Cortex
AII	Secondary Auditory Cortex
AVCN	Anterior Ventral Cochlear Nucleus
CF	Characteristic Frequency
DCN	Dorsal Cochlear Nucleus
EPSC	Excitatory Postsynaptic Current
GABA	γ-aminobutyric acid
IC	Inferior Colliculus
IHC	Inner Hair Cell
IPSC	Inhibitory Postsynaptic Current
ISI	Interspike Interval
LIN	Lateral Inhibitory Network
LL	Lateral Lemniscus
OAE	Otoacoustic Emission
OHC	Outer Hair Cell
PVCN	Posterior Ventral Cochlear Nucleus
SOAE	Spontaneous Otoacoustic Emission
SOC	Superior Olivary Complex
SPL	Sound Pressure Level

List of Symbols

Symbol	Variable Represented
a	lateral inhibition scaling factor
c	scaling factor of conductances
C	membrane capacitance [F]
$C_{IHC}[i]$	inner hair cell scaling constant
$C_{OHC}[i]$	outer hair cell scaling constant
E_E	excitatory synaptic reversal potential [V]
E_I	inhibitory synaptic reversal potential [V]
$g_{E,i}(t)$	excitatory, unitary synaptic conductance [S]
$g_{I,i}(t)$	inhibitory, unitary synaptic conductance [S]
$G_E(t)$	excitatory conductance [S]
$G_I(t)$	inhibitory conductance [S]
$i_{EPSC,i}(t)$	excitatory postsynaptic current injection [A]
n	number of neurons
Q_{10}	measure of sharpness of a frequency tuning curve
R	membrane leakage resistance [Ω]
$s_{in,i}$	input spike train to the i^{th} neuron
$\boldsymbol{s}_{in}$	matrix of input spike trains
$s_{out,i}$	output spike train of the i^{th} neuron
$\boldsymbol{v}(t)$	column vector of relative membrane potentials [V]
t_{ref}	absolute refractory period [s]
$v_i(t)$	i^{th} neuron's relative membrane potential [V]
$v_m(t)$	membrane potential [V]
v_{rest}	resting potential [V]

Continued on next page ↪

↪ Continued from previous page.

Symbol	Variable Represented
$v_{th,i}(t)$	threshold potential of the i^{th} neuron [V]
$\boldsymbol{V}$	matrix of excitatory synaptic weights
$\boldsymbol{W}$	matrix of inhibitory synaptic weights
$V_{i,j}$	excitatory weight of the j^{th} synapse on the i^{th} neuron
$W_{i,j}$	inhibitory weight of the j^{th} synapse on the i^{th} neuron
α_E	alpha value of $g_{E,i}(t)$
α_I	alpha value of $g_{I,i}(t)$
τ	membrane time constant [s]

Chapter 1

Introduction

Lateral inhibition occurs when a cell projects a signal to its neighbours that suppresses their activity. Such intercellular signals are known to enhance spatial edges in the profiles of cellular activity. Lateral inhibition is thought to exist in many sensory systems. For example, lateral inhibition in the retina is the mechanism by which the optical illusion of Mach bands is created (as will be explained in Section 3.1). Electrophysiological experiments indicate that the tuning curves[1] of neurons in auditory centres of the brain become sharper toward the thalamus. This finding suggests that lateral inhibition contributes to auditory signal processing in the brain. The action of lateral inhibitory processes on abnormal spontaneous-type neural activity has been hypothesized as a mechanism of tinnitus, an auditory condition often referred to as "ringing of the ears". Computational lateral-inhibitory-network (LIN) models were thus developed to investigate the plausibility of a LIN as a central processor of spontaneous and speech-driven auditory nerve activity, and as a central mechanism of tinnitus.

[1]Plots of stimulus intensity required to induce an increase in neural activity versus the acoustic stimulus frequency.

1.1 Scope of Work

The ultimate goal of computational modeling of the human auditory system is to develop a highly realistic model that can completely account for all empirical observations. Obviously, the ultimate goal is an indefinite one that is unrealistic for any discrete project. The intention of the endeavour reported here is not to address this massive problem, but rather to make a small contribution toward the ultimate goal.

Models of the mammalian auditory periphery such as the Bruce et al. model have already been fairly well developed and therefore reflect many of the physiological features observed. In contrast, the development of models of the central auditory system has been limited because the detailed anatomy and physiology of the central auditory system remains largely unknown to date. Nonetheless, computational models based on the little that is known of the central auditory system can and have been developed to evaluate the plausibility of conceptual models. The lateral-inhibitory-network models that will be outlined in Section 3.3 were created with the latter intention.

The objective of this project is to further develop existing auditory LIN models in order to assess the plausibility of a more realistic LIN model of spiking neurons performing contrast enhancement. The approach taken was to analyze the effect of neural and network parameters on contrast enhancement achieved by a LIN model of spiking neurons. It was thought that the model presented in this report might serve as a preprocessor to a model of the auditory cortex that is currently being developed[2].

1.2 Contributions of this Work

The processing of auditory nerve activity by a uniform, single-layer, recurrent LIN model of spiking neurons is novel in that the time instances of action potentials is accounted for. For neural units, this model employs the biologically realistic leaky

[2]A joint project between the departments of Electrical and Computer Engineering and Psychology at McMaster University by Drs. Melissa Dominguez, Suzanna Becker and Ian Bruce.

integrate-and-fire model of a neuron where previous models used non-spiking models of neural activity.

The spiking LIN is able to achieve contrast enhancement in the form of a spurious peak and valley flanking an edge in neural activity only if the values of the neural parameters fall within a very specific and narrow range. Furthermore, the spatial edge in the input has to be high (on the order of hundreds of spikes per second) and steep for contrast enhancement to be observed. A spurious peak of activity at a steep transition between normal and impaired regions of hearing is therefore produced when speech stimulus is presented to the ear at a high intensity. Compared to non-spiking neuron models, it is quite difficult for spiking neurons to achieve contrast enhancement. The difficulty arises due to the relatively short duration of the excitatory and inhibitory postsynaptic potentials (EPSPs and IPSPs) elicited by the input spike trains. As such interaction between the EPSPs and IPSPs can only occur if the input spikes occur close to one another in time and space. The spatial and temporal frequency of spikes must, therefore, be high. Contrast enhancement improves when the duration of disturbances in the inhibitory conductance is long, the neurons are densely connected and have a short refractory period, and the inhibition is of the hyperpolarizing type. Inhibitory interactions are most effective when the threshold potential for spiking is set to just allow a single excitatory input spike to induce an output spike in the absence of inhibitory input, as the presence of any inhibition will prevent an output spike from being generated. This threshold value was around 20 mV above the resting potential of the cell, which is consistent with patch clamp data from the inferior colliculi of mice [Basta & Vater, 2003] and rats [Koch & Grothe, 2003]. For a membrane capacitance of around 8 pF, a membrane time constant and absolute refractory period on the order of 1 ms are required to produce contrast enhancement. The time-course of the disturbance in excitatory conductance due to a single input spike must be similar to the membrane time constant and an order of magnitude shorter in duration than that of the inhibitory conductance to achieve significant contrast enhancement (e.g., 2 ms versus 20 ms, respectively). Non-spiking neurons

achieve contrast enhancement more easily because the graded membrane potential is used to modulate the inhibitory input, which has a much slower time-course. The slower modulation of the inhibition increases the opportunity for interaction between the excitatory and inhibitory inputs, thus facilitating contrast enhancement.

The spiking LIN was found to be capable of enhancing formant frequencies in the speech stimuli presented at high intensities. When the stimulus is presented at lower intensities, no enhancements were observed. However, the same spiking LIN severely degraded the synchrony of the neural activity to the formant frequencies as measured by power ratios of synchronous rates.

A uniform, single-layer structure of spiking neurons is therefore unlikely to be the mechanism of contrast enhancement in the central, subcortical auditory system if the action potential is the primary information carrier. As such, this work does not support the hypothesis of lateral inhibition as a central mechanism of tinnitus. Since non-spiking neurons have not been found in the auditory system, the biological plausibility of this neural network architecture being the mechanism of contrast enhancement is greater if the graded membrane potential of the spiking neurons is the primary modulator of lateral inhibitory activity.

1.3 Thesis Layout

Following this introduction, Chapter 2 contains a description of the anatomy and physiology of neurons and the auditory system, the leaky integrate-and-fire model of a neuron and the acoustics of speech sounds. The knowledge presented in that chapter is fundamental to the understanding of material presented in later sections. Chapter 3 provides context for the development of a LIN of spiking neurons and why the results of this work are significant. A detailed description of the model is then given in Chapter 4. Simulation results are summarized in Chapter 5, followed by a discussion in Chapter 6 that relates these findings to published results of previous models. The latter chapter also reveals limitations of this model and proposes the

investigation of a modified version of this model. In the final chapter, a summary of findings and suggestions for future work conclude this report.

1.4 Related Publications

Parts of this thesis have appeared in the following publications:

Bruce, I. C., Bajaj, H. S., & Ko, J. (2003). "Lateral-inhibitory-network models of tinnitus," in D. D. Feng & E. R. Carson (Eds.), *Proceedings of the 5th IFAC Symposium on Modelling and Control in Biomedical Systems* (pp. 359–363). Oxford, UK: Elsevier Ltd.

Ko, J. and Bruce, I.C. (2004). "Lateral-Inhibitory-Network Models of the Central Processing of Auditory Nerve Activity," in *Abstracts of the 27th ARO Midwinter Research Meeting.*

Ko, J. and Bruce, I.C. (2004). "Lateral-Inhibitory-Network Models of the Central Processing of Auditory Nerve Activity," in *Abstracts of papers presented at the 2004 Meeting on Computational and Systems Neuroscience.*

A journal article is currently being prepared for submission to *Neural Computation.*

Chapter 2

Background

2.1 Anatomy and Physiology of a Neuron

A nerve cell, or neuron, is the fundamental functional unit of the nervous system. The properties and function of neurons differ depending on their location in the nervous system. As shown in Figure 2.1, a typical neuron consists of four main parts: The cell body, or soma, contains the cell nucleus and delineates the input (or information-gathering) parts from the output (or information-transmitting) parts of the cell; the dendrites are tree-like structures that make connections with other cells to collect input; the axon serves as a transmission medium along which electrical signals propagate; and, the telodendria are the branches of the end of the axon where connections are made with other cells to transmit the electrical signal generated by the neuron.

A neuron processes and transmits information via electric currents. An electric current is generated and propagated in neurons by the movement of ions. Ions that are commonly found in the cytoplasm and extracellular fluid include Na^+, K^+, Ca^{2+} and Cl^-. At equilibrium, there is a higher concentration of Na^+, Ca^{2+} and Cl^-, and a lower concentration of K^+ outside the cell than inside. An electric potential is created across the semi-permeable cell membrane by a difference in electric charge. This time varying potential difference is referred to as the transmembrane potential

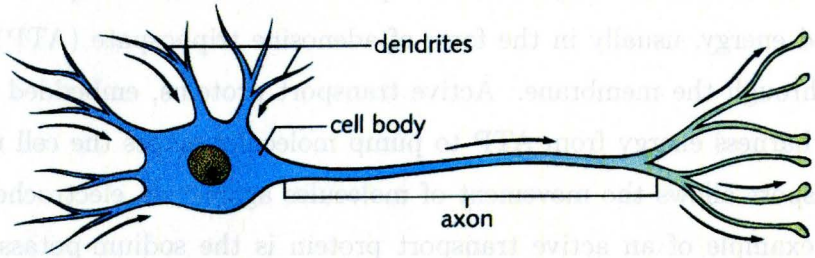


Figure 2.1: Illustration of a typical neuron. (Reprinted with permission from Levine & Miller, 1994, Fig.43.1, p.896.)

(or simply the membrane potential) and is denoted by $v_m(t)$. The membrane potential that is characteristic of the cell at equilibrium is known as the resting potential, v_{rest}, and is typically around -70 mV (intracellularly negative). When $v_m(t)$ becomes more positive or negative than v_{rest} due to ionic movement, the cell is said to be depolarized or hyperpolarized, respectively.

Ions and other molecules may pass through the semi-permeable membrane by means of passive or active transport mechanisms. Diffusion, simple or facilitated, is one means by which molecules may pass through the cell membrane. Diffusion is driven by concentration and electrical potential gradients, thereby making it a form of passive transport, not requiring the use of stored energy. Simple diffusion allows small, lipid soluble molecules to pass through the lipid bi-layer that makes up the cell membrane, governed only by the laws of diffusion. Diffusion may also be facilitated by carrier proteins that form pore-like channels through the cell membrane, allowing only specific molecules to pass through. The forces that drive diffusion may oppose one another. For example, since Cl^- is concentrated outside the cell, the concentration gradient across the cell membrane drives Cl^- into the cell. At the same time, the usually negative $v_m(t)$ tends to push Cl^- out of the cell. An equilibrium is reached when these two opposing forces are balanced. The electrical potential required to exactly counterbalance the concentration gradient for a particular ion is known as an equilibrium or reversal potential. Another means by which molecules may pass

through the cell membrane is by active transport. This form of transport requires the use of stored energy, usually in the form of adenosine triphospate (ATP), to pump molecules through the membrane. Active transport proteins, embedded in the cell membrane, harness energy from ATP to pump molecules across the cell membrane. Active transport allows the movement of molecules against an electrochemical gradient. An example of an active transport protein is the sodium-potassium pump present in all cells. This lipoprotein transports three Na^+ ions out of the cell for every two K^+ ions that are transported into the cell. The same lipoprotein acts as the enzyme ATPase to release the energy required to cleave K^+ from the complex at the inner surface of the cell membrane. These sodium-potassium pumps prevent the depletion of Na^+ from the extracellular fluid and K^+ from the intracellular fluid that would otherwise result after a neuron's transmission of action potentials (which will be described in a paragraph to follow).

The ability of neurons to generate and transmit electrical signals is facilitated by gated ion channels. Ionic channels have only two states: open and closed. In the closed state, there is no current flow through the ionic channel. In the open state, a few nanoamperes of current flows through the ionic channel. With a constant input, the current through a channel switches randomly between the open and closed states. There are two kinds of ionic channels: electrically controlled and chemically controlled channels. The rate of open and closed states of electrically gated channels varies systematically with the membrane potential. Most electrically controlled channels remain closed most of the time when $v_m(t)$ is close to v_{rest}, thereby allowing very few ions to pass through the cell membrane. The membrane can therefore be thought of as having low ionic conductances at rest. However, when $v_m(t)$ increases, the probability of channels opening increases, thereby allowing many ions to pass and the ionic conductances to increase. For chemically gated channels, the probability of the channel being open is dependent on the binding of neurotransmitter molecules to the channel complex, which is a synaptic receptor.

The majority of neurons carry information by means of action potentials, which

are brief, but dramatic, positive disturbances (thus also referred to as spikes) in $v_m(t)$ from v_{rest} that sweeps along the neuron. An action potential is initiated by a depolarization to threshold of some part of the neuron (① of Figure 2.2), usually in a dendrite or the soma. When the threshold potential, v_{th}, is reached, the probability of electrically controlled Na^+ channels opening increases, allowing Na^+ to flood into the cell causing $v_m(t)$ to increase rapidly up to a peak on the order of 40 mV (i.e., 110 mV above v_{rest}). This local current causes adjacent parts of the neuron to depolarize and thus a wave of disturbance to travel down the axon. The probability of Na^+ channels being open decreases shortly after opening, thereby halting the influx of Na^+ (②). Simultaneously, many electrically controlled K^+ channels open (③), allowing K^+ to rush out of the cell (④) causing $v_m(t)$ to repolarize and sometimes even hyperpolarize before returning to v_{rest} (⑤). Once an action potential is initiated, the Na^+ and K^+

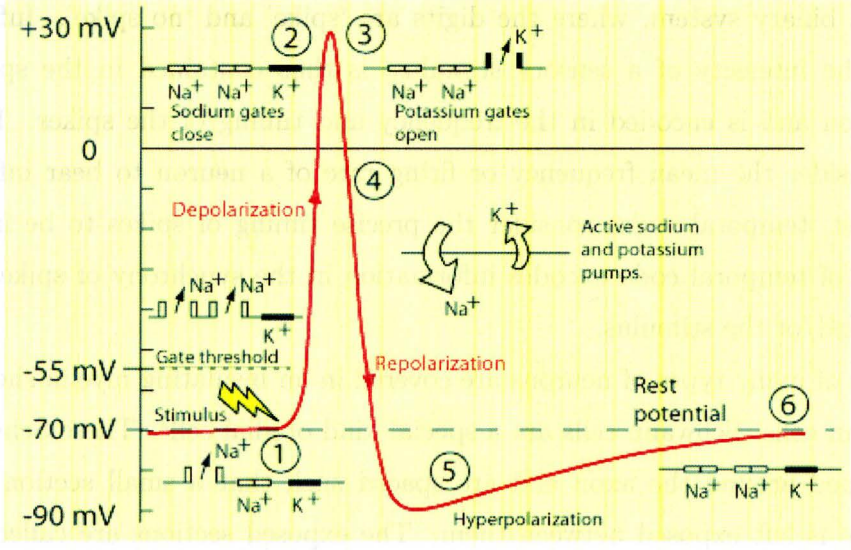


Figure 2.2: Main features of an action potential. The membrane potential is shown along the ordinate and time on the abscissa. (Reprinted with permission from Charand, 2000.)

channel dynamics dominate the neural activity. Due to the interplay between the Na^+ and K^+ conductances, a refractory period follows the spike. For a short period

immediately following the spike, no action potential can be generated. This period typically lasts for about one millisecond and is referred to as the absolute refractory period. As the ionic conductances return to their resting levels, action potentials can be induced only by strong input, i.e., v_{th} is elevated. This phase following the absolute refractory period is known as the relative refractory period and has a duration of a few milliseconds. Action potentials are all-or-nothing events because the electrically gated channels of a neuron all have very similar thresholds. Hence, when a depolarization causes the conductance of one channel to increase, all the channels in the vicinity will undergo the same increase. By the same token, once the threshold is reached, all action potentials of a neuron are of nearly the same magnitude, shape and speed, regardless of the stimulus strength.

Since action potentials are nearly identical to one another, they are only the medium by which information is conveyed. Spikes can be thought of as forming the basis of a binary system, where the digits are 'spike' and 'no spike'. Information, such as the intensity of a sensory stimulus, is thus contained in the spike train[1] of a neuron and is encoded in the frequency and timing of the spikes. Frequency codes consider the mean frequency or firing rate of a neuron to bear information. In contrast, temporal codes consider the precise timing of spikes to be important. One type of temporal code encodes information in the synchrony of spikes between neurons and/or the stimulus.

Axons of many types of neurons are covered in an insulating myelin sheath made of Schwann cells. Schwann cells are a special kind of glial cell[2]. These Schwann cells are wrapped around the axon and are spaced such that a small section of neural membrane is left exposed between them. The exposed sections are called nodes of Ranvier. This myelin sheath alternating with nodes of Ranvier allows for a much faster transmission of spikes down the axon. The speedy transmission occurs because the myelin insulation forces the action potential to jump from one node of Ranvier to the next (i.e., saltatory conduction) in order to close the current loop. Without

[1]Spike trains are records of spike instances in time.
[2]Glial cells provide neurons with structural and metabolic support.

the myelin sheath, an action potential travels down an axon as a wave of disturbance due to very short-range, local current loops. Hence, the delay between the location of spike generation and the terminating synapse is reduced by the myelin sheath to about one tenth of that without the insulation.

Neurons form connections with each other and with other cells via synapses. There are two types of synapses: electrical and chemical. Electrical synapses are the simplest type of connection between cells. Also known as gap junctions, electrical synapses allow the free flow of ions between the cells, thus making the effective function of the cell pair a single electrical unit. Transmission of action potentials across electrical synapses is therefore extremely fast, but they possess little ability to process information. Hence, electrical synapses are useful in situations that require extremely rapid response, but are not well suited for learning or other types of more complex behaviour. A chemical synapse is illustrated in Figure 2.3. This type of synapse dif-

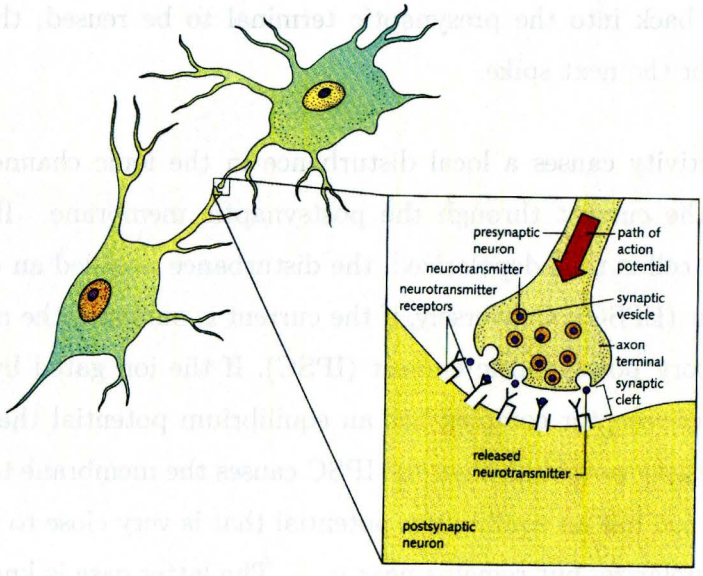


Figure 2.3: Illustration of a typical chemical synapse between two vertebrate neurons. (Reprinted with permission from Levine & Miller, 1994, Fig. 43.10, p.905.)

fers from electrical synapses in that the electrical signal is converted into a chemical one, then back into a electrical signal in the postsynaptic neuron. When an action potential in the presynaptic neuron reaches the axon terminal, it triggers an increase in the probability of electrically gated Ca^{2+} channels opening. This causes an influx of Ca^{2+}. The presence of Ca^{2+} inside the axon terminal causes synaptic vesicles containing neurotransmitter molecules to fuse with the cell membrane. Neurotransmitters are thus released into the synaptic cleft. Some of the neurotransmitters successfully diffuse across the synaptic cleft and bind to neurotransmitter receptors on the surface of the postsynaptic cell membrane. The type of neurotransmitter dictates whether the synapse is excitatory or inhibitory. A common neurotransmitter of inhibitory synapses in the brain is γ-aminobutyric acid (GABA), while glutamate is a common excitatory neurotransmitter. A neurotransmitter-receptor complex triggers the opening of chemically gated channels to allow specific ions to flood the postsynaptic neuron resulting in a local disturbance in current through the cell membrane. The neurotransmitters remaining in the synaptic cleft are either deactivated by enzymes or transported back into the presynaptic terminal to be reused, thus clearing the synaptic cleft for the next spike.

Synaptic activity causes a local disturbance in the ionic channel conductances and therefore the current through the postsynaptic membrane. If the current is inward and the cell is thus depolarized, the disturbance is called an excitatory postsynaptic current (EPSC). Conversely, if the current is outward, the neuron is said to have an inhibitory postsynaptic current (IPSC). If the ion gated by the inhibitory neurotransmitter-receptor complex has an equilibrium potential that is more negative than the resting potential, then the IPSC causes the membrane to hyperpolarize. However, if the ion has an equilibrium potential that is very close to v_{rest}, then $v_m(t)$ does not hyperpolarize, but remains near v_{rest}. The latter case is known as shunting or silent inhibition. In both cases, the neuron is inhibited from firing an action potential by keeping its membrane potential away from its spiking threshold potential. An example of hyperpolarizing inhibition is the operation of the $GABA_B$ receptor.

This receptor controls K^+-selective channels. When activated, K^+ channels tend to open, thereby increasing the membrane conductance and causing an IPSC. Because the reversal potential of K^+ is lower than the resting potential of the membrane, hyperpolarization occurs as K^+ leaves the cell. An example of shunting inhibition is the operation of the $GABA_A$ receptor. When in complex form, Cl^- channels tend to open. However, because the reversal potential of Cl^- is very close to the resting potential of the membrane, hyperpolarization does not occur (hence the name silent inhibition). In the presence of excitatory activity, the increase in Cl^- conductance facilitates the shunting of excitatory currents, thereby greatly reducing the excitatory postsynaptic potential. Postsynaptic potentials from numerous synapses on a dendrite are integrated by the neuron.

Neuromodulators are compounds that regulate synaptic activity by modulating neurotransmitter synthesis, storage or release. They do so by altering the ability of neurotransmitters to bind to their receptors or by altering the fate of neurotransmitters in the axon terminal or in the synaptic cleft. Neuromodulators thus govern the efficacy of synapses. Changes in the efficacy of synapses is called plasticity and facilitates complex behaviour such as learning and memory.

2.2 Leaky Integrate-and-Fire Model of a Neuron

The idea of the leaky integrate-and-fire model was first introduced by Lapicque in the early 1900s despite a lack of understanding, at the time, of the mechanism of neural activity. Lapicque's model [1907] (described in detail by Tuckwell [1988]) is a simple electric circuit consisting of a capacitor in parallel with a resistor and battery to represent the membrane capacitance, the leakage resistance across the cell membrane and the resting potential (C, R and v_{rest} in Figure 2.4). Although this simple circuit cannot generate action potentials, Lapicque postulated that an action potential would be generated when the capacitor was charged to a certain threshold

potential. Once the threshold was reached and the action potential generated, the capacitor would discharge and thereby reset the membrane potential to its resting state. Modern versions of the leaky integrate-and-fire model do exactly that to produce action potentials. Another non-linear feature that is often incorporated into the model is the refractory period [Stein, 1967; Jack et al., 1975]. Although Hodgkin and Huxley later came up with a much more complex model that explained the mechanism of neural activity [Hodgkin & Huxley, 1952], the leaky integrate-and-fire neuron remains more widely used in neural network models because it is fairly simple, yet it captures the main features of a neuron's function, including the action potential. Stein [1967] has shown that the simple leaky integrate-and-fire neuron can approximate features of spike trains, generated by empirical means or by complex models such as that of Hodgkin and Huxley, that are crucial to neural coding (e.g., the frequency-current relation).

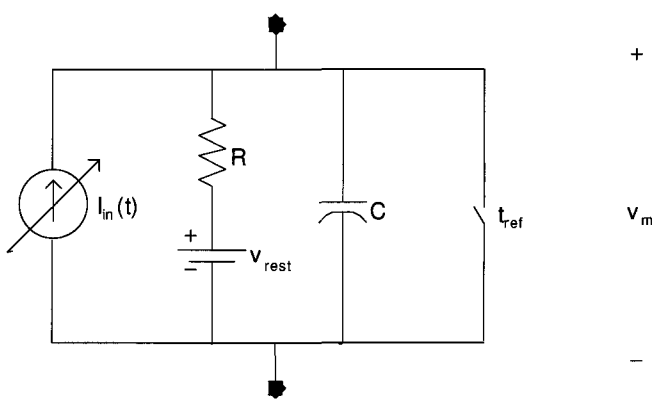


Figure 2.4: Circuit Diagram of the leaky integrate-and-fire model. I_{in}, R, C, and v_{rest} represent the injection current, leakage resistance, membrane capacitance and resting potential, respectively. The switch labelled t_{ref} is used to implement the refractory period.

The operation of the leaky integrate-and-fire model can be divided into two parts: linear and non-linear. The linear portion involves all of the circuit elements other than the switch. Using basic circuit theory, one can derive the following differential

equation for the linear operation governing all subthreshold activity of the model:

$$\tau \frac{dv_m(t)}{dt} + v_m = I_{in}(t) \cdot R + v_{rest} \tag{2.1}$$

The non-linear operation of the model is represented by the switch in the electric circuit. Once the membrane potential, $v_m(t)$, reaches a threshold potential, $v_{th}(t)$, the model's function that immediately follows is non-linear. As soon as the membrane potential reaches the threshold potential, an instantaneous spike of arbitrary height above the threshold potential is issued to represent the occurrence of an action potential. Immediately following the spike, the membrane potential is held at the resting potential for the duration of the absolute refractory period, t_{ref}.

The leaky integrate-and-fire model is a point model because the whole cell is represented by one circuit. A limitation of this type of model is that there is no representation for position of input or delay of signal transmission across the neuron. As a result, interaction between inputs at various locations on the cell membrane cannot be simulated. However, due to the high speed of transmission along myelinated axons, the inherent inability of the integrate-and-fire models to simulate delays is of little consequence in many applications.

2.3 The Auditory System

The auditory system can be divided into two main sections: peripheral and central. The peripheral auditory system is the ear, from the pinna to the auditory nerve (shown in Figure 2.5). The auditory nerve provides the only input from the ears to the central auditory pathways of the brain. Hearing is a mechanoreceptive sense in that the ear responds to mechanical vibrations of air. The intensity of these mechanical vibrations of the air are quantified by a relative measure of the air pressure waves. This measure is known as the sound pressure level (SPL) that has units of dB and is given by Equation 2.2, where P is the impinging sound pressure in units of Pa and 2×10^{-5} Pa is the reference. The reference is associated with the minimum pressure fluctuation

that is detectible by the ear for vibrations of 1000 Hz.

$$L_p = 20 \log_{10} \left(\frac{P}{2 \times 10^{-5} \text{ Pa}} \right) \tag{2.2}$$

Sound pressure waves travelling through the air are amplified in the ear canal. These sound pressure waves cause the tympanic membrane to vibrate, which in turn moves the middle ear bones. These bones are attached to the oval window. The oval window then induces pressure waves through the fluid that fills the cochlea of the inner ear. The cochlea has three tubes that are coiled in parallel: the vestibular, central

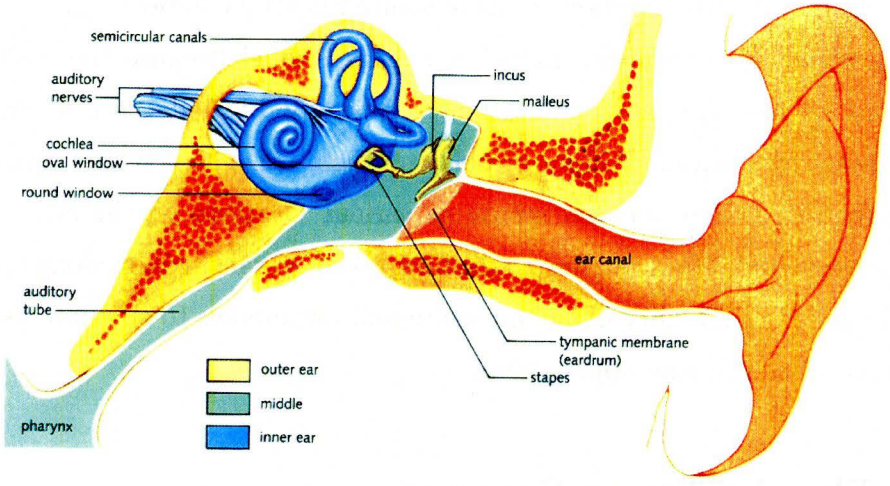


Figure 2.5: General anatomy of the human ear - the peripheral auditory system. (Reprinted with permission from Levine & Miller, 1994, Fig. 44.16, p.936.)

and tympanic canals. The central and tympanic canals are separated by the basilar membrane. As illustrated in Figure 2.6, the organ of corti that sits on the basilar membrane houses the inner and outer hair cells (IHCs and OHCs) that transduce the mechanical signal into neural activity. When the basilar membrane vibrates, it causes the stereocilia to brush against the tectorial membrane and thus bend. The bending of the stereocilia causes K^+ channels within the stereocilia to open. K^+ is then

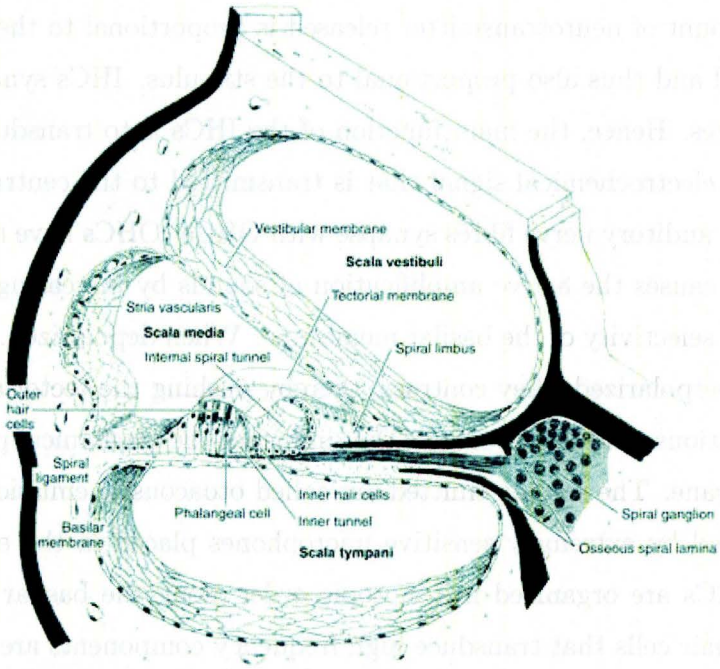


Figure 2.6: Illustration of a trans-section through an unwound cochlea, showing the anatomy of the organ of Corti. (Reprinted from Junqueira & Carneiro, 2002, Fig. 24-24, p.486.)

able to flood the hair cell via the stereocilia because the extracellular fluid potential is approximately +80 mV while the intracellular potential is approximately −50 mV. This large potential difference across the cell membrane drives diffusion of K^+ into the cell, thus depolarizing the cell. This depolarization is proportional to the intensity of the stimulus. Depolarization of the hair cell activates electrically gated ionic channels that allow K^+ to leave the cell and Ca^{2+} to enter the cell. The increased concentration of Ca^{2+} within the cell then triggers the release of neurotransmitters into the synaptic cleft. The amount of neurotransmitter released is proportional to the depolarization of the hair cell and thus also proportional to the stimulus. IHCs synapse with auditory nerve fibres. Hence, the main function of the IHCs is to transduce the acoustic signal into an electrochemical signal that is transmitted to the central auditory system. Very few auditory nerve fibres synapse with OHCs. OHCs have an electromotile response that causes the active amplification of sounds by increasing the sensitivity and frequency selectivity of the basilar membrane. When depolarized, OHCs shorten, and when hyperpolarized, they contract, thereby pushing the tectorial membrane to generate vibrations or low-level sounds that influence the mechanical properties of the basilar membrane. The sounds emitted are called otoacoustic emissions (OAEs) and can be recorded by extremely sensitive microphones placed in the ear canal. Both IHCs and OHCs are organized in tonotopic order along the basilar membrane. In other words, hair cells that transduce high frequency components are located toward the base (near the oval window) of the basilar membrane, while hair cells that transduce low frequency components are found toward the apex (deep within the coil) as illustrated in Figure 2.7. In mammals, this tonotopic ordering is established by the physical properties of the basilar membrane. The basilar membrane is narrower, thinner and stiffer at its base than towards its apex. This means that the resonant frequency of the membrane changes from the base to the apex. The basal end of the basilar membrane vibrates with the greatest amplitude at high frequencies. High frequency vibrations die out quickly as they travel beyond their resonant location toward the apex. Conversely, low frequency vibrations grow in amplitude as they travel

along the basilar membrane until they peak near the apex. Each of the auditory nerve fibres is associated with the same frequency components as those transduced by its convergent hair cells. The auditory nerve processes the information it receives from the hair cells in order to transmit the sensory information efficiently to the central auditory system in the brain.

