ADAPTIVE PSYCHOACOUSTIC EXPERIMENTS FOR COCHLEAR IMPLANTS

USING THE PDA RESEARCH PLATFORM

by

Nageswara Rao Gunupudi

APPROVED BY SUPERVISORY COMMITTEE:

____

Dr. Philipos C. Loizou, Chair

____

Dr. Nasser Kehtarnavaz

____

Dr. Murat Torlak

To my parents and friends

ADAPTIVE PSYCHOACOUSTIC EXPERIMENTS FOR COCHLEAR IMPLANTS USING THE PDA RESEARCH PLATFORM

by

NAGESWARA RAO GUNUPUDI, B.Tech

THESIS

Presented to the Faculty of

The University of Texas at Dallas

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

THE UNIVERSITY OF TEXAS AT DALLAS

May, 2010

ACKNOWLEDGEMENTS

I would like to sincerely thank my advisor, Dr. Philipos C. Loizou, for giving me an opportunity to work under his esteemed guidance. His constant motivation, support and invaluable suggestions during course of my work enhanced my research skills. I would like to express my gratitude to Dr. Nasser Kehtarnavaz and Dr Murat Torlak for serving on my committee.

I am thankful to members of the Speech Processing Laboratory at The University of Texas at Dallas for helping me throughout my stay here and reviewing the thesis document. I am thankful to Dr. Arthur Lobo for his help and support during this work. I am also thankful to Dr. Ray Goldsworthy for his support during experiments and data analysis of this work. Needless to say, I am greatly thankful to my friends at here, who made my life here a memorable, and fun filled stay. I am extremely grateful to my family for their constant support to keep me focused on my studies and research work.

This research work was funded by the National Institutes of Health (Contract No. NIH-NO1-DC-6-0002).

January, 2010

ADAPTIVE PSYCHOACOUSTIC EXPERIMENTS FOR COCHLEAR IMPLANTS USING THE PDA RESEARCH PLATFORM

Publication No. ___________________

Nageswara Rao Gunupudi, M.S.
The University of Texas at Dallas, 2010

Supervising Professor: Dr. Philipos C. Loizou

The growth and progress in cochlear implants is largely motivated and driven by research in many areas of cochlear implant technologies such as speech science, bioengineering, signal processing and psychophysiology. Continuous evaluation of the performance and development of new strategies is essential for cochlear implant speech processors as the performance of cochlear implant patients varies over time. Thus, a flexible research platform that facilitates easy design, upgrade of existing algorithms and continuous study of cochlear implant is needed for the advancement in research and development of cochlear implants. This thesis work focused on providing software controlled environment to design and conduct psychoacoustic experiments using a PDA research platform for cochlear implants. The psychoacoustic experiments are used to evaluate the speech perception, intelligibility, and source localization performance of cochlear implants patients. Transformed adaptive up-down methods are implemented as part of the software to use in the psychoacoustic

experiments. Using the software research platform, binaural cues such as interaural level difference and interaural time difference are evaluated using normal-hearing listeners. The results are presented and compared against with prior psychoacoustic results for validation of the experimental software.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The study of cochlear implants is a multi-disciplinary subject that covers many fields including speech science, bioengineering, signal processing, and psychophysiology. This thesis focuses primarily on signal processing and psychophysiology aspects of cochlear implants. As described in next chapter, cochlear implants are characterized by various parameters and strategies such as number of electrodes, electrode configuration, type of stimulation (analog or pulsatile), and transmission link (transcutaneous or percutaneous) and signal processing strategies based on information representation or feature extraction. The variety of choices and parameters involved in cochlear implants and speech processor design provides researchers numerous challenges in developing better and efficient methods to restore near normal hearing capability in hearing impaired people. The main challenge in developing an efficient cochlear implant lies in deriving an optimal electrical stimulus that can generate auditory sensations close to normal speech. It is also known that the hearing and speech intelligibility behavior of implant user depends on the duration for which implants are in use. This demands continuous evaluation of the performance and development of new strategies. Thus, the growth and progress in cochlear implant technology is largely motivated and driven by research. Therefore, a flexible research platform that facilitates continuous study of cochlear implant is essential for the advancement in the research and development of cochlear implants. It is known that the performance of CI users vary (improve or change) while in constant use for a period of time, which necessitates development of new strategies and novel algorithms are evaluated after long-term use of the device. Thus, a portable, flexible and easy to use research processor is highly needed. A personal digital assistant (PDA) based research platform developed by Loizou et., al, [15] aims to provide such flexibility and ease.

This thesis involves the design of a simple and flexible software tool for conducting and analyzing psychoacoustic experiments with CI and normal hearing (NH) subjects. The software tool is integrated with the PDA research platform [15] to provide researchers and patients with an easier and flexible way to carry out psychoacoustic experiments. Both monaural and binaural experiments are incorporated into the software. The graphical user interface (GUI) setup guides the subject under test (SUT) to run through the test without much training and experimenter's intervention.

This thesis work involves design and development of psychoacoustic experiments and GUI based software tool integrated with the PDA-based research platform for cochlear implants. The software, developed in MATLAB, provides an easy interface for researchers to study the behavior of cochlear implant users or normal hearing subjects for various kinds of stimuli. Simple and easy graphical user interface allows normal hearing and CI subjects to carry out experiments in less time without any programming knowledge, and helps researchers study and analyze the behavior of subjects. A set of monaural and binaural psychoacoustic adaptive experiments like pitch discrimination, pitch ranking, loudness balancing, speech reception thresholds, interaural level difference (ILD), and interaural time difference (ITD) are designed. A set of flexible transformed up-down adaptive methods [33] are implemented to provide the researcher with easy and quick setup to conduct the psychoacoustic experiments and study the behavior of subjects using the PDA research platform. Simple and modulated sinusoidal tones and a predefined set of spoken words are included in the software tools to be used as reference acoustic stimuli. These stimuli are encoded using CIS and ACE cochlear implant strategies in Matlab and sent to the PDA which in turn sends them to CI users (or Implant-in-Box for research purposes) using a field programmable gate array (FPGA) integrated circuit based board. The performance and functionality of the software tool is verified for its usability along with the PDA research platform by conducting all psychoacoustic experiments with normal hearing subjects. Also a study was carried out to analyze NH subject's response to binaural cues. The results obtained are compared with other state of the art studies.

1.1 Thesis Organization:

Chapter 2 briefs about basic principle of cochlear implants and speech processing strategies used in today's cochlear implants. Chapter 3 describes the development of PDA based research platform and its operation for use of cochlear implants and on its research. Chapter 4 describes transformed adaptive methods used for psycho experimental in them. Chapter 5 provides an overview of useful monaural and binaural cues and their significance with cochlear implants. Chapter 6 provides details of the graphical user interface developed and its interface with the PDA research platform. Chapter 7 describes psychoacoustic procedures and their implementations for measuring few hearing cues. Chapter 8 presents the psychoacoustic experiments carried out to evaluate the software platform. It also presents analysis of the results and general comparison of results of monaural and binaural experiments with the state of art results. Chapter 9 provides a conclusion of the work and summarizes the results.

CHAPTER 2

COCHLEAR IMPLANTS

Cochlear Implants (CI) are electronic devices that are designed to bypass the damaged parts in the inner ear to restore hearing in partially and profoundly deaf people by direct electrical stimulation of the auditory neurons [1]. In the past decade the use of cochlear implants grew rapidly from 12,000 in 1995 to over 100,000 and interest in the study of cochlear implants progressed rapidly in improving speech perception and intelligibility. This thesis focuses on developing software platforms that help in the study of various features and parameters using psychoacoustical methods to improve CI's capabilities. This chapter is dedicated to describe the basic functional units of cochlear implants and strategies for signal processing in cochlear implants. It also describes the need for research platforms in the advancement of cochlear implants. This chapter also presents the organization of the thesis and the goals of this work.

2.1 Cochlear Implants Design Characteristics

Any cochlear implant based hearing prosthesis system consists of an external sound processor that captures the acoustic signal, using a microphone, and analyzes the signal and generates a stimulus signal to excite a series of electrodes placed in the cochlea through medical implantation. An RF signal is transmitted to an implanted receiver stimulator that produces the electric stimuli. These electric stimuli are sent to designated electrodes in the implanted electrode array to stimulate different regions of the cochlea.

The cochlear implant devices are characterized based on, 1) Electrode design (e.g., number of electrodes, electrode configuration), 2) Type of stimulation - analog or pulsatile, 3) Transmission link - transcutaneous or percutaneous, 4) Signal processing - waveform representation or feature extraction.

2.1.1 Electrode design

The design of electrodes for cochlear implants plays an important role in hearing of intelligible speech. Majorly, electrode design deals with (1) electrode placement, (2) number of electrodes and spacing of contacts, (3) orientation of electrodes with respect to the excitable tissue, and (4) electrode configuration.

Today's popular cochlear implant devices such as Nucleus, Clarion, Med-El employ different implementations to address the electrode design issues mentioned above. For example, Ineraid and Med-El employ monopolar electrodes, while Nucleus devices employ bipolar electrodes and and Clarion devices provide both types of electrodes. The Nucleus CI device uses 22 electrodes with 0.75 mm of inter electrode spacing and in bipolar mode, the electrodes are spaced 1.5 mm apart. The Med-El device uses 12 electrodes in monopolar configuration.

2.1.2 Type of stimulation

Based on the kind of electrical stimulation to the electrodes, there are two types of stimulation used in cochlear implants, 1) Analog stimulation, where the electrical signal of acoustic waveform is presented in analog form 2) Pulsatile stimulation, where the electrical signal is presented in pulses. Often analog stimulation suffers from channel interactions during the simultaneous stimulation of electrodes. But in pulsatile stimulation, the acoustic waveform is delivered to the electrodes using a set of narrow pulses, which allows delivering pulses in a non-overlapping (i.e., non-simultaneous) fashion, thereby minimizing channel interactions.

The rate at which these pulses are delivered to the electrodes, i.e., the pulse rate, has also been found to affect speech recognition performance [9]. And several studies reported that higher pulse rates yield better performance than low pulse rates [10], [12].

2.1.3 Transmission link

The acoustic signal in electrical form is transmitted from the external processor to the implanted electrodes in two different ways: (1) through a transcutaneous connection and (2) through a percutaneous connection as shown in figure 2.1. The transcutaneous system uses a radio link to transmit the stimuli. In this system, an external transmitter is used to encode the stimulus information for radio-frequency transmission from an external coil to an implanted coil. The internal receiver decodes the signal and delivers the stimuli to the electrodes. The transmitter and the implanted receiver are held in place on the scalp by a magnet. Most of the today's cochlear implant devices (e.g., Nucleus, Clarion, Med-El) use transcutaneous connections. The percutaneous system transmits the stimuli to the electrodes directly through plug connections as shown figure 2.1. In percutaneous links, no other electronic circuitry is implanted other than the electrodes. The percutaneous system offers high flexibility and signal transparency, which makes it ideal for research purpose. The Ineraid device uses percutaneous connectors. Nucleus, Clarion and Med-El devices employ a transcutaneous based transmission link.

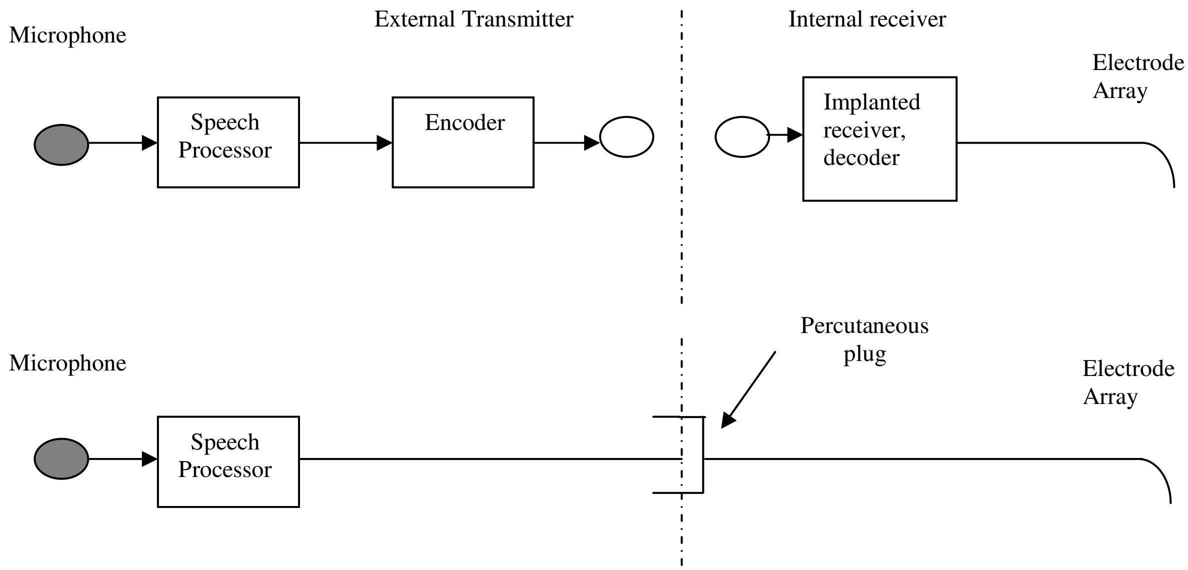

Figure 2.1 Cochlear Implants Signal Transmission Model (Loizou, 1999 [1])

2.1.4 Signal Processing

The most important module of the cochlear implant and the one that differentiates design in different devices is the signal processing strategy used for transforming the speech signal to electrical stimuli. Several signal processing techniques have been developed in the past which can be broadly categorized as 1) techniques that preserve waveform information, 2) techniques that preserve envelope information, and 3) techniques that preserve spectral features (e.g., formants). Here, we are interested in channel vocoder based techniques which make use of envelopes of acoustic signal to generate electrical stimulation. Section 2.2 details about vocoder based speech processing techniques.

2.2 Speech Vocoder for Cochlear Implants

The most recent cochlear implant speech processors are based on the simple principle of channel vocoder, which is used for speech communication over telephone lines with much less bandwidth than that required for transmitting the unprocessed speech signal. A channel vocoder [5], [6], [7] consists of a speech analyzer and a speech synthesizer. The state of art cochlear implant speech processors uses the analyzer block of the channel vocoder. As shown in Figure 2.2, the analyzer of a vocoder filters the incoming speech signal into a number of contiguous frequency channels using a bank of band-pass filters.

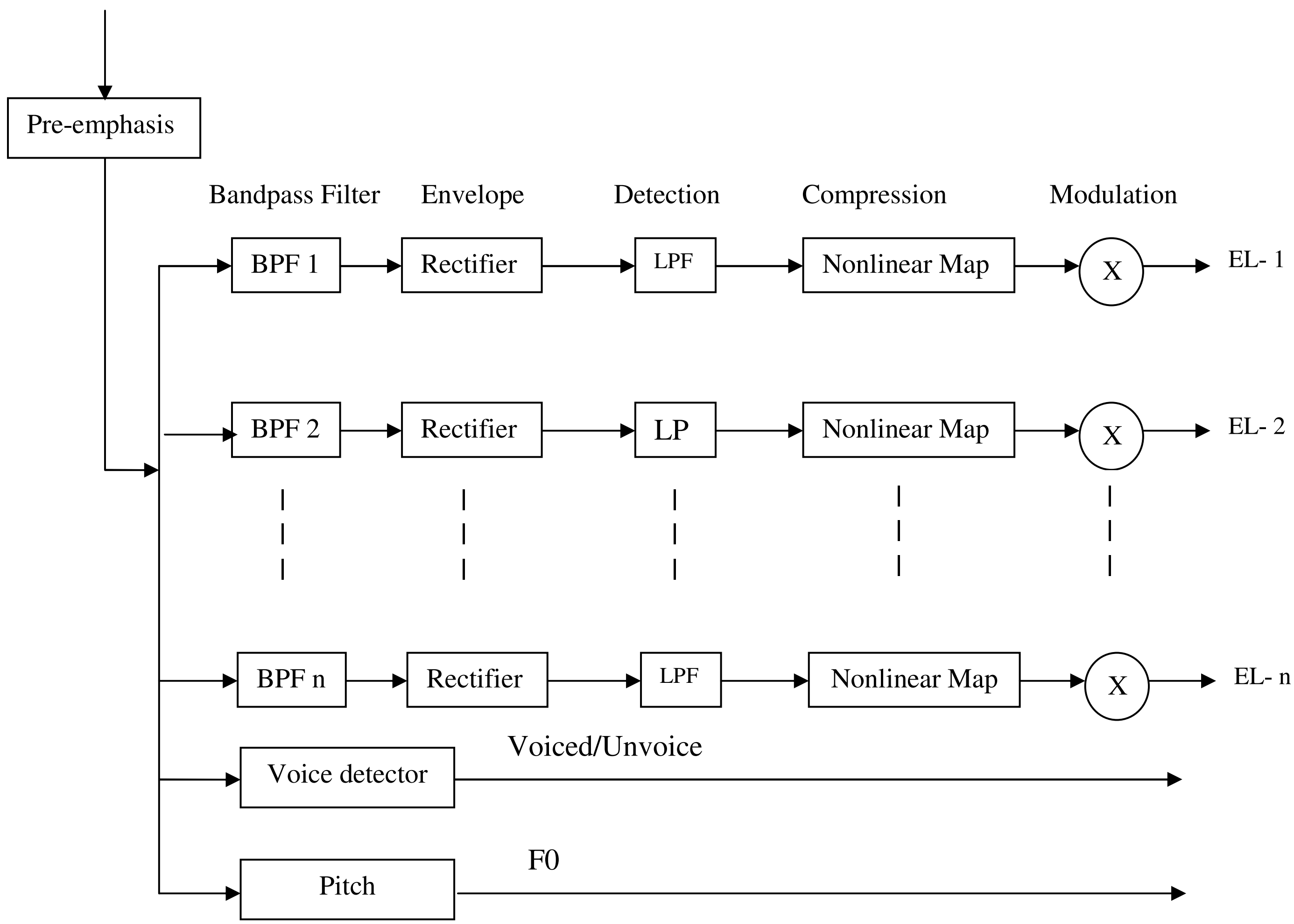

Figure 2.2 Channel Vocoder Analyzer for Cochlear Implants

The envelope of the signal in each channel is estimated by full-wave rectification and low-pass filtering (the envelopes are further down- sampled and quantized for transmission in communication applications). In addition to envelope estimation, the vocoder analyzer makes a voiced/unvoiced decision and estimates the vocal pitch (F0) of the signal. These two pieces of information are transmitted alongside the envelope information. The envelopes, estimated using the vocoder analyzer are transmitted to the individual electrodes in the cochlear implant to stimulate the neurons electrically. The performance of cochlear implants highly depends on various factors such as patient's age, history of deafness/hearing, number of surviving spiral ganglion cells, number of electrodes, speech processing strategy, etc [1]. To combat various issues related to any class of user, various speech processing strategies have been developed that exploit different design characteristics like number of electrodes, stimulation process, etc. CIS and ACE are two such strategies that deal with two specific problems and are used in this work to study the cochlear implants. They are discussed in the following sections.

2.3 Continuous Interleaved Sampling (CIS) Strategy

In general, the filtered envelopes of all channels are used to stimulate all the electrodes in the implant simultaneously. A major problem with simultaneous stimulation is the interaction between channels caused by the summation of electrical fields from individual electrodes. Thus neural responses to stimuli from one electrode may be significantly distorted by stimuli from other electrodes. These interactions distort speech spectral information and therefore degrade speech intelligibility. Continuous Interleaved Sampling (CIS) strategy proposed by Research Triangle Institute (RTI) addresses this electrode interaction problem by use of non simultaneous interleaved pulses [10], [11]. They proposed modulating the filtered waveforms by trains of biphasic pulses that were delivered to the electrodes in a non-overlapping interleaved fashion, so that at any time only one electrode is stimulated. To generate the envelope for each channel, the vocoder method described earlier is employed. The signal is first pre-emphasized and then applied to a bank of bandpass filters. The envelopes of the outputs of these bandpass filters are then full-wave rectified and low-pass filtered (typically with 200 or 400 Hz cutoff frequency). The envelopes of the outputs of the bandpass filters are finally compressed and used to modulate biphasic pulses. To ensure the envelopes that are stimulating the electrodes fit the patient's dynamic range, a non-linear compression (such as logarithmic) is applied on the envelopes. The compressed envelopes are used to modulate a series of biphasic pulse carrier signal at a constant rate and biphasic pulses are delivered to the electrodes in a non-overlapping fashion.