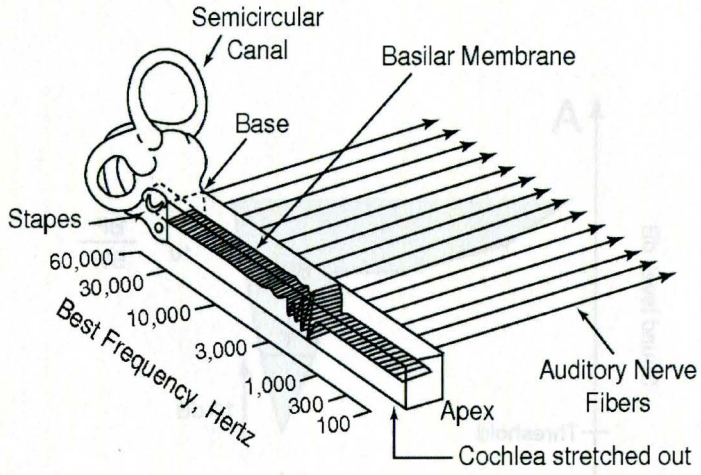


Figure 2.7: Illustration of how the tonotopic map is established by the basilar membrane. (Reprinted from Sachs et al., 2002, Fig.1B, p.158.)

The frequency response of a neuron is reflected in a tuning curve - a plot of threshold for tonal stimuli versus frequency of the tone, illustrated in Figure 2.8. A common measure used to quantify the sharpness of the tip of the tuning curve is the Q_{10} index, which is the ratio of the characteristic frequency to the bandwidth 10 dB above the lowest threshold. The frequency of maximal sensitivity (i.e., the tip of the tuning curve) defines the characteristic frequency (CF) of the neuron.

The main nuclei of the central auditory system are shown in Figure 2.9. The tonotopic map is maintained throughout the central auditory system all the way up to the auditory cortex. The auditory nerve transmits the neural signal into the brainstem where the auditory neurons synapse with the cochlear nucleus. Information from the

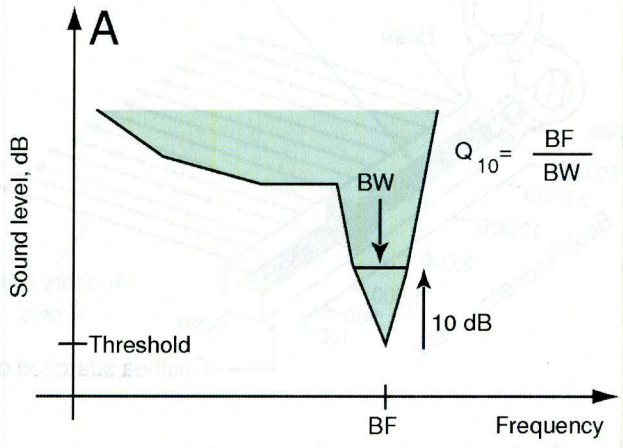


Figure 2.8: A tuning curve shows the sensitivity (represented as a threshold to tonal stimuli) of a neuron as a function of frequency. The tip of the tuning curve defines the characteristic or best frequency (CF=BF) of the neuron. The Q_{10} index quantifies the sharpness of the tuning curve. (Reprinted from Sachs et al., 2002, Fig.2A, p.159.)

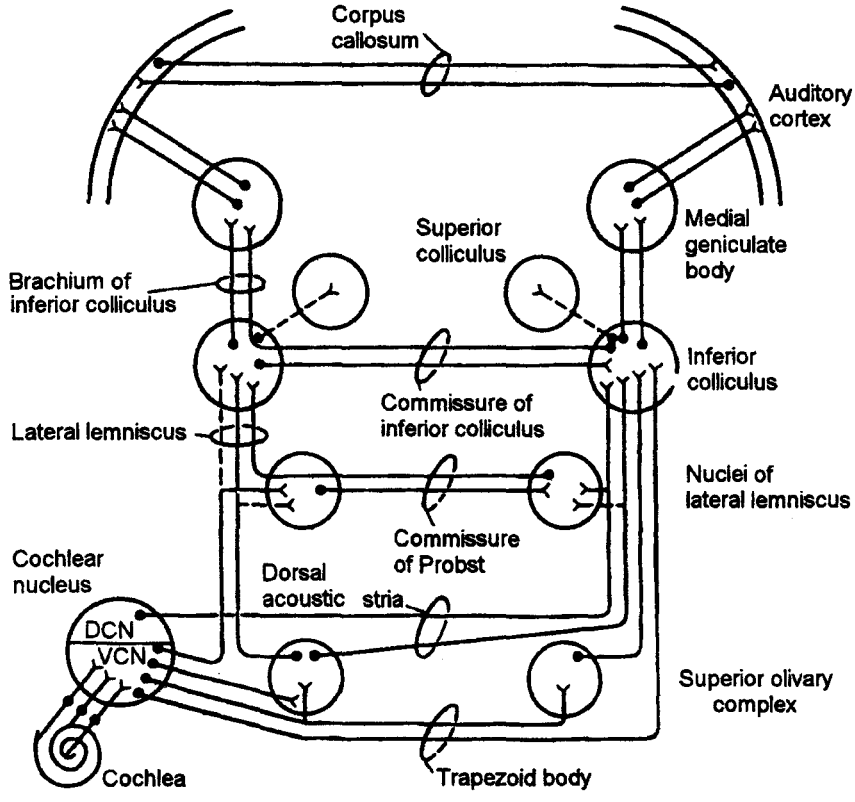


Figure 2.9: Diagram outlining the major central auditory pathways. (Reprinted from Ehret & Romand, 1997, Fig.4.2.)

auditory nerve is split into three streams: one going to the dorsal cochlear nucleus (DCN), another going to the anterior ventral cochlear nucleus (AVCN) and a third to the posterior ventral cochlear nucleus (PVCN). Auditory nerve fibres whose target cells are in the AVCN have very large axon terminals that form tight synapses. These synapses allow the preservation of timing of spikes to the order of microseconds. It is believed that this pathway conducts timing-critical, binaural processing tasks such as spatial localization. The AVCN provides input to the superior olivary complex (SOC). From the SOC, this binaural pathway projects to the inferior colliculus (IC) via the lateral lemniscus (LL). The PVCN is part of the intermediate brainstem pathway whose functionality we know very little about. The intermediate brainstem pathway starts at the PVCN, connects to the periolivary nuclei, then to the nuclei of the LL and finally to the IC. The monoaural contralateral pathway starts at the DCN. The complex neural circuitry through the DCN includes descending input and allows fine frequency discrimination. It is therefore possible that this pathway contributes to the perception of the quality of sound. Furthermore, the DCN integrates vestibular and somatosensory information into its representation of spectral cues from the auditory stimulus. It has therefore been hypothesized that the DCN contributes to sound source localization by correcting the spectral cues for the position and movement of the head and pinnae and for eliminating neural activity related to self-generated noise [Oertel & Young, 2004]. The DCN provides direct input to the IC via the LL.

All three auditory pathways, along with other sensory pathways, provide input to the IC where signals are integrated and routed. The IC's sensitivity to spectral changes could theoretically explain neural responses to spatio-temporal patterns that would be necessary in speech perception. The IC is known to play a key role in spatial localization and provides input to the thalamus. The primary nucleus responsible for auditory function in the thalamus is the medial geniculate body (MGB). The thalamus serves as a relay site to the cortex and is also capable of performing intensity and duration comparisons. The MGB projects to the auditory cortex where several tonotopic and nontonotopic maps exist in parallel, each performing a specific function

in the demodulation of sounds. The auditory cortex is part of the temporal lobe and is functionally divided into a number of parts. The tonotopic cortex has four divisions: the primary auditory cortex (AI), the anterior auditory field (AAF), the posterior auditory field, and the ventral posterior auditory field. The nontonotopic cortex includes the secondary auditory cortex (AII) that receives input from both AI and MGB. The auditory cortex is important in the discrimination of sound patterns and the interpretation of the meaning or significance of the sounds heard.

The signals from both ears are transmitted through central auditory pathways of both sides of the brainstem (i.e., bilateral transmission of both signals). However, transmission through the contralateral pathway is predominant. Cross-overs between sides occur at the trapezoid body, the commissure of probst, and the inferior collicular commissure. Due to the redundancy in transmission, the source of hearing impairment is rarely in the central auditory system.

Upon briefly reviewing the literature, it becomes quite obvious that relatively little is known about the central auditory pathways compared to the auditory periphery. Currently, a lot of experimental research is actively being conducted to improve our understanding of the central auditory system. However, due to the complexity and often inaccessibility of the brain, progress is slower in this area than for peripheral systems.

2.4 Speech Sounds

Speech sounds may be classified in a number of ways, one of which is by the type of source. If periodic puffs of air originate from the source, the speech sound is said to be voiced. Otherwise, the speech sound is called unvoiced. Subclasses of speech sounds include: fricatives, which are generated by a constriction in the oral tract that causes a noisy source; plosives, which have an impulsive source in the oral tract; and, whispers, which result from a partially closed glottis causing a noisy source. Examples of voiced sounds are vowels and examples of unvoiced sounds are /f/ (a

fricative), /h/ (a whisper), and /p/ (a plosive).

For voiced sounds, pitch is determined by the period of the puffs of air from the source. The inverse of the pitch period is the first harmonic and fundamental frequency, F_0. Peaks in the spectral envelope are called formants. These formant frequencies are the resonant frequencies of the vocal tract. Formant frequencies are important features of speech because they convey information about the vocal tract configuration over time and allow us to distinguish between voiced sounds.

There are a number of ways to analyze speech sounds. One method is to capture the time-varying frequency content of the sound in order to observe the spectral energy shifts. The spectrogram is just such a graphical display of the magnitude of the time-varying spectral characteristics. A spectrogram is the magnitude squared of the short-time Fourier transform of the sound waveform, $p(n)$, as given by:

$$S(\omega, \tau) = \left| \sum_{n=-\inf}^{\infty} w[n, \tau] p[n] e^{-j\omega n} \right|^2 \tag{2.3}$$

The short-time Fourier transform involves taking the Fourier transform of pieces of the waveform. The waveform is broken into pieces by a sliding window, $w[n, \tau]$, that is tapered at its ends (such as the Hamming window used in this project) to prevent distortion of the spectrum. A short window results in a wideband spectrogram characterized by fine temporal resolution and poor spectral resolution. In contrast, a narrowband spectrogram can be created with a longer window which results in poor temporal resolution and fine spectral resolution.

For more detailed information on speech production, classification and signal processing, the reader is referred to Quatieri's book [2002].

Chapter 3

Motivation

3.1 Lateral Inhibition in Sensory Systems

Knowledge of inhibitory processes in the nervous system has been documented since the 1600s: In 1662, Descartes described how antagonistic muscles facilitate limb movement[1] in his work entitled *De Homine*. Ernst Mach, an Austrian physicist, philosopher and psychologist, was probably the first to describe lateral inhibition in a sensory system, namely Mach bands in vision[2]. Mach first noticed this phenomena by chance in his experiments involving colour wheels (rotating discs with coloured sections) [Mach, 1865]. When a disc with various patterns of black and white sections painted on it is spun, the patterns blur or fuse to form annular rings of different shades of grey. Mach observed that for ring patterns that possessed a darker ring toward the outer edge of the disc and a bright centre with a gradient between these two areas is perceived to also have a brighter ring between the white centre and the gradient and a darker ring between the darker ring and the gradient. A stylized example of this phenomenon is shown in Figure 3.1. Mach found that the brighter

[1] To move a limb, one muscle must contract while the antagonistic muscle must relax.

[2] A comprehensive review of Mach's life and work on Mach bands can be found in a book written by Ratliff [1965]. This book contains the English translations of Mach's original papers.

and darker bands were perceived regardless of how the spatial distribution of illumination was generated (e.g., by fusion of moving patterns or shadow patterns). He also found that the bright and dark bands are not symmetrical. As is evident in Figure 3.2, the overshoot is usually much larger in magnitude than the undershoot, especially when there is a large stimulus gradient and a sudden step [Ratliff, 1965; von Békésy, 1967a]. Although Mach bands occur in all patterns of illumination, they

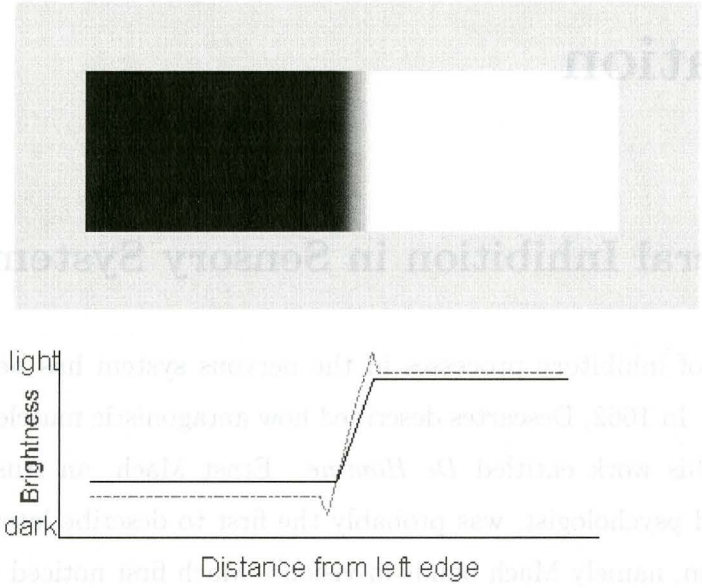


Figure 3.1: A black and white visual stimulus is shown in the upper half of this diagram. In the lower plot, the solid black line traces the actual brightness of the stimulus, while the dashed line indicates the perceived brightness. Note that the amplitudes of brightness are arbitrary and should not be compared. (Reprinted from Kaiser, 2002).

are more pronounced in simple, rectilinear patterns. Despite his inability to study neural events directly at the time, Mach was able to deduce from the results of his psychophysical experiments that the phenomenon must be attributed to reciprocal inhibition of neighbouring units - lateral inhibition - in the retina. Mach bands result from contrast enhancement achieved by lateral inhibition in the retina as long as the

inhibitory influence has some spatial extent. Lateral inhibition results in contrast enhancement because the inhibitory influence from neighbouring regions diminishes as excitation to these regions decreases. In vision, at the boundary between brightly and darkly illuminated regions, the processing unit on the bright side of the edge receives less inhibition from its neighbours that lie in the dark region than do other units that are completely surrounded by equally illuminated units in the bright field. Similarly, units on the dark side of the edge receive more inhibition from their neighbours in the bright region. These differential effects result in a peak and a valley in the response of the units on the bright and dark sides of the boundary, even though no such maximum or minimum occurs in the stimulus. As shown in Figure 3.2, if the intensity of illumination in the bright and dark regions is held constant as the width of the gradient between the two regions is varied, it is evident that the steeper the gradient, the higher and deeper the peak and valley. To explain the sharpness of images perceived, Mach suggested that a by-product of contrast enhancement is sharpening of edges [Mach, 1865]. Inhibition is a negative process in that information is suppressed as it is transmitted and any information lost cannot be recovered. However, by selectively discarding less significant information via inhibition, the more significant information that remains appears enhanced [Hartline, 1974].

It was not until three-quarters of a century later that electrophysiological measurement techniques became developed enough to discover inhibitory interaction in the retina[3] of the eye. It was largely the work of Hartline and colleagues [1949] that renewed interest in sensory inhibition, for which he won a Nobel Prize [Hartline, 1967]. Their work was focused on the optic nerve of the lateral eye of the *Limulus polyphemus* (the horseshoe crab) because it is much simpler than that of vertebrates. Using electrophysiological recordings of action potentials, they found that sensory elements in this eye, called ommatidia, inhibited each other. The magnitude of the inhibition was affected positively by: increased intensity of the inhibiting illumination, increased area (or number of ommatidia) covered by the inhibiting illumination,

[3]The retina lines the dorsal, interior wall of the eye and is responsible for transducing lightwaves into neural signals.

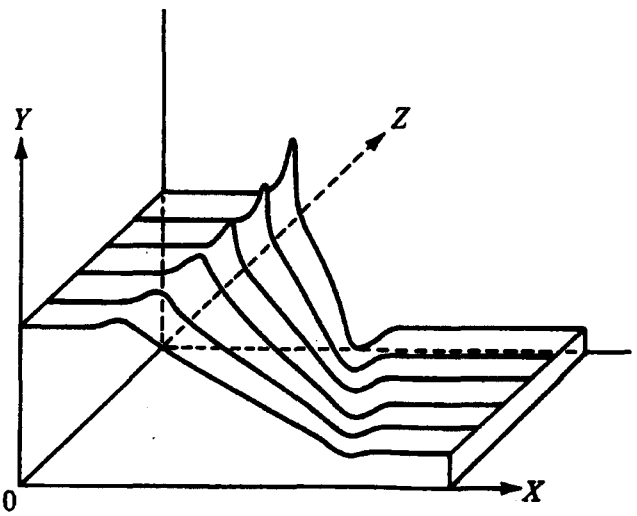


Figure 3.2: The effect of the boundary gradient width on the height and depth of Mach band peaks and valleys. The X-axis represents the spatial location of the neuron, the Y-axis represents the perceived illumination, and the Z-axis represents the trials with varying gradient widths. (Reprinted from Ratliff, 1965, Fig. 2.13, p.59.)

and by the proximity of the illuminated inhibiting ommatidia [Hartline et al., 1956]. They also found that the mechanism of the inhibition was a repolarization of the membrane potential that was depolarized by illumination [Tomita, 1958]. Hartline and Ratliff [1958] deduced from their experiments that inhibitory influences on a single ommatidium combine by simple addition according to:

$$r_p = e_p - \sum_{\substack{j=1 \\ j \neq p}}^{n} K_{pj}(r_j - r_{pj}^0) \tag{3.4}$$

where r_p is the response of an optic nerve fibre, measured by frequency of discharges in its axon [spikes/s]; e_p is the excitation that the ommatidium would receive in response to being the only receptor illuminated; K_{pj} is the inhibitory coefficient representing the inhibitory influence of neuron j on neuron p; r_j is the response of neuron j; and, r_{pj}^0 is the threshold frequency that must be exceeded before a receptor can exhibit inhibition. Similar results were found in the retina of other animals such as frogs [Barlow, 1953] and cats [Kuffler, 1953].

It is now known how Mach bands arise from lateral inhibitory mechanisms of the retina. The four classes of interneurons in the retina are: horizontal, bipolar, amacrine and ganglion cells. Photoreceptors (rods and cones) synapse with bipolar cells and horizontal cells, as illustrated in Figure 3.3. As their name suggests, horizontal cells transmit signals horizontally to bipolar cells. The bipolar cells transmit signals vertically from photoreceptors and horizontal cells to ganglion and amacrine cells. Amacrine cells transmit signals both horizontally and vertically between other interneurons. Finally, the ganglion cells transmit the pre-processed signal to the brain via the optic nerve. It should be noted that the ganglion cell is the first level in the visual system to transmit information by action potentials. All other neurons in the retina transmit signals by electrotonic conduction. The significance of electrotonic conduction is that it allows transmission of a graded signal (instead of a binary, all-or-nothing train of action potentials). Thus when photoreceptors hyperpolarize in response to light, the hyperpolarization is proportional to the intensity of impinging

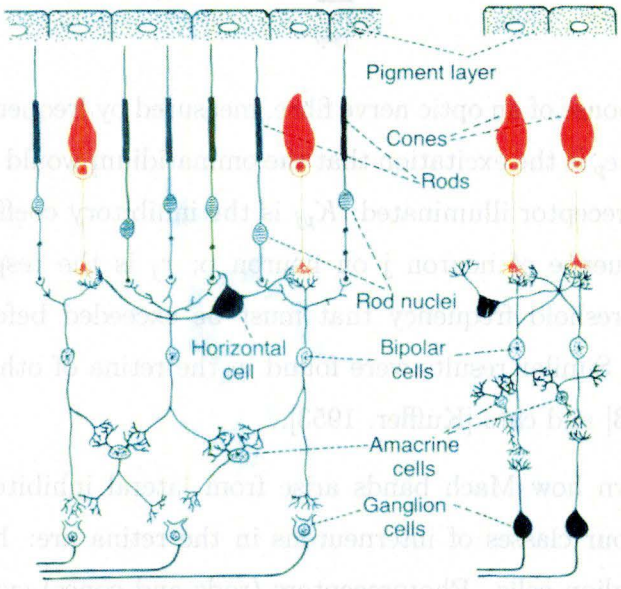


Figure 3.3: Neural organization of the retina. The peripheral area shown on the left side of the illustration contains both cones and rods. The foveal area only contains cones and is illustrated on right side. (Reprinted from Guyton & Hall, 2000, Fig. 50-11, p.586 with permission from Elsevier.)

light. This graded signal is then transmitted through the neural network of the retina to the ganglion cell. The ganglion cell then converts the graded signal into a neural code of action potentials for long-haul transmission to the brain. The interneurons of the retina are the only non-spiking neurons known to exist in the mammalian nervous system. For a comprehensive description of the anatomy and physiology of the eye, the reader is referred to standard texts such as those by Levine & Miller [1994]; Guyton & Hall [2000].

The complexity of the neural circuitry in the retina facilitates some signal processing in the peripheral visual system, including edge-detection or contrast enhancement. Horizontal cells are all inhibitory, i.e., when they are excited, they produce an inhibitory effect on postsynaptic neurons to which they connect. An example of an inhibitory neural arrangement is illustrated in Figure 3.4. When light falls on a photoreceptor, the photoreceptor excites a bipolar cell. The bipolar cell or photoreceptor then excites the horizontal cells connected to it. These horizontal cells inhibit the activity of neighbouring bipolar cells, which in turn lowers the firing rate of the ganglion cells associated with those neighbouring bipolar cells. These inhibitory horizontal connections provide the mechanism for lateral inhibition in vision, exactly as Mach had deduced.

Von Békésy, who is best known for winning the 1961 Nobel Prize in medicine for his discovery of the workings of the basilar membrane [von Békésy, 1961], also extensively studied lateral inhibition in sensory systems. He classified lateral inhibitory architectures into four types: simple, forward, backward and central. These architectures are shown in Figure 3.5. Von Békésy demonstrated by means of psychophysical experiments that Mach band-like contrast enhancement exists in sense organs other than the eye (e.g., skin sensation, hearing, and taste) [von Békésy, 1967b,a]. Others have also attempted to demonstrate the existence of lateral inhibition in hearing via psychophysical experiments (e.g. Houtgast [1972]), but the most convincing evidence is provided by Katsuki and Suga [1995]. Using electrophysiological techniques, Katsuki and colleagues [1958; 1959] measured frequency-tuning curves of single neurons

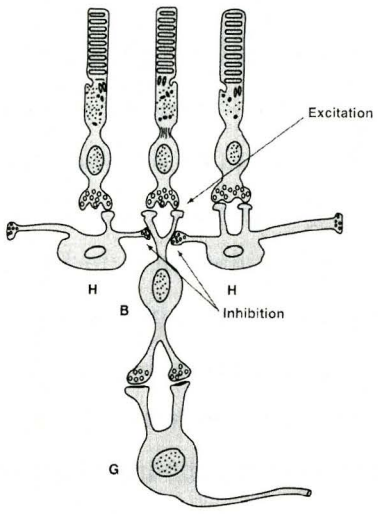


Figure 3.4: An illustration of a typical arrangement of cells in the peripheral area. Excitatory synapses exist between the rods and the horizontal cells (H), while inhibitory synapses connect horizontal cells and bipolar cells (B). The graded signal from the bipolar cell is then converted to action potentials transmitted by the ganglion cell (G). (Reprinted from Guyton & Hall, 2000, Fig. 50-14, p.589 with permission from Elsevier.)

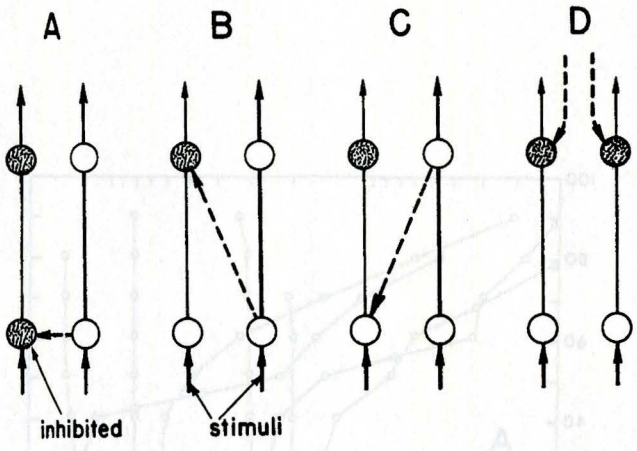


Figure 3.5: Illustration of four types of lateral inhibitory architectures. Configuration A is the simple type of lateral inhibition, where a neuron inhibits its neighbour directly. B the forward type, C the backward type and D the central type. (Reprinted from von Békésy, 1967b, Fig. 24, p.41.)

at different levels of the auditory system of cats. They found that the frequency-tuning curves became sharper toward higher centres of the central auditory system. Figure 3.6 shows a comparison of tuning curves at the level of the auditory nerve and the IC. It is quite obvious from the plots that the tuning curves of neurons from the auditory nerve have much wider tails (or skirts as Suga calls them), which means that the neurons are not sharply tuned for high intensity stimuli. The tuning curves of neurons in the IC are pencil-shaped and thus sharper for high intensity stimuli. Suga and his colleagues have called such a pencil-shaped frequency-tuning curve a 'level-tolerant' sharp frequency tuning curve. Some people have argued that these results do not show that tuning curves are sharpened since the Q_{10} values remain the same. In rebuttal, Suga suggests that the Q_{10} measure is a poor indication of tuning curve sharpness as it does not account for a neuron's response to high intensity stimuli. The sharpening of tuning curves is commensurate with observations that animals (including humans) are able to perform fine frequency discrimination even at high

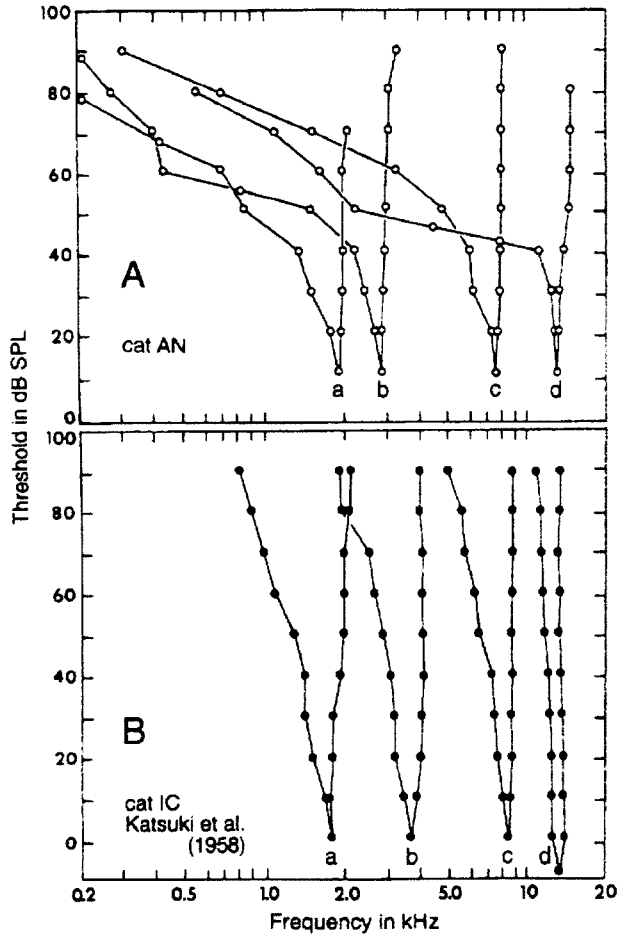


Figure 3.6: Frequency-tuning curves of four auditory nerve fibres (A) and four IC neurons (B) of cats. (Reprinted from Suga, 1995, Fig. 2, p.289.)

sound pressure levels (SPLs). Such fine frequency discrimination cannot be explained by the tuning curves at the level of the auditory nerve alone. The tuning curves were sharpest at the MGB, but were broad in the auditory cortex. Frequency discrimination must therefore be made level-tolerant by some subcortical mechanism. Like von Békésy, Katsuki, Suga and their coworkers believe that this subcortical mechanism is lateral inhibition [Katsuki et al., 1959; Suga, 1995]. Results from Yang and colleagues support the latter hypothesis. Yang and colleagues [1992] used bicuculine, an antagonist specific to the inhibitory neurotransmitter, GABA, to show that tuning curves in the IC of the mustache bat broaden with disinhibition. It may then be inferred that inhibition exists in the IC and plays a key role in the sharpening of tuning curves. However, the architecture of the inhibitory neural circuits that produces such sharpening of tuning curves is unknown. Lateral inhibition or on-CF inhibition[4] could theoretically produce these results. Neurophysiologic studies suggest the existence of on-CF inhibition at the level of the cochlear nucleus [Caspary et al., 1994] and IC [Palombi & Caspary, 1996] in chinchilla. The findings of Yang and colleagues are in agreement with Gerken's [1996] conclusion from his electrophysiological studies on cats that lateral inhibition occurs in the vicinity of the IC. This conclusion was based on results that showed hyper-responsiveness of the cochlear nucleus, SOC, IC and MBG in normal hearing cats in the presence of a sustained sound and in cats with induced hearing impairment. The higher the nucleus in the auditory system up to the IC, the greater the hyper-responsiveness.

Although the neural circuitry that gives rise to the experimental observations described above is not precisely known, there appears to be convincing evidence for the existence of LINs in the central auditory system. LIN models would provide a means of evaluating the plausibility of various network configurations and their functions in the auditory system.

[4]On-CF inhibition is said to occur when a neuron's inhibitory field is aligned with its target. On-CF inhibitory tuning curves have elevated tips and may have slightly broader tails compared to the excitatory tuning curves of the target cells.

3.2 Tinnitus

Tinnitus is an auditory disorder where a sound is perceived in the absence of a corresponding acoustic stimulus and is often referred to as ringing of the ears. For a more comprehensive overview of tinnitus the reader is referred to the recent review written by Lockwood and colleagues [2002]. According to the Tinnitus Association of Canada, 360,000 Canadians suffer from tinnitus. A broad range of sounds may be perceived by tinnitus sufferers. These sounds range from tones to white noise. Tinnitus sufferers can often modulate the phantom sounds somatically [Levine et al., 2003; Cacace, 2003], such as by changing their direction of gaze or clenching their jaw [Pinchoff et al., 1998]. Tinnitus can be acute or chronic in nature. Acute tinnitus has sudden onset and is a temporary condition. Chronic tinnitus usually has a gradual onset and is a permanent condition. In some cases, tinnitus can be very debilitating, yet treatment is severely limited due to a lack of understanding of the underlying mechanisms. However, considering that our understanding of tinnitus is still in its early stages and the increase in public recognition the condition has received [Geary, 1998], there is hope that with greater knowledge of the underlying mechanisms of tinnitus, effective treatments will arise. The work reported in this thesis aims to contribute to this end.

Animal behavioural models have been developed to show that tinnitus can be induced and tested for in human and animal subjects. Kaltenbach [2000] provides an excellent review of these models. Some of these models involve training animal subjects to indicate by some identifiable behaviour that they hear a conditioning sound. When this behaviour is then exhibited in the absence of a stimulus, the animal is presumed to be experiencing tinnitus. Using this general method, and surveying clinical experience, it has been found that tinnitus can be induced in subjects (both animal and human) by chemical and mechanical agents. Such chemical agents include high doses of the commonly used drugs quinine or salicylate. Quinine is a drug that was originally made from the bark of the Cinchona tree and is best known for its use in the treatment and prevention of malaria. Various forms of salicylate (e.g.,

sodium salicylate, acetylsalicyclic acid a.k.a. Aspirin), are used to relieve pain, reduce fever and inflammation, and improve circulation. Tinnitus induced by these chemical agents is usually of the acute kind that subsides when the drug therapy is ceased, and is perceived as a high pitched tone or narrow-band noise. However, extremely high doses of these drugs are known to cause permanent hearing loss and chronic tinnitus. The mechanical agent is exposure to intense sound, i.e., very loud sound for an extended duration. Exposure to explosive sounds (loud, but short-lived sounds) are likely to only cause acute tinnitus. However, intense sounds, such as those present in noisy industrial work environments, usually result in chronic tinnitus. Obviously, there is a greater focus on furthering our understanding of the chronic tinnitus than the acute form due to its persistence and therefore greater impact on patients' lives. Even though methods of inducing tinnitus have been found, it is not well understood how these agents cause tinnitus.

A number of neurophysiologic models of tinnitus have been hypothesized to suggest the anatomical source of tinnitus, which in fact remains unknown. Since the tinnitus in many patients who underwent surgical transection of their auditory nerves remained after surgery and, in some cases, was worsened by the surgery, both peripheral and central mechanisms are implicated [Kaltenbach, 2000]. Peripheral mechanisms that have been proposed include IHC, OHC, and auditory nerve dysfunction. IHC dysfunction may result from trauma in the cochlea that causes an increase in the ion conductance across the cell membrane. The increased flux of ions would trigger an increase in the spontaneous release of neurotransmitters, which in turn causes overactivation of the postsynaptic neuron [Zenner & Ernst, 1995]. The increased activity could cause a spurious perception of sound. A fraction of the population produce measurable OAEs in the absence of acoustic stimulus, which is known as spontaneous OAEs (SOAEs). Some believe that these SOAEs could be a source of tinnitus [Penner, 1992; Plinkert et al., 1990]. Since OAEs are associated with OHC function, this mechanism of tinnitus suggests OHC dysfunction. Auditory nerve function can be hampered if the cranial nerves become microvascularly compressed. Moller [1995]

hypothesized that compression of the auditory nerve causes a decrease in excitatory input to the central auditory system, which results in decreased inhibition of certain neurons. When neurons experience less inhibition, they become hyperactive. Tinnitus patients who experienced tinnitus for a shorter period of time benefit more from vascular decompression than those who had tinnitus for a long time. This clinical data is consistent with Moller's hypothesis, for long term tinnitus might cause plastic reorganization in the central neural circuits. Such reorganization might lead to a chronic state of hyperactivity that could be perceived as tinnitus since sound is believed to be encoded in the auditory system as a higher rate of neural firing.