2.4 ACE: Spectral-Maxima Strategy

It is believed that speech can be well understood when only the peaks in the short-term spectrum are used in the speech synthesis. Advanced Combination Encoders (ACE) strategy on Cochlear Corporation devices [11] uses this principle to use only the first few spectral maximum channels to stimulate (spectral-maxima strategy). It is similar to the “n-of-m” strategies in other devices [12], [13]. The spectral-maxima strategy is similar to the CIS strategy except that the number of electrodes stimulated is smaller than the total number of analysis channels. In this strategy, the signal is processed through m bandpass filters from which only a subset n ($n<m$) of the envelope amplitudes are selected for stimulation. More specifically, the n maximum envelope amplitudes are selected for stimulation. The spectral-maxima strategy is sometimes called the “n-of-m” strategy or peak-picking strategy and is available in both Med-El and Nucleus-24 devices. One advantage with selecting only fewer

maxima spectral bands is that the rate of stimulation can be increased to provide more temporal information. Several studies compared the performance of spectral-maxima and CIS strategies [12], [13], [17], [18]. Cochlear implant simulation studies by Dorman et al, [69] indicated high performance with the spectral-maxima strategy even when a small number of maxima were selected in each cycle.

2.4.1 Strategy Design Parameters

The encoding strategies involved in speech processors for cochlear implants can be configured in a number of ways by varying design parameters (e.g., filter spacing, envelope cut-off frequencies, etc) of the vocoder. These parameters include, among other things, the envelope detection method, stimulation rate (i.e., the number of pulses delivered to the electrodes per second), shape of compression function, and maximum comfortable levels (MCL). The main challenge of present day research is to optimally determine these parameters for each individual patient. To address this problem, research platforms are required to be developed, and that is one of the aims of this thesis.

2.5 Bilateral Cochlear Implants

Although unilateral cochlear implants are highly successful and enable deaf people to achieve good speech intelligibility, they cannot utilize the advantage of binaural cues that normal hearing uses in localizing sounds. Interaural level differences (ILD) and interaural time differences (ITD) play an important role in improving speech understanding in noise and localizing sound sources with high accuracy [43]. Recently, study on bilateral implants gained great focus and researchers exploring benefits and issues with two implants.

CHAPTER 3

PDA RESEARCH PLATFORM FOR COCHLEAR IMPLANTS

Chapter 2 described the need for a flexible research platform for sustained growth in the field of cochlear implant research. Many research speech processors developed in the past provide researchers a platform to develop and evaluate new signal processing algorithms, most of them are suffering from their restricted use in the laboratory environment and lack of availability of technical resources. Although the research processors made by Cochlear Corporation and HearWorks [28], [29], [30], [55], [56] are portable and wearable, they are not easy to update and reprogram and they require highly skilled programmers to change the algorithms and strategies. The main drawback of such research processors is that they do not allow to implement new novel algorithms in later stages of cochlear implants use. As mentioned in chapter 2, the performance of CI users changes gradually with time, which necessitates updating the existing algorithms and strategies and investigating novel algorithms from time to time. Such evaluations would give us a more realistic assessment of the performance of new algorithms and new experimental methods. In addition to being portable, the research processor needs to be flexible, easy to use, so that it can be used by both clinicians and researchers without requiring advanced programming skills. The PDA research platform [15] demonstrates all these benefits for research in cochlear implants which requires minimal investment for development and upgrading to new needs. Hence this work extends the use of PDA research platform to provide psychoacoustic experiments. This chapter provides an overview of the PDA-based research platform and its application for this thesis.

3.1 PDA Based Research Platform: Highly Portable, Flexible and Ease to Program

The PDA is a portable programmable device that offers several advantages to use it as speech processors for cochlear implants. PDA offers high flexibility to programmers to develop software solutions using many programming languages like C, C++, VB, Pre-developed libraries like Intel IPP [20], Assembly level programming, and even LabView [21]. As the present day PDAs possess powerful computing capability with operating frequencies ranging from 400MHz to 625MHz, it allows to implement several complex real time speech processing algorithms for study in cochlear implants. Also PDAs are inexpensive devices that provide variety of input/output ports to interface with other external devices like auxiliary microphone, recording devices, etc. Excellent wireless connectivity with technologies such as Wi-Fi, Bluetooth, and Infrared, makes the PDA a universal device allowing patients to interact with doctors remotely and upload/ download their medical documents. Furthermore PDAs are easily adaptable for new and emerging technologies so that users need not to change software or hardware components. All these benefits make the PDA an intelligent choice to use as a research platform for the study of cochlear implants.

3.2 PDA-Research Platform Components

To provide more features and flexibility to study the behavior of cochlear implants, this thesis focuses on developing a software platform for psychoacoustic experiments. The following sections discuss various components of this platform, which can be broadly divided into three blocks: PDA, the cochlear implant transmission coil, and the Secure Digital Input Output (SDIO) interface board.

3.3 Personal Digital Assistant (PDA)

A PDA with supporting secure digital (SD) port and sufficient data block size and transfer rate is an important requirement for PDA research platform. Though, any PDA or smart phone with SD port can be employed, HP iPAQ PDA powered Windows Mobile OS 5.0 and Intel PXA270 (ARM-based processor) as core running at frequency of about 625MHz is employed to demonstrate the research platform developed by Loizou et.al [15].

3.4 Cochlear Implant Transmission Coil

Cochlear implants are intended to mimic the operation of a healthy cochlea [1]. All cochlear implants have microphones to pick up sound, and a sound processor to process the sound and generate electrical signals which excite the electrodes placed in the cochlea through a transmission coil. The processor is responsible for splitting the signals into different frequency bands and sending envelope information to the different electrodes. These operations are similar to that what is carried out by a healthy cochlea. For the PDA-based research platform, the signal processing is carried out in the PDA. Therefore, an interface has to be provided for the PDA to communicate with the transmission coil. This is done by the SDIO interface board.

3.5 SDIO Interface Board

After processing the speech signal using speech processing techniques like vocoder, the series of envelope amplitudes are required to be converted into a pulse sequence for each electrode on the cochlear implant, and an RF communication transmitter is required (for transcutaneous connection based cochlear implants) to present the pulses to the implant using a wireless link. To implement these functions, a reprogrammable FPGA based circuit board is designed [15], which is the interface to the PDA through SDIO protocol using the SD card port on the PDA, thus making the PDA along with SDIO-FPGA board a flexible speech processor for cochlear implants, The SDIO-FPGA board developed by Loizou, et. al [15]has two input channels for receiving speech signals from two microphones located in the left and right BTEs (behind the ear) and two corresponding output channels to transmit the processed output (envelopes) of the speech signals. The SDIO board communicates with the PDA through a 4-bit communication protocol and uses an embedded protocol [31] to stimulate the implant. Figure 3.1 shows the block diagram of PDA-Research platform that shows connections between PDA, SDIO-FPGA board, BTE, and Transmission coils. It also shows the test-setup using implant-in-box and an oscilloscope.

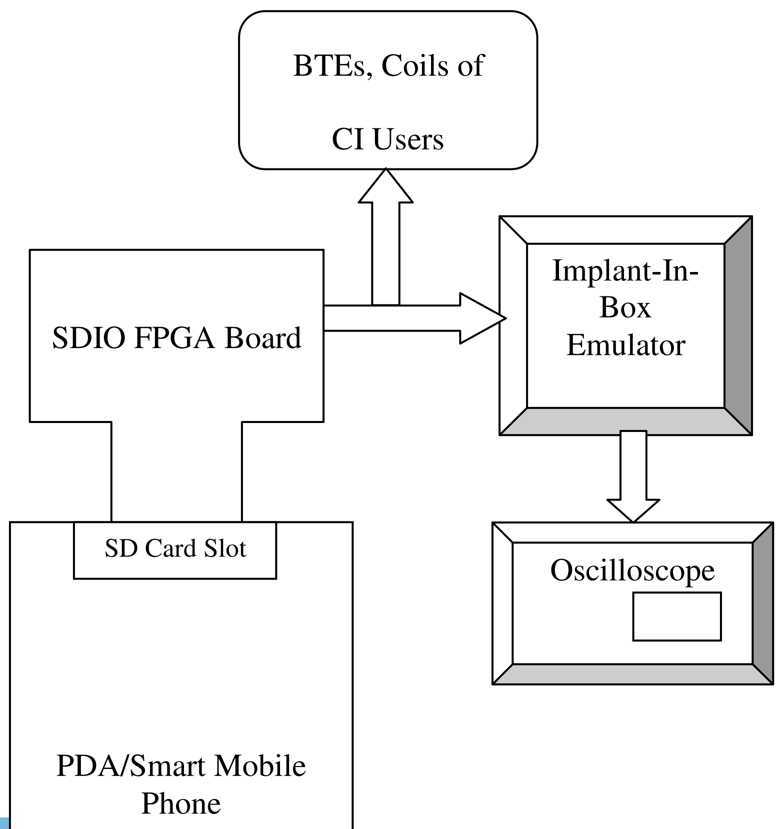

Figure 3.1 Block Diagram of PDA Research Platform with SDIO-FPGA Board

Figure 3.2 shows picture of the set-up for the PDA-based research platform to show the connection between the PDA, the BTE (Behind the ear) units and the SDIO interface board. The SDIO-FPGA board consists of the following components [19]:

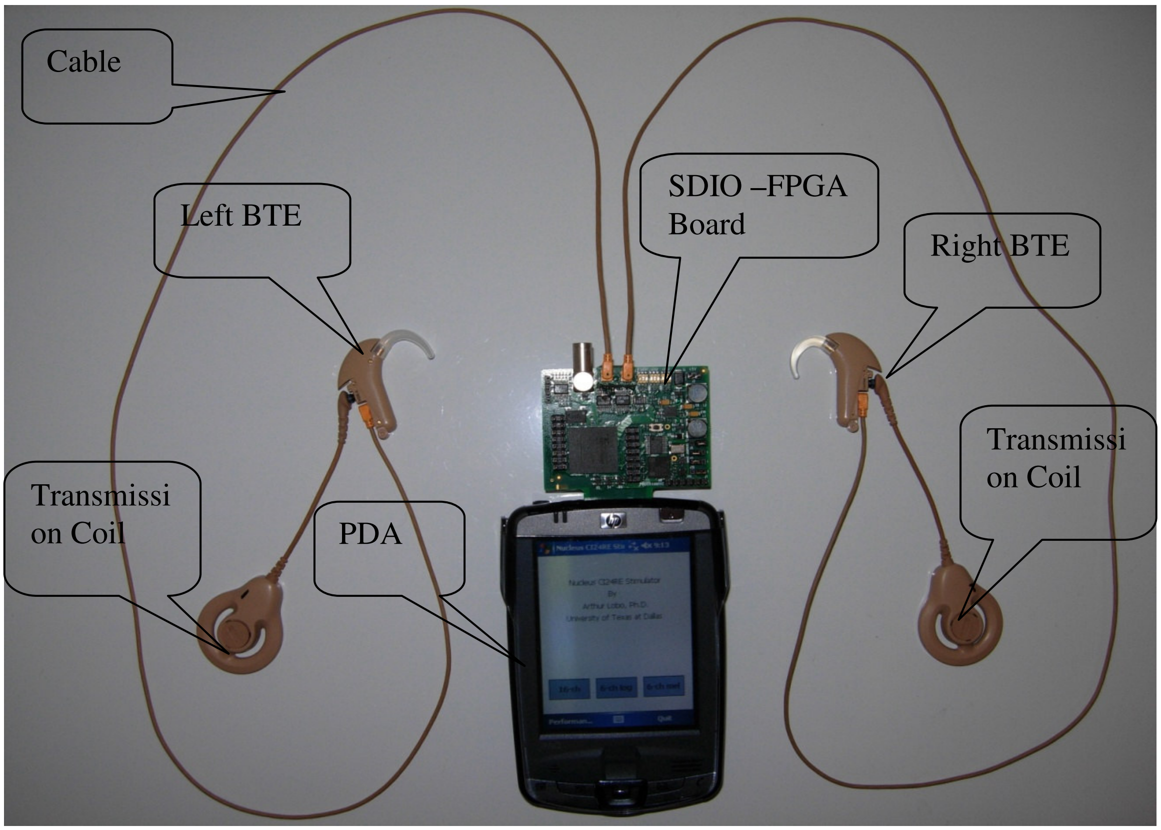

Figure 3.2. Typical connections setup to use the PDA research platform.

Arasan AC 26000: This application specific integrated circuit (ASIC) is the SDIO controller that implements the SD Physical Layer specification 1.10 [19], [32]. It communicates with the ARM-based processor on the PDA through a command response interface. The signals from the PDA are converted into High Speed Synchronous Peripheral Interface (HSSPI) and send to an on-board FPGA by this ASIC. The ASIC also provides options for parallel data transfer [32].

EEPROM 24LC08B: An EEPROM is provided to hold the initialization parameters which are provided to the ASIC during start-up via the I2C bus.

Xilinx XC3S4000L FPGA: The FPGA receives the HSSPI data from the ASIC and implements the embedded protocol [31] to send signals to the Freedom coil of the implant. The FPGA is programmed using a 'receive-transmit' state machine logic. Since, there are two BTEs connected to the board, the SDIO board can be used for unilateral as well as bilateral cochlear implant studies.

Xilinx XCF04S Flash PROM: The program logic of the FPGA is stored in this PROM. During power on, the FPGA is booted from the PROM.

Linear Technology LTC 6912 preamplifier: There are two inverting amplifiers with a programmable gain. The FPGA controls the gain of these amplifiers.

Linear Technology LTC 1407 A/D converter: This on-board stereo Analog-Digital converter samples the microphone outputs from the BTE connected to the respective Freedom coils. The digital output is received by the FPGA over the Synchronous Peripheral Interface (SPI) bus.

Texas Instruments TPS75003 power management IC: A triple-supply power management IC from Texas Instruments supplies power to the FPGA and the PROM. The IC is a switching regulator type which is on only when power is required. An external battery is used for power source. The SDIO board plugs into the SDIO slot of the PDA.

3.5.1 System Requirements for Speech Processing Applications Development:

After introducing the PDA research platform, it is useful to discuss the development of speech processing strategies for it. The Following hardware and software requirements are needed:

I. **Hardware requirements**:

a. PDA device: (recommended: HP iPAQ 2790b PDA with 624 MHz PXA270 XScale processor).

b. SDIO-FPGA Board: The SDIO-FPGA board has to be plugged into the SD slot.

c. 5Volts Battery Pack: The SDIO-FPGA board is externally powered by the 5V supply.

II. **Software requirements**:

a. Development Environment (IDE): MS Visual Studio 2005 or later versions.

b. OS: Windows Mobile 5.0 for PDA target and Windows XP Professional SP2 for PC development.

c. Compiler: Intel C++ 2.0 for Windows CE, ARM assembler, Intel Integrated Performance Primitives on Intel Personal Internet Client Architecture Processors (compliant with ARM architecture V5TE) version 5.0, Platform Builder for WinCE 5.0.

d. Hardware Device Specific Libraries: ArasanSDIO-CPULike.dll for communication with the AC2600 ASIC.

3.6 Software Programs Development on PDA

To develop an application on the PDA-based research platform, the following procedure is adopted. The prototype of the target application is first developed in MATLAB in floating-point to check for the correctness of the application. Since most of the PDAs lack floating-point arithmetic capabilities, implementation in fixed-point arithmetic is inevitable for efficient and real time implementations. Therefore, the floating-point design developed is converted to an equivalent fixed-point implementation in MATLAB as a reference design for PDA implementation. The equivalent fixed point implementation is carried out in C and ARM assembly programming languages, and the implementation is verified by comparing them with the MATLAB fixed point implementation results using mean square error (MSE) metrics.

The general approach for the development of speech processing techniques for cochlear implants followed in this thesis work involves fixed-point PDA implementation in C using the Microsoft Visual Studio 2005 development environment, and use of Intel Performance Primitives (IPP) to implement basic operations like vector multiplication, FFT, etc [20]. Intel Performance Primitives (IPP) offers an optimized library with a C function interface for primitive signal processing operation for Intel-Xscale processor based PDAs. Thus, they are called much like any other function in C in PDA speech processing solutions. The applications are compiled and built for an ARM-based Pocket PC target. Microsoft Active Sync is used by the Visual Studio environment to deploy the application on the PDA. The noise band vocoder [15] developed by Lobo et.al., describes speech processing applications development on the PDA research platform for Nucleus24M cochlear implant devices.

HP iPAQ hx2000 Pocket PC, and Dell Axim X30 Pocket PC are used in the development of the PDA research platform and speech processing techniques to test the same. These PDAs are powered with Windows Mobile OS 5.0 and Intel PXA270 ARM-based processor at operating frequency of about 540-625MHz.

CHAPTER 4

TRANSFORMED UP-DOWN ADAPTIVE METHODS FOR PSYCHOACOUSTIC EXPERIMENTS

The speech dimensions such as pitch, loudness are perceptual quantities that cannot be determined directly, as they are highly subjective. Traditionally, psychophysical methods are popular to estimate perceptual cues and parameters of sound signals. Psychophysical methods involve presenting series of stimulus signal to listeners and the simple feedback is obtained about the stimulus. Some methods present various alternative forced choices (AFC) of stimulus and listeners are directed to select a specific stimulus signal. A psychometric function that describes the listener's perceptual behavior statistically is estimated using psychophysical methods. Transformed up-down methods proposed by Levitt [33] estimate a point on the psychometric function efficiently and effectively. This chapter details about various psychophysical methods used for speech applications.

4.1 Psychometric Function

A psychometric function describes the relationship between a parameter of a physical stimulus and the responses of a person who has to decide about a certain aspect of that stimulus. Usually, the psychometric functions are plotted with the percentage of correct responses (or an expected value) on the ordinate and the physical stimulus parameter on the abscissa. If the stimulus parameter is very far towards one end of its possible range, the person likely will be able to respond correctly. Around the origin of the stimulus level, the subjects cannot perceive the stimulus appropriately and therefore the probability of correct responses is minimal. Two parameters most generally characterize psychometric functions, namely, mean convergence location on the curve, usually defined as stimulus corresponding to 50% probability of response (X_{50}), and spread of the psychometric curve. For example, the normal distributed psychometric function estimates these parameters as mean and variance of the probability distribution function (pdf), the mean and variance of X_{50} point and difference limen (DL) as shown in figure 4.1. Estimation of a point on the psychometric function is often required to estimate psychophysical thresholds and differential thresholds [33]. Conventionally several procedures can be used to estimate the psychometric functions [37], [38], including the method of constants, up-down methods, and transformed adaptive methods. The following sections describe popular conventional and adaptive methods used to estimate psychometric functions.

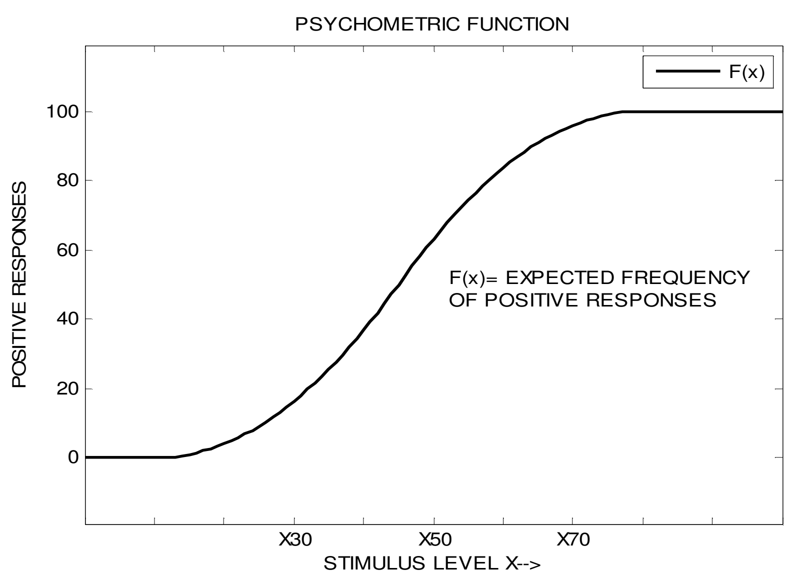

Figure 4.1 Typical Psychometric Function, Levitt, 1971[33]

4.2 Psychophysical Experimental Methods

In psychophysics, experiments seek to determine whether the subject can detect a stimulus, identify it, differentiate between other stimuli, and describe the magnitude or nature of this difference. The subjects are presented with stimulus sequentially and a user interface is provided to capture their response about the presented stimulus. During the experiment, the stimulus level or difference between comparative stimuli is changed from low level to high level or vice-versa. Based on the strategy to update the stimuli in each presentation, and experiment termination method, experimental procedures can be classified as conventional (or traditional) and adaptive psychophysical methods. The adaptive psychophysical methods are sophisticated and simpler to estimate.

4.2.1 Conventional Psychophysical Methods

Psychophysical experiments have traditionally used three methods for testing subjects' perception in stimulus detection and difference detection experiments: the method of limits, the method of constant stimuli and the method of adjustment.

Method of Limits: With the method of limits, some property of the stimulus is gradually changed from relatively large value or small value. In ascending method of limits, the stimulus is started from a low level which is not detectable by listener, and gradually increased until the listener reports that it is detected. For example, if the experiment is to determine the minimum amplitude of sound that can be detected, the sound begins too quietly to be perceived, and is made gradually louder. In the descending method of limits, the stimulus level is gradually decreased from a high value. The threshold is estimated as the level of the stimulus at which it is just detected.

This method suffers from the error of habituation (the subject may become accustomed to reporting that they perceive a stimulus and may continue reporting the same way even beyond the threshold) and the error of expectation (the subject may also anticipate that the stimulus is about to become detectable or undetectable and may make a premature judgment).

Method of Constant Stimuli: Instead of being presented in ascending or descending order, in the method of constant stimuli the levels of a certain property of the stimulus are chosen and presented randomly. This reduces errors of habituation and expectation as the subjects cannot predict the level of the next stimulus easily. This method allows for full sampling of the psychometric function, but requires in large number of trials when several conditions are interleaved.

Method of Adjustment: The method of adjustment (MOA) asks the subject to control the level of the stimulus parameter under examination, instructs them to alter it until it is just barely detectable against the background noise, or is the same as the level of another set of stimuli. This is repeated many times and hence this method is also called the method of average error.

4.2.2 Adaptive Psychophysical Methods

The conventional methods mentioned above often considered as inefficient, because the psychometric threshold is usually unknown and a lot of data has to be collected at points on the psychometric function that provide little information about its shape. The adaptive methods are simple and efficient, thus allowing adjustment of the point convergence on the psychometric function. Adaptation of step size further helps in minimizing the number of data points required to estimate the thresholds or perceptual cues. Also, the adaptive procedures can be designed such that the sampled points are clustered around a point on psychometric threshold.

Staircase Procedures

Staircases usually begin with a high level of stimulus, which is easy to detect, and the level is then reduced until the subject cannot detect it, at which point the staircase reverses and stimulus level is increased until the subject responds positively, to trigger another reversal. The reversals points are averaged to estimate the threshold. There are many different types of staircase, utilizing many different decision and termination rules. Step-size, up-down rules and the spread of the underlying psychometric function determine the convergence point on the psychometric function.

Simple Up-Down Methods:

Simple up-down methods are very popular and relatively efficient methods for estimating the 50% level on the psychometric function. In this method, stimulus level presented to subjects are decreased (down) or increased (up) based on the response from the subject from previous stimulus. After each trial the stimulus level is adjusted by a priori chosen step size. Figure 4.2 shows typical stimuli data during a simple up-down experiment, with + and – symbols on stimulus point representing positive or negative responses respectively. The simple up-down technique estimates most of the observation around X_{50}-point, which estimates the 50%-point on the psychometric function. The disadvantage with this method is it cannot estimate other than the 50% point on the psychometric function. Furthermore, the subject can anticipate the next stimulus level given a sequential adjustment rule. The data points estimated during the test are needs to be analyzed to estimate the X_{50}-point more accurately. A popular technique to analyze the data points proposed by Wetherill [39] averages the peaks and valleys to estimate the stimulus thresholds.

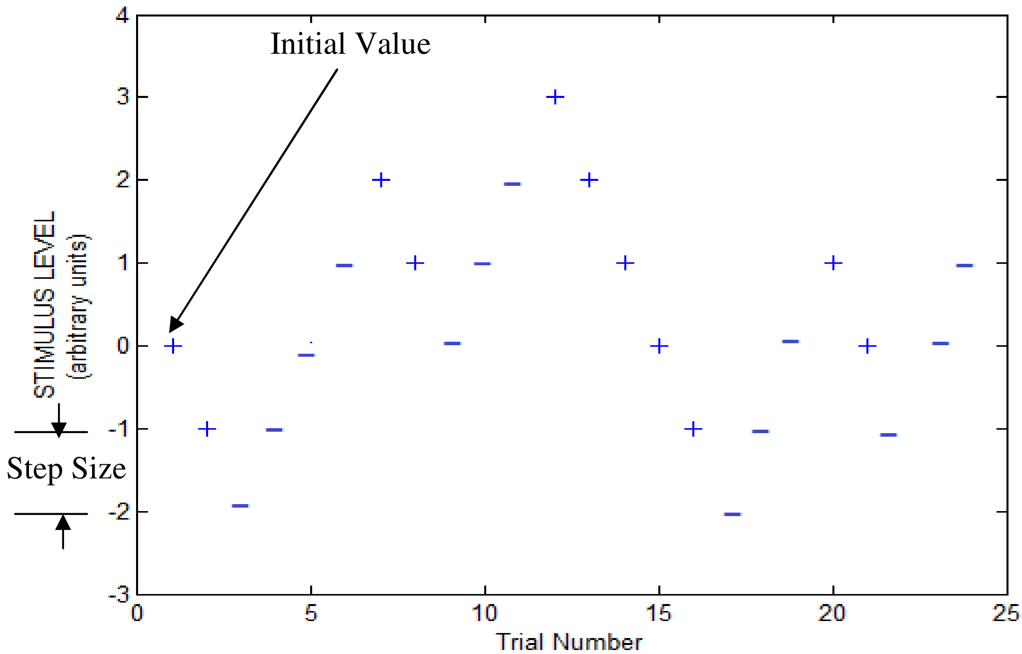

Figure 4.2 Typical data for simple up-down procedure using fixed step size, (+/- indicates positive or negative response. (Levitt, 1971 [33])

Transformed Up-Down Methods (Levitt 1971 [33]):

As mentioned above, the simple up-down procedures estimate only the X_{50}-point on the psychometric function, and is not suitable to estimating points other than the X_{50}-point. Transformed up-down methods provide generalized techniques to estimate other points on psychometric function. The transformed up-down methods, classifies the observations into two mutually exclusive groups, namely the UP group and DOWN group. The rule for the controlling the stimulus level is similar to the simple up-down rule, except that the stimulus level is changed only after a sequence of observations belonging to either UP or DOWN group is obtained. For example 1-up, 2-down procedure uses two consecutive positive responses in down group and one negative response in up group, where, the stimulus level is decreased only after two consecutive positive responses and it is increased in step size after every negative response from the subject. The number of observations in each group determines the estimation point on psychometric function. Using the transformed up-down methods, the stimulus level convergence point on psychometric function corresponds to the probability where the down response sequence is equal to the up response sequence. Thus for 1-up, 2-down method converges at the 70.7% point on the psychometric function.

Given the simplicity and efficiency with the transformed up-down methods, this thesis implemented the transformed up-down adaptation procedure with two or three alternative forced choice algorithms (2AFC/3AFC) in the software tool designed. Various monaural parameters Speech Reception Threshold (SRT), Just Noticeable Difference (JND) in frequency and binaural cues (ILD, ITD) estimations procedures are developed using these methods. In each trial, 3 sets of stimulus intervals are presented, where one interval among the three consists of an adaptive parameter (for example intensity for SRT, interaural time for ILD tests) which is varied, while other intervals contain the same stimulus. The interval with the adaptive stimulus parameter is assigned randomly. The details of the experiments are presented in Chapter 7.

CHAPTER 5

HEARING CUES AND BINAURAL HEARING

Normal hearing listeners make use of spectral, intensity, time and pitch cues for sound perception, localization and speech intelligibility. The ability to localize, detect or isolate origin of sound source and motion in the sound source is a key benefit that normal hearing listeners enjoy when using two ears. Interaural spectral content, interaural time difference (ITD), and interaural level difference (ILD) are important cues that play vital role in sound source localization and speech intelligibility in noise [49], [50]. While the spectral information such as spectral notches at high frequencies are mainly used for identifying the sound source elevation [50], the ILD and ITD cues are primarily used for localization of the sound sources in the horizontal plane. To restore normal hearing abilities in cochlear implants users, the speech processing techniques and stimulation devices should be able to preserve important cues involved in sound localization, speech perception and intelligibility. Therefore, the study of perception of those cues and parameters in normal hearing and cochlear implant users would help in the design of speech processing techniques for bilateral cochlear implants. Since perception of sound cues involves the psychological behavior of users, psychoacoustic experiments are essential for their study. As this thesis focuses on developing a platform for psychoacoustic experiments, it is appropriate to discuss the important psychoacoustic parameters used in speech perception and sound localization. The rest of this chapter discusses about the importance of prominent binaural and other speech cues in cochlear implants.

5.1 Hearing Cues in Unilateral Cochlear Implants

In the past, profoundly hearing impaired patients were implanted in only one ear (unilateral cochlear implant), and reserving the other ear for the future use [1]. The unilateral cochlear implants have proven to be highly effective in enabling deaf people to achieve good speech understanding in quiet. Several studies were carried to explore spectral, loudness cues, and pitch cues in unilateral cochlear implants to improve the speech perception [1], [51], [61], [62]. Loudness balancing, pitch ranking, pitch discrimination, and speech reception thresholds are few examples of speech perception cues explored for unilateral cochlear implants.

5.1.1 Pitch Ranking

The ability to perceive pitch increase or decrease determines the speech and music perception in listeners. For cochlear implant users music perception is poorer than speech perception and worse than normal hearing listeners. Pitch raking procedures study the ability of listeners to judge increase or decrease in pitch of a stimulus. Pitch ranking procedures present a sequence of stimuli with increasing or decreasing pitch and listener's feedback is taken to determine their ability to recognize varying pitch.

5.1.2 Frequency Discrimination

The main goal of this study was to determine the subject's sensitivity to frequency differences. The minimum difference in the frequency that can be detected by the subjects is estimated using frequency discrimination tests. Frequency discrimination can be measured using variety of psychoacoustic adaptive methods mentioned in Chapter 4. The listeners are presented with two stimuli with frequency difference between them changed from large perceivable value to a small value adaptively.

5.1.3 Loudness Balance

It has been known that loudness may increase with frequency and stimulation rate in cochlear implants. In order to eliminate possible loudness cues with different stimulation parameters (especially with rate), loudness has to be balanced using adaptive procedures. The reference and target stimuli at two different rates are presented to listeners and loudness of the target stimulus is balanced till the loudness difference between the reference and target are not perceived by adaptively changing the loudness difference between two stimulus signals from a prominently perceivable difference (in dBs) to a minimum detectable loudness difference (few dBs).

5.1.4 Binaural Hearing

Though monaural hearing using single implant achieves good sound perception and speech understanding, it does not provide sound localization and efficient speech intelligibility in various environments like speech in babble noise, reverberation, etc. Recently, researchers and doctors have tried to provide patients with the advantages of binaural hearing as in normal hearing by using bilateral cochlear implants. Bilaterally implanted patients are reported to be benefitted in understanding speech in noisy situations [70], which is attributed to the fact that the speech and noise are spatially separated, and hence different signal-to-noise ratio (SNR) are present at the two ears. Also, the head-shadow effect causes different spectral information at each ear that further benefits the listener by attending to the ear with the better SNR or better spectral cues. Additionally, binaural unmasking benefits referred to as binaural squelch are available from the combined processing of the signals at both ears. Studies of Durlach and Colburn, (1978) [68] reported that the binaural hearing improves sound localization, especially using the interaural cues, such as interaural time differences (ITD) and interaural level differences (ILD). As a consequence, studies on understanding the influence of CI processing on interaural cues and comparing interaural cues in CI users with normal hearing listeners gained attention to understand sound localization benefits using bilateral implants. Several studies were conducted by van Hoesel, Clark and others, [43], [47], [62], [65], [67] to measure just noticeable differences (JND) for ILD and ITD in cochlear implant users and normal hearing subjects. Those studies indicate that ILD sensitivity of CI users are consistent and comparable with that of normal hearing results, but ITD sensitivity of CI users is inconsistent and far higher than normal hearing subjects. However, they reported consistent improvement with use of binaural hearing of CI users in sound-localization testsand speech perception in noise [49], [71].

Given the importance and interest in interaural cues (ILD, and ITD), the psychoacoustic software platform developed in this thesis concentrates on these interaural cues to measure and verify with state of art results reported in literature. The following sections details ILD and ITD and other cues and their role in sound localization.

5.1.5 Interaural Level Difference (ILD):

Interaural level difference (ILD) describes the amount of change in the perceived loudness from one ear to other ear because of, head shadow effect and different incidence angles of source on each ear. When a sound source is on the extreme right of the head of a listener, the right ear perceives higher loudness than the left ear, which is shown as an example in figure

5.1. The sensitivity to ILD is measured as the just noticeable difference (JND)-ILD, and for normal hearing subjects, it ranges from 0.6 dB to 1 dB [44], [62], [71]. In cochlear implant users, the ILD sensitivity is influenced by different stimulation and speech processing parameters, and it is highly desirable to study their influence to maximize benefits of bilateral cochlear implants.

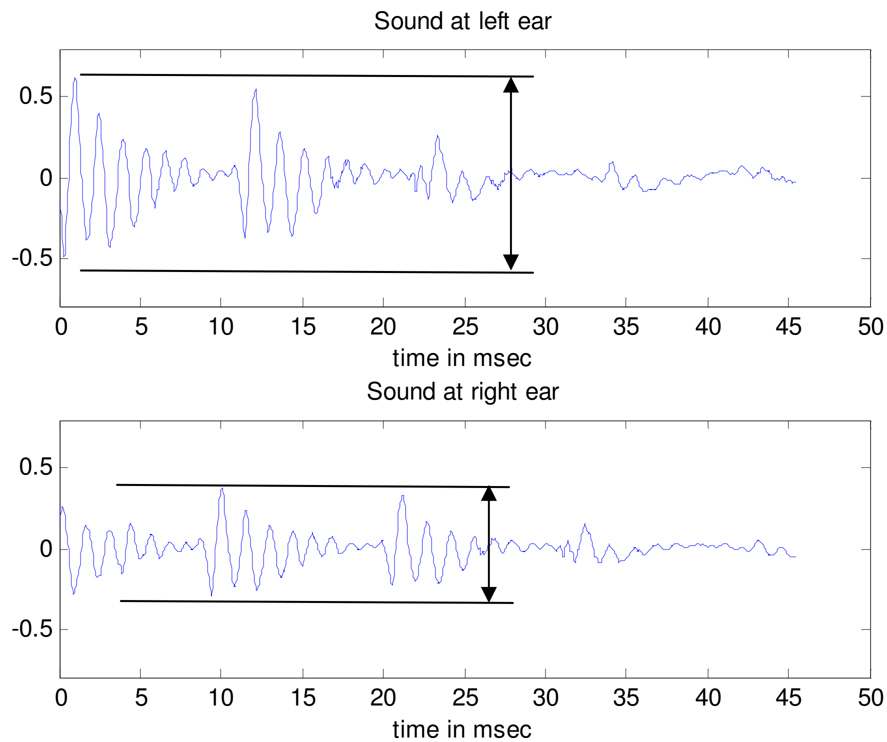

Figure 5.1 Sound signals at left and right ears with ILD

5.1.6 Interaural Time Difference (ITD)

As described in section 5.2, sensitivity to the interaural time differences (ITD) assesses the ability to locate sounds in the horizontal plane. The ITD is defined as the time delay of sound signal reaching an ear (that is located far from the sound source compared to the other ear) after reaching the other ear of a listener as illustrated in the figure 5.2. The sensitivity to ITD is generally measured as the just noticeable difference (JND)-ITD, and the measured ITDs reported in literature show, that the JND-ITD values are highly subjective and are not consistent. The sensitivity varies with frequency and other parameters. In cochlear implant users, the ITD sensitivity will be influenced by stimulation rate and other speech processing parameters. Studies show that sensitivity to ITDs in cochlear implants users is far below that of normal hearing listeners [49].

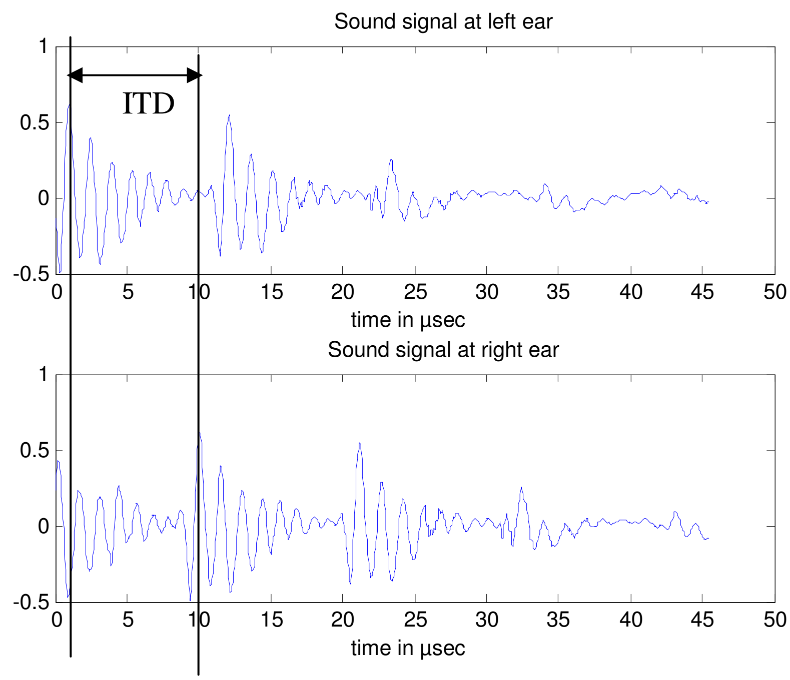

Figure 5.2 Sound signals at Left and Right ears with ITD

5.1.7 Spectral Cues

The spectral cues of a sound arriving at each ear depend on the location of the source relative to the listener due to head shadow and the shape of the pinna. Spectral cues are thought to be important for conveying the elevation of the sound source; left-right or lateral angle is thought to be conveyed by interaural difference. It is also known that spectral cues furnished by the outer ear are important for accurate localization.

5.1.8 Head-related Transfer Functions

When a sound wave travels from the source to the eardrums of a listener, it passes through the listener’s head, neck, and especially through the outer ear (pinnae), which results in a spectral transformation of the sound wave before it is perceived by the inner ear. Moreover, the outer ear's (or pinnae) effect on the spectrum of incoming sounds depends on the angle of incidence of the sound relative to the head [49]. This spectral transformation of sound waves is unique to each listener, and is generally described by head-related transfer functions (HRTFs). The HRTF is defined as a ratio of the complex spectrum of the sound reaching the eardrum and the complex spectrum of the sound source. Recent research indicates that the HRTFs provide important cues in sound localization because they depend on the angle of the sound source relative to each ear. Interaural cues and spectral cues also contribute to HRTF measurements.

Given the details about various hearing cues used in speech understanding, sound perception, and sound localization, a detailed study on these parameters is highly desirable for cochlear implants users to improve the quality of hearing in bilateral cochlear implant users, and a software platform to study these issues is highly needed for the cochlear implants research community. To study these issues, a PDA based software platform is designed to provide flexible tools to conduct psychoacoustic experiments. The Chapters 7 describes the experimental procedures and the Chapter 8 provides the analysis on the results obtained and comparison with other studies.