Central mechanisms of tinnitus that have been hypothesized include hyperactivity, changes in spike patterns, and reorganization of the tonotopic map. Hyperactivity, an increase in spontaneous activity, of some neurons may be perceived as a phantom sound. Experiments involving chemical induction of tinnitus in animal models provide supporting evidence for this hypothesis. Moderate doses of salicylate were found to increase spontaneous discharge rates in AII neurons tuned to high frequencies, but decrease rates in AI and AAF [Eggermont & Kenmochi, 1998]. Increased spontaneous activity in neurons with high characteristic frequencies was also found at the level of the IC in rats [Chen & Jastreboff, 1995]. Hence, the hyperactivity may originate from subcortical areas. On the other hand, experiments involving intense sound to induce tinnitus provided evidence at the auditory nerve level that the impairment caused a decrease in activity. This led to the notion that tinnitus might not be the result of hyperactivity itself, but rather contrast enhancement at the edge between regions of normal and low activity due to lateral inhibition [Eggermont, 2003; Gerken, 1996; Kral & Majernik, 1996]. However, Kaltenbach has demonstrated that intense sound exposure induces hyperactivity in the DCN of rodents [Kaltenbach, 2000]. A correlation between electrophysiologic recordings of activity in the DCN and behavioural evidence of tinnitus is strong. Kaltenbach therefore suggests that DCN activity may serve as an index of the severity of the tinnitus. Findings of hyperactivity in the IC of different animals due to sound exposure is commensurate with hyperactivity in the

DCN as the activity in the DCN is relayed to the IC. Studies have also revealed an increase in activity in the auditory cortex and IC of tinnitus sufferers who "turn on" their tinnitus by somatosensory means [Cacace, 2003]. Again, this hyperactivity may originate in subcortical structures and be relayed up to the cortex.

Moller [1984; 1995] proposed that another hypothesis for a central mechanism of tinnitus is an alteration of spike train patterns. Normally, interspike intervals (ISIs) of spontaneous discharges are highly irregular. In the presence of a stimulus, ISIs become highly regular, thereby creating periodicities in spike trains. These periodicities are usually phase-locked with the stimulus for frequencies lower than 6 kHz. The spike trains of a group of neurons thus appears to have bursts of spikes while the waveform cycles are in their positive phase (i.e., neurons are synchronously phase-locked to the stimulus). Phase-locking is believed to be important in encoding the lower frequency components of sounds in normal hearing. When tinnitus is induced by either salicylate or noise exposure, bursting that looks like synchrony to a stimulus has been observed in the auditory nerve in the absence of a real sound stimulus [Liberman & Kiang, 1978]. Administration of salicylate or quinine causes increased synchronization of spontaneous activity between neurons in AI [Ochi & Eggermont, 1996, 1997]. Spurious synchrony might be interpreted by the brain as being phase-locked to a nonexistent sound stimulus, thus causing tinnitus. The advantage of this hypothesis is that it could explain the occurrence of tinnitus in patients without hearing loss.

Changes in the tonotopic map have also been hypothesized to have some relation to tinnitus. As described in Section 2.3, the IC and auditory cortex are organized tonotopically such that each frequency is represented by a strip of neurons. It has been found that when the input to a strip is degraded by some form of hearing loss, adjacent strips expand and take over the function of deafferented neurons. This change indicates that neural plasticity (an adaptive process of change that facilitates learning by means of altering synaptic efficacy) occurs. Cortical reorganization results in a tonotopic map that has wider than normal representations of the frequencies at the edges of the damaged regions [Rajan et al., 1993; Robertson & Irvine, 1989; Irvine

& Rajan, 1997]. Phantom limb pain experienced by amputees is often thought of as a somatosensory equivalent to the tinnitus of the hearing impaired. Indeed, the cortical reorganization of both conditions is very similar. Cortical reorganization has been found to occur in patients who suffer from tinnitus [Muhlnickel et al., 1998]. Seki and Eggermont have shown that cortical regions that have undergone reorganization due to hearing loss exhibit higher rates of spontaneous activity that might be perceived as tinnitus [Seki & Eggermont, 2003].

The peripheral and central mechanisms of tinnitus that have been proposed are not mutually exclusive. In fact, it is very unlikely that any one of these mechanisms alone could generate the percept of tinnitus. For example, from the evidence described above, one could hypothesize that some form of dysfunction in the peripheral auditory system could cause a decrease in spontaneous activity in the auditory nerve. Lateral inhibition might then reconcile the seemingly disparate findings between the decreased spontaneous activity in the auditory nerve and increased spontaneous activity in central neurons with CFs that are presumed to be associated with the induced tinnitus. Lateral inhibition would act on the edge between normal and impaired signals from the auditory nerve to create a spurious peak in the brainstem or midbrain. This peak of activity might be transmitted up specific neural circuits towards the auditory cortex as a higher rate of spikes. Furthermore, these same spikes may occur in bursts that are synchronized between a group of neurons. Since the cortex loses normal input to regions associated with the impaired CFs, plasticity might facilitate the expansion of neighbouring normal regions of input into the region of impairment. However, the spurious peak is located exactly in this neighbouring region that expands into the region of impairment. The perceived spectrum of tinnitus might therefore look like a spatially expanded version of the spurious peak, which Norena and colleagues [2002] have shown to be the case through psychoacoustic tests. Tinnitus in patients with seemingly normal audiograms and auditory brain stem responses (ABR) might be explained by hearing losses in narrow frequency notches that would not show up on tests using the traditional audiometric frequencies. The associated notch in the auditory

nerve response could be processed as described above to generate tinnitus.

Considering a broader scope of effect, Cacace [2003] proposed that the biological basis of tinnitus is not in hearing alone, but rather involves a multimodal neural network that spans other sensory and sensorimotor systems as well as those of attention, cognition and emotion. As Kaltenbach [2000] aptly suggests, the cause of tinnitus must be separated from the epiphenomena in order to reconcile the hypothesized mechanisms with each other. This thesis aims to contribute to the solution of the latter problem.

3.3 Existing Models

Real biological neural networks are very difficult to explore directly. This is particularly true of the central nervous system due to the complexity of the neural connections. Hence, computational models are useful in the analysis of the behaviour of hypothetical neural networks, the results of which might complement empirical studies. Previously developed models that are relevant to this thesis will be described in this section.

3.3.1 The Bruce et al. Model of the Auditory Periphery

The Bruce et al. model of the auditory periphery is a phenomenological, computational model that has been shown to correlate well with physiological data, especially of cats [Bruce et al., 2003]. This model takes an impinging sound pressure wave on the tympanic membrane as input and calculates the response of an auditory nerve fibre of a specific characteristic frequency. The model consists of a series of blocks that represent the middle ear, the tuning of the basilar membrane (influenced by the function of the OHC), the synapse between the IHC and the auditory nerve, and the auditory nerve itself. Refer to [Bruce et al., 2003; Zhang et al., 2001] for details of these blocks. The auditory nerve response is characterized by its simulated spike times.

Despite the high level of development of the Bruce et al. model of the auditory periphery, it alone cannot explain the mechanism of tinnitus, for physiological experiments have shown that tinnitus involves the central auditory system. Furthermore, the auditory periphery only accounts for the first steps of speech feature extraction: mainly the physiological origin of the tonotopic map and synchrony of neural activity to lower harmonics. A model of the central auditory system would therefore be desirable as an extension to models of the ear.

3.3.2 The Kral and Majernik Model of Central Auditory Processing

Kral and Majernik are supporters of the idea that lateral inhibition plays a significant role in the central mechanism of tinnitus [Kral & Majernik, 1996]. They agreed that an edge created by a decrease of activity in a subpopulation of neurons acted on by lateral inhibition would give rise to a spurious peak that could be perceived as tinnitus. However, they also suggested that spontaneous activity of auditory nerve fibres processed by some form of LIN might account for tinnitus that nearly everyone perceives when in an acoustically shielded chamber. The spontaneous activity is Poisson-like noise that is masked by ambient acoustic noise that we usually find ourselves in. Since most ambient acoustic noise spans a wide frequency band, a LIN should be able to suppress it.

The model that Kral and Majernik implemented consisted of nine layers of one thousand processing elements connected in a nonrecurrent, forward-type LIN architecture [Kral & Majernik, 1996]. The diagram in Figure 3.7 illustrates the pattern of connectivity in their model.

A neuron only receives excitatory input from the neuron directly below it and inhibitory input from M neurons to either side from the layer below, weighted uniformly. The activity of the i^{th} neuron-like processing element in the $(k+1)^{\text{th}}$ layer is

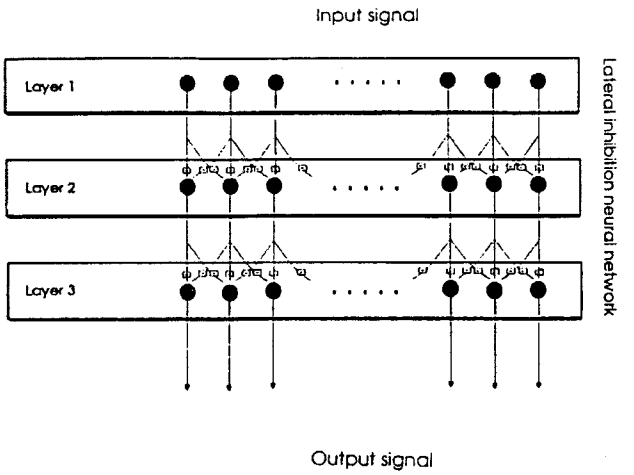


Figure 3.7: Illustration of the architecture of Kral and Majernik's non-recursive LIN model. (Reprinted with permission from Kral & Majernik, 1996, Fig. 1, p.113.)

given by the following equation:

$$\Psi_{k+1}[i] = f\left(\sum_{i-M}^{i+M} w[x]\Psi_k[x] - \Theta\right) \tag{3.5}$$

where $w[x]$ are the weights (with $w[i]$ being positive for the excitatory input and all other weights being negative for lateral inhibitory input), Θ is the neuron's threshold value and f is the linear threshold input-output function of a processing element shown in Figure 3.8.

A single excitatory input was chosen because it was known that convergent excitatory input opposes the sharpening effect of lateral inhibition. Furthermore, they found no physiological evidence of such convergence. The processing elements are non-spiking, but Kral and Majernik claim that the input and output curves represent an analogy to rate-place profiles (spike rate versus characteristic frequencies). The input was considered representative of auditory nerve activity.

Not only was this model able to sharpen an input excitation curve with a single

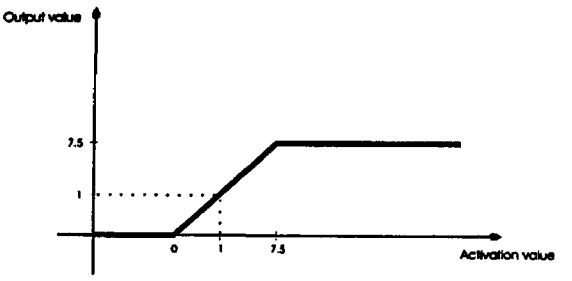


Figure 3.8: Input-Output Function of Kral and Majernik's Processing Element. (Reprinted with permission from Kral & Majernik, 1996, Fig. 2, p.114.)

maximum, it was able to separate the two maxima of a smooth M-shaped profile of input activity, i.e., perform contrast enhancement. They also showed that this model suppresses white noise, as expected. To simulate sensorineural hearing loss, Kral and Majernik used a step-like fall out of activity in a band of neurons amongst others that are noisy. Again, the contrast enhancement feature of lateral inhibition appears at the borders of the defect so long as the defect is wider than M. The spurious peaks at the edges could be perceived as tinnitus. Spontaneous activity from the auditory nerve is a Poisson-like process which has an uneven probability density function of ISI across the neurons. When the response to this type of input was simulated, parts of the excitation curve were amplified such that there appeared to be a number of narrow peaks. These spurious peaks could explain tinnitus in normal hearing subjects.

This model is unrealistic in a number of aspects. Firstly, the processing elements do not capture key features of neural processing, especially the timing information contained in trains of action potentials. The temporal aspect of input integration is therefore not addressed. Secondly, the nonrecursive nature of the network architecture requires there to be many layers in order for the lateral inhibitory effects to materialize. Thirdly, although the inhibition is proportional to the neural activity, the use of uniform inhibitory weights is unrealistic. It may be inferred from single-unit electrophysiological studies such as those performed by Ramachandran

and colleagues [1999] that various inhibitory architectures exist. A pattern where inhibitory effects decrease with distance is clear [von Békésy, 1967b; Hartline, 1974].

3.3.3 The Gerken Model of Central Auditory Processing

Gerken's model also relates tinnitus to lateral inhibition in the central auditory system. It relies upon the processing of unevenness in the auditory nerve response by central mechanisms of lateral inhibition to generate a spurious peak. The LIN model that he uses is inspired by the LIN in the visual system and has a single layer of two hundred neural units. As shown in Figure 3.9, a neural unit (e.g. N2) of a given characteristic frequency receives excitatory input from a lower level in the auditory system. This neural unit also has lateral projections that inhibit the activity of its neighbours, but not itself. Because the collateral projections are of the simple-type architecture, this is a recurrent network. The strength of lateral inhibition to each of the neighbouring neurons is governed by the number of inhibitory connections. The pattern and range of lateral inhibition used is shown in the insert of Figure 3.9. The neural units are non-spiking entities that process mean input firing rates. (Details of the computations were not provided.) A baseline input represents spontaneous activity from the auditory nerve. Cochlear injury is simulated by lowering the input firing rate, while acoustic stimulation is simulated by increasing the input firing rate.

Gerken's results for two types of input are shown in Figure 3.10. The normal input consists of uniform spontaneous activity and a response to tonal stimulus from the auditory nerve. The impaired input reflects peripheral damage at the higher characteristic frequencies with lower spontaneous activity in the impaired region. The tonal stimulus is within the impaired region of hearing. The output for the normal case shows amplification of the response to the tone, flanked by valleys. The output for the impaired case shows contrast enhancement at the edge of the impaired region and a larger response to the tone than in the normal case. The tonal response is stronger in the impaired case than in the normal case because lateral inhibition is reduced in the impaired region since the strength of the inhibition is proportional to

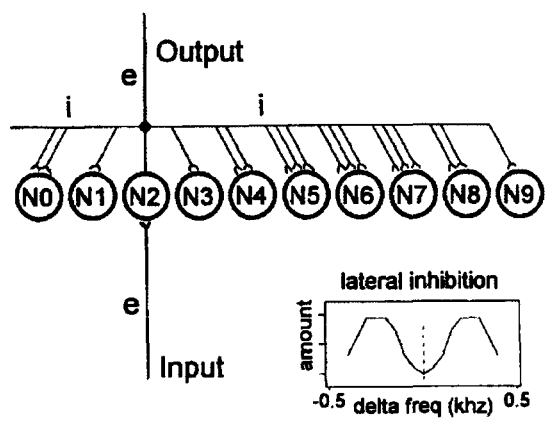


Figure 3.9: Illustration of a single neural unit's connections in Gerken's recursive LIN model. All neurons in the model are actually connected as N2 is here to create a web of mutual inhibition. Input and output connections are excitatory, while lateral connections are inhibitory. The inlay shows the amount and extent of the lateral inhibition used in the simulations. (Reprinted from Gerken, 1996, Fig. 3, p.79.)

the activity of the neural units. This would not be the case *in vivo* since the auditory nerve response to the tonal stimulus would be lower from an impaired ear. Again, the spurious peak at the edge of the impaired region was postulated to give rise to tinnitus.

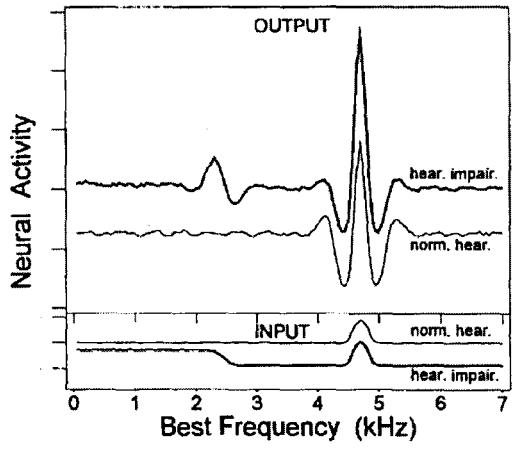


Figure 3.10: The response of Gerken's model to normal and impaired input. The input from the auditory nerve contains spontaneous activity and a tonal stimulus at 4.7 kHz. (Reprinted from Gerken, 1996, Fig. 4, p.79.)

Gerken's model is more realistic than Kral and Majernik's model because of its spatial distribution of inhibitory strength and its recurrent architecture. However, the model is still unrealistic because the neural units are too simple. Temporal aspects are lost due to the non-spiking nature of the neural units.

3.3.4 The Shamma Model of Central Auditory Processing

To emphasize the importance of both time and space in neural codes, Shamma [1985] distinguished between three classes of auditory processing algorithms: spatial, temporal and intermediate. Spatial processing schemes are based on the mean spike rate of fibres. The spatial distribution of this average measure may be interpreted as a

spectral estimate of the stimulus. However, this scheme performs spectral extraction inadequately, for it is known that temporal cues play a key role in the encoding of lower frequencies (i.e., less than 5 kHz). Temporal processing schemes extract spectral estimates from the periodicities in the spike trains. An example of an intermediate scheme that Shamma gives is that of Young and Sachs [1979]. The spatial aspect of the Young and Sachs scheme is that it retains a tonotopic representation, but it uses the synchronous rate as the information carrier, which is temporal in nature.

To evaluate the hypothesis that a LIN constitutes the central auditory processor, Shamma [1985] presents results for a single-layer, nonrecurrent LIN model of non-spiking neurons, though he also describes the implementation of a recurrent network. The dynamics of the neural units or processing elements that Shamma uses in his model is comparable to the linear portion of the leaky integrate-and-fire model. Figure 3.11 shows the model of a neuron that Shamma uses, from which the governing equation can be derived using basic circuit theory:

$$\tau \frac{dY(t)}{dt} + Y(t) = E(t) \tag{3.6}$$

where $Y(t)$ represents the membrane potential, $E(t)$ represents the input to the neuron in volts, and τ is the membrane time constant. Shamma then applies a non-linear transformation G to $Y(t)$ in order to approximate the instantaneous firing rate of the neuron, $Z(t) = G(Y(t))$. G incorporates the effects of saturation of the firing rate (due to the refractory period) and threshold potential in a sigmoid type function.

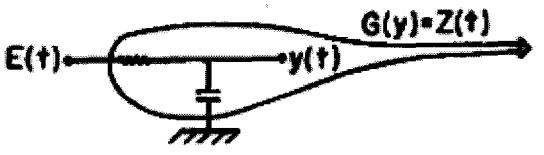


Figure 3.11: Shamma's model of a neuron. (Reprinted with permission from Shamma, 1985, Fig. 1(a), p.1624.)

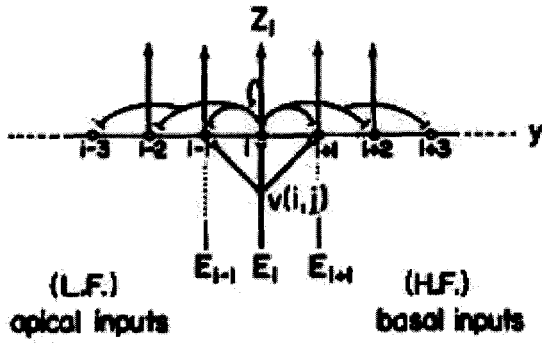
Figure 3.12 illustrates the single-layer recurrent and nonrecurrent LINs that Shamma explored. The recurrent LIN architecture is similar to Gerken's LIN except that Shamma's includes the possibility of divergence of input across a spatial range of neurons. The divergence of the excitatory and inhibitory input is governed by the weighting matrices $\boldsymbol{v(i,j)}$ and $\boldsymbol{w(i,j)}$ respectively, where each weight represents the effectiveness of the input or output at i on neuron j. Combining the excitatory and inhibitory inputs with Equation 3.6 results in Equation 3.7 for processing elements in the recurrent LIN and Equations 3.8 and 3.9 for the nonrecurrent LIN. Thus, the instantaneous firing rate of neuron i is given by Z_i for the recurrent LIN and X_i for the nonrecurrent LIN.

$$\tau\frac{dY_i}{dt} + Y_i = \sum_j v(i,j)E_j - \sum_j w(i,j)G(Y_j), \qquad i = 1,\ldots,N \tag{3.7}$$

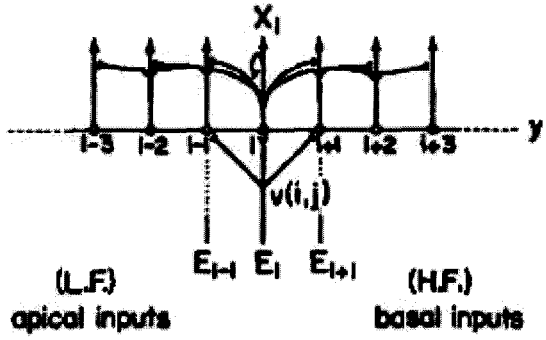
$$\tau\frac{dY_i}{dt} + Y_i = \sum_j v(i,j)E_j, \qquad i = 1,\ldots,N \tag{3.8}$$

$$X_i = G(Y_i) - \sum_j w(i,j)G(y_j), \qquad i = 1,\ldots,N \tag{3.9}$$

Shamma found that the single-layer, nonrecurrent LIN with direct excitatory input (i.e., no divergence of excitatory input) is primarily a spatial pattern processor. This LIN is capable of sharpening spatial edges of both synthetic and real auditory nerve activity. As a result, Shamma's LIN was able to enhance the lower harmonics and formants of speech that evoked auditory nerve activity in a cat. Shamma's model is more realistic than those of Gerken, and Kral and Majernik as it incorporates some membrane potential dynamics. However, like the other existing models, the neural units that Shamma uses do not spike. Inhibition is thus modulated directly by the membrane potentials of neighbouring neurons, in a fashion akin to the mechanism of lateral inhibition in the retina. Since non-spiking neurons have not been found in the auditory system, Shamma's model is unrealistic.



(a) Recurrent LIN



(b) Nonrecurrent LIN

Figure 3.12: Schematic diagrams of Shamma's recurrent and nonrecurrent LINs. Each hollow circle represents a neural unit. Presynaptic terminals represented by a 'Y'-shaped symbol are excitatory, while 'T'-shaped synaptic terminals are inhibitory. (L.F.= low frequency, H.F.=high frequency.) (Reprinted with permission from Shamma, 1985, Fig. 1(b,c), p.1624.)

Chapter 4

The Model

A biologically inspired, computational model of a neural network was constructed to investigate the plausibility of spiking neurons achieving contrast enhancement and speech feature extraction. The lateral-inhibitory-network (LIN) model considered in this thesis consists of a single, uniform layer of leaky integrate-and-fire neurons with synaptic input. The synaptic input is represented by excitatory and inhibitory conductances that are modulated by spike trains. The input spikes are generated by a Bernoulli approximation of a Poisson process [Edwards & Wakefield, 1990] to represent spontaneous neural activity, and in other simulations by the Bruce et al. model [2003] of the auditory periphery in response to speech-like input. The output is a series of spike trains from the leaky integrate-and-fire neurons. All simulations were implemented in MATLAB [1].

4.1 Complexity of the Model

Choosing the appropriate level of detail to incorporate into a model is nontrivial. As Protopapas and colleagues [1999] pointed out, there exists a continuum of models from realistic to what they call demonstration models. Demonstration models are

[1] A programming language for technical computing released by The MathWorks, Inc., Natick, MA, USA

highly theoretical and "are primarily intended to provide support for a particular preexisting theory or functional point of view" [Protopapas et al., 1999, p.462], while realistic models are more closely tied with empirical data. There is, at the same time, too much and too little information about the neurons in the central auditory system to create an ideal model. For example, there exists an abundance of electrophysiological data that correlates with animal behaviour, but it is often not well understood where this signal originated. Since relatively little is known about the details of the interconnection of neurons in the central auditory system, it would be impossible to incorporate a high level of detail specific to these neurons in any model, which eliminates the possibility of designing a highly realistic model for this project. So long as there exists a limitation on computing resources, the tradeoff between the possible size of the network to be modelled and the amount of detail incorporated will always exist. This tradeoff further justifies forgoing the use of a highly realistic model of neural units in this network. Instead, the leaky integrate-and-fire model of a neuron, that Arbib [2002] describes as being the simplest realistic model, has been selected, for the goal here is to capture the general behaviour of a fairly large network of neurons; namely, to show how contrast enhancement might be achieved. Expenditure of computational resources or complexity is thus balanced between modelling interactions of many neurons and modelling biological details of the neurons themselves.

4.2 Neural Network Architecture

The architecture of the LIN model is based on the recurrent LINs proposed by Gerken [1996] and Shamma [1985]. The types of possible connections for a neuron in this LIN are shown in Figure 4.1. The structure of the whole LIN would consist of each neuron having the same lateral connections, thus forming a web of interconnections. Neurons of the single layer of n leaky integrate-and-fire neurons only receive a small number of convergent excitatory projections from the layers below. Each neuron has

lateral projections that inhibit the activity of its neighbours (i.e., reciprocal inhibition) in the same layer, making this model a recurrent network.

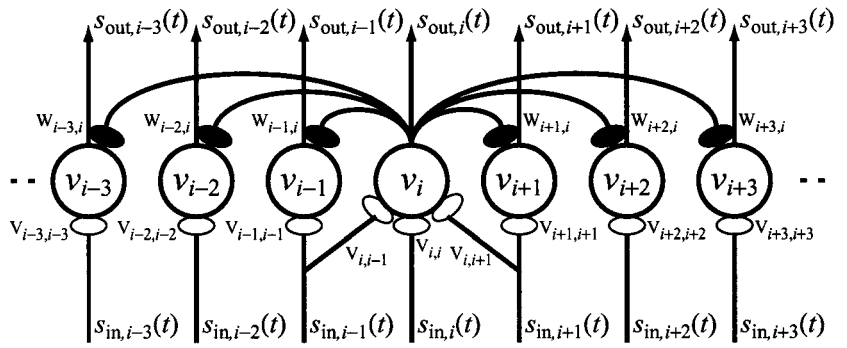


Figure 4.1: In this schematic diagram of the lateral inhibitory network the large circles represent neural bodies whose behaviour is simulated with the integrate-and-fire model described in section 2.2. Hollow ellipses represent excitatory synapses while filled ellipses represent inhibitory synapses. The membrane potential, input and output spike trains, and excitatory and inhibitory weights are represented by v_i, $s_{in,i}$, $s_{out,i}$, $V_{i,j}$ and $W_{i,j}$ respectively. (Reprinted from Bruce et al., 2003, Fig.1, p.360.)

The convergent excitatory input is scaled by $V_{i,j}$ that follows a Gaussian function; the input from a source of the same characteristic frequency is most heavily weighted (i.e., centre of Gaussian function), while the attenuation of lateral input increases with spatial distance from the centre frequency. Figure 4.2 shows a sample of a possible distribution of excitatory synaptic weights. Inhibitory synaptic weights, $W_{i,j}$, are governed by two Gaussian functions of length k; one on each side of the centre frequency (see Figure 4.3 for an example; c.f. [Gerken, 1996, Fig.3]). This implies that the neuron at the centre frequency is not inhibited by its own activity. Each weight, $V_{i,j}$ or $W_{i,j}$ thereby specifies the effectiveness of the j^{th} synapse on the i^{th} neuron. These weights are constant in time such that plasticity is not simulated.

An arbitrary scaling factor, a, is multiplied with the inhibitory synaptic weights to scale the inhibition relative to the excitation. The model parameter a is varied in

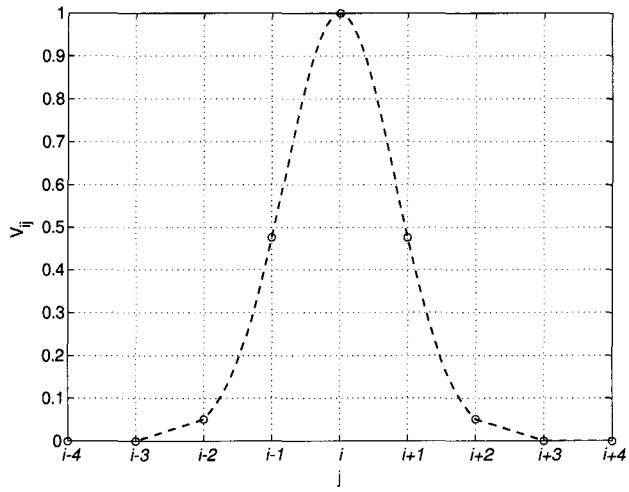


Figure 4.2: Excitatory synaptic weights, $V_{i,j}$ relating the i^{th} input spike train's ($s_{in,i}$) effect on the j^{th} neuron.

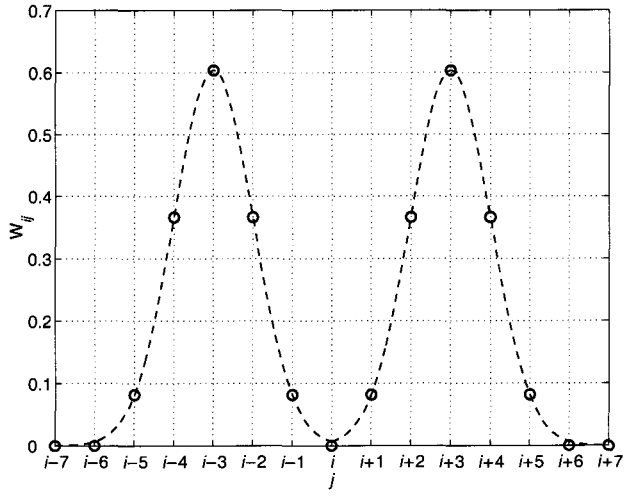


Figure 4.3: Inhibitory synaptic weights, $W_{i,j}$ relating the effect of the i^{th} neuron's output on the j^{th} neuron. (Reprinted with permission from Bruce et al., 2003, Fig.3, p.361.)

this investigation.

4.3 The Spiking LIN Model

For the purposes of this investigation, the network elements used were leaky integrate-and-fire models of a neuron with excitatory and inhibitory synaptic input, and an absolute and a relative refractory period. Note that the resting potential, v_{rest}, was set to 0 volts for this project such that the membrane potential, $v_m(t)$, is always a relative measure from the resting potential. To accentuate the latter point, the relative membrane potential of the i^{th} neuron will be indicated by $v_i(t)$.

Lumped excitatory and inhibitory synaptic conductances were used instead of direct current injections to make the model more realistic. The use of synaptic conductances also facilitates the evaluation of the effect of shunting inhibition versus hyperpolarizing inhibition. Synapses can be represented in the electric circuit of the leaky integrate-and-fire model as a variable conductance in series with a synaptic reversal potential or battery [Koch, 1999]. As shown in Figure 4.4, the excitatory and inhibitory synaptic input of this lateral-inhibitory-network are modelled in this fashion.

The excitatory synaptic reversal potential, E_E, is set to a positive value greater than the threshold potential, v_{th}, so that the synaptic current will flow inward to generate a depolarization toward or past v_{th}. In contrast, the inhibitory synaptic reversal potential, E_I, is set equal to or less than the resting potential so that the synaptic current will flow outward to pull the membrane potential downward towards the resting potential or a hyperpolarization, respectively.

To obtain the time course of the excitatory and inhibitory variable conductances, $G_{E,i}(t)$ and $G_{I,i}(t)$, spike trains, $s_{in,i}(t)$ and $s_{out,i}(t)$, are convolved with unitary conductances, $g_E(t)$ and $g_I(t)$. The time course of the unitary excitatory and inhibitory

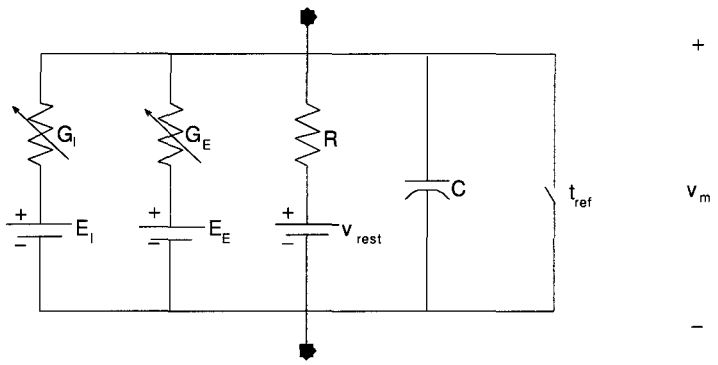


Figure 4.4: Circuit Diagram of the leaky integrate-and-fire model with synaptic input. R, C, and v_{rest} represent the resistance, membrane capacitance and resting potential. $G_E(t)$, $G_I(t)$, E_E and E_I are the excitatory and inhibitory conductances and reversal potentials respectively. The switch labelled t_{ref} is used to implement the refractory period. v_m is the membrane potential and output.

conductances are alpha functions of the form:

$$G(t) = c \cdot \left(\frac{\alpha}{10\tau}\right)^2 te^{-\frac{\alpha t}{10\tau}} \tag{4.10}$$

where $G(t)$ is the conductance in Siemens, t is time in seconds, α is a constant that governs the width of the function, and $\tau(= R \cdot C)$ is the membrane time constant in seconds. c is a scaling factor to bring the alpha function within a realistic order of magnitude for synaptic conductances. Examples of the time courses of unitary excitatory and inhibitory conductances are shown in Figure 4.5. Alpha functions, coined by Jack and colleagues[2], are often used to approximate observed trajectories of EPSPs and IPSPs in order to avoid significantly increasing the complexity of the phenomenological model with details such as individual channel dynamics. Here, alpha functions are used to represent conductance trajectories in order to acquire realistically shaped EPSPs and IPSPs. The squared constants in Equation 4.10 ensure

[2]See [Jack et al., 1975, Eqn.3.50]. This book provides a comprehensive review of the electrical properties of excitable cells.

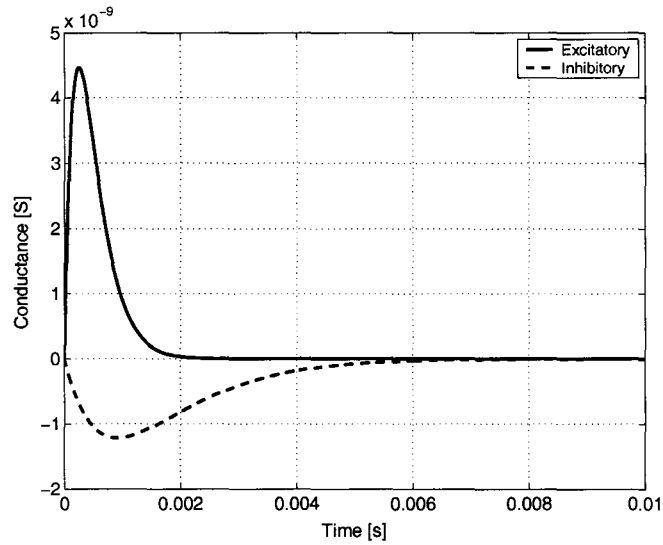


Figure 4.5: The unitary excitatory and inhibitory conductances shown here have alpha values of 11 and 3 respectively, a membrane time constant of 2.75 ms and a scaling factor of 0.30365 nS. Note that the negative of the unitary inhibitory conductance is plotted.

that the area under the curve remains constant as α and τ are varied. The area of 0.01 is arbitrary, but the function is further scaled such that the resulting conductances are on the order of nanosiemens to maintain biological plausibility. The model incorporates both excitatory and inhibitory conductances whose time courses are governed by different alpha values. The alpha value of the inhibitory conductance should be much smaller than that of the excitatory conductance so that the time course of the inhibitory conductance is much slower than that of the the excitatory conductance as illustrated in Figure 4.5. The excitatory and inhibitory alpha values are parameters that are varied in this investigation. Incorporating these conductances into the leaky integrate-and-fire model described in Section 2.2 gives the following equation for the linear operation of the model:

$$\frac{d\boldsymbol{v}(t)}{dt} = \boldsymbol{V} \cdot \frac{\boldsymbol{g_E}(t)(E_E - \boldsymbol{v}(t))}{C} + \boldsymbol{W} \cdot \frac{\boldsymbol{g_I}(t)(E_I - \boldsymbol{v}(t))}{C} - \frac{\boldsymbol{v}(t)}{\tau} \tag{4.11}$$

Note the similarity between Equation 4.11 and the one deduced by Hartline and Ratliff from their experiments on the horseshoe crab (Equation 3.4). Both equations involve subtracting the sum of weighted activity from neighbouring neurons from the excitatory input to the neuron in order to calculate the activity of a neuron. This similarity provides an indication of the biological plausibility of the configuration under consideration.