CHAPTER 6

A SOFTWARE TOOL FOR PSYCHOACOUSTIC EXPERIMENTS

Recently, many experimental interfaces have been developed to have control over the timing parameters and amplitude of each stimulus, along with the option of sending a predefined set of stimuli in succession to work with different speech processing strategies. But, only few experimental interfaces provide researchers access to the programs within the speech processor of a cochlear implant [54], [55]. The PDA research platform [15] is one such experimental platform that gives developers access to program the signal processing strategies along with the flexibility to use different simulation parameters and strategies. It also eases the usage of CI devices by the implanted patient with the help of the PDA. Furthermore, the implementation of a software tool to design psychophysical experiments with a simple and flexible interface would enhances the utility of the PDA research platform.

This chapter describes MATLAB based graphical user interface software to design psychoacoustic experiments for the PDA research platform. Section 6.1 explains the crux of the GUI software tool for the PDA platform. Section 6.2 briefs about the MATLAB-GUI features. Section 6.3 gives the details about designing the psychoacoustic experiments with the software tool. Section 6.4 describes the WINSOCK interface that is used to interface MATLAB GUI and PDA programs and section 6.5 explains the working with different parameters for stimulation through PDA using GUI software.

6.1 MATLAB – PDA Based Psychoacoustic Experimental Software

As explained in chapter 3, highly flexible and programmable software for the study and research of cochlear implants is essential for sustained growth in the field cochlear implants. The PDA based research platform developed by Loizou, et.al, [15] is used to exploit the benefits of portable and flexible design approaches in developing simple and programmable software for conducting psychoacoustic experiments. Since psychoacoustic experiments needs involvement of patients and other subjects who are generally inexperienced in the use of software programs. A Graphical User Interface (GUI) based approach is popular for designing psychoacoustic experiments. The adaptive psychoacoustic methods and GUIs are implemented in MATLAB for obvious benefits of fast and simple design with popular programming language MATLAB to researchers. As explained in the following section, windows sockets APIs are used to interface MATLAB experiments with the PDA research platform.

The detailed setup for MATLAB and PDA research platform based psychoacoustic experiments is given in figure 6.1 that includes various software and hardware units. This GUI software architecture for psychoacoustic experiments provides two kinds of stimulation 1) electrical stimulation 2) acoustic stimulation, which facilitates conducting experiments with cochlear implant and normal hearing subject and study the effects of speech processing for cochlear implants, with various stimulation parameters. A set of psychoacoustic experiment procedures are developed in MATLAB to study monaural and binaural cues in CI and NH subjects. During the psychoacoustic experiments the cochlear implant users listen to a set of stimulus signal through the PDA research platform and are asked to give feedback about the stimulus set through the MATLAB GUI. The psychoacoustic procedures running in

MATLAB estimates the next set of stimulus based on the subjects feedback using up-down adaptive methods. The normal hearing subjects will then listen to stimulus speech signals through MATLAB interface only without the use of PDA research platform. The adaptive variable during the experiment is tracked to analyze the results of the experiments. The detailed description of MATLAB GUI and psychoacoustic experiments procedures are given in the following sections.

The software also provides a series of electric stimulus signals for cochlear implant users using ACE or CIS encoding strategies through the PDA research platform.

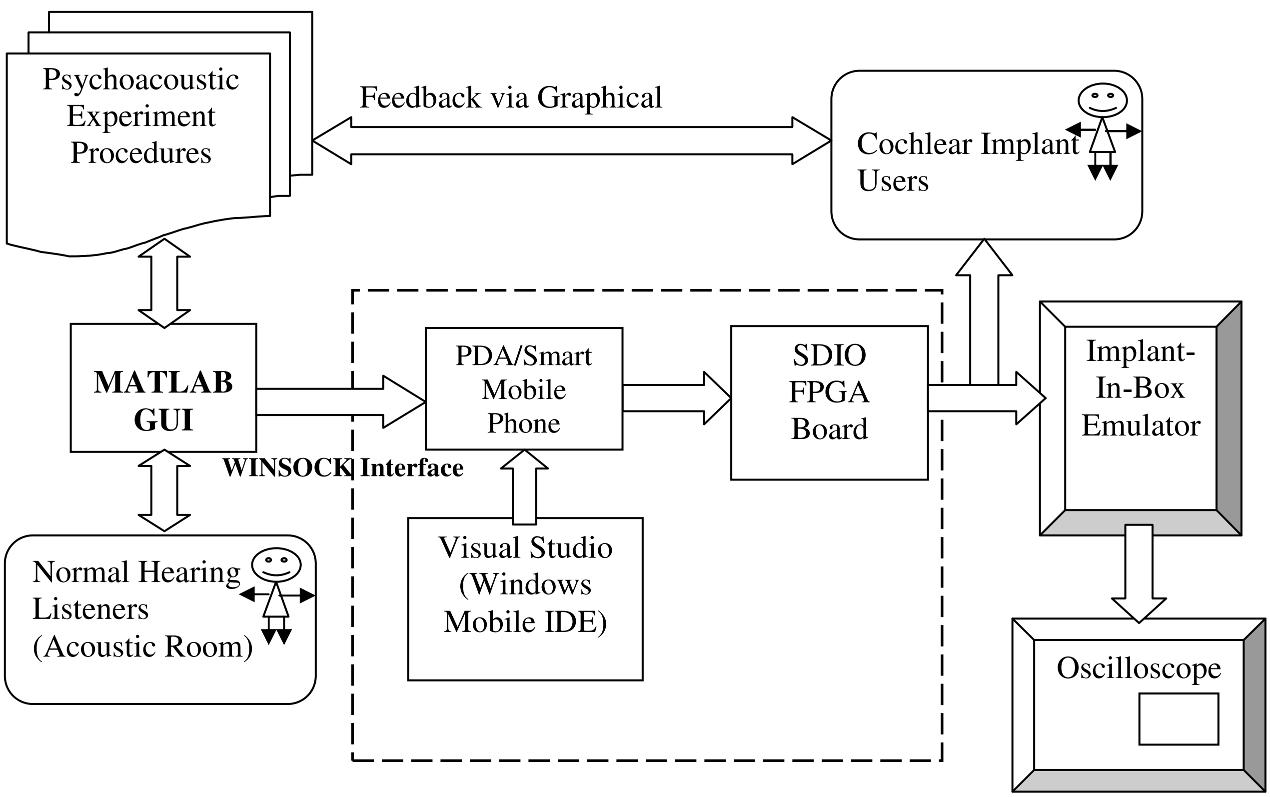

Figure 6.1 Psychoacoustic Experiments Software for PDA research platform

6.2 MATLAB GUI

A MATLAB based GUI software has been developed and various psychoacoustic adaptive methods were designed and integrated. This psychoacoustic software presents a series of acoustic stimulus signal to normal hearing subjects and provides a graphical interface to obtain feedback from the subject under test (SUT). Based on the feedback, the integral adaptive methods estimates the next set of stimulus signals to be presented in order to estimate psychoacoustic features of the subject based on the series of stimulus signals and feedback from the subject. A set of options to set up the experiments can be seen in the figure.

As discussed above, the MATLAB GUI unit in figure 6.1 can provide interactive user interface between the psychoacoustic methods and cochlear implant users. This GUI enables researchers to program monaural and binaural experiments using different up-down adaptive methods with various preselected set of stimuli. Figure 6.2 presents a screen shot of MATLAB GUI window. The *Experiment Setup* panel consists of two list boxes that allow researchers to set up various monaural and binaural experiments. The procedures list box configures the adaptive method used for the experiment, while the stimulation type allows selecting electrical or acoustic stimulation to specify the test subjects as CI user or NH subjects receptively.

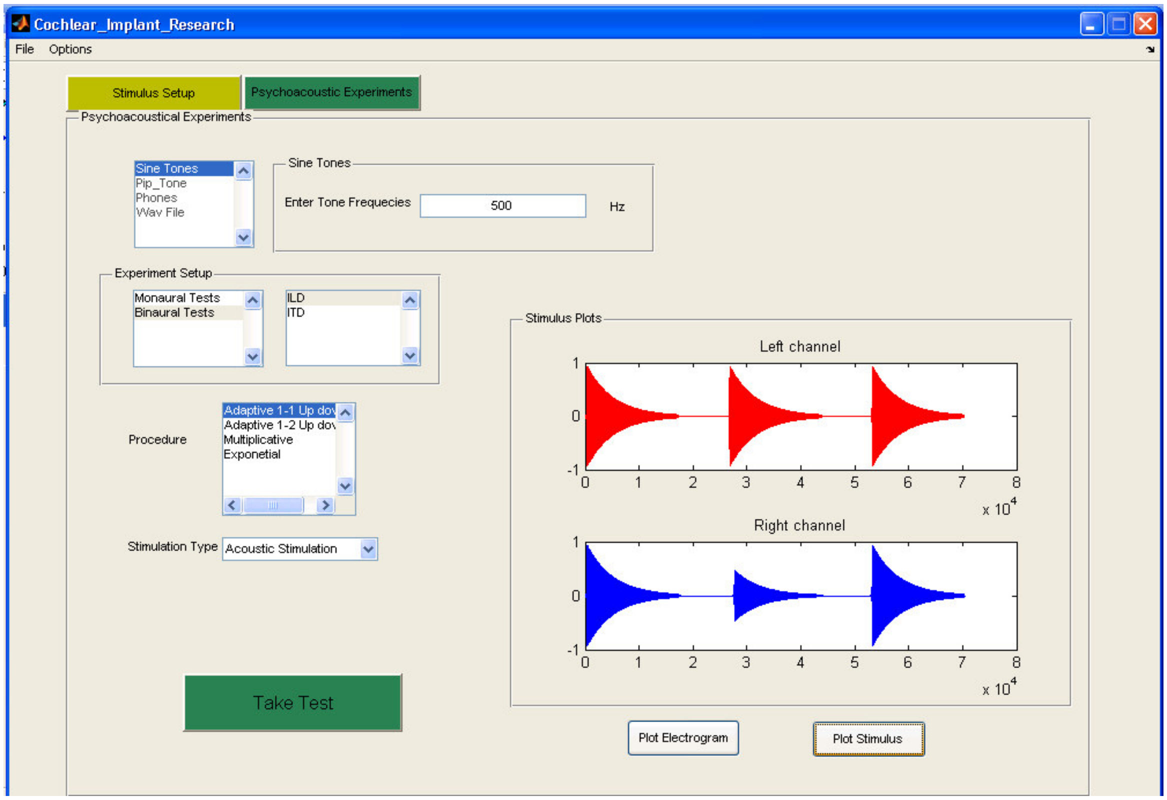

Figure 6.2 MATLAB GUI for Psychoacoustic Experiments showing acoustic stimulus

The time waveform and electrogram of the speech signal used as target stimulus in the psychoacoustic experiment can be viewed in the *stimulus plots* area by clicking the appropriate button. Figure 6.2 shows a GUI window with time waveform of 3 interval stimulus signals while Figure 6.3 shows a GUI window with electrogram of 3 interval stimulus signal encoded with ACE strategy. Standard stimulus parameters for a biphasic pulse include implant type, mode of stimulation, stimulation electrode, reference electrode, pulse width (µs), interphase gap (µs), pulse rate (*pps*) and stimulus amplitude (current amplitude).

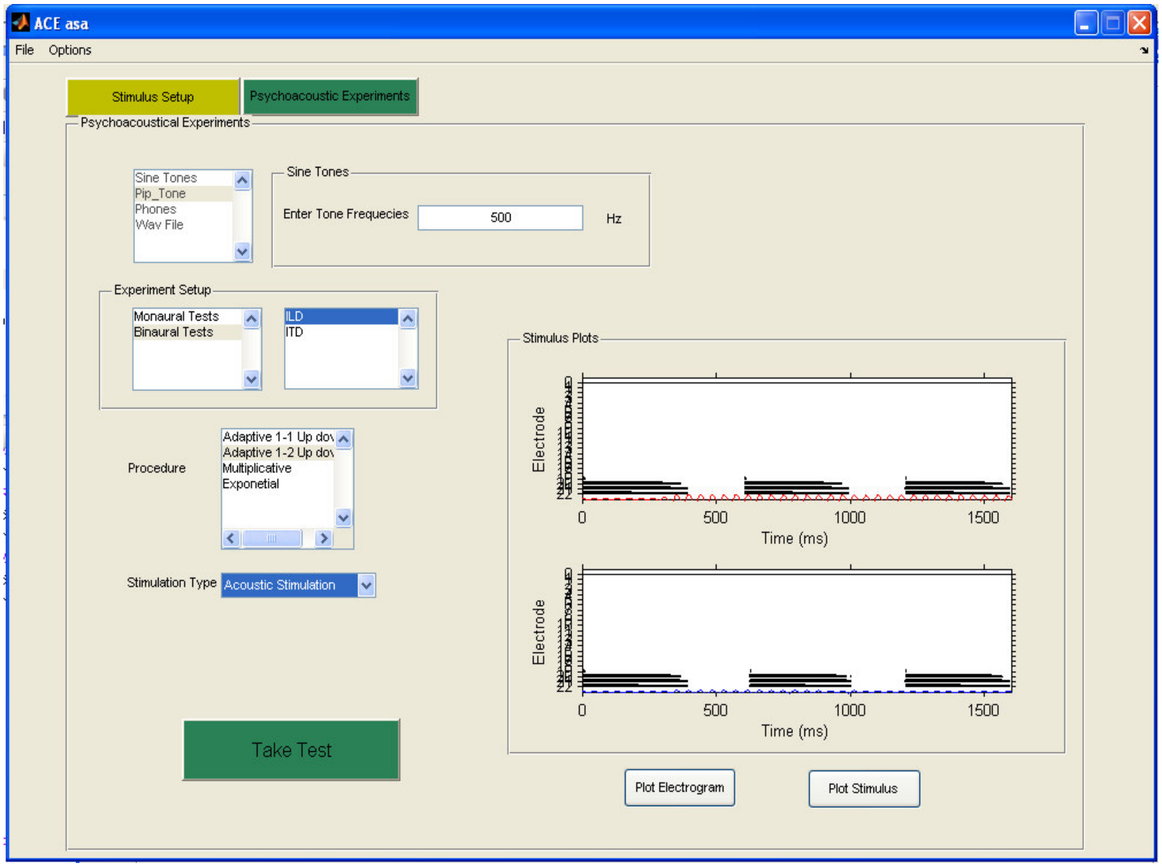

Figure 6.3 MATLAB GUI for psychoacoustic Experiments showing Electric stimulus

Stimuli:

The software tool allows using different kinds of speech signals to generate stimulus for CI and NH subjects. Sine waves with appropriate rise and fall times, and pip tones are incorporated for use in the experiments. The pip tones are generated as sine tones modulated with decaying exponentials, Fig 6.4, shows an example of a 3 interval pip tone stimuli, which contains two reference stimuli and one target stimuli that deviates from the reference stimulus either in pitch, loudness, interaural level difference or interaural time difference. Specific vowel/consonant phone sounds, and simple speech sentences (read from wav files) are also provided for loudness balancing ILD, ITD tests. The amplitude level, pitch, ILD, or ITD is varied adaptively for target stimulus.

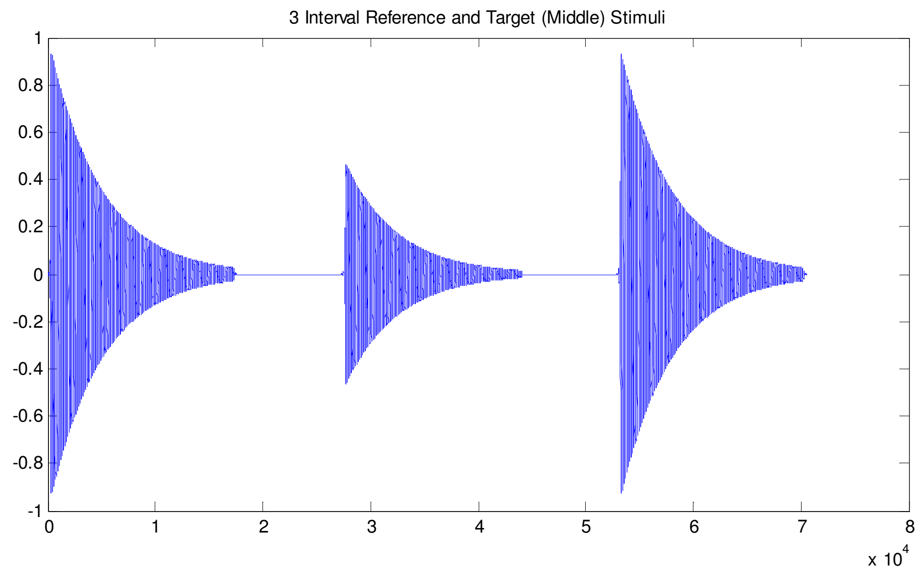

Figure 6.4 Stimulus series for 3 Interval AFC experiment: The loudness of the Target (Middle) stimulus is used as adaptive parameter

For CI users, the above mentioned speech signals are encoded using ACE/CIS strategy to generate a series of stimulus amplitudes for each electrode. These amplitudes along with the parameters are sent to PDA platform using the WINSOCK interface (section 6.4). The PDA platform generates the biphasic pulses. The encoding strategy parameters like implant type, mode of stimulation, electrode configuration, pulse width (µs), pulse rate (*pps*), THR and MCLs are configured with *Stimulus Setup* tab in MATLAB GUI. The configuring parameters using MATLAB GUI is explained in section 6.5.

6.3 Implementation of Psychoacoustic Experiments

Monaural and Binaural psychophysical experimental methods are implemented and experiments are conducted in this study verify its use with PDA research platform and its functionality. Pitch discrimination, loudness balancing, Pitch Ranking, speech reception threshold (monaural experiments) ITD and ILD (binaural experiments) are among the psychoacoustic experiments that are verified with the software tool developed in this work. Various adaptive alternative forced choice algorithms with n-up, m-down adaptation strategies are implemented for psychoacoustic experiments to present the stimulus based on feedback from the subject. A set of sinusoidal pip tones at different frequencies (500, 1000, 2000Hz) are used to test the monaural parameters (SRT, MCL, JND in Frequency) and binaural cues (ILD, ITD). In each trial, 3 sets of stimulus intervals are presented, where one interval among the three consists of an adaptive parameter (e.g., interaural time for ILD tests) which is varied based on the user response to measure the users' perception with respect to that parameter, while the other intervals consists of the same stimulus. The interval with adaptive stimulus parameter is assigned randomly. The subjects are asked to identify which interval contains the stimulus, by clicking on the stimulus numbers buttons provided in the GUI window as shown in figure 6.5. Initially highly and easily discriminating stimulus set (one is discriminating from other two) is selected using the large adaptive parameters, and then the parameter is adjusted in every trial to present a more difficult set if the user responded correctly to the previous trial. The adaptive step size used to change the adaptive variable is adjusted after 2 consecutive reversals.