For the non-linear operation of the model, t_{ref} is held as a constant such that adaptation is not simulated. Theoretically, the threshold potential, $v_{th,i}$, should be set to infinity while the membrane potential is allowed to vary even though it is not permitted to spike. However, during preliminary tests it was found that the membrane potential could wander unrealistically high above the threshold potential during the short duration of the refractory period. To avoid this problem and maintain simplicity of the model, the membrane potential is held at rest as described above. After the absolute refractory period, a relative refractory period is implemented that allows the threshold potential to decay exponentially from infinity (approximated by some value above the spike value, here 5 V) to the fixed threshold:

$$v_{th,i} = 5e^{-3.5\frac{t_{th}}{2\text{ ms}}} \tag{4.12}$$

where t_{th} represents the time since the end of the absolute refractory period, for at most 2 ms. Simultaneously, the model re-enters the linear operation mode such that the membrane potential begins to vary again.

4.3.1 Input

The input to the lateral-inhibitory-network model is an array of spike trains; one spike train for each neuron. The input generated is stored in an $n \times t$ matrix, $\boldsymbol{s}_{in}$, where n is the number of neurons and t is the number of Δ-second time bins (i.e., $t = \frac{\text{duration of impinging sound wave}}{\Delta}$). For this project, input spike trains, that are representative of auditory nerve activity, were generated in one of two ways. The first method is to send an impinging sound wave (with units of Pascals) into the Bruce et al. model of the auditory periphery that was briefly described in Section 3.3.1. The output from the Bruce et al. model (which would be the input to the LIN model) is an array of spike trains. However, the time steps from the Bruce et al. model is 2 µs, which is much smaller than is necessary in the LIN model (Δ =0.02 ms). To resolve this discrepancy, every ten time bins from the Bruce et al. model are summed to make longer time steps. In the case that multiple spikes are found in any bin, the spike count for that bin is reset to one. Since the probability of there being a spike in any given time bin is so low, there is very little chance of the latter case being found. The Bruce et al. model can be configured to respond as a normal ear (i.e., without impairment), or an impaired ear. The impairment is implemented by setting the scaling constants, $C_{OHC}[i]$ and $C_{IHC}[i]$, of the outer and inner hair cells respectively to a value between zero and one, where one represents normal function and zero a complete loss. Figure 4.6 shows the only set of scaling factors used for an impaired case in this project. The C_{OHC}s are designed to fit the 50^{th} percentile of the Q_{10} data from Miller and colleagues [1997] while the C_{IHC}s account for the

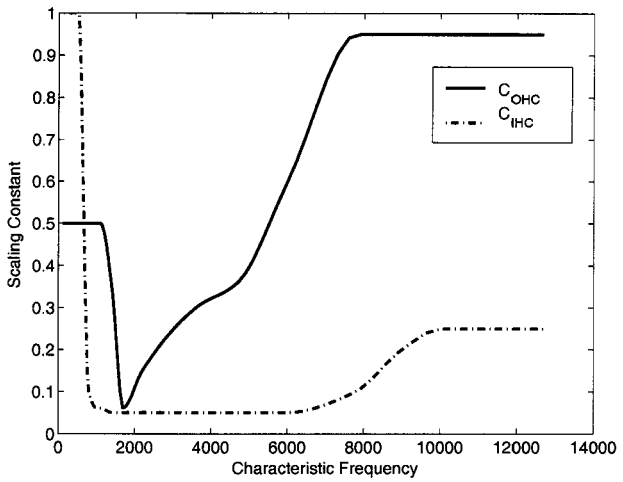


Figure 4.6: Scaling constants for inner and outer hair cell impairment.

minimum threshold shift in the auditory nerve population that cannot be explained by the outer hair cell impairment, as explained in [Bruce et al., 2003].

The second method was to use the Bernoulli approximation of a Poisson process (described in Appendix A) to simulate spontaneous activity from the auditory nerve when there is no impinging sound wave, i.e., in the perfectly quiet environment of a sound-proof, anechoic chamber. This spontaneous activity was generated by multiplying a firing rate by the time step and comparing this probability to uniformly distributed values between zero and one generated by MATLAB's pseudo-random number generator. If the randomly generated number is less than or equal to the specified probability, a spike is generated. All time bins are filled this way. The emulated spontaneous activity is not uniform across all of the neurons. Neurons with lower characteristic frequencies receive a higher rate of spontaneous activity, while those at the upper end of the tonotopic map receive a lower rate of spontaneous activity. The use of this type of input is intended to grossly represent sensorineural hearing loss in the high frequencies and to evaluate the plausibility of spiking neurons achieving contrast enhancement. The results of this method would be identical to

what would have been acquired from the Bruce et al. model if a constant sound wave of 0 Pa were sent through. Since the Bruce et al. model would blindly apply the numerous filters to the null input, the second method was used to save computation time. High frequency, tone-like input could be generated in a similar fashion by creating a spatially narrow, Gaussian-shaped peak of a higher rate of spontaneous activity.

4.4 The Non-Spiking LIN model

To evaluate the effects of the spiking behaviour of neural units on contrast enhancement, a non-spiking LIN model was created for comparison. The only difference between this model and the spiking LIN model is the neural units. For the non-spiking LIN model, Lapique's model (a simple R-C circuit) described in Section 2.2 was used. The operation of this model is completely linear (i.e., no threshold potentials, action potentials or refractory periods) and conductance-based synapses were not used. Instead, $s_{in,i}(t)$ was convolved with the same alpha function as $G(t)$, but was scaled to represent an excitatory current injection, $i_{EPSC,i}(t)$. The inhibition is modulated directly by the membrane potential of neighbouring neurons. The governing equation of this model is:

$$\tau \cdot \frac{d\boldsymbol{v}(t)}{dt} = \boldsymbol{V} \cdot \boldsymbol{i}_{EPSC}(t) - \boldsymbol{W} \cdot \boldsymbol{v}(t) - \boldsymbol{v}(t) \tag{4.13}$$

$\boldsymbol{V}$ and $\boldsymbol{W}$ obey the same weighting functions as in the spiking model, except that $V_{i,j}$ now have units of Ohms.

4.5 Simulations

All simulations were coded and run in MATLAB version 6.5, release 13, on machines described in Table 4.3. A sample set of scripts used can be found in Appendix C. Notice that for some of the longer simulations, the job was broken into parts in order to be run in parallel on the GRID server. The data collected then had

to be concatenated for analysis.

Type:	**Personal Computer**	**Server**	**64-Node Grid Blade Server**
Operating System:	Windows 2000	Red Hat Linux 8.0 (Psyche), kernel 2.4.18-18.8.0smp	Red Hat Linux 7.3 (Valhalla), kernel 2.4.20-20.7.0smp
Processor(s):	Intel Pentium IV, 2.00GHz	Intel Pentium IV Xeon, dual 2.20GHz	64 x Intel Pentium IV Xeon, dual 2.40GHz
RAM:	1GB	1GB	1GB

Table 4.3: Computing Resources

For a network of n neurons, the excitatory and inhibitory synaptic weights, $V_{i,j}$ and $W_{i,j}$, are stored in $n \times n$ matrices, $\boldsymbol{V}$ and $\boldsymbol{W}$, whose functions are described in Section 4.2. However, the neurons at and near the extremities of the network (i.e., $i \leq k$ or $i \geq n-k$) must have fewer inhibitory synapses and possibly fewer excitatory synapses coming from the outer side. To mitigate the confounding edge effect that would result, the weights of the synapses that do exist are increased such that the sum of the excitation and inhibition are equal to those of the middle neurons. The linear equation (either Equation 4.11 or 4.13) is then solved using a 4^{th}-order Runge-Kutta algorithm (described in Appendix B.2) with a fixed time step Δ. Preliminary versions of the model employed Euler's method (described in Appendix B.1) to solve the linear equations, and intracellular current injections (excitatory and inhibitory postsynaptic currents) instead of synaptic conductances. The column vector $\boldsymbol{v}(t)$ of length n that results from all of these methods is the relative membrane potential or output at a given time. In the case of the spiking model, if any $v_i(t)$ is greater than $v_{th,i}(t)$, then that $v_i(t)$ is set to an arbitrarily high value (0.15 V unless stated otherwise) and $s_{out,i}(t)$ is set to 1. $\boldsymbol{v}_{th,i}(t)$ is also updated to include the relative refractory period described in Section 2.2. If a spike occurred within the last t_{ref} seconds, then $v_i(t)$ is set to zero, as is $s_{out,i}(t)$, regardless of the potential computed using Equation 4.11.

Although minimal effort was spent on maximizing the computational efficiency of the code, a considerable amount of time was spent on making the code comprehensible and on minimizing computation time. The rationale for this investment of effort was that the simulations employing the Bruce et al. model were lengthy. Since the Bruce et al. model was the source of a significant portion of the computations and it was beyond the scope of this project to modify that model, there appeared to be little advantage in optimizing the code for the LIN model. Note that a modified version of the Bruce et al. model is now available and should run more quickly. Furthermore, the Bruce et al. model is currently being re-written in C-language to maximize computational efficiency while adding functional features to make the model more realistic[3]. However, as mentioned in Section 4.3.1, use of the Bruce et al. model was avoided for investigating spontaneous-type input in an effort to improve efficiency. Computation time was reduced significantly by compiling the computationally heavy portions of the code in C-language (namely earandlin.m) and by running some simulations in parallel on a 64-node computing grid.

[3]Development of the Bruce et al. model is a project of the Auditory Engineering Laboratory of McMaster Unversity.

Chapter 5

Results

The effect of various neural and network parameters on the network's mean output discharge rate of each neuron was evaluated first with the spontaneous activity type input (for which the method of generation is described in Section 4.3.1). The model's processing of speech-driven activity was then examined using the mean spike rate measure as well as a measure of synchrony to the formant frequencies.

5.1 Spontaneous Activity as Input

Unless indicated otherwise, the high and low rates of spontaneous activity are 200 and 20 spikes/s respectively, and the drop from the high to the low rates occurs with one intermediate step, i.e., the ramp width is across three neurons with input rates of 200, 110, 20 spikes/s. The higher rate represents normal spontaneous activity in the auditory nerve, while the lower rate models spontaneous activity in a region of peripheral impairment.

First a comparison between the use of Euler's method and the 4^{th}-order Runge-Kutta algorithm shows that there is a slight difference in the membrane potentials computed. Figure 5.1 shows the time course of the membrane potentials computed using the two methods on the same input. The membrane potential traced by the dashed, red line was computed using Euler's method, where the solid, blue line was

computed using the 4^{th}-order Runge-Kutta algorithm. The difference in results warrants the use of the more computationally intensive, but more accurate 4^{th}-order Runge-Kutta algorithm henceforth.

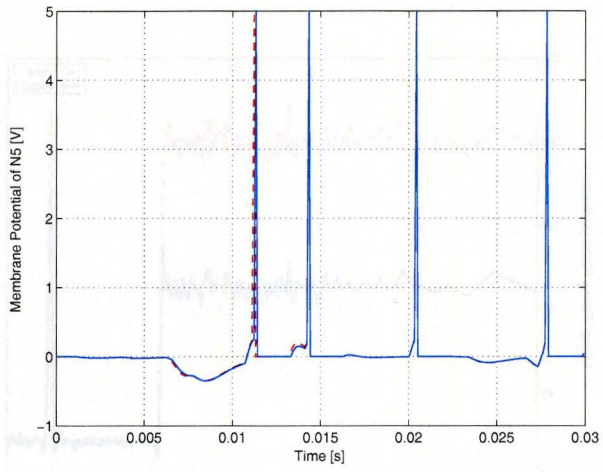


Figure 5.1: Membrane potential of the fifth neuron in the preliminary 100-neuron LIN computed using Euler's method (dashed, red line) and the 4^{th}-order Runge-Kutta algorithm (solid, blue line). Inhibitory connections span 6 neurons to either side of any given neuron, i.e., a total span of 13 neurons. Other parameter values were set as follows: $t_{step} = 0.1$ ms, $\tau = 1$ ms, $a = 2$, $v_{th} = 0.25$ V, $t_{ref} = 2$ ms, $\alpha_{EPSP} = 5$, $\alpha_{IPSP} = 2$.

In Figures 5.2 and 5.3, it can be seen that the neural activity in regions receiving the high rate of spontaneous activity is decreased, in large part due to the lateral inhibition. Figure 5.2 also shows that neural activity is decreased in the region receiving the low rate of spontaneous activity, but this decrease is minute because the inhibition here (proportional to the neural activity of neighbouring neurons) is weak. Hence, the LIN suppresses regions of fairly uniform activity. There is clearly a peak and valley at the edge between high and low rates of spontaneous input. The tone-like peak of input activity in the impaired region is sharpened by the LIN and flanked by

two small valleys. It is therefore evident that a LIN model of spiking neurons is capable of achieving the contrast enhancement that is characteristic of lateral inhibition. Figure 5.3 shows that without large, rapid spatial changes in the input excitation pattern, no spurious peaks or valleys are generated by the LIN processing.

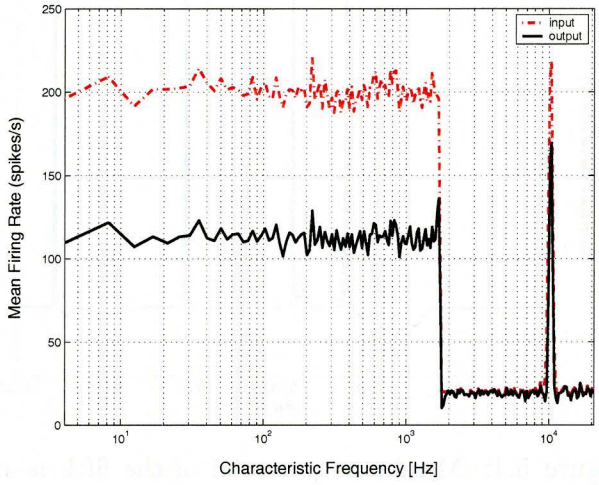


Figure 5.2: Contrast enhancement achieved by the preliminary LIN model of spiking neurons. Neural and network parameters are the same as those for Figure 5.1.

Contrast enhancement is very difficult to achieve with shunting inhibition. v_{th} must be set such that one input spike is just able to cause an output spike without lateral inhibition. Such a configuration was used to generate the data shown in Figure 5.4(a) that clearly shows small peaks and valleys. If E_I is set to hyperpolarize, as in Figure 5.4(b), then contrast enhancement is much stronger. Figure 5.5 shows that contrast enhancement cannot be achieved with shunting inhibition if v_{th} is set any lower than described above. The mean output spike rates produced with hyperpolarizing inhibition is more realistic than those generated with shunting inhibition when compared to the control group results of Koch & Grothe [2003] and Basta & Vater [2003].

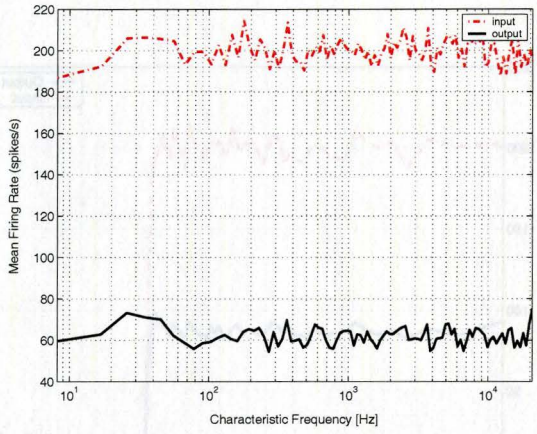
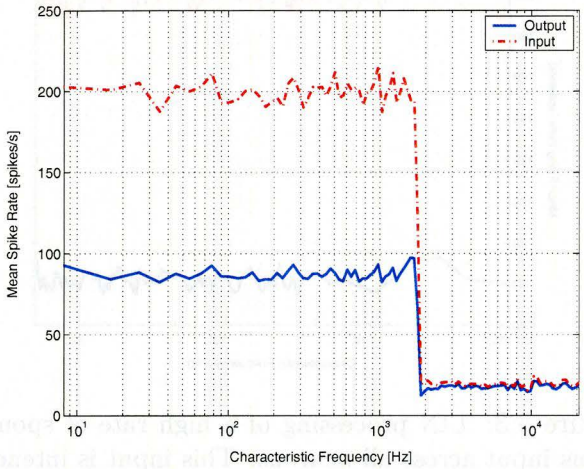
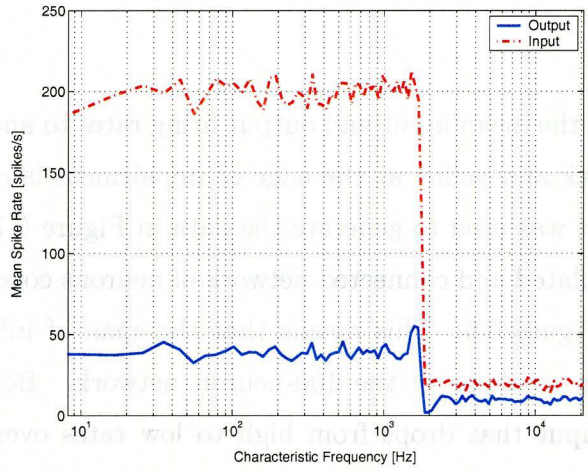


Figure 5.3: LIN processing of a high rate of spontaneous input across all neurons. This input is intended to simulate the normal auditory periphery response to silence. The same model as in Figure 5.2 was used.

Figure 5.6 shows the mean input and output firing rates to and from a 100-neuron LIN. A spurious peak and valley at the edge of impairment is clearly evident. The 500-neuron LIN that was used to generate the data in Figure 5.7 may be considered a more densely populated and connected network of neurons covering the same range of CFs as that of Figure 5.6. This means that the span of inhibitory connections cover the same CFs as those of the 100-neuron network. Both networks receive spontaneous-type input that drops from high to low rates over three neurons. In comparing Figures 5.6 and 5.7, it can be seen that the peak and valley in the 500-neuron network have greater magnitudes than those of the 100-neuron network. In the 500-neuron network the range of CFs represented by three neurons is narrower than in the 100-neuron network. If the ramp width of the otherwise identical input presented to the 500-neuron network is made to span the same range of CFs as three neurons in the 100-neuron network, Figure 5.8 results. From the latter result, it can be seen that increasing the ramp width of spontaneous input decreases the magnitude of the spurious peak and valley slightly, but not to the level of the 100-neuron network.



(a) Shunting Inhibition: $E_I = 0$ V



(b) Hyperpolarizing Inhibition: $E_I = -0.02$ V

Figure 5.4: Shunting inhibition. A 100-neuron network with lateral inhibitory connections spanning 6 neurons to either side of each neuron. Other parameter values were set as follows: $t_{step} = 0.02$ ms for 5 s, $\tau = 1.5$ ms, $C = 8$ pF, $a = 32$, $v_{th} = 0.0231$ V, $t_{ref} = 2$ ms, $\alpha_{GE} = 11$, $\alpha_{GI} = 0.5$, and $E_E = 0.1$ V.

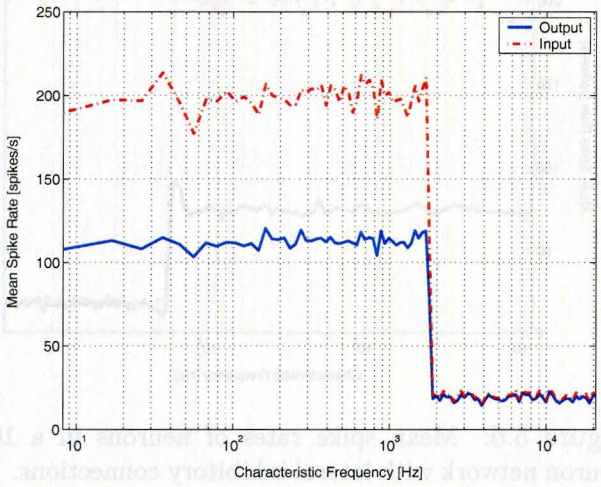


Figure 5.5: Shunting inhibition with the same parameters as in Figure 5.4(a) except $v_{th} = 0.0200$ V. Contrast enhancement is lost.

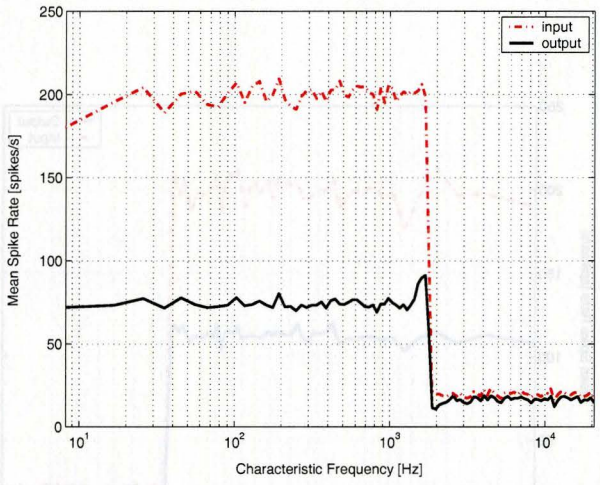


Figure 5.6: Mean spike rates of neurons in a 100-neuron network with lateral inhibitory connections. Inhibitory connections span 6 neurons to either side of any given neuron, i.e., a total span of 13 neurons. Other parameter values were set as follows: $t_{step} = 0.1$ ms for 7 s, $\tau = 1$ ms, $C = 7.5$ pF, $a = 2$, $v_{th} = 0.02$ V, $t_{ref} = 2$ ms, $\alpha_{GE} = 10$, $\alpha_{GI} = 3$, $E_E = 0.1$ V and $E_I = -0.02$ V.

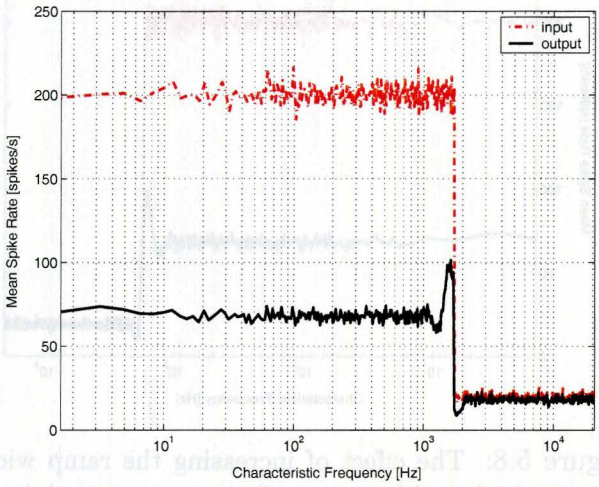


Figure 5.7: Average spike rates of neurons in a 500-neuron network with lateral inhibitory connections. Inhibitory connections span 32 neurons to either side of any given neuron, i.e., a total span of 65 neurons covering the same CF range as the case with only 100 neurons. Other parameter values are the same as for Figure 5.6

The effect of the span of inhibitory connections on contrast enhancement was quantified by a series of simulations. A measure of contrast enhancement to be defined is the index of edge-enhancement given by:

$$\text{Index}_{\text{EE}} = \frac{r_{\text{peak}} - r_{\text{highsp}}}{r_{\text{highsp}} \cdot (r_{\text{highsp}} - r_{\text{lowsp}})} \tag{5.14}$$

where r_{peak} is the peak value of the spurious peak, r_{highsp} is the average of the normal region, and r_{lowsp} is the average of the impaired region in the mean firing rate of the LIN. This measure provides an indication of the significance of the peak with respect to the magnitude of the edge and the high spontaneous input rate. Figure 5.9 shows the results of these simulations. The straight line fit highlights the upward trend in the significance of the peak as the lateral spread of inhibition is increased.

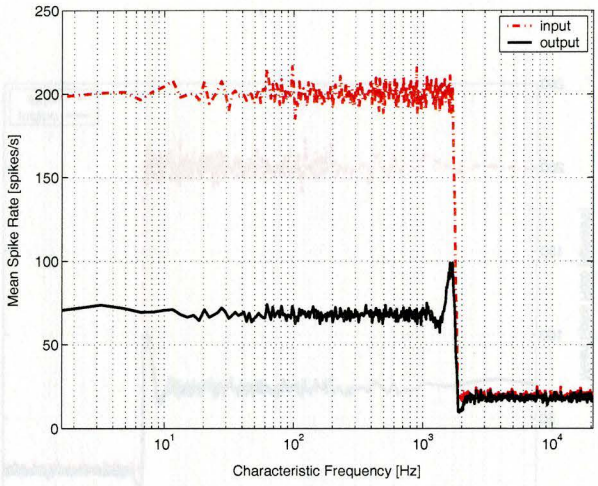


Figure 5.8: The effect of increasing the ramp width between high and low rates of spontaneous activity to match the CF span of the ramp used in the 100-neuron LIN of Figure 5.6. Neural and network parameters are the same as in Figure 5.7.

The effect of lateral excitation is shown in Figure 5.10. Although the spike rates are increased compared to the rates from the level of the auditory nerve, the spurious peak is significantly decreased compared to equivalent results from the network without lateral excitation shown in Figure 5.7. However, the magnitude of the valley is increased. This occurs largely because the the neurons on the normal side of the edge receive less excitation compared to its neighbours further into the normal region combined with an increase in inhibition. For neurons on the impaired side of the edge, the increased inhibition due to the increase in neural activity of neurons in the impaired region dominates over the increase in excitatory input, thus causing a deeper valley.

To further explore the effect of the ramp width from high to low rates of spontaneous input, the width was varied systematically and compared with the the non-spiking model. Input to the spiking and non-spiking models is made identical for all simulations in this set, except at the ramps, by seeding the pseudo-random number

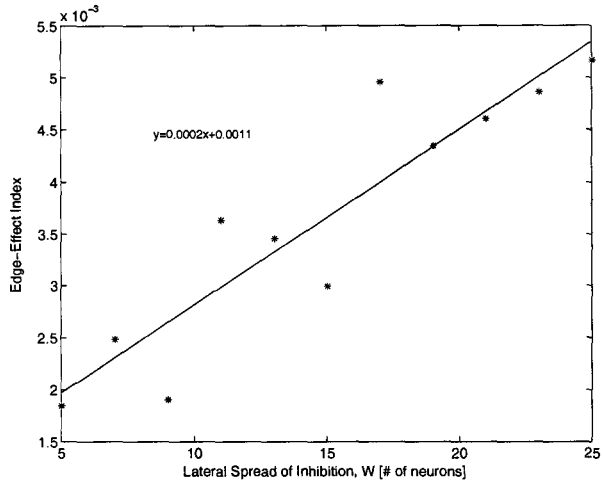


Figure 5.9: The effect of varying the span of inhibitory connections on contrast enhancement. Neural and network parameters values were set as follows: $t_{step} = 0.05$ ms for 5 s, $\tau = 1$ ms, $C = 7.5$ pF, $a = 2$, $v_{th} = 0.02$ V, $t_{ref} = 2$ ms, $\alpha_{GE} = 10$, $\alpha_{GI} = 3$, $E_E = 0.1$ V and $E_I = -0.02$ V.

generator in MATLAB, thereby causing it to produce the same set of input spike trains in each simulation. The input is the same between simulations run with the spiking and non-spiking models for the same ramp width. It is evident from Figure 5.11 that as the ramp width is increased, contrast enhancement is decreased in both the spiking and non-spiking models. This result is consistent with Mach's findings in the visual system that are shown in Figure 3.2. For the non-spiking model, the mean membrane potential was used as an equivalent measure to the mean spike rate of the spiking model. The results of the simulations shown in Figure 5.11 suggest that greater contrast enhancement is achieved by the non-spiking model than the spiking model, especially at the valleys. To quantify the change in contrast enhancement as the ramp width of the spontaneous input is varied, a series of simulations were run and the spurious peak height measured. A peak index is defined to be the difference between the peak height from the average of the mean spike rate in the normal

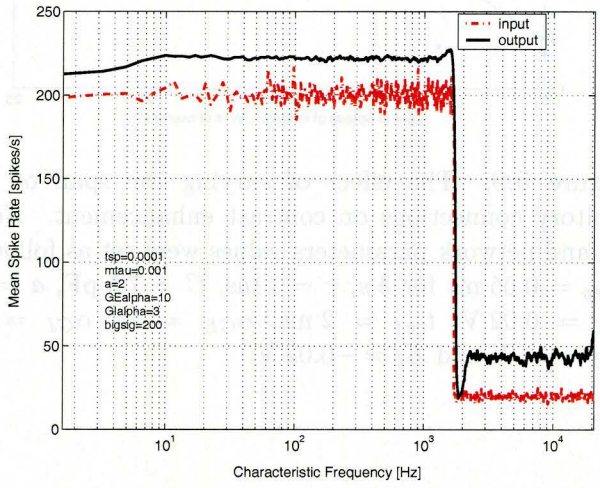
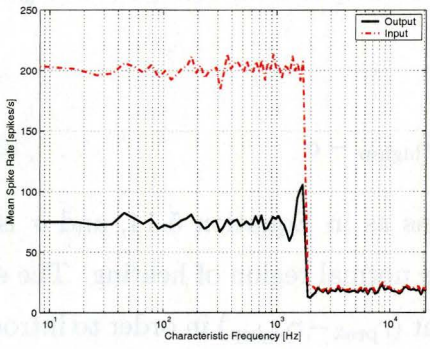
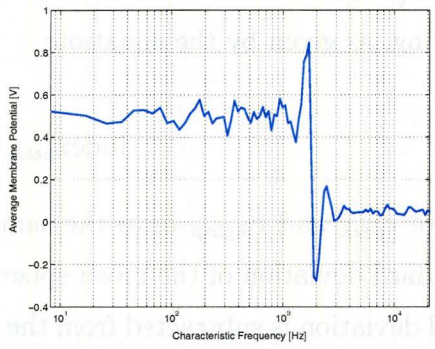


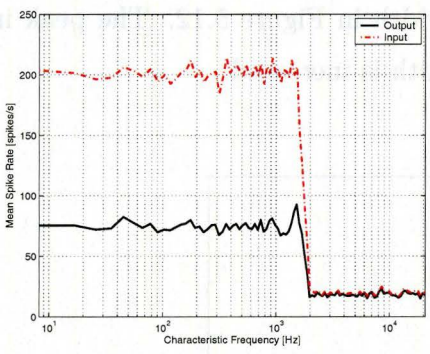
Figure 5.10: Average spike rates of neurons in a 500-neuron network with both lateral excitatory and inhibitory connections. Excitatory connections span 17 neurons while inhibitory connections span 32 neurons to either side of any given neuron. Other parameter values are the same as those in Figure 5.7.



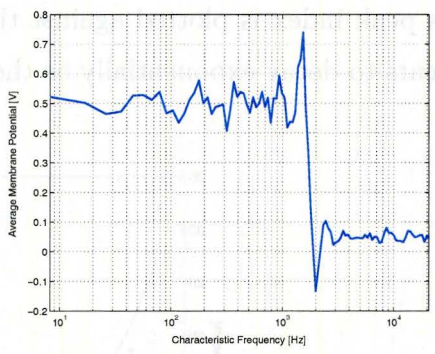
(a) Ramp across 3 spiking neurons.



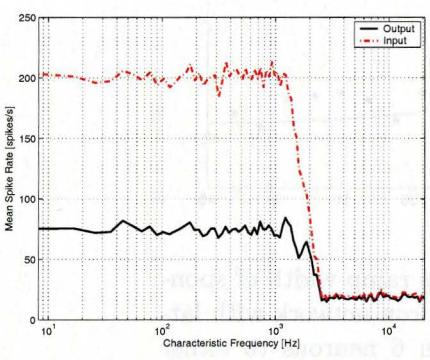
(b) Ramp across 3 non-spiking neurons.



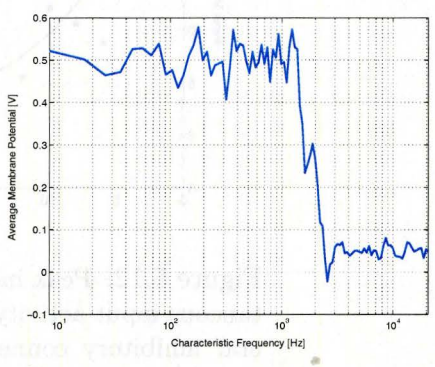
(c) Ramp across 6 spiking neurons.



(d) Ramp across 6 non-spiking neurons.



(e) Ramp across 15 spiking neurons.



(f) Ramp across 15 non-spiking neurons.

Figure 5.11: Effect of varying the ramp width of the rate of spontaneous input. LINs consisted of 100 neurons with inhibitory connections spanning 6 neurons to either side. Other parameter values for the spiking models were set as follows: $t_{step} = 0.02$ ms for 5 s, $\tau = 1$ ms, $C = 7.5$ pF, $a = 32$, $v_{th} = 0.015$ V, $t_{ref} = 2$ ms, $\alpha_{GE} = 11$, $\alpha_{GI} = 3$, $E_E = 0.1$ V and $E_I = -0.02$ V. Parameters for the non-spiking models were: $t_{step} = 0.02$ ms for 5 s, $\tau = 1$ ms, $a = 1$, and $\alpha_I = 11$.

hearing region and the standard deviation of the output from the normal region of hearing, as given by the equation:

$$\text{Index}_{\text{peak}} = r_{\text{peak}} - r_{\text{highsp}} - \sigma \tag{5.15}$$

where r_{peak} and r_{highsp} have the same definitions as in Equation 5.14, and σ is the standard deviation of the mean spike rate in the normal region of hearing. The standard deviation is subtracted from the peak height ($r_{\text{peak}} - r_{\text{highsp}}$) in order to introduce a penalty for large fluctuations in a region with a uniform rate of spontaneous input. This peak index is plotted against the ramp width in Figure 5.12. The peak index appears to decay exponentially as the ramp width is increased.