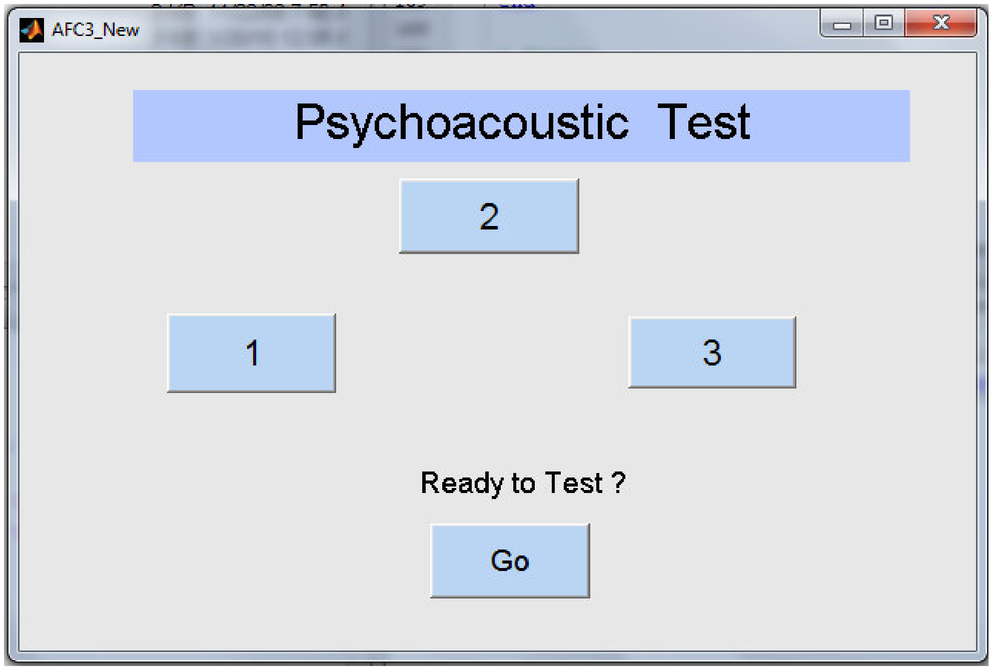

Figure 6.5 GUI Interface window for user to identify stimulus and give feedback

6.4 WINSOCK Interface between MATLAB and PDA

While the PDA research platform provides excellent ease of use for patients and flexible programmability for researchers to experiment with cochlear implants encoding strategies, use of MATLAB software for generation of variety of speech stimuli and design of adaptive methods for psychoacoustic experiments is further convenient for the researcher, given the rich mathematical function library support and simple programming style in MATLAB. Thus MATLAB software was chosen for speech signal generation, implementing the psychoacoustic experiments, and designing graphical user interface for conducting experiments. The PDA research platform is used to process the stimulus signal according to required encoding strategies and transmission to the cochlear implant using the RF communication link. The PDA research platform is interfaced using WINSOCK APIs (WINdows SOCKets API) to establish the communication path between MATLAB GUI and PDA platform [60].

WINSOCK is an acronym of Windows Sockets (Microsoft, 1997) [59], that is used as the interface between TCP/IP protocol stack and Windows. WINSOCK is essentially a technical specification that defines how Windows network software should access network services, especially TCP/IP. It interfaces a Windows TCP/IP client application the TCP/IP protocol stack. The WINSOCK interface APIs are popular with many Internet applications for Microsoft Windows and other popular OS. WINSOCK package is provided as a Dynamic Link Library (DLL), which consists of several APIs that are used to create sockets, establish connection and data transfer, and can be invoked in Windows applications.

In this work WINSOCK interface for MATLAB-to-PDA communication is developed to transfer stimulus parameters and amplitudes of speech stimuli from MATLAB running on PC to the PDA for presenting test stimuli in psychoacoustic experiments. To create a communication link between MATLAB and PDA using WINSOCK APIs, the following three important software components are used:

1) A WINSOCK server running on the PDA
2) A WINSOCK client .mexw32 (DLL) called from a MATLAB command script,
3) MATLAB command script running on PC.

Figure 6.6 shows the transfer of parameters and amplitudes from MATLAB to the PDA and status returned from the PDA to MATLAB.

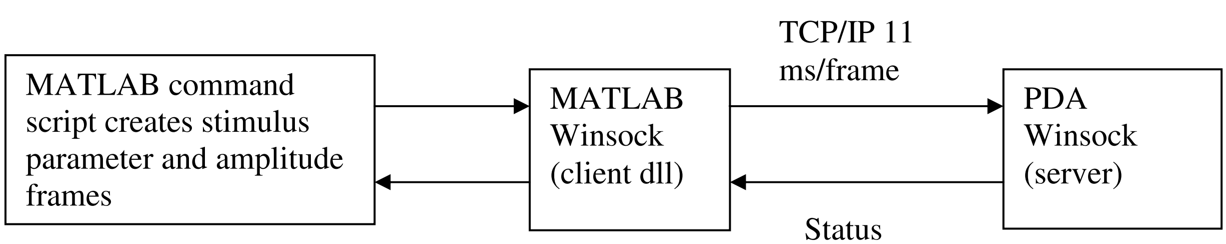

Figure 6.6 MATLAB PDA Interface using WINSOCK [60]

6.5 Establishing WINSOCK Communication and Transmitting Data

The PDA component initializes Winsock, creates a socket, binds the socket, "listens" to the socket, accepts incoming connections, and performs blocking receives to receive the parameter and stimulus amplitude data from the client. The "receive" is performed within a thread in two steps. In the first step, information about the total number of 11 ms frames, denoted as *nframes*, to be transmitted and the number of pulses per frame is received. The server then performs *nframes* receives, each time sending the data to the SDIO board. After *nframes* receives and sends, the server closes the socket and the connection. The PDA component is built as a Windows Mobile 5.0 executable using Visual Studio 2005 Professional. The executable is deployed on the PDA and run from the desktop remotely using the Windows Remote API (RAPI) application *prun*. The server needs to be started on the PDA before the client is run. Furthermore, the connection between the PDA and the desktop is switched to RNDIS Sync Mode from the default USB Serial Sync Mode to allow network connectivity. The MATLAB client dll initializes Winsock, creates a socket, connects to the server and transmits parameters and stimulus amplitude frames. It does this in two steps to match the receive function on the server: first the number of 11 ms frames, *nframes*, and the number of pulses per frame are transmitted. Second, *nframes* frames are transmitted continuously with the time interval between frames set to 11 ms. The dll is compiled from the C source using the MATLAB 7.7 MEX compiler. The third component is the MATLAB script.

Before calling the socclient(.) again with a new set of stimulus data, a minimum pause of 1.3 ms has to be inserted. The PDA server will automatically initialize a new connection for the next incoming stimulus and wait for the client to transfer the corresponding next set of parameter and amplitude frames. In this way the transfer of amplitudes takes place on demand i.e. the transfer is made under the complete control of the user. As illustrated in Figure 6.6 communication from the server back to the client and MATLAB has been implemented so as to confirm to the subject that a complete speech stimulus has been played. In summary, the WINSOCK interface algorithms to initiate and transmits data between MATLAB and PDA involves the following series of steps:

Step 1: Setup the parameters for the required stimulation

Step 2: Create initial parameter frames to generate patient's parameter files

Step 3: Generate stimulus signal and form frames of 11msec

Step 4: Start the Socket client from the MATLAB mex framework,

Step 5: Start the socket server on the PDA from MATLAB using WINSOCK APIs

Step 6: Start transmitting the parameter frames and stimulus amplitude frames

6.6 Flexible Parameter Interface using MATLAB GUI

The GUI interface also allows the user to enter or edit various parameters such as pulse width, stimulation rate, etc. If an erroneous value or an out-of-range value is entered, an error message is issued. Error checking is performed for all parameters entered to ensure that the parameters fall within the range supported by the commercially available Nucleus Freedom (or older generation) processor. Finally, the *activate* button is used for downloading the patient file to the PDA. A graphical user interface (GUI) was built to allow the user to import, save and modify patient parameter files (MAP) to be used in the PDA code.

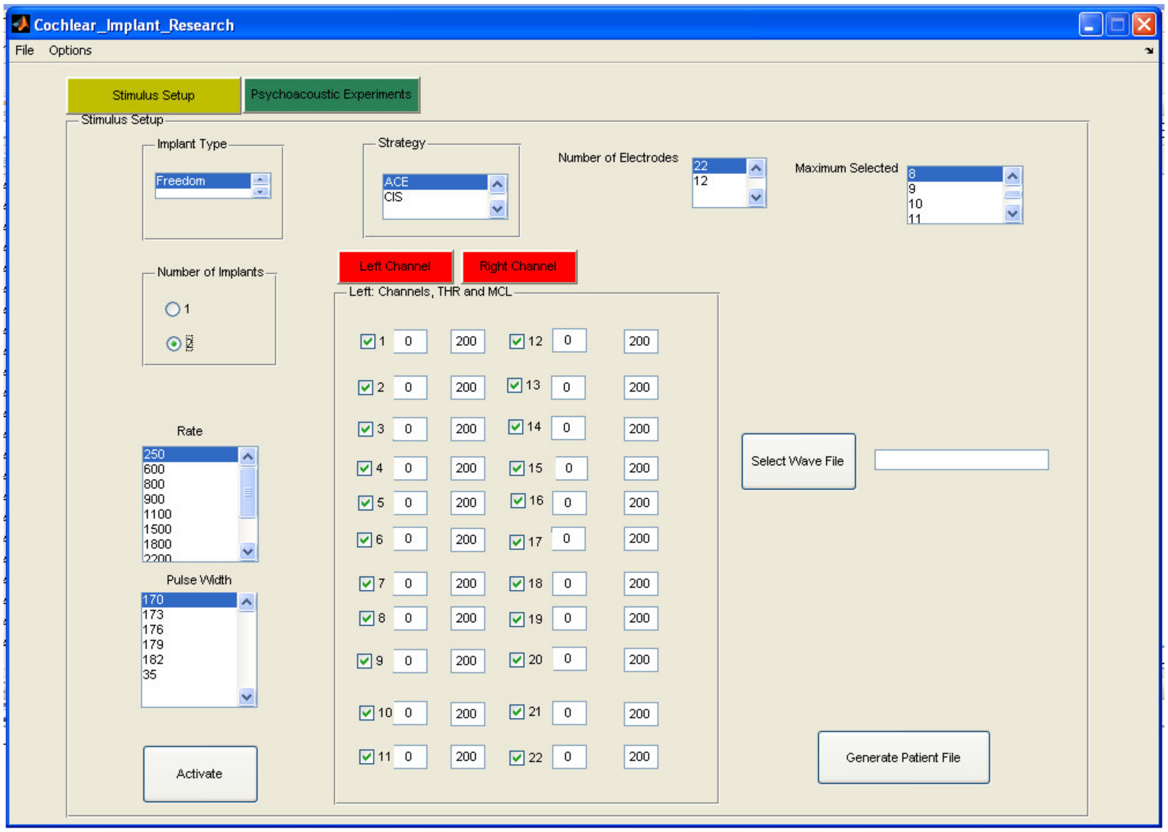

Figure 6.7 GUI Interface to setup parameters to PDA research platform

The GUI allows the user to create a new patient file, as well as load existing patient files. Figure 6.6 shows an example snapshot of the developed GUI. The parameters included in the patient file are listed in table 1 with their valid range of values. To accommodate bilateral users, the GUI has two panels, one for the left implant and one for the right implant. The user has the flexibility to turn on/off individual channels, by checking the appropriate box located left to each channel.

Table 6.1. The list of parameters used for stimulation in research platform

Parameter	Applicable Strategy	Value Range
Strategy, Left Implant:	[ACE, CIS]	0 or 1
Strategy, Right Implant	[ACE, CIS]	0 or 1
Number of maxima, Left	[ACE]	0 to 22
Number of maxima, Right	[ACE]	0 to 22
Electrode configuration, Left [MP1, MP2, MP1+2]	[ACE, CIS]	1,2,3
Electrode configuration, Right [MP1, MP2, MP1+2]	[ACE, CIS]	1,2,3
Stimulation rate per channel, Left	[ACE, CIS]	200-1200
Stimulation rate per channel, Right	[ACE, CIS]	200-1200
Pulse width- Left	[ACE, CIS]	12-150
Pulse width- Right	[ACE, CIS]	12-150
Threshold values, 22, Left	[ACE, CIS]	0-200
Most comfortable levels, 22, Left	[ACE, CIS]	50-255
Threshold values, 22, Right	[ACE, CIS]	0-200
Most comfortable levels, 22, Right	[ACE, CIS]	50-255

The next chapter (Chapter 7) presents general descriptions of psychoacoustic experiments and procedures to setup and conduct the experiments using the GUI software presented in this chapter.

CHAPTER 7

PSYCHOACOUSTIC EXPERIMENTS IMPLEMENTATION

To verify the usability of the software tool described in previous chapter in different areas of research in cochlear implants, various psychoacoustical experiments were implemented and conducted using the PDA research platform and GUI software. In this Chapter, sections 7.2 and 7.3 provide the general description of setting up a psychoacoustic experiment and using the software to conduct an experiment, and section 7.4 describes the psychoacoustic tests proposed to verify the software.

7.1 Psychoacoustic Experiment Definitions:

Before discussing the experiments' procedures, it is useful to introduce some common terms that are used in this document with respect to adaptive psychoacoustic experiments. This section introduces the definitions of the terms related to adaptive psychoacoustic experiments.

Stimulus: Stimulus is a speech signal either processed for use with cochlear implants to stimulate cochlear implants in CI subjects or directly used for normal hearing subjects.

Target Stimulus: A stimulus signal whose frequency, amplitude, phase parameters are adjusted adaptively in psychoacoustic experiments in comparison with a reference stimulus signal.

Reference Stimulus: A stimulus signal whose frequency parameters are fixed and allows the subjects to differentiate target stimulus (whose parameters are adjusted adaptively).

Trial: Trial is a series of steps in a test that involves of displaying a GUI window, presenting series of stimulus to the subject under test through a device (cochlear implant or speaker) and obtaining the response from the subject.

Procedure: Procedure is an algorithm that controls the flow of an experiment, to determine the next stimulus for the next trial.

Experiment: Experiment consists of a combination of procedures and the definition of all modules that are necessary to conduct those procedures.

Response: Response is a feedback from the subject that is associated with a trial; usually it is an integer that identifies the index of the target stimulus recognized by the subject in the series of single target and multiple reference stimuli.

Result: Result is associated with an experiment that captures the responses and analysis in a predefined way to generate the outcome of an experiment.

Reversal: Reversal in adaptive variable is defined as the point at which the variable changes its direction. The reversal points are used to evaluate average of adaptive variable of the experiment.

Run: Run is a series of trials that are taken by subjects until a particular number of reversals occurred (or fixed number of trials) in estimating parameters adaptively. For each test, a few runs are conducted to estimate the effective result of the test as the average of all runs.

Adaptive Variable: A parameter of the target stimulus signal that is varied after the trial based on correctness of the response of the subject using an adaptive rule.

Step size: The step size is the amount of change permitted to the adaptive variable after every trial.

Direction: The direction of an adaptive variable at each next trial represents whether it increases (positive direction) or decreases (negative direction).

7.2 Psychoacoustic Experiments Setup

The software platform that enables the researchers to conduct psychoacoustic experiments is designed and implemented in MATLAB. The GUIDE toolbox in MATLAB was used to develop a simple and easy to use graphical interface. Before starting the experiments, the researcher is required to setup the experiments. With the flexible and easy interface provided in MATLAB GUI window, setting up an experiment is an easy job for researcher or doctor who is interested in studying the behavior of cochlear implants users. After the GUI program is started, the following steps are taken to setup the required experiment:

1. Click on *psychoacoustic experiments* tab on top of the GUI window,
2. Select the stimulus type from the options provided in the *stimulus setup* list box and based on the selected stimulus type key in frequency if the stimulus for a sinusoidal or pip tone, or browse for the waveform file to be used as stimulus signal.
3. From the *experiment setup* panel, select the experiment type as either monaural or binaural experiment and then select the intended experiment from the next list box. The experiments' list boxes shows Frequency Discrimination, Loudness Balance, Speech Reception Threshold, and etc for monaural experiments, and ILD, ITD for binaural experiments.
4. Select the adaptive method used for the experiment by selecting one of the options from the *procedure* list box. Adaptive 1 up 1 down, 1 up 2 down, and etc are among the procedures supported.

5. The stimulation type specifies acoustic stimulus for normal hearing subjects and electrical stimulus for cochlear implant subjects.
6. If the electrical stimulus is used with CI subjects, setup the PDA platform by turning on the PDA and insert the SDIO board. Select *Stimulus Setup* tab to setup encoding strategy parameters such as strategy type, pulse rate, width, thresholds (THRs) and most comfortable levels (MCLs) and etc. When the test is started, the WINSOCK interface continuously transfers encoded stimulus amplitudes to PDA platform.

Figure 7.1 gives an example experiment setup in the psychoacoustic experiments GUI window for *Loudness Balance* experiment for sinusoidal pip tones of 500Hz as test stimulus signals.

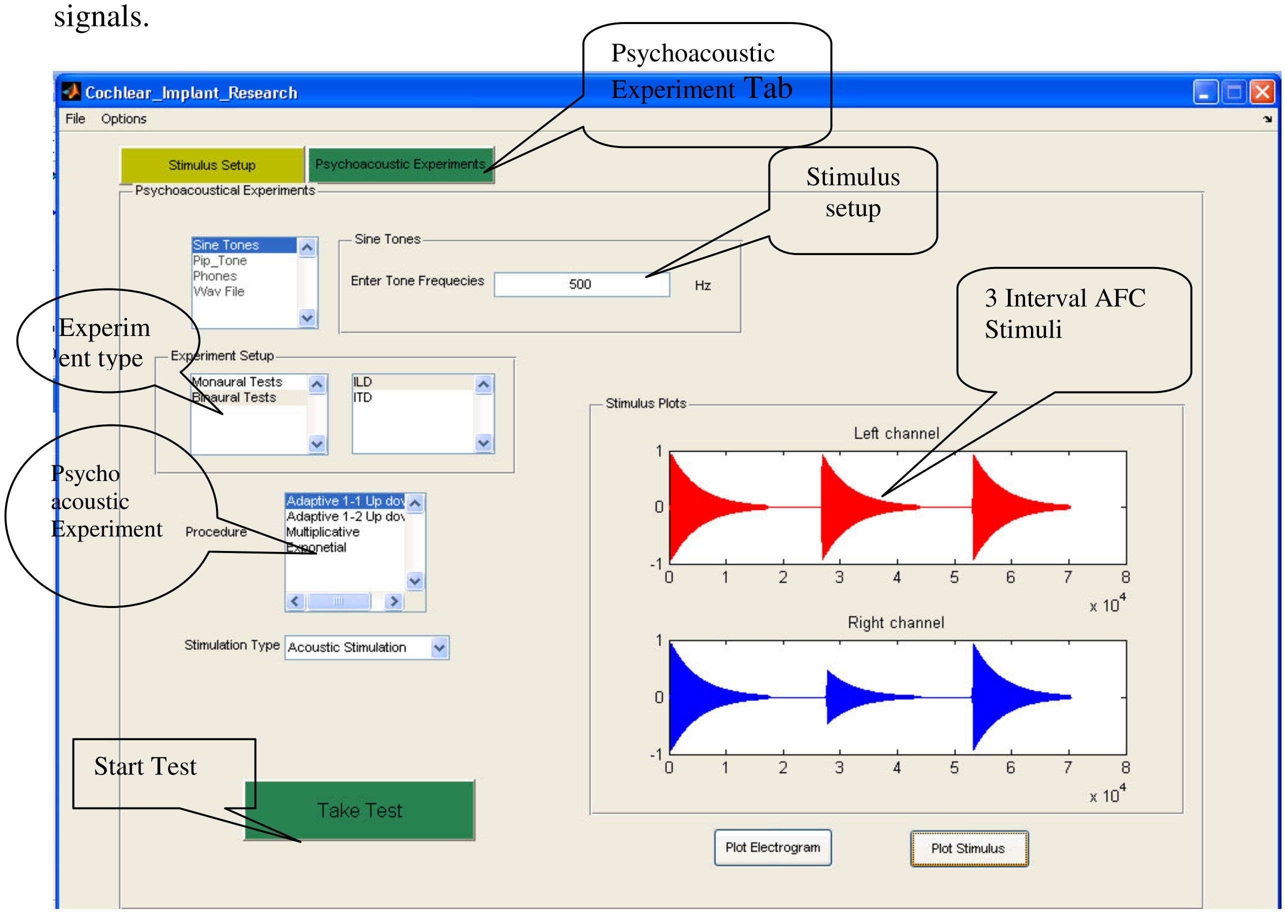

Figure 7.1 Graphical Unser Interface Window to setup and run a psychoacoustic experiment

7.3 Working Flow of Psychoacoustic Experiments

After setting up the GUI software as mentioned in the previous section, a psychoacoustic experiment can be conducted as shown in figure 7.2. During the experiment the subjects under test is guided by the GUI software as shown on the right part of the figure 7.2.