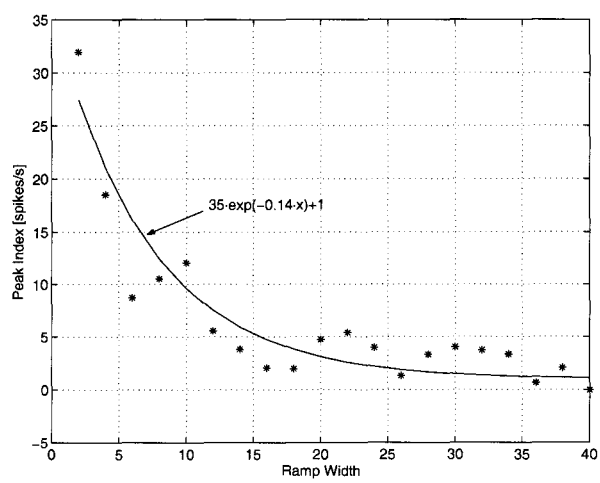


Figure 5.12: Peak indices versus ramp width of spontaneous input activity of 100-neuron network with lateral inhibitory connections span 6 neurons to either side of any given neuron. Other parameter values were set as follows: $t_{step} = 0.02$ ms for 5 s, $\tau = 1.5$ms, $C = 8$ pF, $a = 32$, $v_{th} = 0.015$ V, $t_{ref} = 2$ ms, $\alpha_{GE} = 11$, $\alpha_{GI} = 0.5$, $E_E = 0.1$ V and $E_I = -0.02$ V. The exponential curve was fit to the data by inspection.

The effect of numerous model variables on contrast enhancement was evaluated by

running several series of simulations. Each series of simulations considers the effect of varying two parameters. The range of values considered were chosen to remain close to being biologically realistic where relevant and to span the values that give the greatest contrast enhancement. The results of the simulations are summarized in contour plots.

In the first series of simulations, the rates of spontaneous input activity were varied between 20 and 400 spikes/s. As can be seen in Figure 5.13, the greater the difference between the rates, i.e., the greater the magnitude of the edge, the better the contrast enhancement. The peak index measure was unnecessary in this case because none of the neural or network parameters were varied. Hence, the peak height itself is plotted. In simulations with an input edge height of 380 spikes/s, the average firing

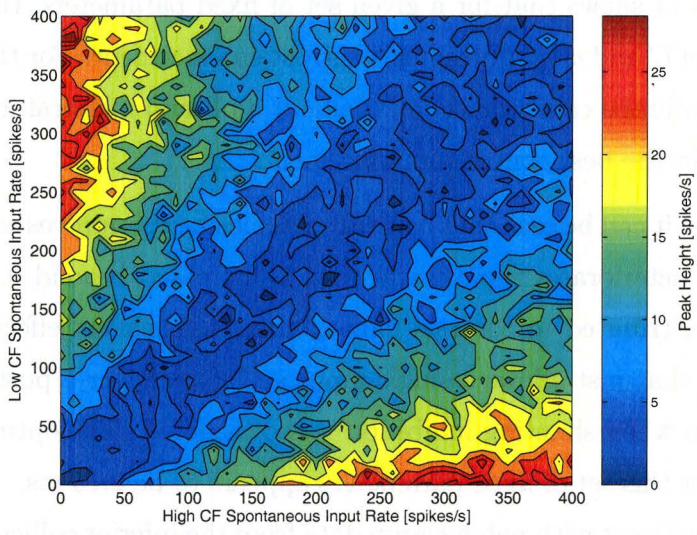


Figure 5.13: Effect of varying the high and low spontaneous rates. The 100-neuron network has lateral inhibitory connections that span 6 neurons to either side of each neuron. Other parameter values were set as follows: $t_{step} = 0.05$ ms for 5 s, $\tau = 1$ ms, $C = 7.5$ pF, $a = 2$, $v_{th} = 0.02$ V, $t_{ref} = 2$ ms, $\alpha_{GE} = 10$, $\alpha_{GI} = 3$, $E_E = 0.1$ V and $E_I = -0.02$ V. Amplitudes of the inhibitory and excitatory conductances were normalized to 10^{-8} before they were scaled by a.

rate in the normal region of output was about 115 spikes/s. This rate is unrealistically high, for electrophysiologic recordings from guinea pigs [Torterolo et al., 2002] and cats davis have shown that the spontaneous spike rate of neurons in the IC are well below 50 spikes/s. However, it is this high rate of input and output spikes that facilitates the coincidence between excitatory and inhibitory activity necessary for contrast enhancement to occur. The high rate of input might be achieved *in vivo* by the convergence of excitatory input projected from a few neurons. Lower input spike rates result in lower output spike rates. In the latter case, the sparseness of input and output spikes makes the coincidence of excitatory and inhibitory activity rare. Contrast enhancement is therefore not visible, which is shown in Figure 5.13 as the dark blue-coloured diagonal band across the plot.

The membrane capacitance, C, and the lateral inhibition factor, a, were then varied. Figure 5.14 shows that for a given set of fixed parameters, there is a narrow range of values of C and a that optimize contrast enhancement. For the configuration presented, a membrane capacitance, C, of about 8 pF and a lateral inhibition factor, a, of 32 results in the best contrast enhancement.

In Figure 5.15 it can be seen that contrast enhancement improves as v_{th} is increased to 0.015 V, but deteriorates rapidly when v_{th} is increased beyond 0.02 V. This can presumably be attributed to contrast enhancement being most effective when v_{th} is close to a value that just allows a single input spike to trigger a postsynaptic action potential as seen when shunting inhibition was considered. The optimum membrane time constant for this set of fixed parameters appears to be 2.75 ms. These values for v_{th} and τ are consistent with patch clamp data from the inferior colliculi of mice Basta & Vater [2003] and rats Koch & Grothe [2003].

When the alpha values of the excitatory and inhibitory conductances were varied, it was found that α_E had to be fairly large and α_I very small to obtain contrast enhancement. From Figure 5.16, the optimum alpha values are 11 for α_E and 0.5 for α_I. Contrast enhancement deteriorates rapidly as α_I is increased from 0.5. Contrast enhancement is less sensitive to the value of α_E, though values below 6 yield little to

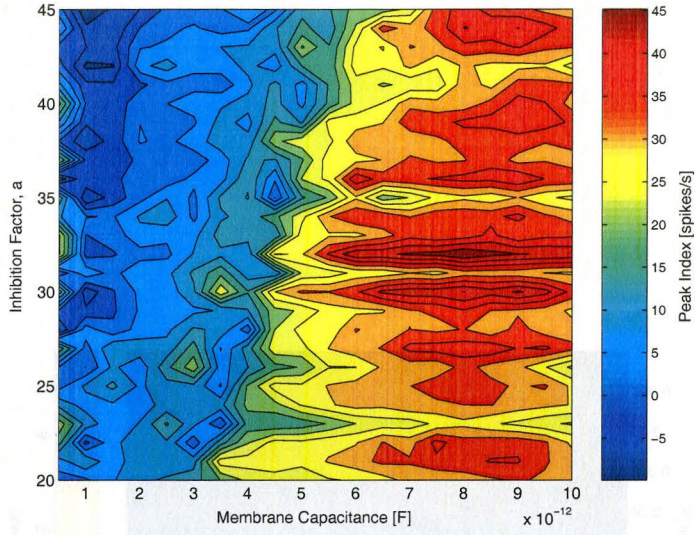


Figure 5.14: Effect of varying the membrane capacitance, C, and the lateral inhibition factor, a. The 100-neuron network has lateral inhibitory connections that span 6 neurons to either side of any given neuron. Other parameter values were set as follows: $t_{step} = 0.02$ ms for 5 s, $\tau = 1$ ms, $v_{th} = 0.02$ V, $t_{ref} = 2$ ms, $\alpha_{GE} = 11$, $\alpha_{GI} = 0.5$, $E_E = 0.1$ V and $E_I = -0.02$ V.

no contrast enhancement in combination with any value of α_I.

Finally, the duration of the absolute refractory period, t_{ref}, was varied. For each of the simulations shown in Figure 5.17 the input used is identical. From Figure 5.17 it is evident that as the duration of the absolute refractory period is increased, contrast enhancement decreases. This result would be expected since a longer refractory period makes coincidence between excitatory and inhibitory activity of the neurons less likely.

5.2 Response to Synthesized Speech Stimulus

As described in Section 4.3.1, synthesized speech processed by the Bruce et al. model was used as input to the LIN model of spiking neurons. The stimulus was

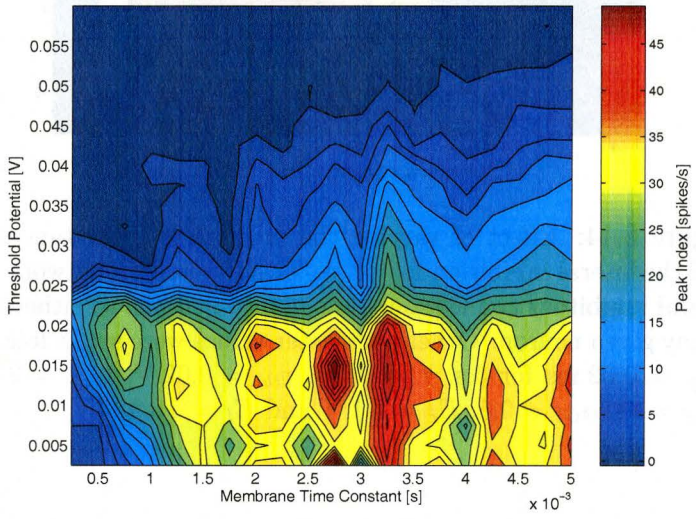


Figure 5.15: Effect of varying the membrane time constant, τ, and the threshold potential, v_{th}. The 100-neuron network has lateral inhibitory connections that span 6 neurons to either side of any given neuron. Other parameter values were set as follows: $t_{step} = 0.02$ ms for 5 s, $C = 7.5$ pF, $a = 36$, $t_{ref} = 2$ ms, $\alpha_{GE} = 11$, $\alpha_{GI} = 0.5$, $E_E = 0.1$ V and $E_I = -0.02$ V.

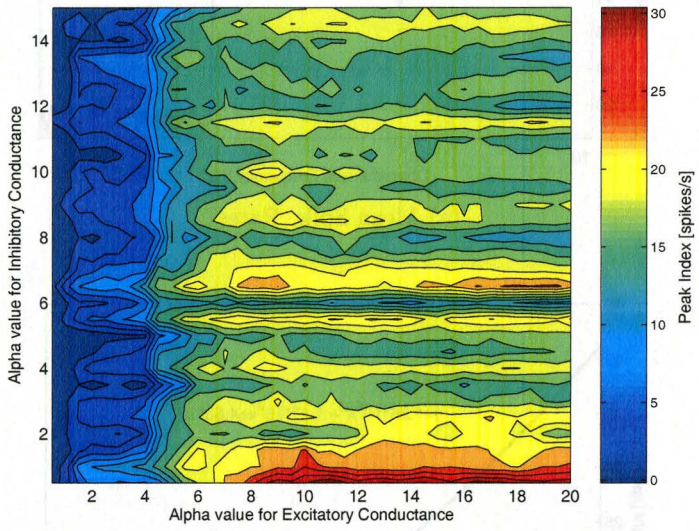
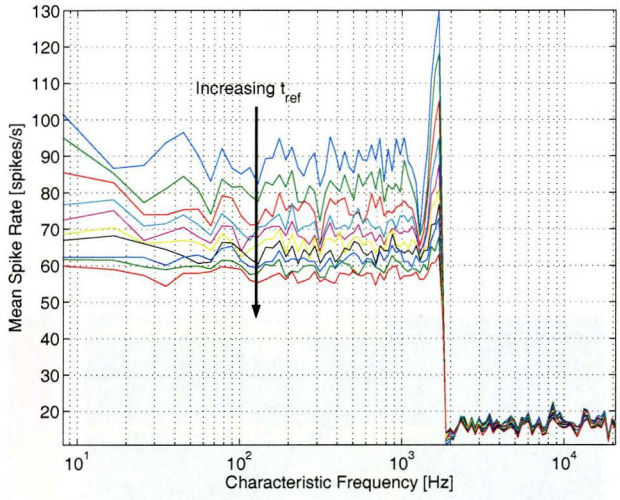
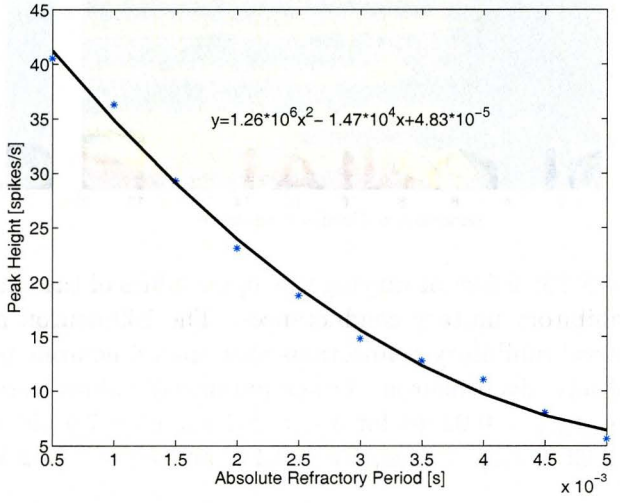


Figure 5.16: Effect of varying the alpha values of the excitatory and inhibitory unitary conductances. The 100-neuron network has lateral inhibitory connections that span 6 neurons to either side of any given neuron. Other parameter values were set as follows: $t_{step} = 0.02$ ms for 5 s, $\tau = 1$ ms, $C = 7.5$ pF, $a = 36$, $v_{th} = 0.02$ V, $t_{ref} = 2$ ms, $E_E = 0.1$ V and $E_I = -0.02$ V.



(a) Mean Spike Rates



(b) Peak Height

Figure 5.17: Effect of varying the absolute refractory period, t_{ref}. The LIN consisted of 100 neurons with inhibitory connections spanning 6 neurons to either side. Other parameter values for the spiking models were set as follows: $t_{step} = 0.02$ ms, $\tau = 1$ ms, $C = 7.5$ pF, $a = 32$, $v_{th} = 0.015$ V, $\alpha_{GE} = 11$, $\alpha_{GI} = 3$, $E_E = 0.1$ V and $E_I = -0.02$ V.

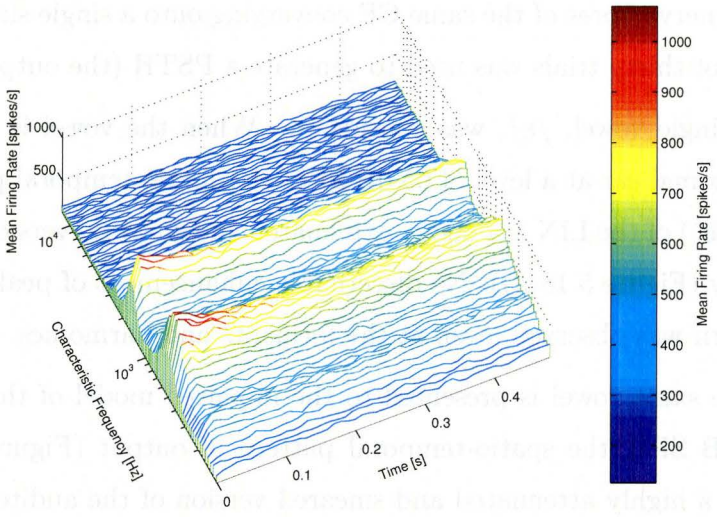
processed by the model of the ear five times, then summed to simulate the effect of five auditory nerve fibres of the same CF converging onto a single subcortical neuron. The average of thirty trials was used to generate a PSTH (the output).

First, a single vowel, /ɛ/, was considered. When the vowel is presented to the model of a normal ear at a level of 65 dB SPL, the spatio-temporal pattern of output (Figure 5.18(b)) of the LIN is a highly attenuated and smeared version of the auditory nerve activity (Figure 5.18(a)). No discernable enhancement of peaks or edges in the spatial pattern was observed, even at the formants and harmonics.

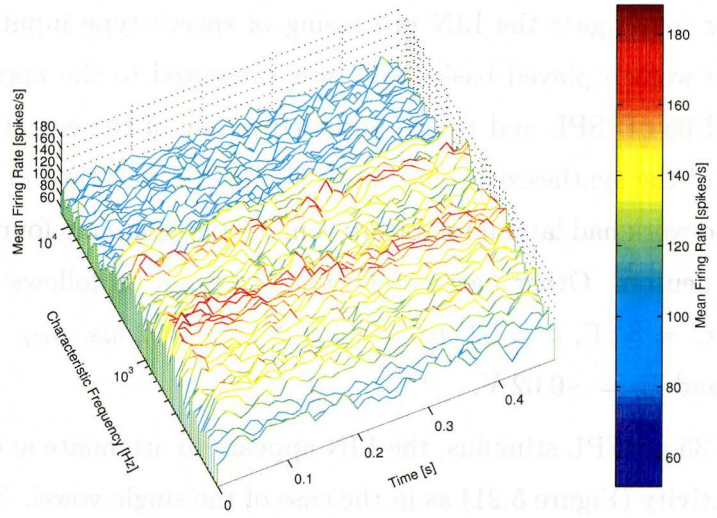
When the same vowel is presented to the impaired model of the ear at a higher level of 95 dB SPL, the spatio-temporal pattern of output (Figure 5.19(b)) of the LIN is again a highly attenuated and smeared version of the auditory nerve activity (Figure 5.19(a)). From the spatio-temporal plots, there appear to be no enhancements at the transitions between the normal and impaired regions of hearing.

To further investigate the LIN processing of speech-type input, the synthesized sentence, "five women played basketball" was presented to the normal model of the ear at 35 and 65 dB SPL and to the impaired model of the ear at 95 dB SPL. The spectrogram of the synthesized sentence at 65 dB SPL is shown in Figure 5.20. The 100-neuron network had lateral inhibitory connections spanning four neurons to either side of each neuron. Other parameter values were set as follows: $t_{step} = 0.02$ ms, $\tau = 1.5$ ms, $C = 8$ pF, $a = 32$, $v_{th} = 0.015$ V, $t_{ref} = 2$ ms, $\alpha_{GE} = 11$, $\alpha_{GI} = 0.5$, $E_E = 0.1$ V and $E_I = -0.02$ V.

With the 35 dB SPL stimulus, the LIN appears to attenuate and smear the auditory nerve activity (Figure 5.21) as in the case of the single vowel. The main features of the spatio-temporal pattern from the auditory nerve are preserved, but formant, harmonics and other such spatial features do not appear to be enhanced. This latter point is clarified by comparing the plots of the mean spike rates over a 10 ms-window in Figure 5.22. The window begins at 1.1s into the sentence and isolates part of the first 'a' of the word 'basketball'. Notice that with this low intensity stimulus, even the auditory nerve response appears to have an upward-shifted second formant frequency.

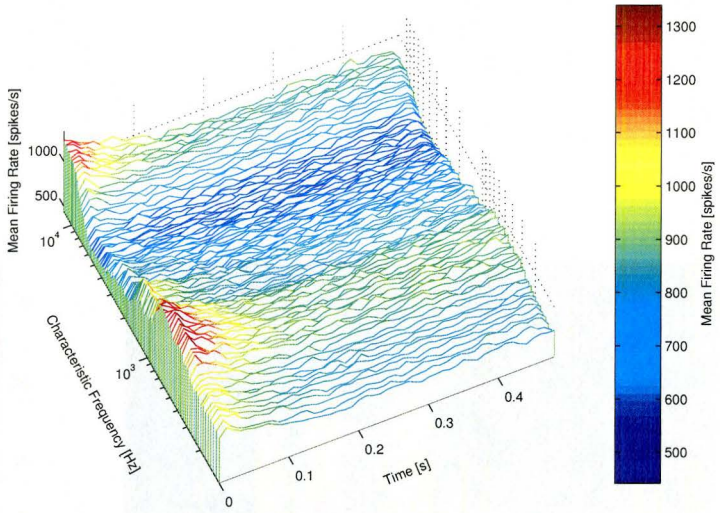


(a) Auditory Nerve Activity

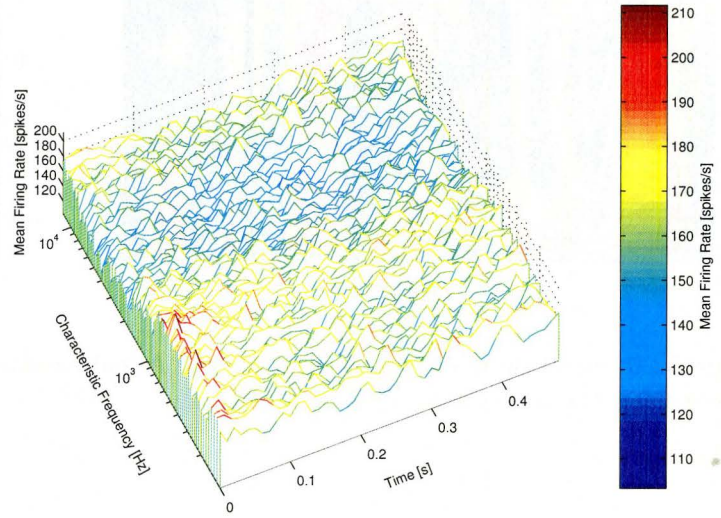


(b) Output from the LIN

Figure 5.18: LIN processing of normal auditory periphery response to synthesized voiced speech /ε/ presented at 65db SPL. The LIN consisted of 100 neurons with inhibitory connections spanning 4 neurons to either side. Other parameter values for the spiking models were set as follows: $t_{step} = 0.01$ ms, $\tau = 3$ ms, $C = 30$ pF, $a = 2$, $v_{th} = 0.08$ V, $t_{ref} = 2$ ms, $\alpha_{GE} = 10$, $\alpha_{GI} = 3$, $E_E = 0.1$ V and $E_I = -0.02$ V.



(a) Auditory Nerve Activity



(b) Output from the LIN

Figure 5.19: LIN processing of impaired auditory periphery response to synthesized voiced speech /ɛ/ presented at 95 dB SPL. Parameters of the LIN are the same as those of Figure 5.18.

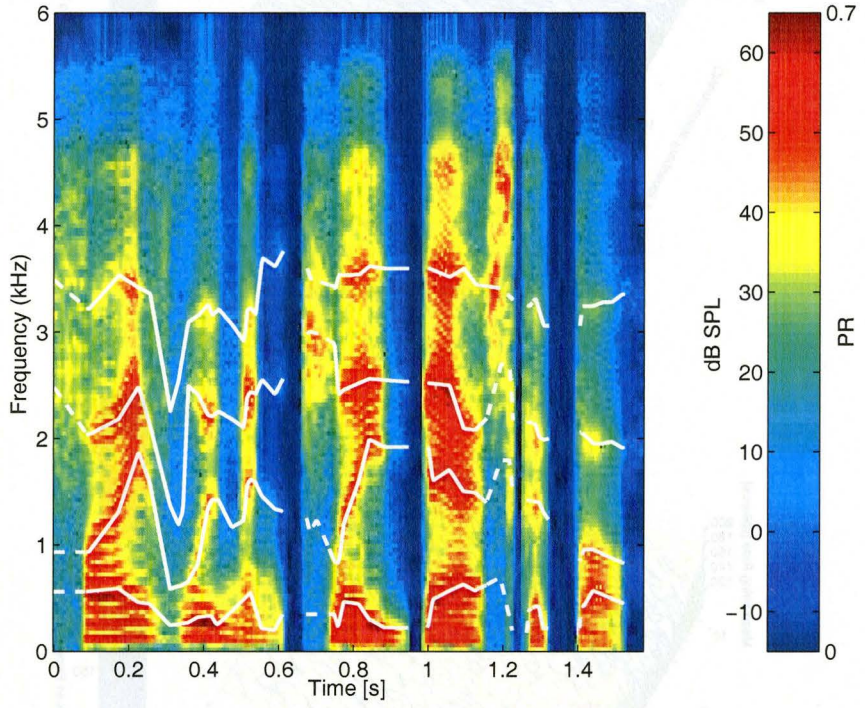
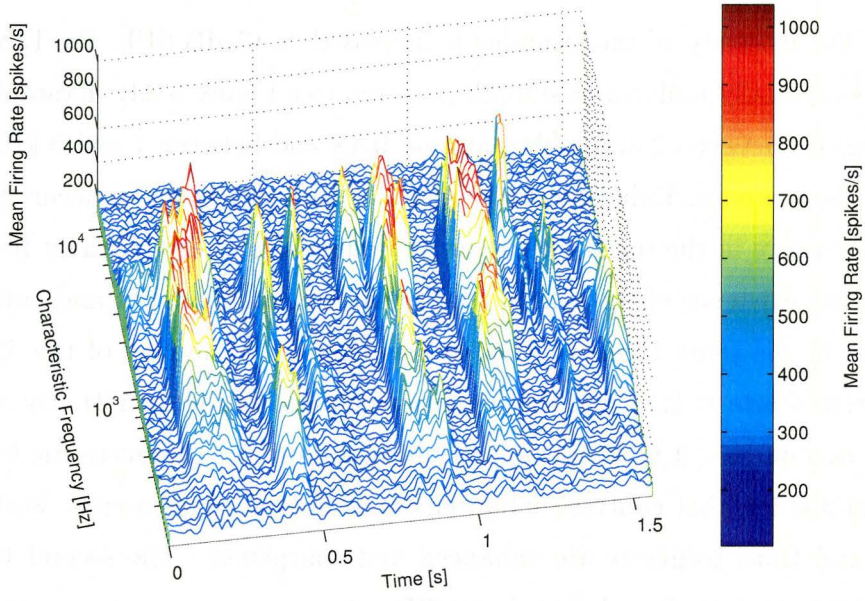
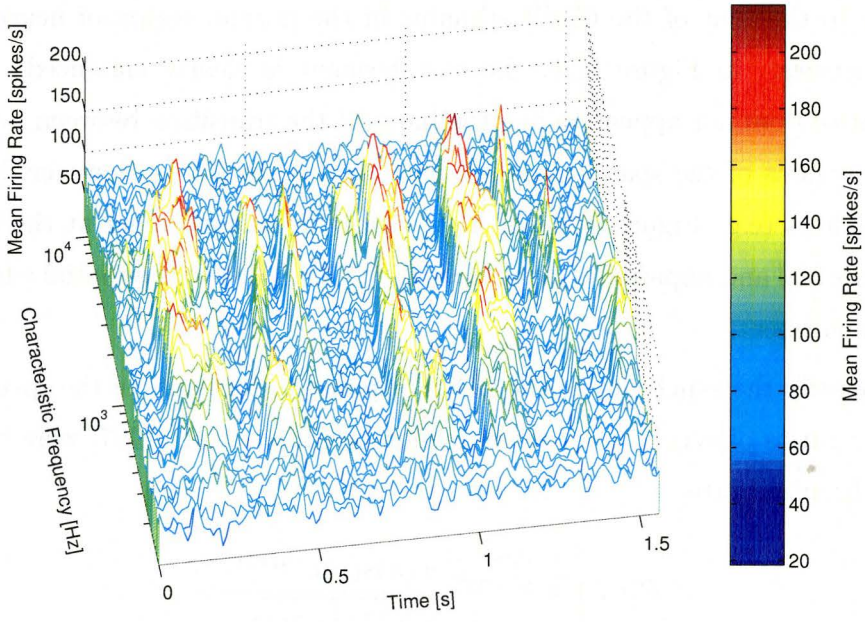


Figure 5.20: Spectrogram of synthesized speech: A male speaker saying, "Five women played basketball" at 65 dB SPL.



(a) Auditory Nerve Activity



(b) Output from the LIN

Figure 5.21: LIN processing of speech presented to a normal ear at 35 dB SPL.

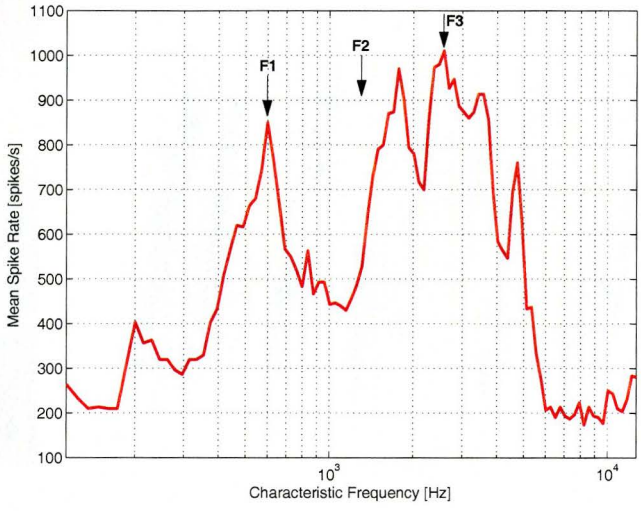
When the intensity of the stimulus is increased to 65 dB SPL, the LIN appears to suppress spatially uniform or smooth patterns (see Figure 5.23). Examples of this are the regions between 2 and 4 kHz, around 0.8 s and between 1 and 5 kHz, around 0.2 s into the sentence. Enhancement of spatial edges also seems to occur. Such edge enhancements are in the regions near the fundamental and first formant frequencies. To clarify the existence of the spatial edge enhancement, the input and output mean spike rates in the same 10 ms-window as was shown for the case of the 35 dB SPL stimulus that starts at 1.1 s into the sentence is plotted in Figure 5.24. By comparing the input and output, it is obvious, particularly around the characteristic frequencies of 150 and 300 Hz, that contrast enhancement at spatial edges occurs. Furthermore, the first and third formants are enhanced and sharpened. The second formant is sharpened but is shifted to slightly lower CFs.

The LIN response to the auditory nerve activity evoked by the 95 dB SPL stimulus is similar to the that of the 65 dB stimulus in the normal region of hearing at the low frequencies. In Figure 5.26, the first formant is clearly enhanced. Contrast enhancement does not appear to be prominent at the transition between normal and impaired regions of the spatio-temporal pattern (Figure 5.25). However, looking at 10 ms-windows (e.g. Figure 5.26), there is clearly an enhancement at the transition between normal and impaired regions (around 1.5 kHz) due to the spatial edge created by the impairment.

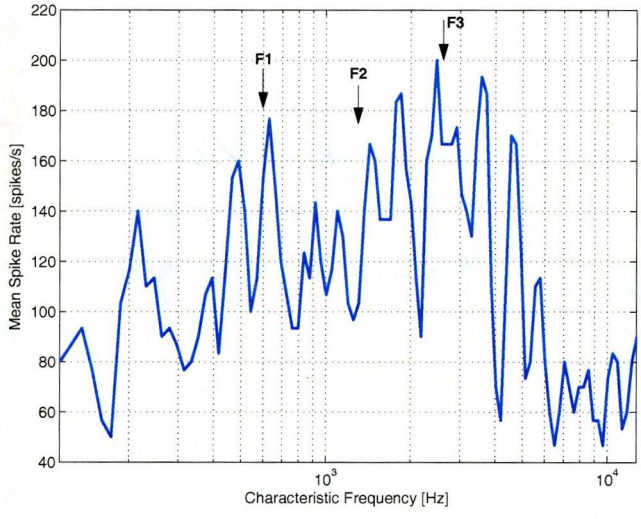
To quantify the synchrony of the LIN's response to formants of the speech stimulus, power ratios of synchronous rates defined by Miller et al. [1997] were computed. The synchronized rate:

$$R(kf_t) = \frac{|\sum_{n=0}^{N-1} w(n)s(n)e^{-j2\pi kn/N}|}{\sqrt{N\sum_{n=0}^{N-1} w^2(n)}} \tag{5.16}$$

is the Fourier transform of the Hamming windowed PSTH, $w(n) \cdot s(n)$. N is the length of the window and f_t is the frequency resolution of the calculation, i.e., the inverse

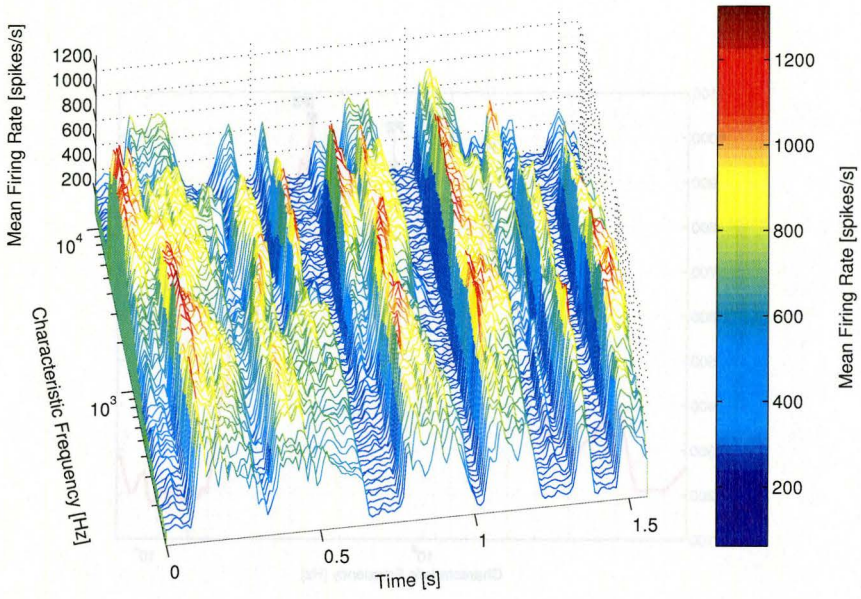


(a) Auditory Nerve Activity

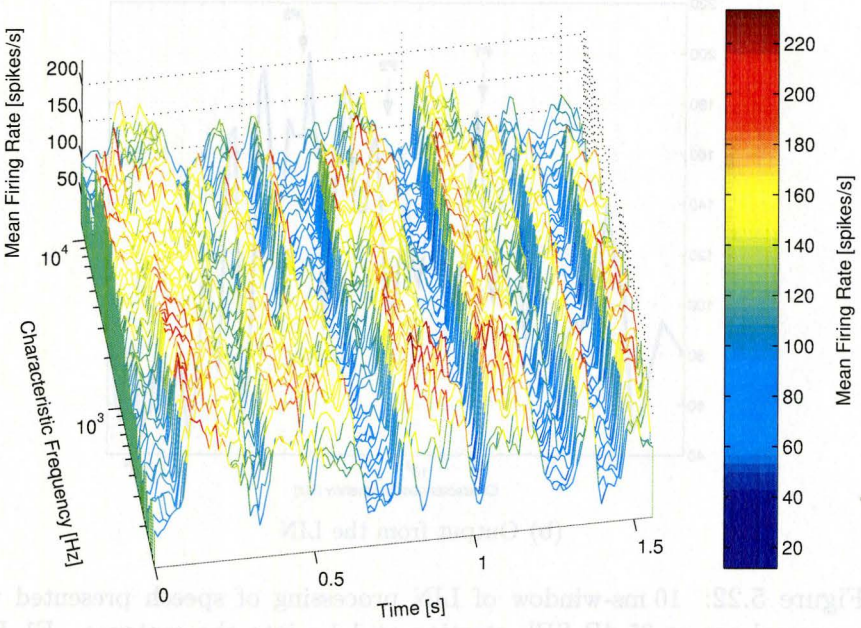


(b) Output from the LIN

Figure 5.22: 10 ms-window of LIN processing of speech presented to a normal ear at 35 dB SPL starting at 1.1 s into the sentence. F1–F4 indicate the first through to fourth formant frequencies.

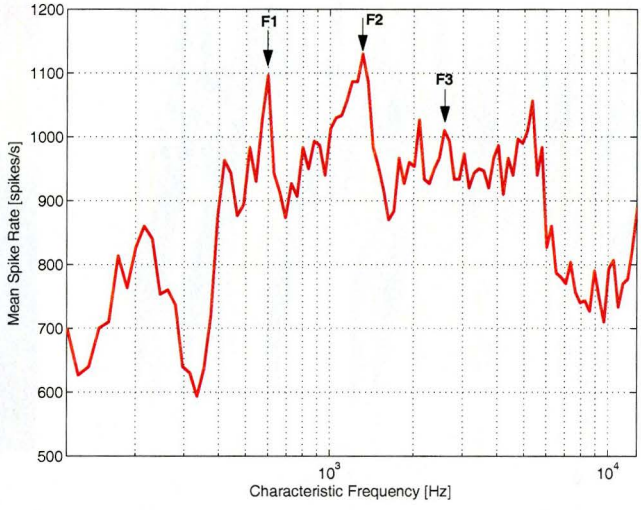


(a) Auditory Nerve Activity

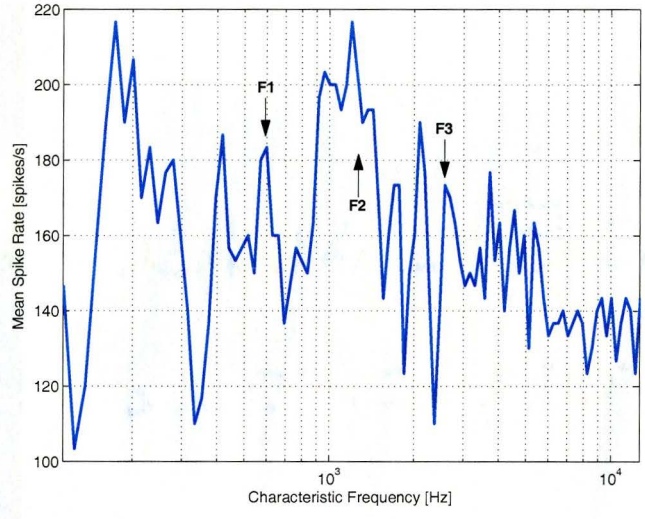


(b) Output from the LIN

Figure 5.23: LIN processing of speech presented to a normal ear at 65 dB SPL.

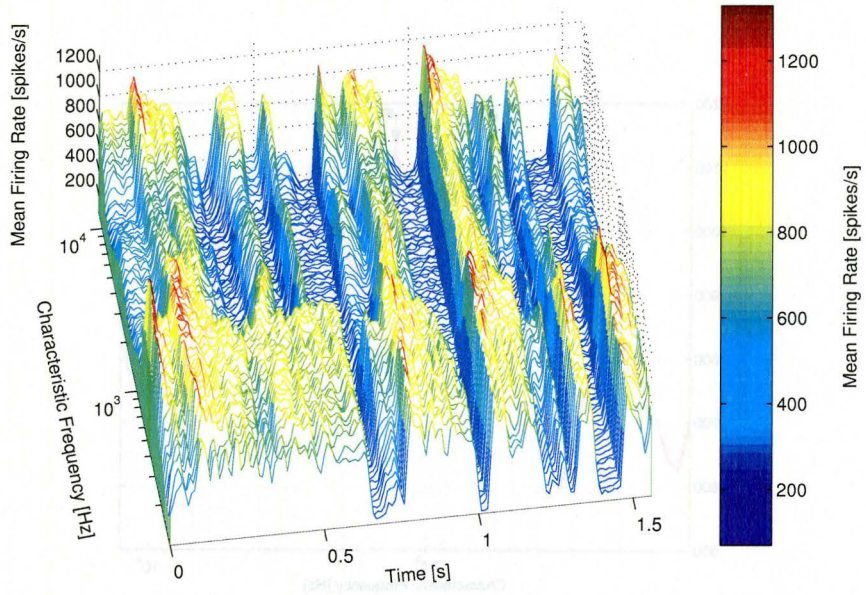


(a) Auditory Nerve Activity

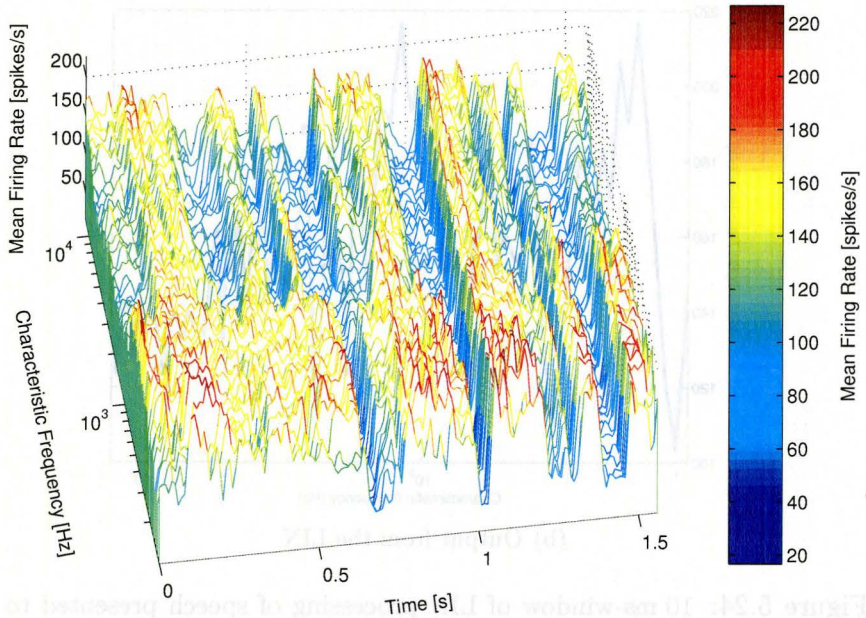


(b) Output from the LIN

Figure 5.24: 10 ms-window of LIN processing of speech presented to a normal ear at 65 dB SPL starting at 1.1 s into the sentence.

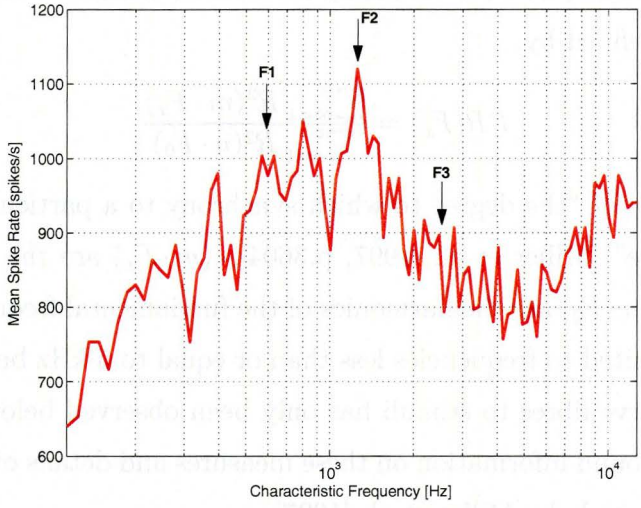


(a) Auditory Nerve Activity

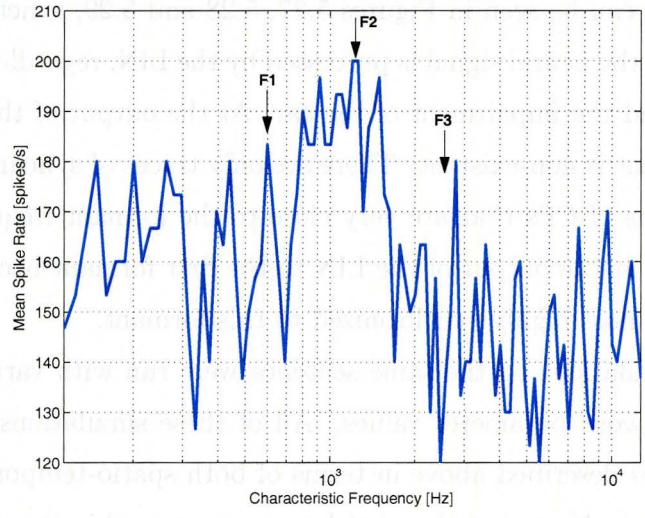


(b) Output from the LIN

Figure 5.25: LIN processing of speech presented to a normal ear at 95 dB SPL.



(a) Auditory Nerve Activity



(b) Output from the LIN

Figure 5.26: 10 ms-window of LIN processing of speech presented to a normal ear at 95 dB SPL starting at 1.1 s into the sentence.

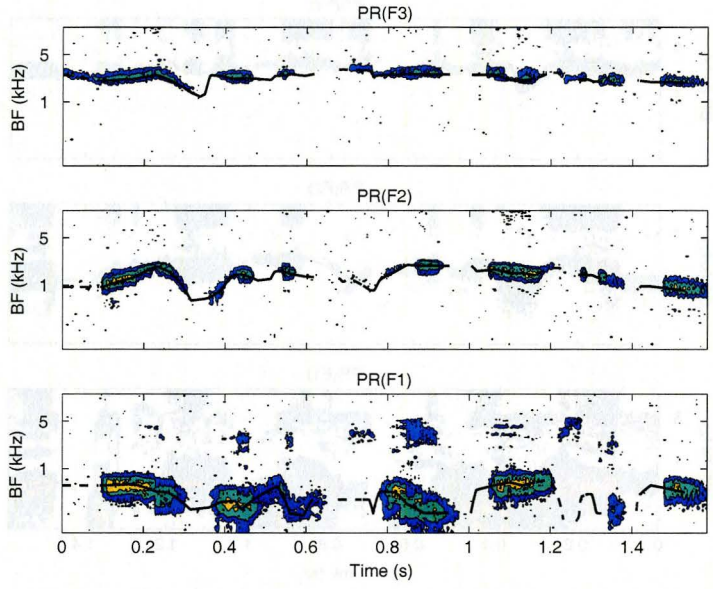
of the window duration. For the purposes of this investigation, a 12.75 ms-Hamming window was applied. The denominator in Equation 5.16 corrects the attenuation incurred by the window, $w(n)$, normalizing the Fourier transform to units of spikes/s. The power ratio defined by:

$$PR(F_x) = \frac{\sum_{m=1}^{u} R^2(m \cdot F_x)}{\sum_{n=1}^{v} R^2(n \cdot F_0)} \tag{5.17}$$

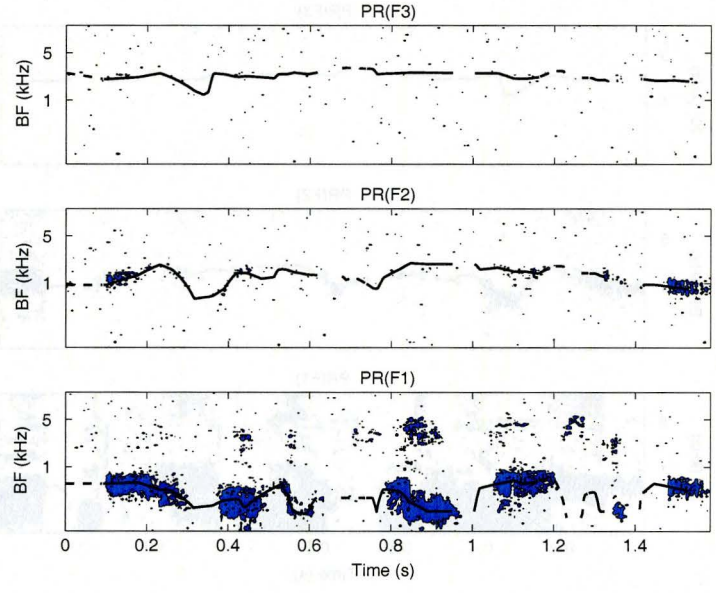
provides a measure of "the degree to which synchrony to a particular formant dominates the response" [Miller et al., 1997, p.3604]. $(m \cdot F_x)$ are the harmonics of the formant, F_x, and $(n \cdot F_0)$ are the harmonics of the fundamental frequency, F_0. $(u \cdot F_x)$ and $(v \cdot F_0)$ are limited to frequencies less than or equal to 5 kHz because phase locking of auditory nerve fibres to stimuli has only been observed below that frequency. For further background information on these measures and details of them, the reader is referred to the article by Miller et al. [1997].

Figures 5.27, 5.28 and 5.29 were generated from the same data as Figures 5.21, 5.23 and 5.25. As can be seen in Figures 5.27, 5.28 and 5.29, synchrony to formants deteriorates when the neural signal is processed by the LIN, regardless of the intensity of the stimulus and the impairment of the ear. At the output of the LIN, synchrony to the third formant is non-existent. There are only traces of synchrony to the second formant in neurons of CFs that are very close to the formant frequency. Synchrony is only significant at the output of the LIN to the first formant because the auditory nerve response is most highly synchronized to this formant.

Numerous simulations on the same sentence were run with various combinations of neural and network parameter values. All of these simulations produced results similar to the ones described above in terms of both spatio-temporal spike rate patterns and synchrony. Note that the model parameters used in the synthesized speech simulations presented herein were those that generated the greatest contrast enhancement in the spontaneous-type input simulations. Hence, suppression of uniform input and contrast enhancement, characteristic of lateral inhibition, that were found was expected in the spatio-temporal patterns of output from the LIN.

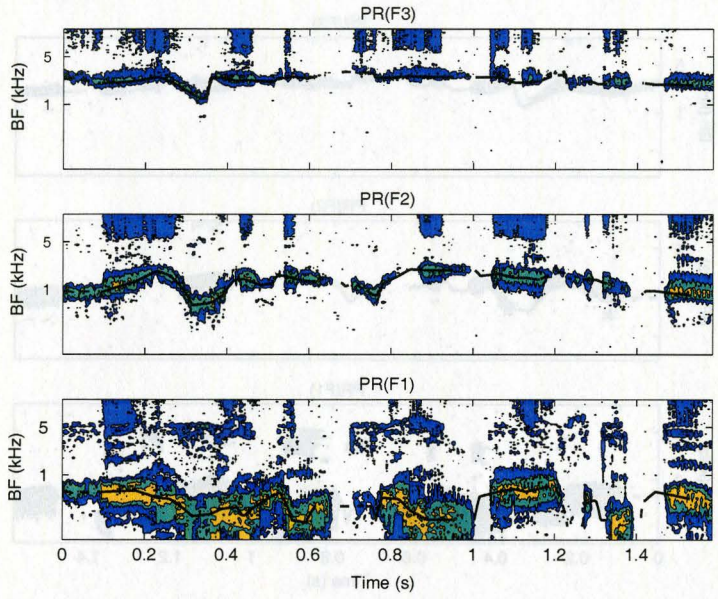


(a) Auditory Nerve Activity

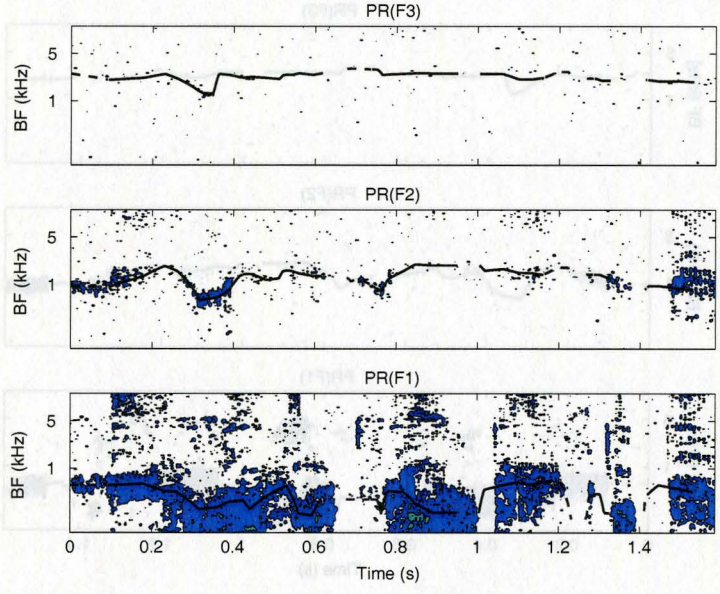


(b) Output from the LIN

Figure 5.27: Power ratios as a measure of synchrony to the LIN's response to speech presented to a normal ear at 35 dB SPL.

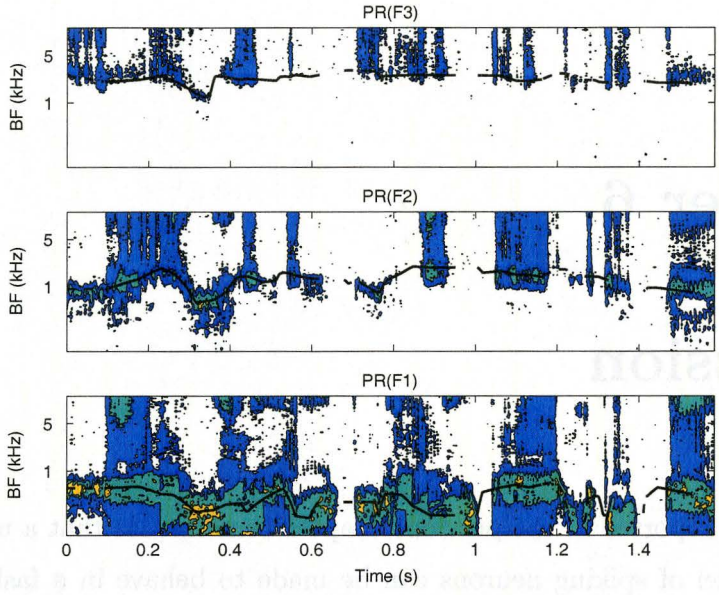


(a) Auditory Nerve Activity

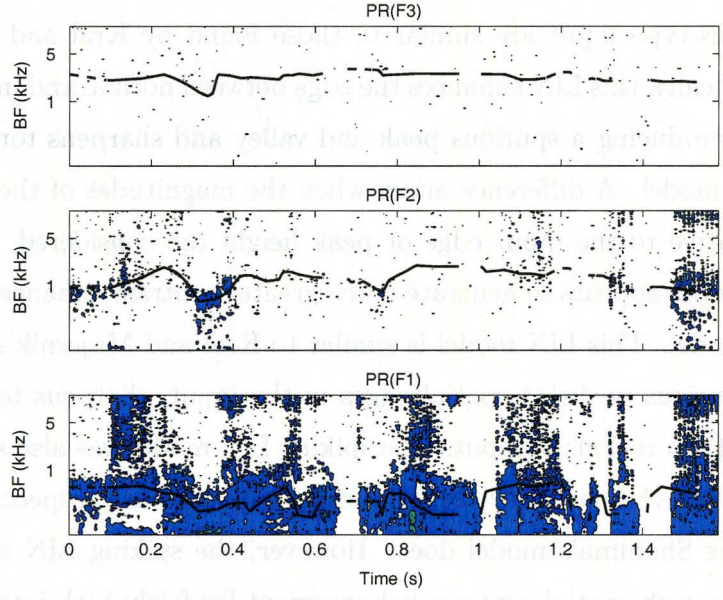


(b) Output from the LIN

Figure 5.28: Power ratios as a measure of synchrony to the LIN's response to speech presented to a normal ear at 65 dB SPL.



(a) Auditory Nerve Activity



(b) Output from the LIN

Figure 5.29: Power ratios as a measure of synchrony to the LIN's response to speech presented to an impaired ear at 95 dB SPL.

Chapter 6

Discussion

The results reported in the previous chapter demonstrate that a uniform, single-layer LIN model of spiking neurons can be made to behave in a fashion similar to that of the simpler LIN models of Kral and Majernik, Gerken, and Shamma. For sets of parameters that are able to perform contrast enhancement, the results for the spontaneous-type input are similar to those found by Kral and Majernik, and Gerken. Specifically, this LIN enhances the edge between normal and impaired regions of hearing by producing a spurious peak and valley and sharpens tone-like input as does Gerken's model. A difference arises when the magnitudes of the spurious peak and valley relative to the input edge or peak height are considered. Gerken's non-spiking LIN model appears to generate much greater contrast enhancement than this spiking LIN model. This LIN model is similar to Kral and Majernik's model in that it produces spurious peaks at spatial edges in the input, sharpens tonal input, and suppresses uniform regions of input. The spiking LIN model was also able to enhance large, rapid spatial changes, the lower harmonics and formants in speech-driven neural activity, just as Shamma's model does. However, the spiking LIN model was only able to perform such spatial contrast enhancement for fairly high intensity stimuli.

Although most of the model configurations presented did perform contrast enhancement, this behaviour was difficult to achieve as it requires a very narrow range of neural and network parameters and a lot of coincident input. In pilot tests that

involved stimulating the Bruce et al. model of the ear with band-passed noise, it was found that without convergent auditory nerve fibres, contrast enhancement could not be achieved by the LIN due to the low levels of input activity. At low levels of input activity, the frequency of spikes is low, which then elicits few output spikes. Because the frequency of both excitatory and inhibitory inputs are low, the chances of them occurring close enough in time for the inhibitory action to have any effect on the excitatory action is slim. Without such coincidence of excitatory and inhibitory activity, lateral inhibition is ineffective. To increase the likelihood of coincident inhibitory activity, the input activity had to be increased. Input activity was increased by simulating convergent input: five auditory nerve fibres for the results presented. The difficulty of the spiking LIN model achieving contrast enhancement is also manifested in the smaller spurious peaks and valleys, and the need for higher intensity stimuli as compared to the existing non-spiking models and the results shown in Figure 5.11. This latter point is magnified by the fact that Kral and Majernik, and Shamma's models are nonrecurrent models while the model presented herein is recurrent. Recurrent networks may be thought of as performing a similar function to an infinite number of nonrecurrent, feed-forward network layers. In the context of LINs, the greater the number of layers in a nonrecurrent network, the more pronounced the lateral inhibitory effect, as is evident in Kral and Majernik's [1996] results. Hence, it would be expected that a recurrent LIN would achieve greater contrast enhancement than a single-layer, nonrecurrent network.

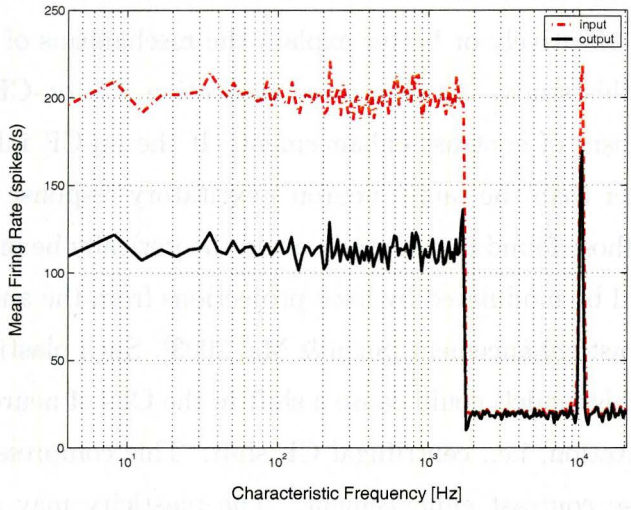
The foregoing explanation is reasonable because non-spiking LIN models use the graded membrane potentials to convey information. Since the time course of graded potential responses is much longer than that of action potentials, the coincidence of excitatory and inhibitory input over time is less critical. Hence, contrast enhancement is achieved more easily with a network of non-spiking neurons. This might explain why the retina performs contrast enhancement so effectively. All the cells and neurons peripheral to the optic nerve transmit information bearing signals via graded potentials and it is a network of these non-spiking cells that performs contrast

enhancement (as described in Section 3.1).

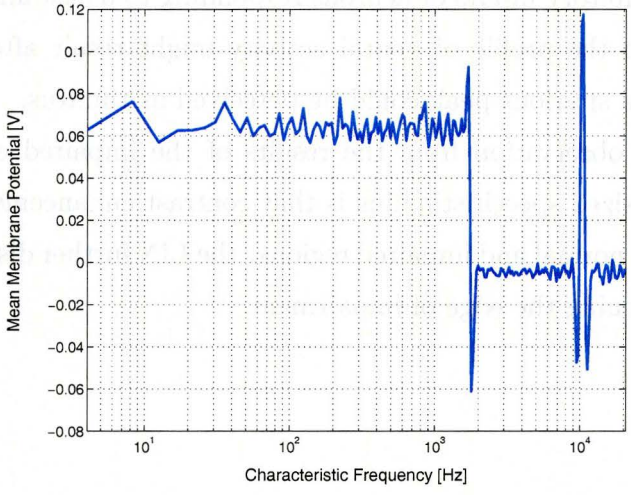
For synthesized speech stimuli, the synchrony of LIN neural activity to the formant frequencies is highly degraded compared to that of the auditory nerve. A possible explanation for this observation is that the LIN model attenuates the neural activity so much that synchrony is largely lost. Even though the LIN model does enhance formant frequencies via contrast enhancement in the mean spike rates, the function of the LIN as a speech feature extractor is dependent on what the neural code is. The current trend in thought, based on psychophysical and neurophysiologic evidence, is that synchrony of neural activity to the stimulus is important at low frequencies, while average spike rate is the only code for high frequencies [Shamma, 1985]. Due to its inability to maintain or enhance synchronous responses, this LIN model could only be, at best, a weak speech feature extractor.

Spiking neurons organized in a single-layer, recurrent LIN architecture is therefore unlikely to be how contrast enhancement and speech feature extraction is realized in the central auditory system if the action potential is the primary information carrier. Consequently, this particular model does not lend strong support to the idea that a subcortical LIN of spiking neurons contributes to a central mechanism of tinnitus.

Krezberg and colleagues [2004] have found that some neurons in the visual system of flies and the central nervous system of leeches use both the graded potential and action potential of spiking neurons to transmit information synaptically. Since non-spiking neurons have not been found in the auditory system, it is conceivable that auditory neurons use the graded, subthreshold portion of the membrane potential to achieve contrast enhancement while the spikes are primarily used to convey temporal cues. By comparing the plots in Figure 6.1, it can easily be seen that the membrane potential conveys contrast enhancement far better than the spikes alone, even when the spike trains and not the graded portion of the membrane potential is used for inhibitory interactions. One can imagine that if the graded portions of the membrane potentials were used to modulate lateral inhibition, the contrast enhancement achieved would be more like the results that Gerken and Shamma presented.



(a) Mean Spike Rate



(b) Mean Membrane Potential

Figure 6.1: Comparison of the graded potential and action potentials as information carriers for contrast enhancement. The preliminary LIN model of 100-neuron network with lateral inhibitory connections was used. Inhibitory connections span 6 neurons to either side of any given neuron i.e. total span of 13 neurons. Other parameters values were set as follows: $t_{step} = 0.1$ ms for 7 s, $\tau = 3$ ms, $a = 2$, $v_{th} = 0.25$ V, $t_{ref} = 2$ ms, $\alpha_{EPSP} = 5$, $\alpha_{IPSP} = 2$.

It should be noted that other neural network architectures could perform contrast enhancement more effectively or better explain the mechanisms of tinnitus than the one considered in this study. As mentioned in Section 3.1, on-CF inhibition could be another mechanism of contrast enhancement. If the on-CF inhibitory fields are considerably broader than the target neuron's excitatory response area, then results that are similar to those found in this study could theoretically be produced. Synaptic plasticity that could be modulated by back projections from the auditory cortex may also facilitate contrast enhancement [Suga & Ma, 2003]. Such plasticity (that was not incorporated into this model) could cause a shift in the CFs of neurons that results in a reduced representation, i.e., centrifugal CF shift. This compressed reorganization would thus increase contrast enhancement. The plasticity may occur in response to input sustained over some period of time. If the balance between the efficacy of excitatory and inhibitory inputs of neurons responding to a sustained tone is skewed, a spurious peak in the profile of neural activity might result after the stimulus is terminated. Such a spurious peak could be perceived as tinnitus.

An interesting observation from the results of the impaired ear processing the 95 dB SPL synthesized speech stimulus is that contrast enhancement does effect the transition between normal and impaired regions: the LIN further distorts the impaired response by sharpening the edge of impairment.

Chapter 7

Conclusions

7.1 Summary

To verify the validity of existing models, which employ non-spiking neural units to represent actual spiking neurons in a LIN, serving as a mid-brain auditory processor that enhances spatial changes or features of speech signals, a more biologically realistic model was developed. By running numerous computer simulations it was found that a uniform, single-layer recurrent LIN of leaky integrate-and-fire neurons is able to perform contrast enhancement under very specific conditions and configurations. Contrast enhancement can be achieved by a spiking LIN with slow inhibitory conductance dynamics, a densely connected network of neurons with hyperpolarizing-type inhibition and a short refractory period. Inhibitory interactions are most effective when the threshold potential is set close to 20 mV above the resting potential of the cell, which just allows a single excitatory input spike to induce an output spike in the absence of inhibitory input, since the presence of any inhibition will prevent an output spike from being generated. For a membrane capacitance of 7.5 pF, a membrane time constant of around 3 ms is required to produce the greatest contrast enhancement. The time-course of the disturbance in excitatory conductance due to a single input spike must be similar to the membrane time constant and an order of magnitude shorter in duration than that of the inhibitory conductance to achieve

significant contrast enhancement (e.g., 2 ms versus 20 ms, respectively). Even if the parameters of the LIN of spiking neurons are optimally set, the input must have large and very steep spatial changes for contrast enhancement to be evident at the output of the LIN model. Non-spiking neuron models proved superior at producing contrast enhancement than spiking neuron models. It may be concluded that spiking neurons organized in a single-layer LIN are unlikely to be the primary mechanism of contrast enhancement in the subcortical region of the central auditory system if the action potential is the primary information carrier. This conclusion contradicts those drawn in previous investigations by Gerken [1996] and Kral & Majernik [1996] that a LIN is likely to be the edge-enhancing central auditory processor. Since non-spiking neurons have not been found in the auditory system, this particular LIN is not supported as a central mechanism of tinnitus.

The LIN model is also a poor speech feature extractor because it is unable to pass or enhance the synchronous responses present at the level of the auditory nerve, and because it requires fairly high (demonstrated at 65 dB SPL and higher) intensity stimuli in order to enhance spike rate representations of formant frequencies and spatial edges. The latter finding shows that Shamma's model [1985] is highly limited in its representation of spiking neural activity. Additionally, the LIN further distorts the impaired response of the auditory nerve by enhancing the edge of impairment when a speech stimulus is presented at 95dB SPL.

The results of this study do not preclude the possibility of other LINs performing contrast enhancement or speech feature extraction in the central auditory system. For example, the recent physiological study by Krezberg and colleagues [2004] has shown the existence of a mixed neural signal of graded potentials and action potentials in flies and leeches. Such a mixed signal could theoretically allow spiking neurons to produce strong contrast enhancement. Further physiological and computational studies would have to be conducted to ascertain the existence and plausibility of such signals in the auditory system.

7.2 Suggestions for Future Work

To clarify the advantages of using a mixed signal instead of action potentials, the LIN model could be modified to incorporate the use of the subthreshold, graded membrane potential in the lateral connections. For example, the hypothesis that the graded potential actually codes for features that were thought to be coded for by the mean spike rate, while spikes serve to amplify temporal fluctuations via synchrony to the stimulus could be tested with a modified model. Simultaneously, neurophysiological studies could be conducted in search of the use of mixed signals by the central auditory system.

A weakness of the current LIN model is that its temporal processing is not entirely realistic. Lateral inhibitory connections are realized in vivo by separate interneurons (e.g., horizontal cells in the retina) while this model uses projections from the neurons themselves to realize inhibitory interactions. Another reason is that neural signal processing delays are not represented. Since the leaky integrate-and-fire model is a point model, there would be little benefit to representing interneurons separately in the current LIN model because the additional delays incurred by such interneurons were not modeled. A significant improvement that could be made to this spiking LIN model would be to incorporate delays to simulate action potentials having to traverse a neuron and across synapses and adding interneurons. These modifications would improve the accuracy of the LIN's processing of temporal cues, thereby making the model more realistic. The modified model would thus improve the accuracy of computational investigations of bursting behaviour in neural activity and synchrony between neurons. Such studies might elucidate Moller's hypothesis on the central mechanism of tinnitus (described briefly in Section 3.2).

Bibliography

Arbib, M. A. (2002). *The Handbook of Brain Theory and Neural Networks* (2nd ed.). MIT Press.

Barlow, H. B. (1953). Summation and inhibition in the frog's retina. *J. Physiol.*, *119*, 69–88.

Basta, D. & Vater, M. (2003). Membrane-based gating mechanism for auditory information in the mouse inferior colliculus. *Brain Research*, *968*, 171–178.

Bruce, I. C., Bajaj, H. S., & Ko, J. (2003). Lateral-inhibitory-network models of tinnitus. In D. D. Feng & E. R. Carson (Eds.), *Proceedings of the 5th IFAC Symposium on Modelling and Control in Biomedical Systems* (pp. 359–363). Oxford, UK: Elsevier Ltd.

Bruce, I. C., Sachs, M. B., & Young, E. D. (2003). An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses. *J. Acoust. Soc. Am.*, *113*(1), 369–388.

Cacace, A. T. (2003). Expanding the biological basis of tinnitus: crossmodal origins and the role of neuroplasticity. *Hear. Res.*, *175*, 112–132.

Caspary, D. M., Backoff, P. M., Finlayson, P. G., & Palombi, P. S. (1994). Inhibitory inputs modulate discharge rate within frequency receptive fields of anteroventral cochlear neurons. *J. Neurophysiol.*, *72*(5), 2124–2133.

Charand, K. X. (2000). Action potentials. `http://hyperphysics.phy-astr.gsu.edu/hbase/biology/actpot.html` accessed May 20, 2004. In HyperPhysics by C.R. Nave at Georgia State University.

Chen, G. D. & Jastreboff, P. J. (1995). Salicylate-induced abnormal activity in the inferior colliculus of rats. *Hear. Res.*, *82*(2), 158–178.

Davis, K. A. (2002). Evidence of a functionally segregated pathway from dorsal cochlear nucleus to inferior colliculus. *J. Neurophysiol.*, *87*, 1824–1835.

Edwards, B. W. & Wakefield, G. H. (1990). On the statistics of binned neural point processes: the bernoulli approximation and ar representation of the pst histogram. *Biol. Cybern.*, *64*, 145–153.

Eggermont, J. J. (2003). Central tinnitus. *Auris, Nasus, Larynx*, *30*, S7–S12.

Eggermont, J. J. & Kenmochi, M. (1998). Salicylate and quinine selectively increase spontaneous firing rates in secondary auditory cortex. *Hear. Res.*, *117*, 149–160.

Ehret, G. & Romand, R. (1997). *The Central Auditory System.* Oxford University Press.

Geary, J. (1998). Fighting phantom noise. *Time Magazine*, *152*(4).

Gerken, G. M. (1996). Central tinnitus and lateral inhibition: an auditory brainstem model. *Hear. Res.*, *97*, 75–83.

Guyton, A. C. & Hall, J. E. (2000). *Textbook of Medical Physiology* (10th ed.). Harcourt Brace and Company Canada, Ltd.