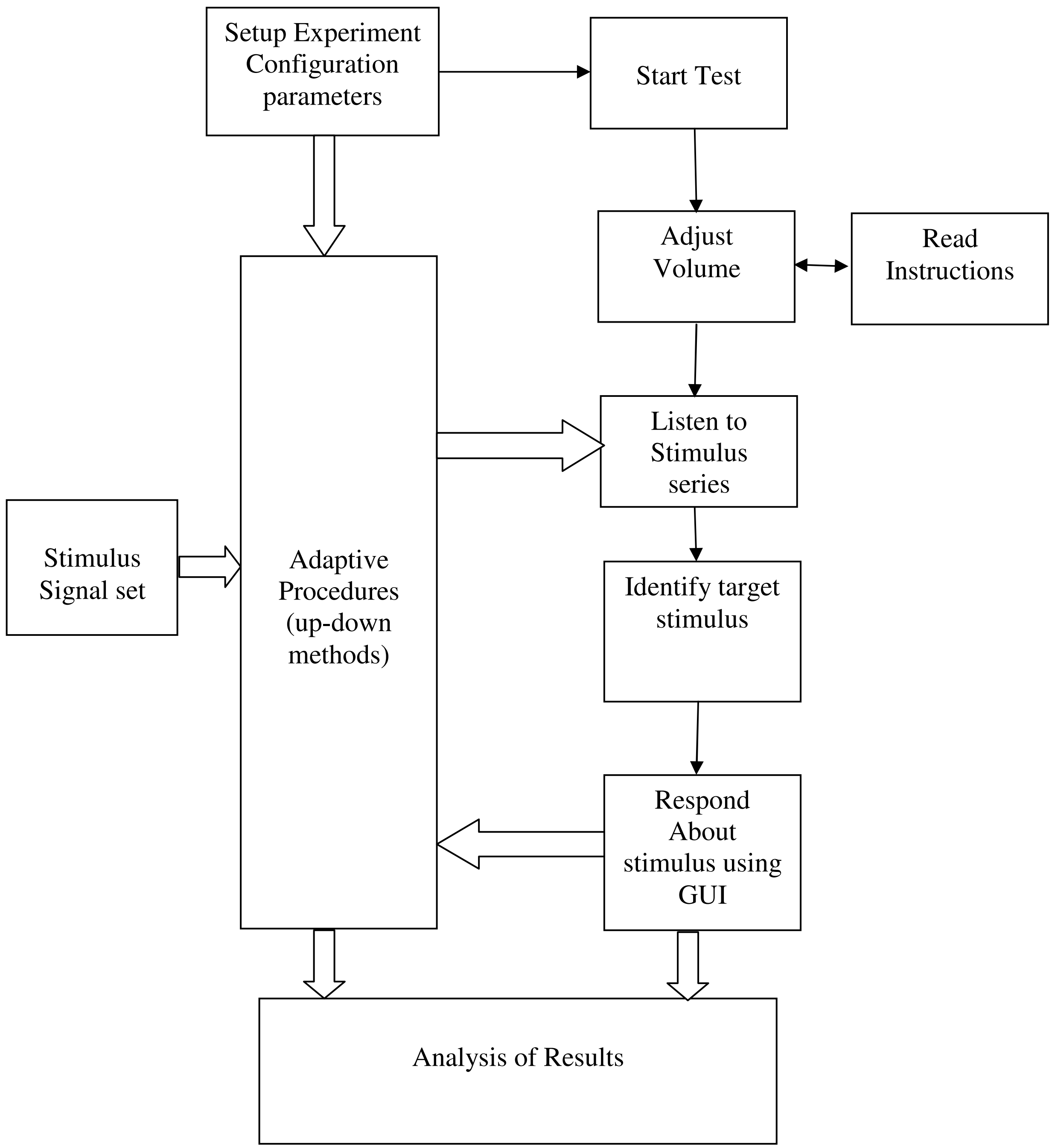

Figure 7.2 Psychoacoustic Experiments working flow

When the test is started, the subject is presented with test stimulus and asked to adjust volume to a comfortable level using the sliding controls provided, and then a series of instructions are (optionally) provided about the experiments. After reading the instructions and, adjusting the volume the adaptive procedure presents the stimuli reference and target stimuli using the forced alternative choice method. In practice, typically 3 Alternative Forced Choice (3AFC) methods are used in which 2 reference stimuli and 1 target stimulus are presented in random order in each trial. The stimuli are numbered in the order they are presented on interactive buttons as shown in figure 7.3. The subjects are required to identify the target stimulus that sounds different in a particular parameter among the 3 stimuli, and it is responded by clicking a button corresponding to the expected target stimulus.

The Adaptive procedure section prepares randomly a set of stimuli and presents them to the subject. It implements transformed up-down adaptive methods, which manipulates an adaptive parameter based on experiment type of the target stimulus. Also these procedures collect and track the response from the subjects in each trial and estimate reversals and evaluates the experiment termination criterion, and finally analyzes the results of the experiments.

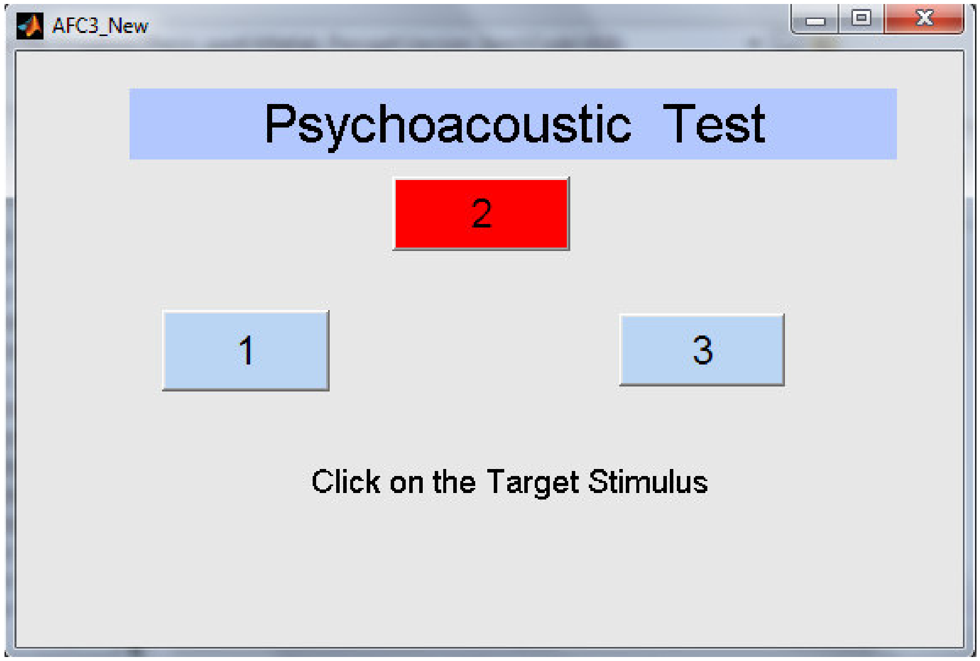

Figure 7.3 3AFC GUI Window: to identify stimulus and give feedback

7.4 Adaptation with Up-Down Method

With transformed up-down adaptive methods, when a user responds correctly as expected, the perceptual discriminating adaptive variable decreases and increases when the user responds incorrectly. The transformed up-down methods estimate the next adaptive variable additively or multiplicatively based on the up-down strategy. The step size and direction determines the adaptive variable. In every trial, the *direction*, *stepsize*, and adaptive variable *v* are updated using a predefined adaptation rule such as additive, multiplicative, and fixed adaptation technique. Following are the steps in updating the adaptive variable.

1. *Obtain the response from the user for the given trial*
2. *Using the up-down adaptation rule, identify if it is a reversal in adaptive variable*
3. *Change the direction (negative of previous direction) of the adaptive variable, if it is a reversal*
4. *If it is a reversal, estimate the stepsize for the next trial, as in Equation 1*

$$stepsize = stepsize * 2^{direction \cdot stepsize} \quad \rightarrow (1)$$

5. *Update the adaptive variable as in equation 2, to generate the stimulus set for the next trial as follows:*

$$\left.\begin{aligned} v_{adapt} &= v_{adapt} + stepsize * v_{adapt} \quad \textit{(Additive Adaptation)} \\ v_{adapt} &= v_{adapt} * v_{adapt}^{stepsize} \quad \textit{(Multiplicative Adaptation)} \end{aligned}\right\} \rightarrow (2)$$

where v_{adapt} is the adapted variable

This procedure is repeated for each trial until certain number of reversals or maximum number of trials is reached.

7.5 Psychoacoustic Experiments to Verify the Software

To verify the usability of the software, adaptive 3 alternative forced choice algorithms (which are estimated as 70.7% point on psychometric function, Levitt, 1971, [33]) were implemented. The experiments present the stimulus after stimulus and collect the feedback from the subject. A set of sinusoidal pip tones at different frequencies (500, 1000, 2000Hz) can be used to test the monaural parameters (SRT, MCL, JND in Frequency) and binaural cues (ILD, ITD). In each trial, 3 sets of stimulus intervals are presented, where one interval among the three consists of an adaptive parameter (e.g., interaural time for ILD tests) which is varied based on the user response to measure the users perception with respect to that parameter, while the other two intervals contains the same stimulus. The interval with adaptive stimulus parameter is assigned randomly. The subjects are asked to respond to indicate which interval contains the stimulus, by clicking on the stimulus numbers buttons provided in the GUI window as shown in figure 7.3. Initially a highly and easily discriminating stimulus set (one is discriminating from other two) is selected using large adaptive parameters, and then the parameter is adjusted in every trial to present more difficult set if the user responded correctly to the previous trial. The adaptive step size used to change the adaptive variable is adjusted after 2 consecutive reversals.

The following hearing experiments are designed and tested to verify the usability of the GUI software developed above.

7.6 Hearing Tests

The monaural tests present the stimulus signal on one ear of the subjects and collects response from the subject at each trial about the series of stimulus. These tests help researchers to understand perceptual cues such as frequency discrimination, audible smallest sounds, etc. Following is the list of monaural tests conducted to verify the software described in chapter 6.

7.6.1 Frequency Discrimination

The main goal of this study was to determine the subject's sensitivity to frequency difference in difference limens (DL). The minimum difference in the frequency that can be detected by the subjects is estimated using this test. Frequency discrimination was measured using a three interval, forced-choice, adaptive procedure. In each trial, a subject is presented with 3 sounds in random order, including two reference stimuli with a fixed frequency and a target stimulus with an adaptive variable frequency that is higher than reference stimulus frequency. The subject was asked to identify the interval with the highest frequency by pressing a button on the computer monitor. Graphical feedback was given after each trial.

7.6.2 Loudness Balancing

The minimum perceivable change in loudness of any speech signal for a subject is estimated using this test. A testing stimulus and reference stimulus signals are presented to the subjects by adaptively changing the loudness difference between two stimulus signals from a prominently perceivable difference (in dBs) to a minimum detectable loudness difference (few dBs).

7.6.3 Binaural Tests

The binaural tests present the stimulus signal to both the ears of the subjects, with one stimulus consisting of binaural cues (like ITD or ILD) and collect a response from the subject at each trial about the series of stimulus. These tests help researchers understand binaural cues that help patients in sound localization. The following provides the list of binaural tests conducted to verify the software described in chapter 6.

7.6.4 Interaural Level Diffence (ILD)

With this test, the minimum perceivable change in the loudness in left and right ears is estimated when speech signal is arriving at same time at both ears (ITD = 0 sec). A testing stimulus and reference stimulus signals are presented to the subjects by adaptively changing the loudness difference between the two ears in the testing stimulus signal while the reference signal is presented with no difference in loudness between the two ears.

7.6.5 Interaural Time Difference (ITD):

With this test, the minimum perceivable change in the arrival time at left and right ears is estimated when speech signal is arriving with same magnitude at both ears (ILD = 0 dB). A testing stimulus and reference stimulus signal are presented to the subjects by adaptively changing the interaural time difference between the two ears in the testing stimulus signal while the reference signal is presented with no difference in the phase between the two ears.

7.6.6 Interaural Frequency Discrimination:

With this test, the minimum perceivable change in the frequency between left and right ears is estimated when speech signal is arriving with the same magnitude and at the same time at both ears (ILD =0 dB, ITD = 0 sec). A testing stimulus and reference stimulus signals are presented to the subjects with adaptively changing frequency difference between the two ears in the testing stimulus signal while the reference signal is presented with no difference in the phase and magnitude between the two ears.

The next chapter (chapter 8) presents details about the subjects, testing environment and results. The detailed analysis of the results obtained and their comparison with corresponding existing experimental results is also provided.

CHAPTER 8

EXPERIMENTAL RESULTS AND ANALYSIS

The software platform and the adaptive procedures developed in this work are evaluated by conducting experiments with normal hearing subjects and analyzing the results for consistency with the other results reported in the literature. This chapter is dedicated to describe experimental procedures used to generate the results, analysis of the results obtained and comparison of those results with the corresponding results reported in the literature.

8.1 Subjects

Though the software tool is aimed to study the behavior of cochlear implant users under different stimulus conditions, this work is dedicate to validate the software with normal hearing subjects only. Six (4 male, 2 female; aged between 20–30 years) normal hearing listeners are involved to estimate the hearing cues such are ITD, ILD, and thereby validate the software and adaptive procedures developed. As all listeners involved for this study lacked previous experience in psychoacoustic experiments, they were trained to use the software. The experiments are conducted in a sound booth (Acoustic systems. Inc) using head phones to listen to the stimulus during the test. The subjects are labeled and referred in following section as S1, S2, S3, S4, S5 and S6.

8.2 Experiment 1: Interaural Level Difference (ILD)

The sensitivity for ILD is measured by an adaptive three-interval three-alternative forced-choice (3AFC) procedure with the 1up/2down adaptation rule that converges at the 70.1 % point on the psychometric function. Sinusoidal signals and piptones at three frequencies (500Hz, 1000Hz, 2000Hz) used as stimuli in this experiment. During the test, three stimuli (one target and two reference stimuli) of 400ms duration with 200ms of pause between the stimuli are played sequentially in each trial, and the order of the target and reference stimuli is chosen randomly. The subjects are asked to indicate stimulus which sounds differently in the left and right ears. Initially, the target stimulus has an ILD of 12dB (amplitude difference between left and right channels) while the reference stimuli consists ILD of 0dB, and after every correct response, the ILD in target stimulus is decreased by step size. Initially, the step size is chosen as 5dB, and after first 3 consecutive correct responses the step size is reduced by half. The stimuli levels are randomly roved in overall level (± 4 dB) to avoid monaural loudness comparison. Six reversals or at least 45 trials are required to terminate for each run in the test. Two runs of the experiment are carried for each condition in the test. At the end of the experiment, the JND-ILD is defined as the mean of the last 4 of 6 reversals in each run.

Results: The adaptive variable (ILD in the target stimulus) is plotted at each trial in the experiment for two subjects (S2 and S5) and is shown in figure 8.1 and 8.2. The plots show 6 runs, of trials 2 for each of the frequency condition (500Hz, 1000Hz, and 2000Hz). The plot shows responses of the subjects at each trial with Δ and ∇ symbols for negative and positive responses indicating the upward or downward direction of the adaptive variable respectively. The reversals are marked with "*" symbol in the plot. The end of the run is indicated as "O" in the figure. As expected the ILD adaptive variable quickly reduced to smaller values, where the subjects tend to respond incorrectly. As the perception is highly subjective, the minimum level of adaptive variable in two subjects is different and from the two figures, it can be inferred that subject S5 shows better sensitivity to the ILD compared to subject S2. The mean error bar in figure shows the average across all the subjects.

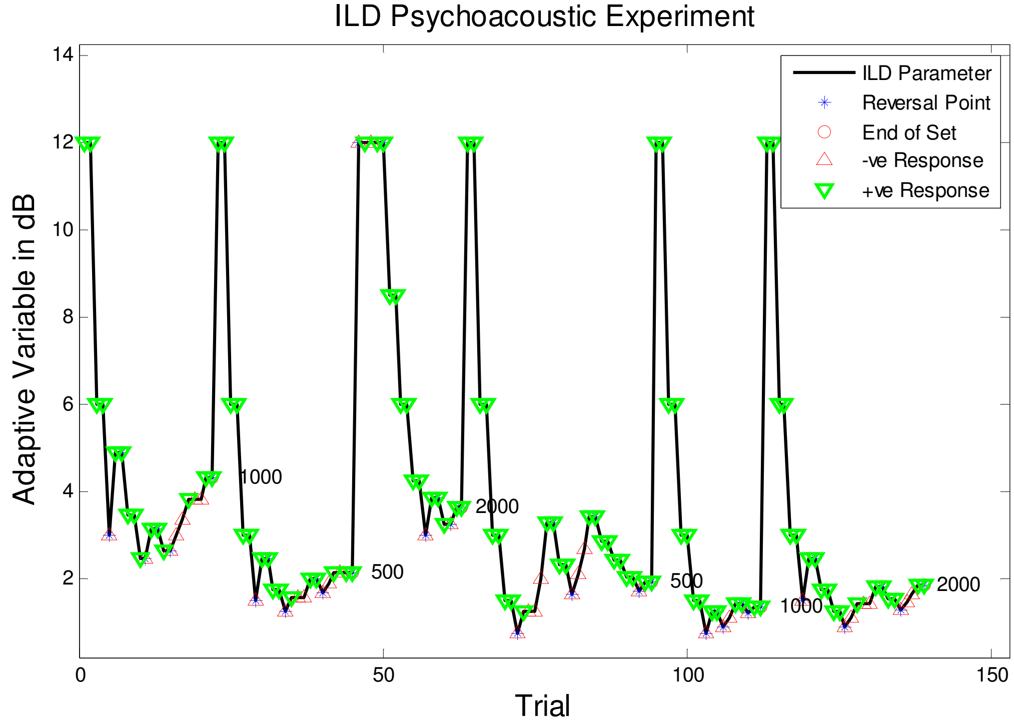

Figure 8.1 Adaptive Variable vs Trials for ILD Task for subject S2

The average JND-ILD and standard deviation of sensitivity to ILD observed in all six subjects are plotted in figure 8.3, 8.4 and 8.5 for condition frequency of 500Hz, 1000Hz and 2000Hz respectively. The subjects showed JND-ILDs ranging from 1.2 dB to 2.3 dB. These plots also include the mean and standard deviations of JND-ILDs across all the subjects as the mean point on it.