Hartline, H. K. (1949). Inhibition of activity of visual receptors by illuminating nearby retinal areas in teh limulus eye. *Federation Proceedings*, *8*(1), 69.

Hartline, H. K. (1967). Visual receptors and retinal interaction. *Nobel Lecture.*

Hartline, H. K. (1974). *Studies in Excitation and Inhibition in the Retina.* New York, NY: The Rockefeller University Press.

Hartline, H. K. & Ratliff, F. (1958). Spatial summation of inhibitory influences in the eye of limulus, and the mutual interaction of receptor units. *J. Gen. Physiol.*, *41*(5), 1049–1066.

Hartline, H. K., Wagner, H. G., & Ratliff, F. (1956). Inhibition in the eye of limulus. *J. Gen. Physiol.*, *39*(5), 651–673.

Hodgkin, A. L. & Huxley, A. F. (1952). A quantitative description of the membrane current and its application to conduction and excitation in a nerve. *J. Physiol. (London)*, *117*, 500–544.

Houtgast, T. (1972). Psychophysical evidence for lateral inhibition in hearing. *J. Acoust. Soc. Am.*, *51*(6, Part 2), 1885–1894.

Irvine, D. R. F. & Rajan, R. (1997). Injury-induced reorganization of frequency maps in adult auditory cortex: The role of unmasking of normally-inhibited inputs. *Acta Otolaryngol. (Stockh.) Suppl.*, *532*, 39–45.

Jack, J. J. B., Noble, D., & Tsien, R. W. (1975). *Electric current flow in excitable cells.* Oxford: Clarendon Press.

Junqueira, L. C. & Carneiro, J. (2002). *Basic Histology* (10th ed.). McGraw-Hill Education (Lange Medical Books).

Kaiser, P. K. (1996–2002). *The Joy of Visual Perception.* Book in progress. `http://www.yorku.ca/eye/machband.htm`.

Kaltenbach, J. A. (2000). Neurophysiologic mechanisms of tinnitus. *J. Am. Acad. Audiol.*, *11*, 125–137.

Katsuki, Y., Sumi, T., Uchiyama, H., & Watanabe, T. (1958). Electric response of auditory neurons in cat to sound stimulation. *J. Neurophysiol.*, *21*, 569–588.

Katsuki, Y., Watanabe, T., & Maruyama, N. (1959). Activity of auditory neurons in upper levels of brain of cat. *J. Neurophysiol.*, *22*, 343–359.

Katsuki, Y., Watanabe, T., & Suga, N. (1959). Interaction of auditory neurons in response to two sound stimuli in cat. *J. Neurophysiol.*, *22*, 603–623.

Koch, C. (1999). *Biophysics of Computation: Information Processing in Single Neurons.* Computational Neuroscience. New York, NY: Oxford University Press.

Koch, U. & Grothe, B. (2003). Hyperpolarization-activated current (i_h) in the inferior colliculus: Distribution and contribution to temporal processing. *J. Neurophysiol.*, *90*, 3679–3687.

Kral, A. & Majernik, V. (1996). On lateral inhibition in the auditory system. *Gen. Physiol. Biophys.*, *15*, 109–127.

Kretzberg, J., Egelhaaf, M., Kristan, W., & Sejnowski, T. (2004). Transmitting more than spikes. In *Computational and Systems Neuroscience* (pp. 105).

Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *J. Neurophysiol.*, *16*, 37–68.

Lapicque, L. (1907). Recherches quantitatives sur l'excitation electrique des nerfs traitee comme une polarization. *J. Physiol. Pathol. Gen.*, *9*, 620–635.

Levine, J. S. & Miller, K. R. (1994). *Biology: Discovering Life* (2nd ed.). Lexington, MA: D.C. Heath and Company.

Levine, R. A., Abel, M., & Cheng, H. (2003). CNS somatosensory-auditory interactions elicit or modulate tinnitus. *Exp. Brain Res.*, *153*, 643–648.

Liberman, M. C. & Kiang, N. Y. (1978). Acoustic trauma in cats. *Acta Otolaryngol. (Stockh.) Suppl.*, *358*, 1–63.

Lockwood, A. H., Salvi, R. J., & Burkard, R. F. (2002). Tinnitus. *New England Journal of Medicine*, *347*(12), 904–910.

Mach, E. (1865). über die wirkung der räumlichen vertheilung des lichtreizes auf die netzhaut. *Classe der kaiserlichen Akademic der Wissenschaften, Wien*, *52*(2), 303–322.

Miller, R. L., Schilling, J. R., Franck, K. R., & Young, E. D. (1997). Effects of acoustic trauma on the representation of the vowel /ɛ/ in cat auditory nerve fibres. *J. Acoust. Soc. Am.*, *101*, 3602–3616.

Moller, A. R. (1984). Pathophysiology of tinnitus. *Ann. Otol. Rhinol. Laryngol.*, *93*, 39–43.

Moller, A. R. (1995). Pathophysiology of tinnitus. In J. A. Vernon & A. R. Moller (Eds.), *Mechanisms of Tinnitus* (pp. 207–218). Boston, MA: Allyn and Bacon.

Muhlnickel, W., Elbert, T., Taub, E., & Flor, H. (1998). Reorganization of auditory cortex in tinnitus. *Proc. Natl. Acad. USA*, *95*, 10340–10343.

Norena, A., Micheyl, C., Chery-Croze, S., & Collet, L. (2002). Psychoacoustic characterization of the tinnitus spectrum: Implications for the underlying mechanisms of tinnitus. *Audiol. Neuro-otol.*, *7*(6), 358–369.

Ochi K. & Eggermont, J. J. (1996). Effects of salicylate on neural activity in cat primary auditory cortex. *Hear. Res.*, *95*, 63–76.

Ochi K. & Eggermont, J. J. (1997). Effects of quinine on neural activity in cat primary auditory cortex. *Hear. Res.*, *105*, 105–118.

Oertel, D. & Young, E. D. (2004). What's a cerebellar circuit doing in the auditory system? *Trends. Neurosci.*, *27*(2), 104–110.

Palombi, P. S. & Caspary, D. M. (1996). Gaba inputs control discharge rage primarily within frequency receptive fields of inferior colliculus neurons. *J. Neurophysiol.*, *75*(6), 2211–2219.

Penner, M. J. (1992). Linking spontaneous otoacoustic emissions and tinnitus. *Br. J. Audiol.*, *26*(2), 91–96.

Pinchoff, R. J., Burkard, F. R., Salvi, R. J., Coad, M. L., & Lockwood, A. H. (1998). Modulation of tinnitus by voluntary jaw movements. *Am. J. Otol.*, *19*(6), 785–789.

Plinkert, P. K., Gitter, A. H., & Zenner, H. P. (1990). Tinnitus associated spontaneous otoacoustic emissions. active outer hair cell movements as common origin? *Acta Otolaryngol., 110*, 342–347.

Protopapas, A. D., Vanier, M., & Bower, J. M. (1999). Simulating large networks of neurons. In C. Koch & I. Segev (Eds.), *Methods in Neuronal Modeling: From Ions to Networks* (2nd ed.). chapter 12, (pp. 461–498). Cambridge, MA: MIT Press.

Quatieri, T. F. (2002). *Discrete-Time Speech Signal Processing: Principles and Practice.* Upper Saddle River, NJ: Prentice Hall.

Rajan, R., Irvine, D. R. F., Wise, L. Z., & Heil, P. (1993). Effect of unilateral partial cochlear lesions in adult cats on the representation of lesioned and unlesioned cochleas in primary auditory cortex. *J. Comp. Neurol., 338*, 17–49.

Ramachandran, R., Davis, K. A., & May, B. J. (1999). Single-unit responses in the inferior colliculus of decerebrate cats: I. Classification based on frequency response maps. *J. Neurophysiol., 82*, 152–163.

Ratliff, F. (1965). *Mach Bands: Quantitative studies on neural networks in the retina.* Psychology. San Francisco, CA: Holden-Day Inc.

Robertson, D. & Irvine, D. R. (1989). Plasticity of frequency organization in auditory cortex of guinea pigs with partial unilateral deafness. *J. Comp. Neurol., 282*, 456–471.

Sachs, M. B., Bruce, I. C., Miller, R. L., & Young, E. D. (2002). Biological basis of hearing-aid design. *Ann. Biomed. Eng., 30*, 157–168.

Seki, S. & Eggermont, J. J. (2003). Changes in spontaneous firing rate and neural synchrony in cat primary auditory cortex after localized tone-induced hearing loss. *Hear. Res., 180*, 28–38.

Shamma, S. A. (1985). Speech processing in the auditory system II: Lateral inhibition and the central processing of speech evoked activity in the auditory nerve. *J. Acoust. Soc. Am., 78*(5), 1622–1632.

Stein, R. B. (1967). The frequency of nerve action potential generated by applied currents. *Proc. R. Soc. London, B, 167*, 64–86.

Suga, N. (1995). Sharpening of frequency tuning by inhibition in the central auditory system: tribute to yasuji katsuki. *Neurosci. Res., 21*, 287–299.

Suga, N. & Ma, X. (2003). Multiparametric corticofugal modulation and plasticity in the auditory system. *Neuroscience, 4*, 783–794.

Tomita, T. (1958). Mechanism of lateral inhibition in eye of limulus. *J. Neurophysiol.*, *21*(5), 419–429.

Torterolo, P., Falconi, A., Morales-Cobas, G., & Velluti, R. A. (2002). Inferior colliculus unitary activity in wakefulness, sleep and under barbiturates. *Brain Res.*, *955*(1–2), 9–15.

Tuckwell, H. C. (1988). *Introduction to theoretical neurobiology: Volume 1, Linear cable theory and dendritic structure.* Cambridge, UK: Cambridge University Press.

von Békésy, G. (1961). Concerning the pleasures of observing, and the mechanics of the inner ear. *Nobel Lecture.*

von Békésy, G. (1967a). Mach band type lateral inhibition in different sense organs. *J. Gen. Physiol.*, *50*, 519–532.

von Békésy, G. (1967b). *Sensory Inhibition.* Princeton, NJ: Princeton University Press.

Yang, L., Pollak, G. D., & Resler, C. (1992). Gabaergic circuits sharpen tuning curves and modify response properties in the mustache bat inferior colliculus. *J. Neurophysiol.*, *68*(5), 1760–1774.

Young, E. D. & Sachs, M. B. (1979). Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. *J. Acoust. Soc. Am.*, *66*(5), 1381–1403.

Zenner, H. P. & Ernst, A. (1995). Cochlear motor tinnitus, transduction tinnitus, and signal transfer tinnitus: Three models of cochlear tinnitus. In J. A. Vernon & A. R. Moller (Eds.), *Mechanisms of Tinnitus* (pp. 237–254). Boston, MA: Allyn and Bacon.

Zhang, X., Heinz, M. G., Bruce, I. C., & Carney, L. H. (2001). A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. *J. Acoust. Soc. Am.*, *109*(2), 648–670.

Appendix A

The Bernoulli Approximation of a Poisson Process

A Poisson process is a point process where the random variable is the number of events (which are spikes in this study) that occur in an interval (size of the time bins), Δ. The random variables are independent so long as the time intervals are non-overlapping. The Bernoulli approximation of a Poisson process is a binomial process that is often used when dealing with discrete time. In the context of this investigation, the Bernoulli approximation assumes that Δ, is small enough that the firing rate, $\lambda(t)$, is constant within the time bins and no more than one spike can occur in one time bin. Each time bin contains a discrete random variable with a probability that a spike will occur in that bin of $\lambda \cdot \Delta$ (a Bernoulli distribution). Equation A.18 (c.f. Edwards & Wakefield, 1990, Eqns.3 and 4) relates the probability of at least one spike occurring in a bin from the Poisson process (left side) to that of the Bernoulli approximation (right side) via a Taylor Series expansion of the decaying exponential.

$$\int_0^{\Delta} \lambda e^{-\lambda\Delta} dt = 1 - e^{-\lambda\Delta} = 1 - [1 - \lambda\Delta + \frac{(\lambda\Delta)^2}{2!} - \frac{(\lambda\Delta)^3}{3!} + \cdots] \approx \lambda\Delta \qquad \text{(A.18)}$$

From Equation A.18 and Figure A.1 it can be seen that the approximation only holds for small $\lambda\Delta$. The expected number of spikes within a bin (the first order moment) is

the same for both the Bernoulli approximation and the Poisson process, regardless of the size of $\lambda\Delta$. However, for higher order statistics such as the variance, the Bernoulli approximation becomes inaccurate as $\lambda\Delta$ increases, as shown in Figure A.2. Hence, very small $\lambda\Delta$s are used in this study.

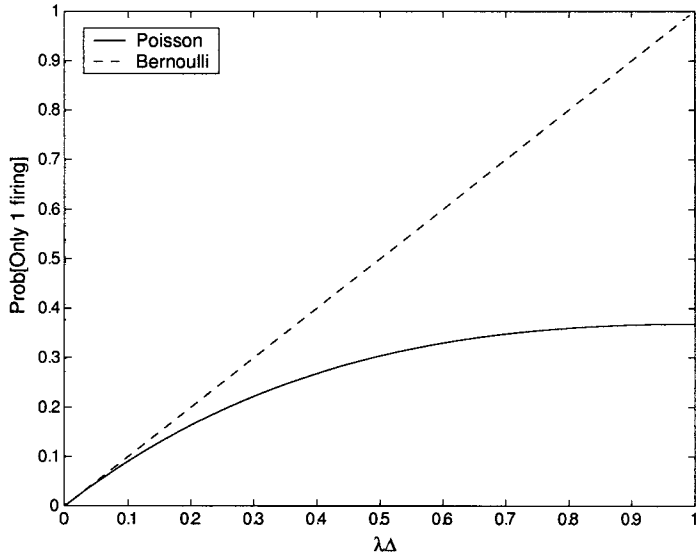


Figure A.1: Comparison of the probability of a single spike occurring within a time bin as derived using the Bernoulli approximation versus the Poisson process.

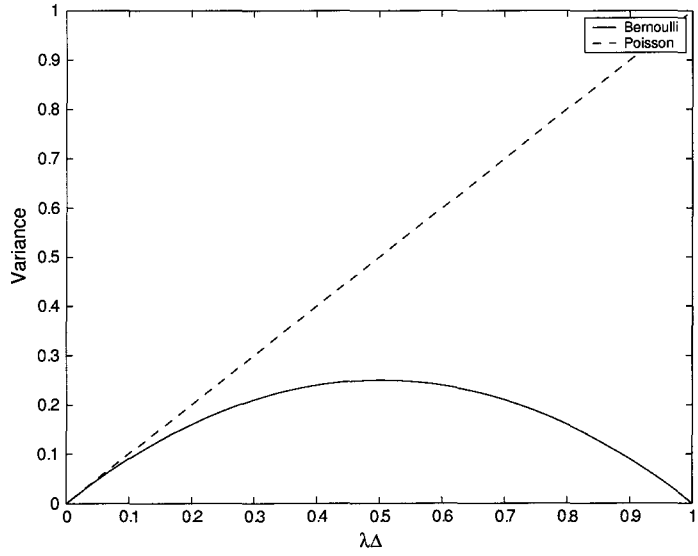


Figure A.2: Comparison of the variance in the number of spikes occurring within a time bin as derived using the Bernoulli approximation versus the Poisson process.

Appendix B

Numerical Methods Employed

An ordinary differential equation $y'(t) = f(t, y(t))$ with the initial condition $y(t_0) = y_0$, has the exact solution, $y(t)$, given by:

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} y'(t)dt \tag{B.19}$$

according to the fundamental theorem of calculus. y_{n+1}, an approximation of $y(t)$, is found by approximating the definite integrals. There are a number of ways to approximate the definite integrals, including the two methods described herein.

B.1 Euler's Method

Euler's method computes y_{n+1} by assuming that the rate of change, $f(t, y(t))$, remains constant over a short period of time, Δ. The definite integral is then approximated by the area of a rectangle giving the equation:

$$y_{n+1} = y_n + f(y_n) \cdot \Delta \tag{B.20}$$

B.2 Fourth-order Runge-Kutta Algorithm

y_{n+1} can be found using Simpson's rule (i.e., approximating the definite integral with the area under a parabola). Hence, acquiring y_{n+1} by means of the fourth-order Runge-Kutta algorithm involves computing the four numbers k_1 through k_4 given by:

$$\begin{aligned} k_1 &= \Delta f(y_n) \\ k_2 &= \Delta f(y_n + \frac{1}{2}k_1) \\ k_3 &= \Delta f(y_n + \frac{1}{2}k_2) \\ k_4 &= \Delta f(y_n + k_3) \end{aligned}$$

$$y_{n+1} = y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \tag{B.21}$$

These four numbers are then used to approximate the definite integral of Equation B.19 to give y_{n+1} as shown in Equation B.21.

The main advantage of using the fourth-order Runge-Kutta method over lower order methods such as the Euler or Improved Euler methods is that greater accuracy can be achieved with the same size of time steps. The disadvantage of higher order methods is that they are more computationally expensive. The fourth-order Runge-Kutta method seems to provide a good balance between these two considerations.

Appendix C

MATLAB Code

A sample of the set of scripts with which simulations were run is provided in this appendix. In all of the simulations, three functions were called: weight.m to generate $\boldsymbol{V}$; weight_2gauss.m to generate $\boldsymbol{W}$; and Greenwood.m to calculate the characteristic frequencies associated with the spatial location of the neurons along the basilar membrane or vice versa. The first section contains the scripts for uneven spontaneous-type input: the preliminary version of the model, spiking.m; the 'base case', EEopt.m; the non-spiking version of the model, NONspiking.m; and the functions that they call (weight.m, weight_2gauss.m, and Greenwood.m). The scripts used to generate the contour plots shown in Section 5.1 were looped versions of the 'base case'. An example of such a script run on a GRID server is given in Section C.1.2. The second section contains the scripts for the processing of speech-like input. These simulations were run by nested function calls. The main script, hear.m, loads the sound pressure wave and initializes the parameters, then calls the function earandlin.m to compute the response of the ear and LIN. earandlin.m then initializes the fixed parameters and calls ear.m that passes only the desired input and output data while discarding the unnecessary information returned by the Bruce et al. model. The function ear.m also sums the results of repeated calls of the Bruce et al. model to acquire a representation of convergent input of the same characteristic frequency to the spiking LIN model. Similar to ear.m, earandlin.m returns the desired results while discarding the fixed

parameters and environment variables that need not be stored. Finally, the main script contains commands to store the results and produce figures for analysis.

C.1 Scripts for Simulations with Spontaneous-type Input

C.1.1 spiking.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% spiking.m                                                                      %
% -- preliminary version of the recurrent LIN (spiking) model in the  %
%    central auditory pathways                                                   %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear all;
close all;
clk=clock;
h=waitbar(0,'Starting...','Name','spikingLIN.m Progress');

%% Create time steps [s]
tsp = [0:1e-4:5];  % Max. run time is 7 secs (for my computer)

%% Set number of neurons to be used (>50)
numofneurons = 100;

%% Membrane Time Constant [s]
mtau = 1e-3;

%% Create random input
rand('seed',0);
inputs_uniform = rand(numofneurons,length(tsp));

%% Gerken's mean, spontaneous firing rate
%% [spikes/s]
%% -- impaired hearing when highspon > lowspon
%% -- normal hearing when highspon = lowspon
highspon=200;
lowspon=20;
down=[highspon (highspon+lowspon)/2 lowspon];
wave = gausswin(10,4)*210+lowspon;
bigsig(1:ceil(numofneurons/2)-1)=highspon;
bigsig(ceil(numofneurons/2):ceil(numofneurons/2)+2)=down(1:end);
bigsig(ceil(numofneurons/2)+3:numofneurons) = lowspon;
bigsig(ceil(numofneurons/1.2):ceil(numofneurons/1.2)+length(wave)-1)=wave(1:end);
bigsig=bigsig.';

%% Random input corresponds to Gerken's average firing rate
```

```
inputs = inputs_uniform < repmat(bigsig*tsp(2),1,length(tsp));

%% Define Excitatory Post-Synaptic Current Filter (normalized)
EPSCalpha = 5;
EPSC=(1/mtau*tsp(1:round(0.01/tsp(2))).*exp(-EPSCalpha/mtau*tsp(1:round(0.01/tsp(2)))))/...
    max(1/mtau*tsp(1:round(0.01/tsp(2))).*exp(-EPSCalpha/mtau*tsp(1:round(0.01/tsp(2)))));

%% Define Inhibitory Post-Synaptic Current Filter (normalized)
IPSCalpha = 2;
IPSC=(1/mtau*tsp(1:round(0.01/tsp(2))).*exp(-IPSCalpha/mtau*tsp(1:round(0.01/tsp(2)))))/...
    max(1/mtau*tsp(1:round(0.01/tsp(2))).*exp(-IPSCalpha/mtau*tsp(1:round(0.01/tsp(2)))));

%% Shape excitatory spikes
inputs2 = filter(EPSC,1,inputs,[],2);

%% Effectiveness coeffients of inputs
%% --- widthofwindow=1 b/c no lateral excitation
%% --- selfsyn=1 b/c not using this function to account for inhibition
V = weight(1,size(inputs2,1),1);

%% Coefficients of inhibitory interactions (2 gaussian humps as in Gerken's model)
%% --- a = arbitrary coefficient describing weight of inhibitory compared
%       to excitatory interactions
a = 2;
W = weight_2gauss(13,size(inputs2,1));
W = a*W/max(sum(W,2));

%% Initialization
spikes=zeros(numofneurons,length(tsp));  % keeps track of spikes generated
spikes2=zeros(numofneurons,length(tsp));  % for lateral inhibition
[spikes2(:,1),Zi] = filter(IPSC,1,spikes(:,1),[],2);  % all 0s
y = zeros(numofneurons,length(tsp));  % [V]
y(:,1) = V*inputs2(:,1)*tsp(2)/mtau;  % set initial condition of neurons [V]
ythres= 0.25;  % threshold potential [V]
tref = 2e-3;  % refractory period
trefsamp = round(tref/tsp(2));  % refractory period [# of samples]

%% Calculate output membrane potential [V] using Shamma's rate equation
for lp=2:length(tsp);  %% lp=n+1 index

    fn = [V*inputs2(:,lp-1) - W*spikes2(:,lp-1) - y(:,lp-1)] /mtau;  % rate eqn

    %% Using Euler's method of approximation (rectangles)
    % y(:,lp)=y(:,lp-1)*(1-1/mtau*tsp(2))+V*inputs2(:,lp)/mtau*tsp(2)-...
    %   W*spikes2(:,lp-1)/mtau*tsp(2);
    % y(:,lp)=y(:,lp-1) + tsp(2)*fn; % same eqn

    %% Using the Runge-Kutta method of approximation (parabolas)
    K1 = fn*tsp(2);
```

```
    K2 = (fn - 0.5*K1/mtau)*tsp(2);
    K3 = (fn - 0.5*K2/mtau)*tsp(2);
    K4 = (fn - K3/mtau)*tsp(2);
    y(:,lp) = y(:,lp-1) + (1/6)*(K1 + 2*K2 + 2*K3 + K4);

    %% Determine if 'lp' is in a refractory period.
    %% If so, set membrane potential to zero.
    % Are there any spikes within the last tref secs?
    p=any(spikes(:, max(lp-trefsamp,1):lp-1),2);
    p=p';  % column to row vector
    m=find(p);
    y(m,lp)=0;

    %% If not in refractory period and membrane potential is > threshold,
    %% generate a spike
    I=find(y(:,lp)>ythres);
    y(I,lp)=5;
    spikes(I,lp)=1;

    [spikes2(:,lp),Zi] = filter(IPSC,1,spikes(:,lp),Zi,2);

    waitbar(lp/length(tsp),h,'Computing...');
end

%% Calculate CF associated with each neuron:
f = zeros(1,numofneurons);
for i = 1:numofneurons
    f(i)=Greenwood(1-i/numofneurons, 'freq', 'man');
end

close(h)  % close waitbar

%% Plot Graph:
figure
semilogx(f,sum(inputs,2)./tsp(end),'r-.');
hold on;
semilogx(f, sum(spikes,2)./tsp(end),'k-');
xlabel('Characteristic Frequency [Hz]');
ylabel('Mean Firing Rate (spikes/s)');
grid;
xlim([f(1) f(end)]);
legend('input','output')

%% Print variables used on graph:
var1=strcat('tsp=',num2str(tsp(2)));
var2=strcat('mtau=', num2str(mtau));
var3=strcat('a=',num2str(a));
var4=strcat('EPSCalpha=', num2str(EPSCalpha));
var5=strcat('IPSCalpha=', num2str(IPSCalpha));
```

```
var6=strcat('bigsig=', num2str(bigsig(1)));
var=strvcat(var1,var2,var3,var4,var5,var6);
text(max(f)*0.75, max(mean(y,2))*0.75, var);
```

C.1.2 EEopt.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% EEopt.m      (Revised Base Case)                                          %
% -- recurrent LIN (spiking) in central auditory pathways                   %
% -- poisson input represents spontaneous activity from auditory nerve%
% ** to be run on GRID server as an array job for ptsdown                   %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear all;
close all;
clk=clock;
% h=waitbar(0,'Starting...','Name','EEopt.m Progress');

% ptsdown=[2:2:20];

%% Create time steps [s]
tsp = [0:2e-5:5];  % simulation length

%% Set number of neurons to be used (>50)
numofneurons = 100;

%% Membrane Parameters:
mtau = 1.5e-3;
C=8e-12;  % Membrane capacitance [F]

%% Create random input
rand('seed',0);
inputs_uniform = rand(numofneurons,length(tsp));

%% Gerken's mean, spontaneous firing rate [spikes/s]
%% -- impaired hearing when highspon > lowspon
%% -- normal hearing when highspon = lowspon
highspon=200;
lowspon=20;
down=[highspon (highspon+lowspon)/3 lowspon];
% wave = gausswin(10,4)*210+lowspon;  % tone
bigsig(1:ceil(numofneurons/2)-1)=highspon;
bigsig(ceil(numofneurons/2):ceil(numofneurons/2)+2)=down(1:end);
bigsig(ceil(numofneurons/2)+3:numofneurons) = lowspon;
% bigsig(ceil(numofneurons/1.2):ceil(numofneurons/1.2)+length(wave)-1)=wave(1:end);
bigsig=bigsig.';

%% Random input corresponds to Gerken's average firing rate
inputs = inputs_uniform < repmat(bigsig*tsp(2),1,length(tsp));
```

```
%% Define Excitatory Synapse Parameters:
EsynE=0.1;  % excitatory reversal potential [V]; must be > ythres
GEalpha = 11;
GEdur=0.01;
GE = (GEalpha/mtau/10)^2*tsp(1:round(GEdur/tsp(2))).*...
    exp(-GEalpha/mtau*tsp(1:round(GEdur/tsp(2))));
sfac=3.0365e-010;  % based on GEalpha=9 and GIalpha=0.5
GE=sfac*GE;  % scaled excitatory conductance [S] - scaling factor much match GI

%% Define Inhibitory Synapse Parameters:
EsynI=-0.02;  % inhibitory reversal potential [V]; < resting potential
% EsynI=0;  % Shunting Inhibition reversal potential [V]; ~= yrest
GIalpha=0.5;
GIdur=0.01;
GI = (GIalpha/mtau/10)^2*tsp(1:round(GIdur/tsp(2))).*...
    exp(-GIalpha/mtau*tsp(1:round(GIdur/tsp(2))));
GI=sfac*GI;  % scaled inhibitory conductance [S] - scaling factor must match GE

%% Shape excitatory spikes
inputs2=filter(GE,1,inputs,[],2);  % shaped conductance

%% Effectiveness coeffients of inputs
%% -- widthofwindow=1 b/c no lateral excitation
%% -- selfsyn=1 b/c not using this function to account for inhibition
V = weight(1,size(inputs2,1),1);

%% Coefficients of inhibitory interactions (2 gaussian humps as in Gerken's model)
%% -- a = arbitrary coefficient describing weight of inhibitory compared
%%    to excitatory interactions
% a = 36;  % to get back original amplitude of inhibition as when normalized
a=32;
W = weight_2gauss(13,size(inputs2,1));
W = a*W/max(sum(W,2));

%% Initialization
spikes=zeros(numofneurons,length(tsp));  % keeps track of spikes generated
spikes2=zeros(numofneurons,length(tsp));  % for lateral inhibition
[spikes2(:,1),Zi]=filter(GI,1,spikes(:,1),[],2);  % all 0s
y=zeros(numofneurons,length(tsp));  % Membrane potential [V]
y(:,1) = V*inputs2(:,1)*tsp(2)/mtau;  % set initial condition of neurons [V]
tref=2e-3;  % absolute refractory period
trefsamp=round(tref/tsp(2));  % refractory period [# of samples]
spikevalue=0.15;
% nominal threshold potential [V]:
yth=0.015;
ythres=yth*ones(numofneurons,length(tsp)+2*trefsamp);

%% Calculate output membrane potential [V] using
```

```
%% Shamma's rate equation modified with Koch p19
%%  - modified to account for synaptic conduction
for lp_time=2:length(tsp);  % lp_time=n+1 index

    %% Determine if 'lp_time' is in a refractory period.
    %% If so, set threshold potential higher.
    % Did a spike occur in the last time step?
    p=any(spikes(:,lp_time-1),2);
    p=p';  % column to row vector
    if any(p)
        m=find(p);
        ythres(m,lp_time:lp_time+trefsamp)=5;
        ythres(m,lp_time+trefsamp+1:lp_time+2*trefsamp)=...
            repmat(5*exp(-3.5*[0:trefsamp-1]./trefsamp),length(m),1);
    end % if p

    fn = V*(inputs2(:,lp_time-1).*(EsynE-y(:,lp_time-1))./C)...
        +W*(spikes2(:,lp_time-1).*(EsynI-y(:,lp_time-1))./C)...
        -y(:,lp_time-1)./mtau;    % rate eqn

    %% Using the 4th-order Runge-Kutta method of approximation (parabolas):
    K1 = fn*tsp(2);
    K2 = (fn-0.5*K1/mtau)*tsp(2);
    K3 = (fn-0.5*K2/mtau)*tsp(2);
    K4 = (fn-K3/mtau)*tsp(2);
    y(:,lp_time)=y(:,lp_time-1) + (1/6)*(K1 + 2*K2 + 2*K3 + K4);

    %% Set refractory period after a spike:
    % Are there any spikes within the last tref secs?
    ph=any(spikes(:, max(lp_time-trefsamp,1):lp_time-1),2);
    ph=ph';  % column to row vector
    mh=find(ph);
    y(mh,lp_time)=0;

    %% If not in refractory period and membrane potential is > threshold,
    %% generate a spike:
    I=find(y(:,lp_time)>ythres(:,lp_time));
    y(I,lp_time)=spikevalue;
    spikes(I,lp_time)=1;
    % shaped inhibitory conductance:
    [spikes2(:,lp_time),Zi]=filter(GI,1,spikes(:,lp_time),Zi,2);

%     waitbar(lp_time/length(tsp),h,'Computing...');
end  % for lp_time

%% Calculate CF associated with each neuron:
f = zeros(1,numofneurons);
for i = 1:numofneurons
    f(i)=Greenwood(1-i/numofneurons, 'freq', 'man');
```

```
end

elptime=etime(clock,clk)/3600;  % Time elapsed [hours]
% close(h)  % close waitbar
save EEopt_inslope.mat inputs spikes tsp elptime f a GEalpha GIalpha mtau yth C

disp('Simulation is complete and has been saved.')

%% Plot Graphs:
% figure
% plot(tsp,y(10,:),'k--');
% title('Neuron #10');
% xlabel('Time [s]');
% ylabel('Membrane Potential [V]')

% figure
% % hold on;
% semilogx(f, sum(spikes,2)./tsp(end));
% hold on
% semilogx(f,sum(inputs,2)./tsp(end),'r-.');
% % title(['\bf 4-point edge']);
% xlabel('Characteristic Frequency [Hz]');
% ylabel('Mean Spike Rate [spikes/s]');
% xlim([f(1) f(end)]);
% legend('Output','Input')
% grid;
```

C.1.3 NONspiking.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% NONspiking.m  (Jennifer Ko, March 2004)
% -- recurrent LIN in central auditory pathways
% -- poisson input represents spontaneous activity from auditory nerve
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear all; close all; clk=clock;
h=waitbar(0,'Initializing...','Name','NONspiking.m Progress');

%% Create time steps [s]
tsp = [0:2e-5:5];

%% Set number of neurons to be used (>50)
numofneurons = 100;

%% Membrane Parameters:
mtau = 1e-3; % Membrane time constant [s]

%% Create random input
inputs_uniform = rand(numofneurons,length(tsp));
```

```
%% Gerken's mean, spontaneous firing rate
%% [spikes/s]
%% -- impaired hearing when highspon > lowspon
%% -- normal hearing when highspon = lowspon
highspon=400; lowspon=20; down=[highspon (highspon+lowspon)/2
lowspon]; bigsig=zeros(1,numofneurons);
% wave = gausswin(10,4)*210+lowspon;  % tone
bigsig(1:ceil(numofneurons/2)-1)=highspon;
bigsig(ceil(numofneurons/2):ceil(numofneurons/2)+2)=down(1:end);
bigsig(ceil(numofneurons/2)+3:numofneurons) = lowspon;
% bigsig(ceil(numofneurons/1.2):ceil(numofneurons/1.2)+length(wave)-1) = wave(1:end);
bigsig=bigsig.';

%% Random input corresponds to Gerken's average firing rate
inputs = inputs_uniform < repmat(bigsig*tsp(2),1,length(tsp));

%% Define unitary input
Ialpha=9; dur = 0.005; I =
0.5*(Ialpha/mtau/10)^2*tsp(1:round(dur/tsp(2))).*...
    exp(-Ialpha/mtau*tsp(1:round(dur/tsp(2))));

%% Shape excitatory input
inputs2=filter(I,1,inputs,[],2); % injected voltage [V]

%% Effectiveness coeffients of inputs across BFs
%% -- widthofwindow=1 b/c no lateral excitation
%% -- selfsyn=1 b/c not using this function to account for inhibition
V = weight(1,numofneurons,1);

%% Coefficients of inhibitory interactions (2 gaussian humps as in Gerken's model) across BFs
%% -- a = arbitrary coefficient describing weight of inhibitory compared to excitatory intera
a=1;
W = weight_2gauss(13,numofneurons); W = a*W/max(sum(W,2));

%% Initialization
y=zeros(numofneurons,length(tsp));  % Membrane potential [V]
y(:,1) = V*inputs2(:,1)*tsp(2)/mtau;  % set initial condition of neurons [V]

%% Calculate output membrane potential [V] using Shamma's rate equation (4)
%% with non-linearity (i.e. arctan)
for lp=2:length(tsp);

    fn = (V*inputs2(:,lp-1)-W*y(:,lp-1)-y(:,lp-1))./mtau;    % rate eqn

    %% Using the 4th-order Runge-Kutta method of approximation (parabolas):
    K1 = fn*tsp(2);
    K2 = (fn-0.5*K1/mtau)*tsp(2);
    K3 = (fn-0.5*K2/mtau)*tsp(2);
```