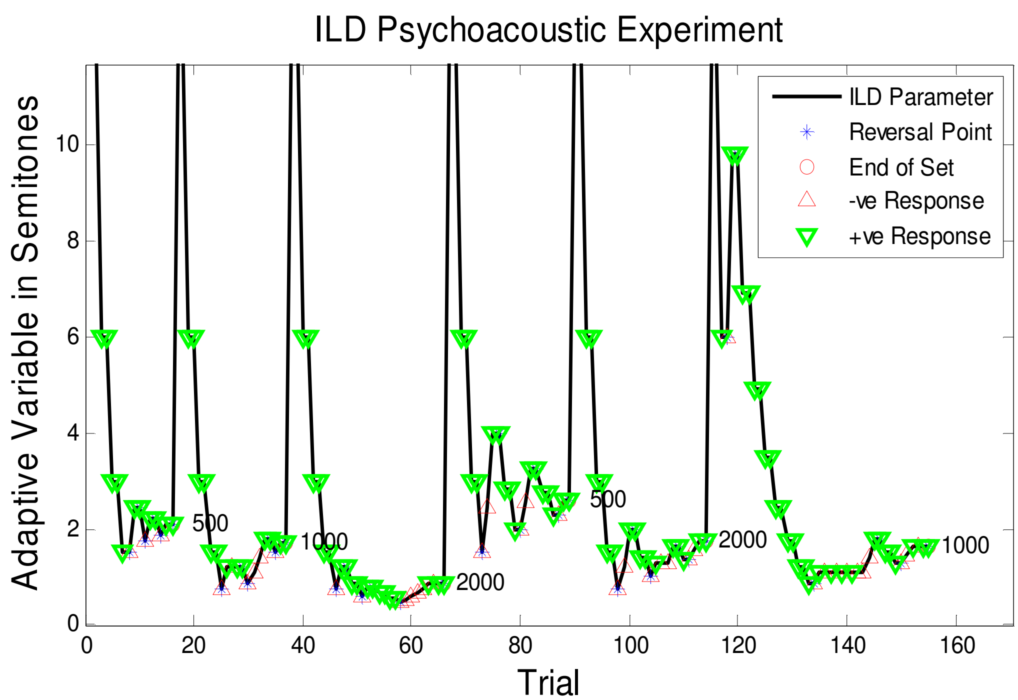

Figure 8.2 Adaptive Variable vs Trials for ILD Task for subject S5

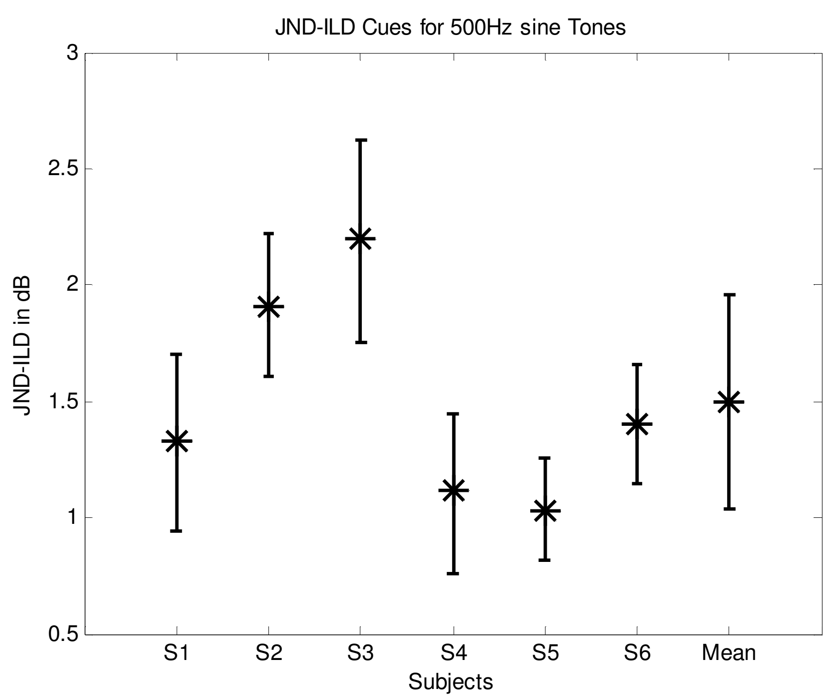

Figure 8.3 JND-ILD in dB for normal hearing subjects with 500Hz sine tones

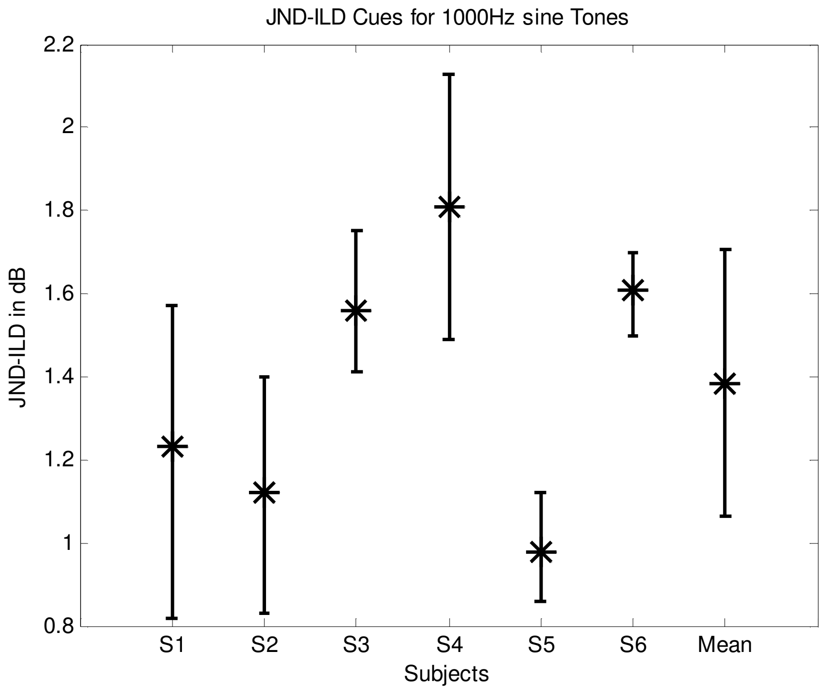

Figure 8.4 JND-ILD in dB for normal hearing subjects with 1000Hz sine tones

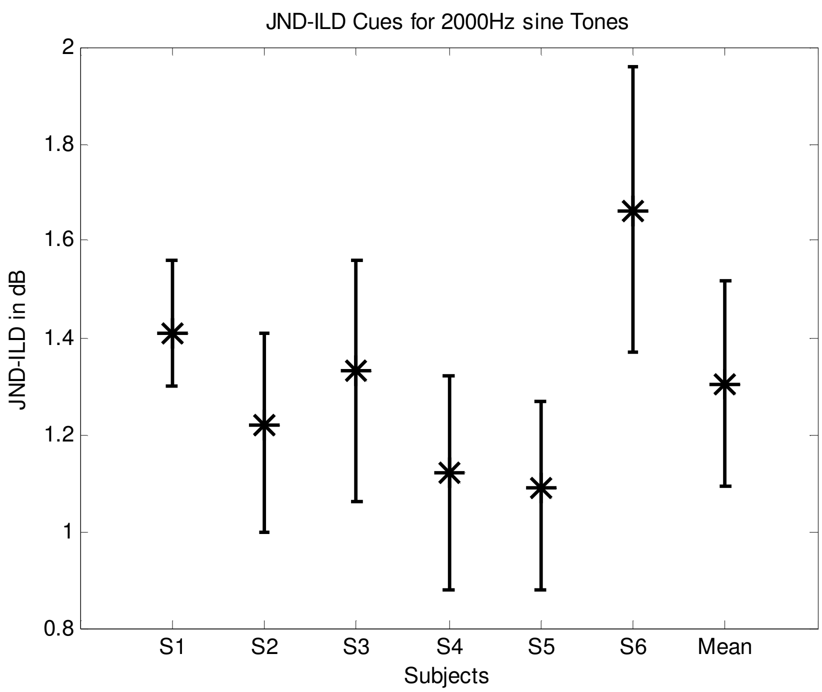

Figure 8.5 JND-ILD in dB for normal hearing subjects with 2000Hz sine tones

8.3 Experiment 2: Interaural Time Difference (ITD)

The JND-ITD was also measured using an adaptive three-interval three-alternative forced-choice (3AFC) procedure with the 1up/2down adaptation rule. Sinusoidal signals and piptones at three frequencies (500Hz, 1000Hz, 2000Hz) used as stimuli for this ITD experiment. During the test, three stimuli (one target and two reference stimuli) of 400ms duration with 200ms of pause between stimuli are played sequentially in each trial, and the presentation order of the target and reference stimuli is chosen randomly. The subjects are asked to indicate only target stimulus in the three stimuli which sounds differently (only target with non-zero ITD is different from the other two reference stimuli) in the left and right ears. Initially, the target stimuli have an ITD of 20ms (the time difference between left and right channels) while the reference stimuli have no ITD (ITD = 0 ms) and after every correct response, the ITD in target stimulus is decreased by the step size. Both target and reference stimuli are with zero ILD. Initially, the step size is chosen as 5 ms, and after first 3 consecutive correct responses the step size is reduced by half and step size is adapted based on the adaptive variable level. The stimuli levels are randomly roved in overall level (± 4 dB) to avoid monaural loudness comparison. Six reversals or at least 45 trials are required to terminate each run in the test. Two runs of the experiment are carried for each condition (frequency of the stimulus tone) in the test. At the end of the experiment, the JND-ITD is measured as the mean of the last 4 of 6 reversals in each run.

Results: Figure 8.6 and 8.7 shows the plot for adaptive variable (ITD in the target stimulus) at each trial in the experiment for two subjects (S1 and S5). The plots show 6 runs of trials, 2 for each of the frequency condition (500Hz, 1000Hz, and 2000Hz). The plot shows responses of the subjects at each trial with Δ and ▼ symbols for negative and positive responses indicating the upward or downward direction of the adaptive variable respectively. The reversals are marked with “*” symbol in the plot. The end of the run is indicated as “O” in the figure. As expected the ITD adaptive variable quickly reduced to smaller values, where the subjects tend to respond incorrectly. From the two figures, it can be inferred that subject S1 shows better sensitivity to the ITD compared to subject S5, and subject S1 exhibits best performance for ITD sensitivity among all the subjects involved in this experiment.

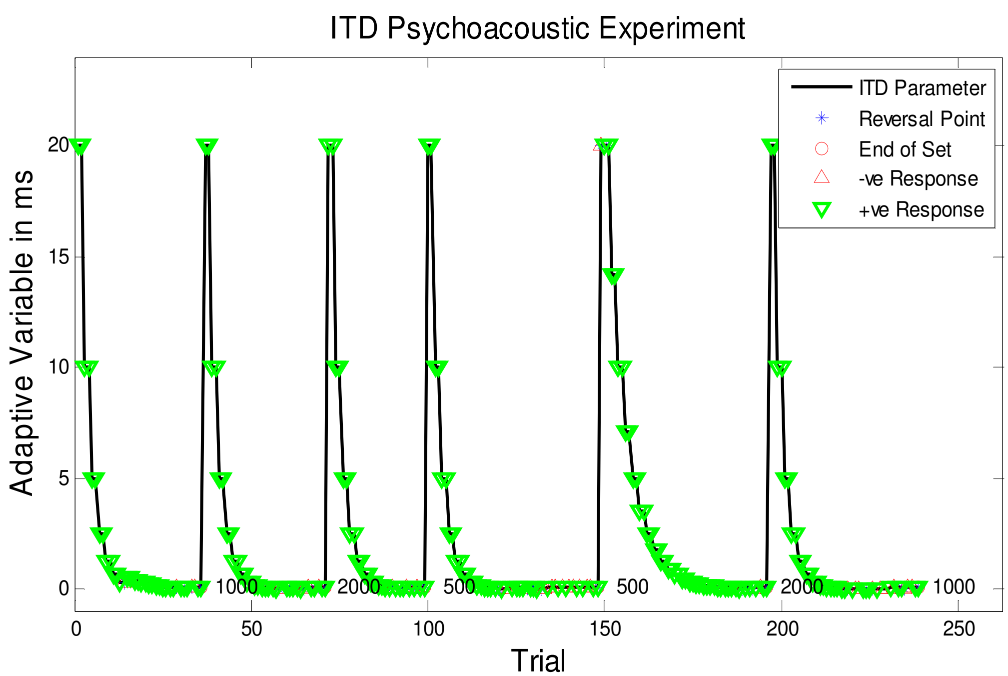

Figure 8.6 Adaptive Variable vs Trials for ITD Task for subject S1

The average JND-ITD and standard deviation of sensitivity to ITD observed in all subjects are plotted as along with error bars in figure 8.8, 8.9 and 8.10 for condition frequency of 500Hz, 1000Hz and 2000Hz respectively. The subjects showed JND-ITDs ranging from 36 µsec to 190 µsec. As seen several studies, the sensitivity for ITDs is highly subjective, and same observation can be made in this experiment from the results shown in figures 8.7, 8.8, and 8.9. These plots also include the mean and standard deviations of JND-ITDs across all the subjects as the mean point on it.

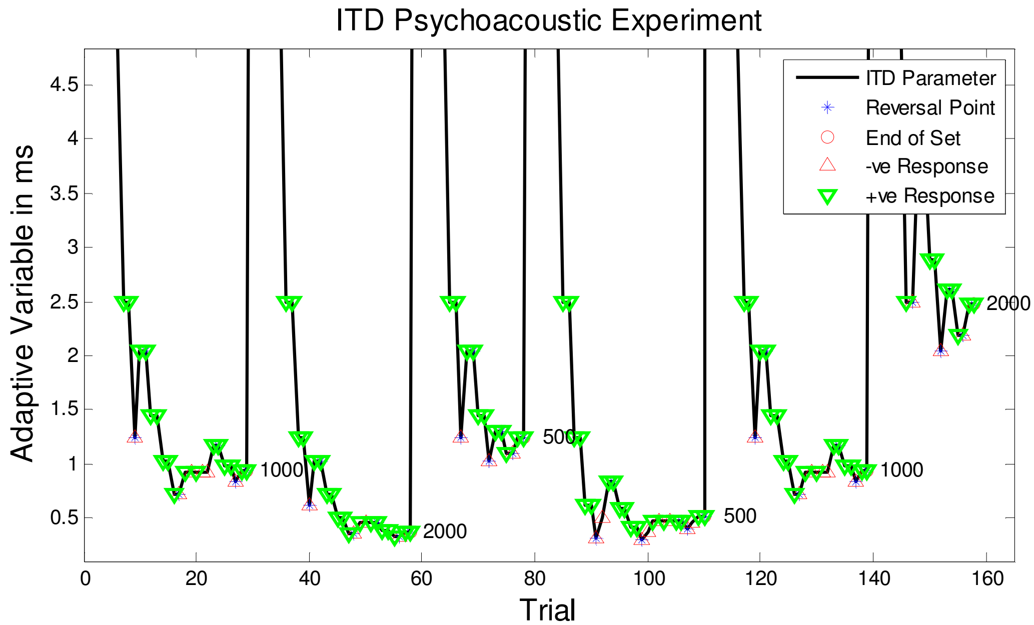

Figure 8.7 Adaptive Variable vs Trials for ITD Task for subject S5

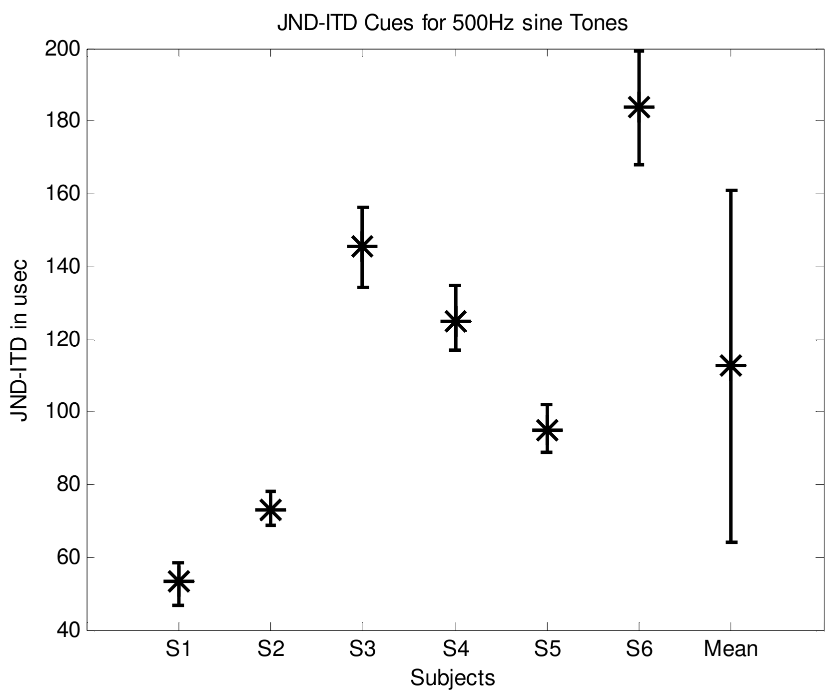

Figure 8.8 JND-ITD in μsec for normal hearing subjects with 500Hz sine tones stimulus

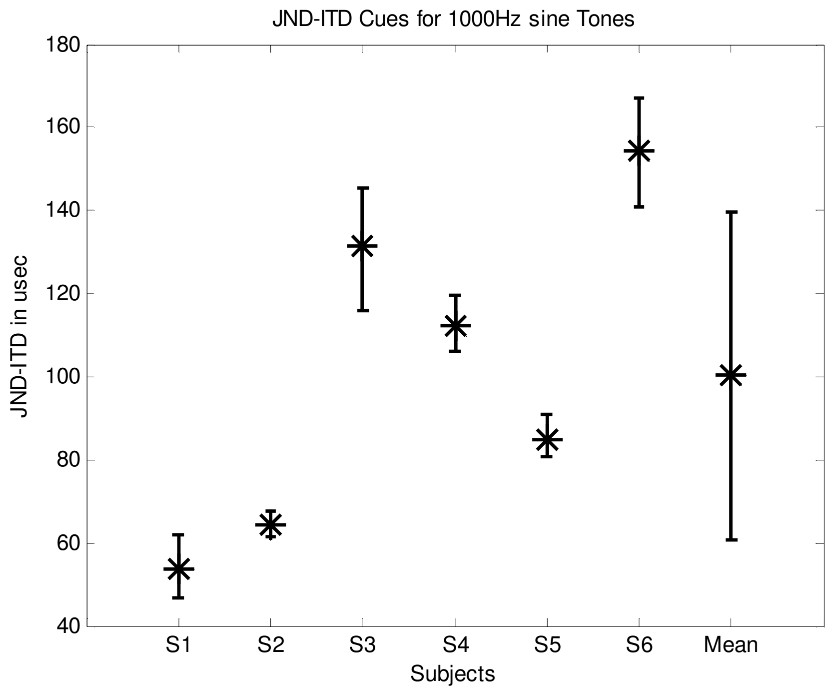

Figure 8.9 JND-ITD in µsec for normal hearing subjects with 1000Hz sine tones stimulus

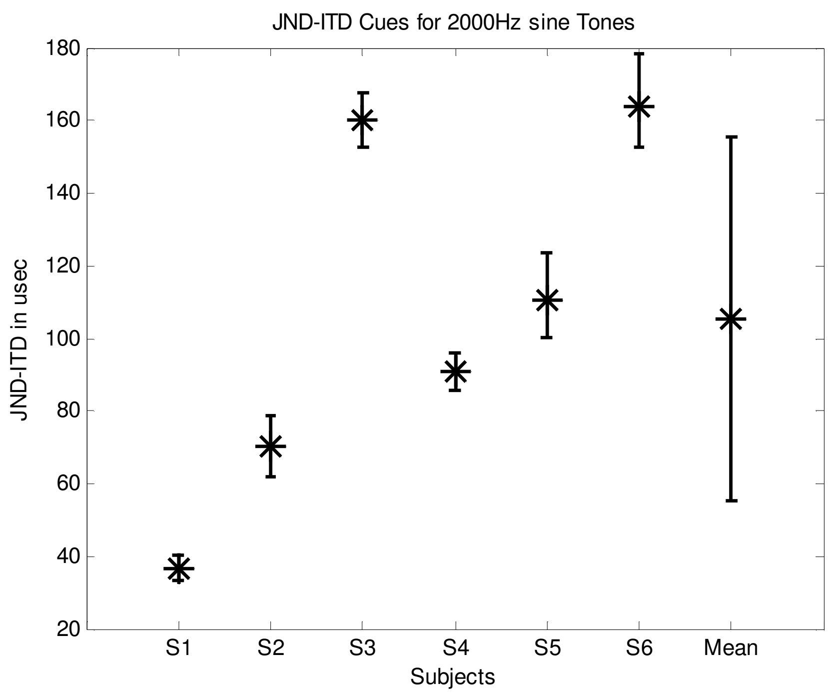

Figure 8.10 JND-ITD in µsec for normal hearing subjects with 2000Hz sine tones stimulus

8.4 Experiment 3: Frequency Discrimination

As described in the previous chapter, the frequency discrimination experimental task presents stimulus of different frequency and subject's feedback is evaluated to estimate frequency discrimination ability. Sinusoidal tones at 500Hz, 1000Hz, and 2000Hz frequencies are used as reference frequencies for the frequency discrimination task. The frequency discrimination cue is measured using an adaptive three-interval three-alternative forced-choice (3AFC) procedure with the 1up/2down adaptation rule. During the test, three stimuli (one target and two reference stimuli) of 400ms duration with 200ms of pause between stimuli are played sequentially in each trial, and the presentation order of the target and reference stimuli is chosen randomly. The subjects are asked to indicate only the target stimulus in the three stimuli which sounds differently (only target differs in frequency while the other two reference stimuli have the same frequency).

The target stimulus is higher in frequency (pitch) than the reference stimulus, with an amplitude determined based on the loudness balancing procedure result. All amplitudes were roved by one standard deviation of the loudness balance data. The adaptation of the target stimulation rate followed the rules of the 2-down, 1-up procedure. Initially the target stimulus frequency is chosen as 25% higher than reference stimulus frequency. And after every correct response, the frequency difference is decreased by a step size. The initial step size was chosen as 1 semitone. After every reversal, the step size is reduced by half. Six reversals or 45 trials are required to complete the frequency discrimination task. The subjects are asked to identify the target among the three stimuli which sounds higher in frequency. The last 4 reversals are averaged to determine the difference limens (DL) for each reference frequency. In the cases where the subject did not reach 6 reversals, all the reversals were averaged. Two runs of the experiment are carried for each condition (frequency of the reference stimulus tone) in the test. The same measurement is done with 2 runs and the frequency difference sensitivity is measured as average of DLs estimated in the two runs.