```
    K4 = (fn-K3/mtau)*tsp(2);
    y(:,lp)=y(:,lp-1)+(1/6)*(K1+2*K2+2*K3+K4);

    waitbar(lp/length(tsp),h,'Computing...');
end

%% Calculate CF associated with each neuron:
f = zeros(1,numofneurons); for i = 1:numofneurons
    f(i)=Greenwood(1-i/numofneurons, 'freq', 'man');
end

elptime=etime(clock,clk)/3600;  % Time elapsed [hours]
close(h)  % close waitbar

%% Plot Graphs:
figure plot(tsp,y(5,:)); title('Neuron #5'); xlabel('Time [s]');
ylabel('Membrane Potential [V]')

% figure
% semilogx(f, mean(y,2));
% % hold on;
% % semilogx(f, mean(inputs2,2),'r:');
% xlabel('Characteristic Frequency [Hz]');
% ylabel('Average Membrane Potential [V]');
% xlim([f(1) f(end)]);
% legend('Output','Input')
% grid on;
```

C.1.4 weight.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% weight.m function                                                     %
% -- calculates weight coefficients determining                        %
%    the effectiveness of the inputs                                    %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function V = weight(widthofwindow,numofneurons,selfsyn)

%% Create a Gaussian wave:
w=gausswin(widthofwindow).';

c=ceil(widthofwindow/2);

%% Determine if self-synapsing or not for inhibition
% if selfsyn==0
%     w(c) = 0;
% end

%% Create convolution matrix for Gaussian filter
```

```
Wtmp = convmtx(w,numofneurons-(widthofwindow-1));

%% Add rows to make convolution matrix square
j=size(Wtmp,1);
d=2;
p=1;

while d <= c
    Wtmp = [Wtmp(p,d:end) zeros(1,d-1);  Wtmp; zeros(1,d-1) Wtmp(j,1:end-p)];
    d=d+1;
    p=p+1;
    j=j+1;
end


%% Account for missing excitation to first and last neurons
rowsum=sum(Wtmp,1);
check=any(rowsum<max(rowsum),1);
ind=find(check);
if sum(ind>0)
    d=2;
    p=1;
    while p<c
        x=[Wtmp(p,d:end) zeros(1,d-1)];
        Wtmp(p,:)=Wtmp(p,:)+x;
        z=[zeros(1,d-1) Wtmp(numofneurons-p+1, 1:end-d+1)];
        Wtmp(numofneurons-p+1,:)=Wtmp(numofneurons-p+1,:)+z;
        d=d+2;
        p=p+1;
    end
end

V = Wtmp;
```

C.1.5 weight_2gauss.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% weight_2gauss.m function                             %
% -- calculates weight coefficients that determine     %
%    inibitory interactions                            %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function V = weight_2gauss(widthofwindow,numofneurons)

%% Create two consecutive Gaussian waves whose spacing depends on odd/even width:
if mod(widthofwindow,2)==0
    w = [gausswin(widthofwindow/2-1).' 0 0 gausswin(widthofwindow/2-1).'];
else
    w = [gausswin(floor(widthofwindow/2)).' 0 gausswin(floor(widthofwindow/2)).'];
```

```
end

c=ceil(widthofwindow/2);

%% Create convolution matrix for a two-Gaussian filter
Wtmp = convmtx(w,numofneurons-(widthofwindow-1));

%% Add rows to make convolution matrix square
j=size(Wtmp,1);
d=2;
p=1;

while d <= c
    Wtmp = [Wtmp(p,d:end) zeros(1,d-1);  Wtmp; zeros(1,d-1) Wtmp(j,1:end-p)];
    d=d+1;
    p=p+1;
    j=j+1;
end

%% Account for missing lateral inhibition to first and last neurons
rowsum=sum(Wtmp,1);
check=any(rowsum<max(rowsum),1);
ind=find(check);
if sum(ind>0)
    d=2;
    p=1;
    while p<c
        x=[Wtmp(p,d:end) zeros(1,d-1)];
        Wtmp(p,:)=Wtmp(p,:)+x;
        z=[zeros(1,d-1) Wtmp(numofneurons-p+1, 1:end-d+1)];
        Wtmp(numofneurons-p+1,:)=Wtmp(numofneurons-p+1,:)+z;
        d=d+2;
        p=p+1;
    end
end

V = Wtmp;
```

C.1.6 Greenwood.m

```
function y = greenwood(x,type,species,param)
%GREENWOOD  Greenwood's cochlear position function.
%   Usage
%     f = greenwood(x)
%     f = greenwood(x,'freq')
%     f = greenwood(x,'freq',species)
%     f = greenwood(x,'freq','other',param)
%     x = greenwood(f,'dist')
%     x = greenwood(f,'dist',species)
```

```
%     x = greenwood(f,'freq','other',[A a k])
%   Arguments
%     x       = normalised distance from base
%     type    = 'freq' (default) or 'dist'
%     species = 'man' (default) or 'cat' or 'other' or 'model'
%     param   = user supplied parameters for Greenwood's equation = [A a k]
%
%   David Au, 1994.

if nargin < 3,
    species = 'man';
end

if nargin < 2,
    type = 'f';
end

switch species
case 'man'
    A = 165.4;
    a = 2.1;
    k = 1;
case 'cat'
    A = 456;
    a = 2.1;
    k = 0.8;
case 'model'
    A = 320.8729543315967;
    a = 2.17938764568596;
    k = 0.49548345352825;
case 'other'
    A = param(1);
    a = param(2);
    k = param(3);
otherwise
    error('unknown species')
end

switch type
case 'freq'
    y = A * ( 10 .^(a*(1-x)) - k );
case 'dist'
    y = 1 - 1/a * log10( x(:)/A + k);
otherwise
   error('unknown type')
end
```

C.2 Scripts for Simulations with Speech Input

C.2.1 hear.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% hear.m                                                                        %
% -- recurrent LIN (spiking) in central auditory pathways                       %
% -- input is from Ian's model of the auditory nerve                            %
% -- conductance based synapses between Ian's model and LIN                     %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear all;
close all;
clk=clock;

% mcc -x earandlin.m

% h=waitbar(0,'Starting...','Name','Progress');

%% Set number of neurons to be used (>=50):
numofneurons = 100;

%% Calculate CF (aka BF) and distance associated with each neuron:
bf=zeros(1,numofneurons);
x=zeros(1,numofneurons);
for i =1:numofneurons
    % Normalized distance from base of basilar membrane
    x(i)=1-(0.1+0.8*i/numofneurons);
    bf(i)=Greenwood(x(i), 'freq', 'man');  % Best frequency
end

%% Generate stimulus:

%% * bandpassed white noise:
% tsp = [0:1e-5:1];  % create time steps [s] for output from LIN
% f=20e3;  % lower sampling rate for generation of input [Hz]
% reptime=tsp(end)*1e3;  % time between stimulus repetitions [ms]
% rand_noise=0.0001*randn(1,1*reptime*f*1e-3);
% [Num,Den]=cheby1(10,0.5,[0.2 0.4]);  % Filter (passband=2-4kHz) coefficients
% pin=filter(Num,Den,rand_noise);

%% * pure tone:
% tsp = [0:1e-5:1];  % create time steps [s] for output from LIN
% ftone=6000;  % frequency of tone [Hz]
% pin=0.0001*sin(2*pi*ftone*tsp);
% f=round(1/tsp(2));

%% * synthesized steady vowel:
% [pin,f,Nbits]=wavread('eh.wav');  % synthesized vowel
```

```
% pin=0.001*pin';  % column to row and scale to ~15dB SPL
% pin=0.0099*pin';  % 35dBSPL
% pin=0.31*pin';  % 65dB SPL
% pin=9.81*pin';  % 95dB SPL
% tsp=[0:1e-5:length(pin)/f];  % create time steps [s] for output from LIN
% tsp=[0:2e-6:length(pin)/f];  % time steps [s] match those of Ian's model

%% * speech (synthesized male speaker: "Five women played basketball"):
filename = 'sent';
stimdb = 95;  % stimulus intensity in dB SPL
f = 40e3;  % sampling rate [Hz]
pin=resample(load(['speech/' filename '500k']),f,500e3);  % downsampled input
pin=10^(stimdb/20)*pin';  % unramped stimulus
frmts_kl.data = load(['speech/' filename '_kl.txt']);
sources.data = load(['speech/' filename '_source.txt']);
frmts.data = load(['speech/' filename '_hl.txt']);
T = round(length(pin))/f;
tsp = 0:1/f:T-1/f; % time vector

%% Plot Spectrogram of input waveform:
% specgram(pin,256,f,hamming(256),128)
% title('\bf Input');

%% Ramp stimulus on and off (Ian Bruce's code):
rt = 10e-3;   % rise time = 10 ms
mxpts = length(pin); irpts = round(rt*f); tind   = 0.0; for lp =
1:irpts
    pin(lp) = pin(lp)*tind/irpts;
    tind = tind + 1.0;
end tind = tind + mxpts-2*irpts; for lp = (mxpts-irpts):mxpts
    pin(lp) = pin(lp)*((mxpts-tind) / irpts);
    tind = tind + 1.0;
end

%% Simplified sparse connection of neurons
%%  - neuron of a given BF only connects with neurons of neighbouring BFs,
%%    not to the same BF
%%  - connections expected to be sparse but occur more randomly in nature
reps=30;    % # of trials to average over for PSTH
impair=1;  % normal (=0) or impaired (=1) ear
output=zeros(numofneurons,length(tsp)-1);
input=zeros(numofneurons,length(tsp)-1); for lp_reps=1:reps
    [spikes,inputs]=earandlin(pin,f,tsp,bf,impair);
    output=output+spikes;
    input=input+inputs;
%     save hear.mat
    display(lp_reps);
%     waitbar(lp_reps/(reps),h,'Computing...');
end  % for lp_reps
```

```
output=output./reps;  % output PSTH
input=input./reps;  % post-convergence input PSTH
clear inputs;
% close(h)  % close waitbar

elptime=etime(clock,clk)/3600;  % Time elapsed [hours]
save speech_EE95.mat

%% Compute Synchronized Rate and Power Ratios:
% [synrate,t_sp] = prvt_Jen(output,tsp,bf,frmts_kl,frmts,sources,pin,stimdb);
% prvt_Jen

%% Segment averaged output into time frames:
Tw=20e-3;  % length of window in time [s]
Nw=round(Tw/tsp(2));  % length of window in indices
if gcd(Nw,2)~=2  % make w even
    Nw=Nw+1;
end % if gcd
numframes=floor((length(tsp)-1)/Nw)*2-1;  % 0.5 frame overlap
frames=zeros(numofneurons,Nw,numframes); for k=1:numframes
    frames(:,:,k)=output(:,k*Nw/2-Nw/2+1:k*Nw/2+Nw/2);
%     frames(:,:,k)=input(:,k*Nw/2-Nw/2+1:k*Nw/2+Nw/2);
end

%% Compute average spike rate [spikes/s]:
avg_spikes=zeros(numofneurons, numframes); for k=1:numframes
    avg_spikes(:,k)=mean(frames(:,:,k),2)./tsp(2);
end

%% Plot Waterfall Spectrogram:
% figure
% waterfall([0:numframes-1].*Tw./2,bf',avg_spikes);
% set(gca,'YScale','log')
% cb=colorbar;
% set(get(cb,'ylabel'),'string','Mean Firing Rate [spikes/s]')
% title('\bf Spatio-temporal pattern of activity for output from the
LIN given normal input');
% ylabel('Characteristic Frequency [Hz]');
% xlabel('Time [s]');
% zlabel('Mean Firing Rate [spikes/s]');
% axis tight
% set(gca,'View', [-8.5 65]);
% % paperfig2(gcf,12)

%% Plot Average Output of LIN:
% figure
% semilogx(bf./1e3',mean(output,2)./tsp(2));
% xlim([0 bf(end)./1e3]);
% xlabel('Characteristic Frequency [KHz]');
```

```
% ylabel('Average Firing Rate [spikes/s]');
% title(['\bf Spiking, recurrent LIN (normal hearing) ', date]);
% % title(['\bf Spiking, recurrent LIN (impaired hearing) ', date]);
% grid on;

%% Plot spike train of a single neuron:
% figure
% subplot(2,1,1);
% plot(tsp(1:length(y)),y(20,:),'g')
% ylabel('Vm [V]');
% xlabel('Time [s]');
% subplot(2,1,2);
% plot(tsp(1:length(y)),inputs(20,:))
% hold on
% plot(tsp(1:length(y)),spikes(20,:),'r')
% ylabel('Spikes');
% xlabel('Time [s]');
% legend('input','output');

%% Plot average spike rate for a 10ms (0.2-0.21 sec of sentence):
% % avgoframe=sum(output(:,8000:8400),2)./0.01;
% % avgiframe=sum(input(:,8000:8400),2)./0.01;
% figure
% semilogx(bf,avgoframe,'b')
% xlim([bf(1) bf(end)])
% grid on
% xlabel('Characteristic Frequency [Hz]')
% ylabel('Mean Spike Rate [spikes/s]')
% title('\bf Output from LIN')
% figure
% semilogx(bf,avgiframe,'r')
% xlim([bf(1) bf(end)])
% grid on
% xlabel('Characteristic Frequency [Hz]')
% ylabel('Mean Spike Rate [spikes/s]')
% title('\bf Auditory Nerve Response')
```

C.2.2 earandlin.m

```
function [spikes,inputs] = earandlin(low,f,tsp,bf,impair)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% earandlin.m                                                                 %
% -- Model of the mammalian auditory system (peripheral and central)  %
% -- Takes sound-pressure wave that impinges on the ear as input      %
% -- Output is a spike train from the LIN                                     %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% Set Parameters:
```

```
numofneurons=length(bf);
fs=500e3; % sampling frequency for Ian's model [Hz]
pin=resample(low,fs,f); % upsample to fs
reptime=tsp(end)*1e3;  % time between stimulus repetitions [ms]
tsp=tsp(1:length(tsp)-1);

%% Membrane Parameters:
C=8e-12;  % Membrane capacitance [F]
mtau = 1.5e-3;  % Membrane Time Constant [s]

%% Define Excitatory Synapse Parameters:
EsynE=0.1;  % excitatory reversal potential [V]; must be > ythres
GEalpha = 11;
GEdur=0.01;  % (IFAC 2003)
GE = (GEalpha/mtau/10)^2*tsp(1:round(GEdur/tsp(2))).*...
    exp(-GEalpha/mtau*tsp(1:round(GEdur/tsp(2))));
sfac=3.0365e-010;  % based on GEalpha=9 and GIalpha=0.5
GE=sfac*GE;  % scaled excitatory conductance [S] - scaling factor must match GI

%% Define Inhibitory Synapse Parameters:
EsynI=-0.02;  % inhibitory reversal potential [V]; < resting potential
% EsynI=0;  % Shunting Inhibition reversal potential [V]; ~= yrest
GIalpha = 0.5; GIdur = 0.015;
GI=(GIalpha/mtau/10)^2*tsp(1:round(GIdur/tsp(2))).*...
    exp(-GIalpha/mtau*tsp(1:round(GIdur/tsp(2))));
GI=sfac*GI;  % scaled inhibitory conductance [S] - scaling factor must match GE

%% Effectiveness coeffients of inputs across BFs
%% -- widthofwindow=1 b/c no lateral excitation
%% -- selfsyn=1 b/c not using this function to account for inhibition
V = weight(1,numofneurons,1);

%% Coefficients of inhibitory interactions across BFs
%% -- 2 Gaussian humps as in Gerken's model
%% -- a = arbitrary coefficient describing weight of inhibitory compared
%%    to excitatory interactions
a=32;
W = weight_2gauss(9,numofneurons); W = a*W/max(sum(W,2));

%% Normal or impaired ear:
if impair==0  %for normal OHC and IHC function
    COHCs = ones(size(bf));
    CIHCs = ones(size(bf));
else  %for "average" cat OHC and IHC dysfunction
    COHCs = interp1([0 1 1.1 1.4 1.7 2.2 3.6 4.7 6.0 7.8 inf],...
        [0.5 0.5 0.5 0.35 0.06 0.15 0.3 0.36 0.6 0.95 0.95],bf/1e3,'cubic');
    CIHCs = interp1([0 0.5 0.8 1 1.1 1.4 1.7 2.2 3.6 4.7 6.0 7.8 9 10 inf],...
        [1 1 0.1 0.06 0.06 0.05 0.05 0.05 0.05 0.05 0.05 0.1 0.2 0.25 1],...
        bf/1e3,'cubic');
```

```
end

%% Get input from Ian's model (deafcat2) of the ear
%% -- nerve.m and nervetone.m are functions that call deafcat2
inputs=zeros(numofneurons,length(tsp));

%% Downsample output from Ian's model of the ear:
for i=1:numofneurons
    [timeout,psth]=ear(bf(i),reptime,pin,fs,COHCs(i),CIHCs(i));
    n=length(tsp);
    m=floor(length(psth)/n);
    inputs(i,:)=sum(reshape(psth(1:m*n)',m,n));
end  % for i

%% Shape excitatory spikes
inputs2=filter(GE,1,inputs,[],2);  % shaped conductance

%% Initialization
spikes=zeros(numofneurons,length(tsp));  % keeps track of spikes generated
spikes2=zeros(numofneurons,length(tsp));  % for lateral inhibition
[spikes2(:,1),Zi]=filter(GI,1,spikes(:,1),[],2);  % all 0s
y=zeros(numofneurons,length(tsp));  % Membrane potential [V]
y(:,1) = V*inputs2(:,1)*tsp(2)/mtau;  % set initial condition of neurons [V]
tref=2e-3;  % absolute refractory period [s]
trefsamp=round(tref/tsp(2));  % refractory period [# of samples]
spikevalue=0.15;
% nominal threshold potential [V]:
ythres=0.015*ones(numofneurons,length(tsp)+2*trefsamp);

%% Calculate output membrane potential [V] using
%% Shamma's rate equation modified with Koch p19
%%  -- modified to account for synaptic conduction
for lp_time=2:length(tsp);  % lp_time=n+1 index

    %% Determine if 'lp_time' is in a refractory period.
    %% If so, set threshold potential higher.
    p=any(spikes(:,lp_time-1),2);  % Did a spike occur in the last time step?
    p=p';  % column to row vector
    if any(p)
        m=find(p);
        ythres(m,lp_time:lp_time+trefsamp)=5;
        ythres(m,lp_time+trefsamp+1:lp_time+2*trefsamp)=...
            repmat(5*exp(-3.5*[0:trefsamp-1]./trefsamp),length(m),1);
    end % if p

    in = V*(inputs2(:,lp_time-1).*(EsynE-y(:,lp_time-1))./C)...
        +W*(spikes2(:,lp_time-1).*(EsynI-y(:,lp_time-1))./C)...
        -y(:,lp_time-1)./mtau;    % rate eqn
```

```
    %% Using the 4th-order Runge-Kutta method of approximation (parabolas):
    K1 = fn*tsp(2);
    K2 = (fn-0.5*K1/mtau)*tsp(2);
    K3 = (fn-0.5*K2/mtau)*tsp(2);
    K4 = (fn-K3/mtau)*tsp(2);
    y(:,lp_time)=y(:,lp_time-1) + (1/6)*(K1 + 2*K2 + 2*K3 + K4);

    %% Set refractory period after a spike:
    % Are there any spikes within the last tref secs?
    ph=any(spikes(:, max(lp_time-trefsamp,1):lp_time-1),2);
    ph=ph';  % column to row vector
    mh=find(ph);
    y(mh,lp_time)=0;

    %% If not in refractory period and membrane potential is > threshold,
    %% generate a spike:
    I=find(y(:,lp_time)>ythres(:,lp_time));
    y(I,lp_time)=spikevalue;
    spikes(I,lp_time)=1;
    % shaped inhibitory conductance:
    [spikes2(:,lp_time),Zi]=filter(GI,1,spikes(:,lp_time),Zi,2);

end  % for lp_time
```

C.2.3 ear.m

```
function [htimeout,hpsth] = ear(bf,reptime,pin,fs,cohc,cihc)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% -- generates input for LIN model using output from Bruce et al. model (2003) %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% Initialize Parameters:
stimtime=reptime-1;  % duration of stimulus [ms] **longer than pin
% t=[0:1/fs*1e3:stimtime];  % time vector [ms]
hnrep=5;  % repetitions for psth representing high spontaneous rate converging neurons
binwidth=0.002;  % bin size [ms] **must be 0.002

%% Get output from auditory nerve:
[htimeout,hmeout,hbmout,htausp,hvihc,hsynout,hpsth] =
deafcat2(pin,bf,hnrep,binwidth, stimtime,reptime,cohc,cihc);
% t=timeout;
% spk=psth;
```

C.2.4 pvrt_Jen.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% prvt_Jen.m                                                                    %
% -- this script calculates the power ratio vs time                            %
```

```
% Source of code: Ian Bruce (described in Miller et al. (1997))       %
% -- modified by Jennifer Ko                                          %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

[b_stim,f_stim,t_stim] = specgram(pin,256,f,hamming(256,'periodic'),192);

%% Downsample to normal hearing range:
fmax=10e3;  % Maximum frequency range of hearing [Hz]
m=max([1 round(1/(2*fmax*tsp(2)))]);
n=floor(length(output(1,:))/m);
numofneurons=length(bf);
psth=zeros(numofneurons,n);
for lp_bf=1:numofneurons
    psth(lp_bf,:)=sum(reshape(output(lp_bf,1:m*n),m,n))./(m*tsp(2));
end
% t_psth=[tsp(2):m*tsp(2):n*m*tsp(2)];


%% Calculate Synchronized Rate (Miller et al. 1997):
win=hamming(256,'periodic');
b_sp = specgram(psth(lp_bf,:),256,1/(m*tsp(2)),win,192); % 75% overlap
synrate=zeros([size(b_sp) numofneurons]);  % [freq,time,numofneurons]
for lp_bf=1:numofneurons
    [b_sp,f_sp,t_sp]=specgram(psth(lp_bf,:),256,1/(m*tsp(2)),win,192);
    % Synchronized rate:
    synrate(:,:,lp_bf)=abs(b_sp/length(win)/sqrt(sum(win.^2)/length(win)));
end
synrate(1:2,:,:) = 0;  % Remove low frequency/bias components

%% Plot synchronized rate:
figure
% subplot(2,1,1);
% semilogx(f_sp, mean(synrate(:,:,24),2));
% xlim([100 10e3]);
% ylabel('Synchronized Rate [/s]');
% xlabel('Frequency [Hz]');
% title('\bf BF=513.4 Hz');
% subplot(2,1,2);
% semilogx(f_sp, mean(synrate(:,:,47),2));
% xlim([100 10e3]);
% ylabel('Synchronized Rate [/s]');
% xlabel('Frequency [Hz]');
% title('\bf BF=1487.1 Hz');


%% Calculate Power Ratios:

% Initialize variables
totalpower = squeeze(sum(synrate.^2,1))';  %% [numofneurons,time]
```

```
f1power = zeros(size(totalpower));
f2power = f1power;
f3power = f1power;
f4power = f1power;

% Time shift:
latency_comp = 5e-3;
t_spcgrm = t_sp+tsp(2)*length(win)/2;
t_dlyd = t_spcgrm-latency_comp;

% Find formants:  %% F = interp1(t,f,T)
F0_kl=interp1q([-100; frmts_kl.data(:,1)*1e-3; inf],...
    frmts_kl.data([1 1:end end],2),t_spcgrm);
F1=interp1q([-100; frmts.data(:,1)*1e-3; inf],frmts.data([1 1:end end],3),t_spcgrm);
F2=interp1q([-100; frmts.data(:,1)*1e-3; inf],frmts.data([1 1:end end],4),t_spcgrm);
F3=interp1q([-100; frmts.data(:,1)*1e-3; inf],frmts.data([1 1:end end],5),t_spcgrm);
F4=interp1q([-100; frmts.data(:,1)*1e-3; inf],frmts.data([1 1:end end],6),t_spcgrm);
AV=interp1q([-100; sources.data(:,1)*1e-3; inf],sources.data([1 1:end end],2),t_spcgrm);
AF=interp1q([-100; sources.data(:,1)*1e-3; inf],sources.data([1 1:end end],3),t_spcgrm);
silence = (AV==0).*(AF==0);

F0dlyd_kl=interp1q([-100; frmts_kl.data(:,1)*1e-3; inf],...
    frmts_kl.data([1 1:end end],2),t_dlyd);
F1dlyd=interp1q([-100; frmts.data(:,1)*1e-3; inf],frmts.data([1 1:end end],3),t_dlyd);
F2dlyd=interp1q([-100; frmts.data(:,1)*1e-3; inf],frmts.data([1 1:end end],4),t_dlyd);
F3dlyd=interp1q([-100; frmts.data(:,1)*1e-3; inf],frmts.data([1 1:end end],5),t_dlyd);
F4dlyd=interp1q([-100; frmts.data(:,1)*1e-3; inf],frmts.data([1 1:end end],6),t_dlyd);
AVdlyd=interp1q([-100; sources.data(:,1)*1e-3; inf],sources.data([1 1:end end],2),t_dlyd);
AFdlyd=interp1q([-100; sources.data(:,1)*1e-3; inf],sources.data([1 1:end end],3),t_dlyd);
silencedlyd = (AVdlyd==0).*(AFdlyd==0);

% Compute power:
for lp=1:length(t_sp)

    %for hlp = 1:floor(min(4,5e3/F1dlyd(lp)))
    for hlp = 1
        f1_higher = min(find(f_sp>hlp*F1dlyd(lp)));
        if isempty(f1_higher)
            f1_higher=length(f_sp)-1;
        end
        f1_lower = max(find(f_sp<hlp*F1dlyd(lp)));
        f1_range = unique(f1_lower-1:f1_higher+1);
        if length(f1_range)>4
            f1_range = f1_range(2:4);
        end
        f1power(:,lp) = f1power(:,lp)+squeeze(sum(synrate(f1_range,lp,:).^2,1));
    end

    %for hlp = 1:floor(min(4,5e3/F2dlyd(lp)))
```

```
    for hlp = 1
        f2_higher = min(find(f_sp>hlp*F2dlyd(lp)));
        if isempty(f2_higher)
            f2_higher=length(f_sp)-1;
        end
        f2_lower = max(find(f_sp<hlp*F2dlyd(lp)));
        f2_range = unique(f2_lower-1:f2_higher+1);
        if length(f2_range)>4
            f2_range = f2_range(2:4);
        end
        f2power(:,lp) = f2power(:,lp)+squeeze(sum(synrate(f2_range,lp,:).^2,1));
    end

    %for hlp = 1:floor(min(4,5e3/F3dlyd(lp)))
    for hlp = 1
        f3_higher = min(find(f_sp>hlp*F3dlyd(lp)));
        if isempty(f3_higher)
            f3_higher=length(f_sp)-1;
        end
        f3_lower = max(find(f_sp<hlp*F3dlyd(lp)));
        f3_range = unique(f3_lower-1:f3_higher+1);
        if length(f3_range)>4
            f3_range = f3_range(2:4);
        end
        f3power(:,lp) = f3power(:,lp)+squeeze(sum(synrate(f3_range,lp,:).^2,1));
    end

    %for hlp = 1:floor(min(4,5e3/F4dlyd(lp)))
    for hlp = 1
        f4_higher = min(find(f_sp>hlp*F4dlyd(lp)));
        if isempty(f4_higher)
            f4_higher=length(f_sp)-1;
        end
        f4_lower = max(find(f_sp<hlp*F4dlyd(lp)));
        f4_range = unique(f4_lower-1:f4_higher+1);
        if length(f4_range)>4
            f4_range = f4_range(2:4);
        end
        f4power(:,lp)=f4power(:,lp)+squeeze(sum(synrate(f4_range,lp,:).^2,1));
    end

end

% Compute power ratio:
f1pr = f1power./totalpower;
f2pr = f2power./totalpower;
f3pr = f3power./totalpower;
f4pr = f4power./totalpower;
```

```
%% Plot:
F1voiced=F1;
F2voiced=F2;
F3voiced=F3;
F4voiced=F4;

F1voiceddlyd = F1dlyd;
F2voiceddlyd = F2dlyd;
F3voiceddlyd = F3dlyd;
F4voiceddlyd = F4dlyd;

F1unvoiced=F1;
F2unvoiced=F2;
F3unvoiced=F3;
F4unvoiced=F4;

F1unvoiceddlyd = F1dlyd;
F2unvoiceddlyd = F2dlyd;
F3unvoiceddlyd = F3dlyd;
F4unvoiceddlyd = F4dlyd;

F1voiced(find((F0_kl==0)|(silence==1))) = NaN;
F2voiced(find((F0_kl==0)|(silence==1))) = NaN;
F3voiced(find((F0_kl==0)|(silence==1))) = NaN;
F4voiced(find((F0_kl==0)|(silence==1))) = NaN;

F1voiceddlyd(find((F0dlyd_kl==0)|(silencedlyd==1))) = NaN;
F2voiceddlyd(find((F0dlyd_kl==0)|(silencedlyd==1))) = NaN;
F3voiceddlyd(find((F0dlyd_kl==0)|(silencedlyd==1))) = NaN;
F4voiceddlyd(find((F0dlyd_kl==0)|(silencedlyd==1))) = NaN;

F1unvoiced(find((F0_kl>0)|(silence==1))) = NaN;
F2unvoiced(find((F0_kl>0)|(silence==1))) = NaN;
F3unvoiced(find((F0_kl>0)|(silence==1))) = NaN;
F4unvoiced(find((F0_kl>0)|(silence==1))) = NaN;

F1unvoiceddlyd(find((F0dlyd_kl>0)|(silencedlyd==1))) = NaN;
F2unvoiceddlyd(find((F0dlyd_kl>0)|(silencedlyd==1))) = NaN;
F3unvoiceddlyd(find((F0dlyd_kl>0)|(silencedlyd==1))) = NaN;
F4unvoiceddlyd(find((F0dlyd_kl>0)|(silencedlyd==1))) = NaN;

prvt_frmts_contour_Jen
```

C.2.5 pvrt_frmts_contour_Jen.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% prvt_frmts_contour_Jen.m                                                    %
% -- this script plots the power ratios vs time                              %
```

```
% Source of code: Ian Bruce (described in Miller et al. (1997))       %
% -- modified by Jennifer Ko                                          %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Formant PR contour plots
figure
sp1=gca;
%imagesc(t_stim,f_stim/1e3,stimdb+db(b_stim*sqrt(2)/256/20e-6));
imagesc(t_stim,f_stim/1e3,db(b_stim*sqrt(2)/256/20e-6));
axis xy
ylim([0 6]);
%caxis([stimdb-60 stimdb])
caxis([stimdb-80 stimdb])
%caxis([stimdb-30 stimdb+50])
set(sp1,'xtick',0:0.2:1.6)
% set(sp1,'xticklabel',[])
% ps=get(sp1,'position');
% set(sp1,'position',ps+[-0.04 0 0 0]);
title('\bf Synthesized male speaker: "Five women played basketball"')
xlabel('Time [s]');
ylabel('Frequency (kHz)')
hold on
plot(t_sp,[F1voiced F2voiced F3voiced F4voiced]/1e3,'w-','linewidth',2)
plot(t_sp,[F1unvoiced F2unvoiced F3unvoiced F4unvoiced]/1e3,'w--','linewidth',2)
ps = get(sp1,'position');
set(sp1,'position',ps.*[1 1 0.85 1])
if impair==0
    title(strcat(['\bf Normal  ', num2str(stimdb), 'dB SPL']));
else
    title(strcat(['\bf Impaired  ', num2str(stimdb), 'dB SPL']));
end
ps = get(sp1,'position');
cb1 = colorbar('vert','peer',sp1);
% set(sp1,'position',ps)
ps = get(cb1,'position');
set(cb1,'position',ps + [0.085 0 0 0])
set(cb1,'YAxisLocation','left');
set(get(cb1,'ylabel'),'string','dB SPL')
axes(cb1)
text(2.3,stimdb,'0.7')
tx1=text(3.5,stimdb-40,'PR');
text(2.3,stimdb-80,'0')
set(tx1,'HorizontalAlignment','center','Rotation',90)
% colormap(1-gray)

figure
sp2=subplot(3,1,1);
% sp2=subplot(4,1,2);
%imagesc(t_sp,bf/1e3,f3pr)
```

```
%imagesc(t_sp,log10(bf/1e3),f3pr)
contourf(t_sp,log10(bf/1e3),f3pr,0.1:0.2:0.7)
set(sp2,'xtick',0:0.2:2.0)
set(sp2,'xticklabel',[])
set(sp2,'ytick',log10([0.1 1 5]))
set(sp2,'yticklabel',[0.1 1 5])
axis xy
caxis([0 0.7])
%ylim(yl)
title('PR(F3)')
ylabel('BF (kHz)')
%xlabel('Time (s)')
hold on
%plot(t_sp,F3dlyd/1e3,'k-','linewidth',1.5)
plot(t_sp,log10(F3voiceddlyd/1e3),'k-','linewidth',1.5)
plot(t_sp,log10(F3unvoiceddlyd/1e3),'k--','linewidth',1.5)

sp3=subplot(3,1,2);
% sp3=subplot(4,1,3);
%imagesc(t_sp,bf/1e3,f2pr)
%imagesc(t_sp,log10(bf/1e3),f2pr)
contourf(t_sp,log10(bf/1e3),f2pr,0.1:0.2:0.7)
set(sp3,'xtick',0:0.2:2.0)
set(sp3,'xticklabel',[])
set(sp3,'ytick',log10([0.1 1 5]))
set(sp3,'yticklabel',[0.1 1 5])
axis xy
caxis([0 0.7])
%ylim(yl)
title('PR(F2)')
ylabel('BF (kHz)')
%xlabel('Time (s)')
hold on
%plot(t_sp,F2dlyd/1e3,'k-','linewidth',1.5)
plot(t_sp,log10(F2voiceddlyd/1e3),'k-','linewidth',1.5)
plot(t_sp,log10(F2unvoiceddlyd/1e3),'k--','linewidth',1.5)

sp4=subplot(3,1,3);
% sp4=subplot(4,1,4);
%imagesc(t_sp,bf/1e3,f1pr)
%imagesc(t_sp,log10(bf/1e3),f1pr)
contourf(t_sp,log10(bf/1e3),f1pr,0.1:0.2:0.7)
set(sp4,'xtick',0:0.2:2.0)
set(sp4,'ytick',log10([0.1 1 5]))
set(sp4,'yticklabel',[0.1 1 5])
axis xy
caxis([0 0.7])
%ylim(yl)
title('PR(F1)')
```

```
ylabel('BF (kHz)')
xlabel('Time (s)')
hold on
%plot(t_sp,F1dlyd/1e3,'k-','linewidth',1.5)
plot(t_sp,log10(F1voiceddlyd/1e3),'k-','linewidth',1.5)
plot(t_sp,log10(F1unvoiceddlyd/1e3),'k--','linewidth',1.5)

gtext('\bf Auditory Nerve Response')
% gtext('\bf Output from LIN')
```