Results: Figure 8.11 and 8.12 shows the plot for adaptive variable (Frequency difference between the target stimulus and reference stimulus) at each trial in the experiment for two subjects (S2 and S3). The plots show 6 runs of trials, 2 for each of the frequency condition (500Hz, 1000Hz, and 2000Hz). The plot shows responses of the subjects at each trial with Δ and ∇ symbols for negative and positive responses indicating the upward or downward direction of the adaptive variable respectively. The reversals are marked with "*" symbol in the plot. The end of the run is indicated as "O" in the figure. From the two figures, it can be inferred that subject S3 shows better frequency discrimination ability compared to subject S2, and also subject S3 exhibits best performance for frequency discrimination sensitivity among all the subjects involved in this experiment. The just noticeable frequency difference in semitones for all the subjects are plotted in figure 8.13 for sine tone stimulus at 1000 Hz. 1.24 to 3.05 semitones of frequency difference observed across all the subjects involved in the test.

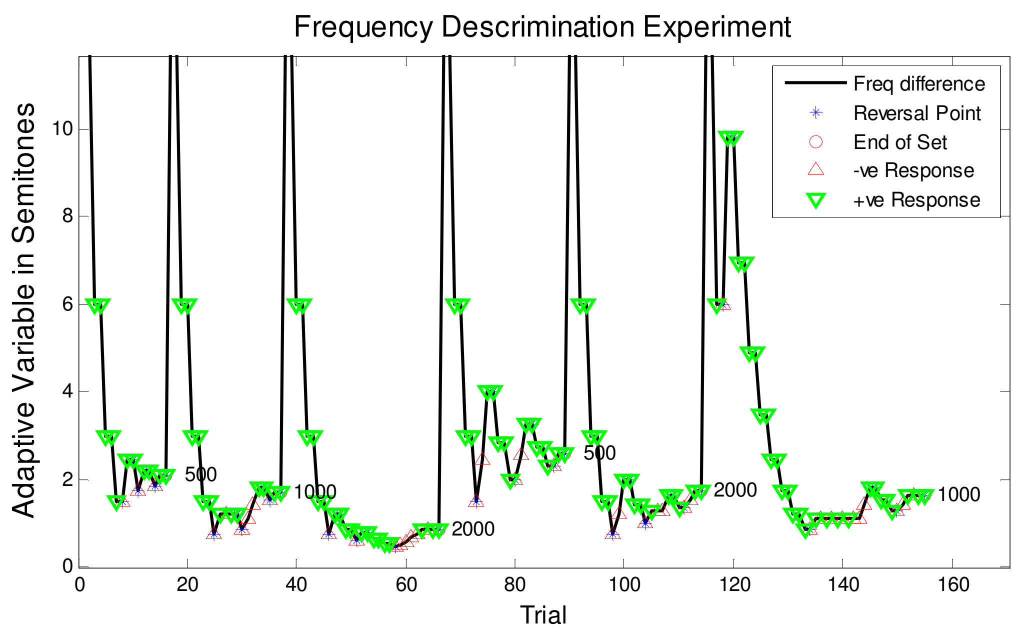

Figure 8.11 Adaptive Variable vs Trials for Frequency Discrimination Task for subject S3

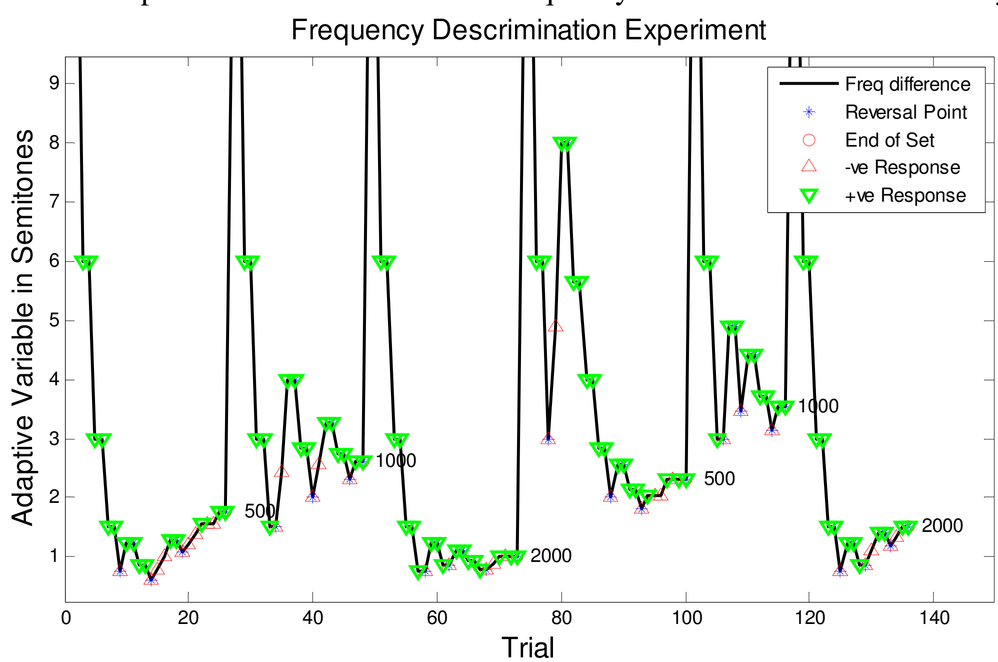

Figure 8.12 Adaptive Variable vs Trials for Frequency Discrimination Task for subject S2

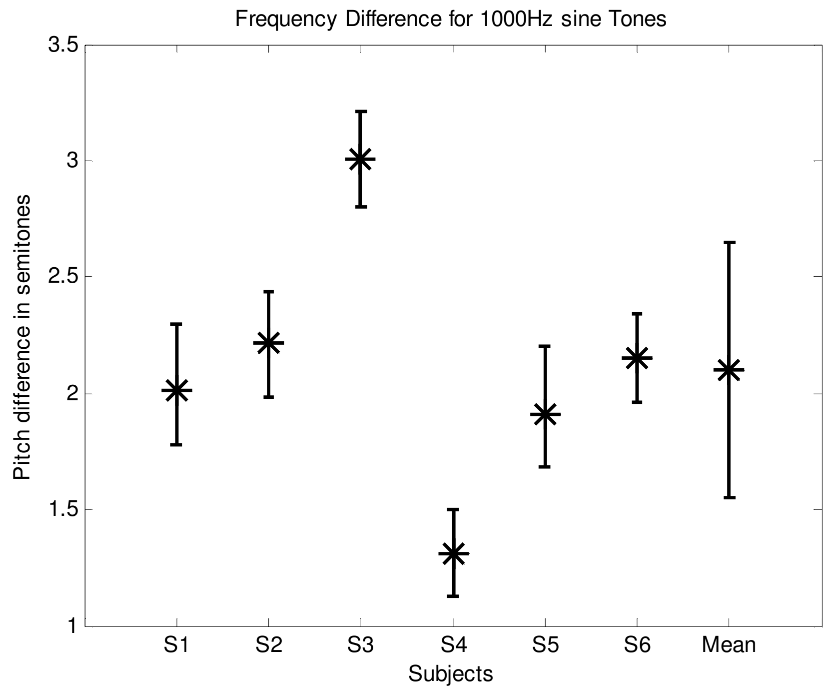

Figure 8.13 Frequency Difference measured in semitones for normal hearing subjects with 1000Hz sine tones stimulus

8.5 Experiment 4: Loudness Balance

The loudness balancing procedures are conducted to avoid possible loudness cues across different frequencies in the binaural experiments mentioned above. The loudness balancing is done for pair of sinusoidal tones at 200Hz, 500Hz, 1000Hz, 2000Hz frequencies. Three stimuli of 400ms duration with 200ms of pause between them are played sequentially in random order during each trial. The loudness balance procedures are implemented as an adaptive three-interval three-alternative forced-choice (3AFC) procedure with the 1-up, 2-down adaptation rule. The loudness of the target stimulus is kept higher than the reference stimulus, and subjects are asked to identify the stimulus that is louder, and the difference is adjusted until the listener could not perceive the difference in loudness. Initially the reference stimulus loudness is chosen as 20dB, and after every correct response, the loudness of the target stimulus is decreased by a step size. The initial step size is chosen as 5dB. After every reversal, the step size is reduced by half.

8.6 Results Reported in Previous Studies

Several studies and research has been carried to measure hearing cues and analyze them in speech perception, intelligibility and localizations ability in normal hearing and cochlear implant users. Several studies were done to measure ILD and ITD and their role in sound localization for cochlear implant users done. It is generally understood that the sensitivity of cochlear implants users to ILD are consistent across different studies, while ITD cues are largely subjective and varies from 100 μs to few milliseconds [43], [47], [48], [49], [50], [62], [63], [67], [70], [64]. Several other studies have shown that the sound localization and ITD cues sensitivity in cochlear implant users is far below that of normal hearing subjects, the ILD sensitivity in CI users is comparable to that of normal hearing subjects. These studies present results for ILD and ITD with different pulse rate, duration of CI usage and other parameters and found that the sensitivity of ITD is also influenced with these parameters. Pelizzone et al., (1990) reported that about 150 μs of JND-ITD in one subjects and about 1000 μs in for some subjects implanted with bilateral cochlear implants. Poon's (1998) [66] data indicated JND-ILDs for cochlear implanted subjects ranging from 320 μs -580 μs.

However, the ITD sensitivity in normal hearing subjects are known to be higher and average JND-ITDs of 10 μs to 70 μs were measured in normal hearing subject by Klumpp and Eady, (1956) [65]. The results presented in this thesis show the JND-ITD ranging from 50 μs to 220 μs (see Fig. 7.8, 7.9 and 7.10) in normal hearing subjects which are consistent with the results reported in literature for normal hearing subjects.

The studies of Durlanch and Colburn, (1978) [68] reported 1 dB to 2.2 dB of JND-ILD in normal hearing subjects, Poon, B., Litovsky, R. Y, et al., (2001) [66] reported 0.4 dB to 3.4 dB of JND-ILD in cochlear implant users. The results presented in this chapter (see Fig. 7.3, 7.4 and 7.5) for JND-ILD are between 0.9 dB to 2.4 dB in the normal hearing subjects involved in these experiments, and consistent with other studies. The experiments and the results obtained using the psychoacoustic software platform developed demonstrated that it can produce results consistent with the literature.

CHAPTER 9

CONCLUSIONS

This thesis presented a software tool designed for adaptive psychoacoustic procedures for cochlear implants. Psychoacoustic experiment procedures with MATLAB GUI are integrated with a PDA research platform using a WINSOCK interface. The GUI allowed to configure the cochlear implants speech processing strategies and design experiments to study particular speech processing parameters in cochlear implants. A software controlled environment is provided to design and conduct psychoacoustic experiments using a PDA research platform for cochlear implants. The psychoacoustic experiments facilitate the evaluation of speech perception, intelligibility, and source localization performance of cochlear implants patients. Transformed adaptive up-down methods are implemented as part of the software to use in psychoacoustic experiments. Using the software research platform, binaural cues such as interaural level difference and interaural time difference are evaluated for normal-hearing listeners. Results are presented and compared with results reported in the literature for validation of the experimental software. A few psychoacoustic experiments are conducted with normal hearing listeners to evaluate the validity of the implemented software. ITD and ILD are measured for 6 subjects and results shows that the JND-ITD varies from 35 μsec to 220 μsec and the JND-ILD varies from 0.8 dB to 2.2 dB. The results reported are compared with previous studies and found to be consistent. A wide range of results show subjects sensitivity to ITD and ILD cues is highly subjective.

9.1 Future Work

The software platform was developed in MATLAB as extension to existing software for PDA research platform. Furthermore, the implementations can be ported into the PDA environment to provide more portability and easy access to conduct the experiments. The psychoacoustic software can be upgraded to with different kinds of cochlear implant devices with different cochlear implant encoding strategies. The psychoacoustic research platform is verified with various basic experiments conducted monaurally and binaurally. Further, the work can be extended to carry out experiments to study the influence of noise, and reverberation on speech recognition by cochlear implant listeners.

BIBLIOGRAPHY

[1] P. Loizou, (1999). “Introduction to Cochlear Implants,” *IEEE Engineering in Medicine and Biology*, 32-42

[2] P. Loizou, (1999). "Signal-processing techniques for cochlear implants," *IEEE Engineering in Medicine and Biology Magazine,* vol. 18, no. 3, 34-46

[3] P. Loizou and O. Poroy, (1999). "A parametric study of the CIS strategy," Conference on Implantable Auditory Prostheses, Asilomar, Monterey, CA.

[4] P. Loizou and O. Poroy, (1999). "Minimum spectral contrast for vowel identification in CIS-type processors," *Conference on Implantable Auditory Prostheses*, Asilomar, Monterey, CA.

[5] P. Loizou, (2006). “Speech Processing in Vocoder-Centric Cochlear Implants,” *Cochlear and Brainstem Implants, (Moller, Ed.), Adv Otorhinolaryngol*, Basel, Karger, vol. 64, 109-143

[6] M. Schroeder, (1966). “Vocoders: Analysis and synthesis of speech**,”** Proc *IEEE,* vol. 54, 720-734

[7] B. Gold and C. Rader, (1967). “The channel vocoder,” *IEEE Trans Audio Electroacoustics* AU, vol. 15, no. 4, 148-161.

[8] P. Loizou, G. Stickney, L. Mishra, and P. Assmann, (2003). “Comparison of speech processing strategies used in the Clarion implant processor,” *Ear and Hearing,* vol. 24, no. 1, 12-19

[9] B. Wilson, D. Lawson, and M. Zerbi, (1995). "Advances in coding strategies for cochlear implants," *Advances in Otolaryngology - Head and Neck Surgery*, vol. 9, 105-129

[10] Hochmair-Desoyer, E. Hochmair, and K. Burian, (1983). "Design and fabrication of multiwire scala tympani electrodes," *Annals of New York Academy of Sciences*, vol. 405, 173-182

[11] J. Patrick, P. Seligman, D. Money, and J. Kuzma, (1990). "Engineering in Cochlear prostheses,” Edinburgh: Churchill Livingstone, 99-124

[12] J. Patrick and G. Clark, (1991). "The Nucleus 22-channel cochlear implant system," *Ear and Hearing*, vol. 12, 3-9

[13] E. Wallenberger and R. Battmer, (1991). "Comparative speech recognition results in eight subjects using two different coding strategies with the Nucleus 22 channel cochlear implant," *British Journal of Audiology*, vol. 25, 371-380

[14] G. Stickney, P. Loizou, L. Mishra, P. Assmann, R. Shannon, and J. Opie, (2006). "Effects of electrode design and configuration on channel interactions." *Hearing Research*, vol. 211, 33-45

[15] Lobo, P. Loizou, N. Kehtarnavaz, M. Torlak, H. Lee, A. Sharma, P. Gilley, V. Peddigari, and L. Ramanna, (2007). "A PDA-based Research Platform for Cochlear Implants", *3rd International EMBS Conference on Neuroengineering,* Hawaii

[16] P. Seligman and H. McDermott, (1995). "Architecture of the Spectra 22 speech processor," *Annals of Otology, Rhinology and Laryngology*, 139-141

[17] A. Boothroyd, (1993). "Profound deafness," *Cochlear Implants: Audiological Foundations* (R. Tyler, ed.), Singular Publishing Group, Inc, 1-34

[18] A. Lobo, P. Loizou, N. Kehtarnavaz, M. Torlak, H. Lee, A. Sharma, P. Gilley, V. Peddigari, and L. Ramanna, (2007). "A PDA-based Research Platform for Cochlear Implants," *3rd International EMBS Conference on Neuroengineering*, Hawaii

[19] A. Lobo, H. Lee, S. Guo, N. Kehtarnavaz, V. Peddigari, M. Torlak, D. Kim, L. Ramanna, S. Cifti, P. Gilley, A. Sharma, and P. Loizou, (2007). "Open architecture research interface for cochlear implants," *5th Quarterly Progress Report, NIH-NO1-DC-6-0002*

[20] Intel Corporation, (2005). "Intel Integrated Performance Primitives on Intel Personal Internet Client Architecture processors Reference Manual," Version 5.0, Doc. No. 278378-025

[21] National Instruments, LabVIEW PDA 8.0 PDA Module

[22] V. Peddigari, N. Kehtarnavaz, and P. Loizou, (2007). "Real-time LabVIEW implementation of cochlear implant signal processing on PDA platforms," *Proc. IEEE Intern. Conf. Acoust. Speech Signal Processing*

[23] R. Ramachandran and P. Loizou, (2007). "Real-time Pitch Detection on the PDA for cochlear implant Applications," Engineering *in Medicine and Biology Workshop, 2007 IEEE Dallas*, 90-93

[24] Cochlear L, (1998). WinDPS Help Version R116.00 Build 445 Cochlear, Ltd

[25] Cochlear L, (2001). "ShaLo: A software loader for the Sharp and Spear sound processors," Cochlear, Ltd., 5.0 edition

[26] L. Geurts, J. Wouters, (2000). "A concept for a research tool for experiments with cochlear implant users," *Journal of the Acoustical Society of America,* vol. 108, 2949–2956

[27] M. Harriso, (2003). "SPEAR3: UserManual CRC for Cochlear Implant and Hearing Aid Innovation," HearWorks Pty. Ltd., 2.0 edition, 384-388

[28] HearWorks PL, (2003). "Seed-Speak Product Brief CRC for Cochlear Implant and Hearing Aid Innovation," HearWorks, Pty. Ltd., 384-388

[29] HearWorks PL, (2003b). "SPEAR3 3rd generation Speech Processor for Electrical and Acoustic Research Cochlear Implant and Hearing Aid Innovation," and HearWorks, Pty. Ltd

[30] HearWorks PL, (2003c). "SPEAR3 Software Development Library and Programmers Guide Cochlear Implant and Hearing Aid Innovation," 1.03 edition.

[31] C. Daly and H. McDermott, (1998). "Embedded data link and protocol," *US Patent 5,741,314*

[32] Arasan Chip Systems Inc, (2005). "Arasan SD Combo HDK", WinCE user's Manual

[33] H. Levitt, (1971). "Transformed Up-Down Methods in Psychoacoustics," *J. Acoust. Soc. Am.*, vol. 49, no. 2, 467-477

[34] G.A. Gescheider, (1997). "*Psychophysics: The fundamentals" (3rd Ed.)*. Lawrence Erlbaum Associates, New Jersey

[35] C. C. Wier, W. Jesteadt, and D. M. Green, (1977). "Frequency discrimination as a function of frequency and sensation level," *J. Acoust. Soc. Am.*, vol. 61, no. 1, 178-184.

[36] S. A. Gelfand, (2004). "Hearing an Introduction to Psychological and Physiological Acoustics," United States of America: Marcel Dekker, 4th edition

[37] D. M. Green, (1990). "Stimulus selection in adaptive psychophysical procedures," *Journal of the Acoustical Society of America*, vol. 87, 2662- 2674

[38] K. Saberi, (1995). "Some considerations on the use of adaptive methods for measuring interaural-delay thresholds," *Journal of the Acoustical Society of America*, vol. 98, 1803-1806

[39] G. B. Wetherill, (1963). "Sequential Estimation of Quantal Response Curves," J. Roy. Statist. Soc. B25, 1-48

[40] CM. McKay, HJ. McDermott, (1998). "Loudness perception with pulsatile electrical stimulation: The effect of interpulse intervals," *Journal of the Acoustical Society of America,* vol. 104, 1061–1074

[41] BE. Pfingst, LA Holloway, SA. Razzaque, (1996). "Effects of pulse separation on detection thresholds for electrical stimulation of the human cochlea," *Hearing Research,* vol. 98, 77–92

[42] RY. Litovsky, PM. Johnstone, SP. Godar, (2006). "Benefits of bilateral cochlear implants and/or hearing aids in children," *International Journal of Audiology,* 78-91

[43] RJ. Van Hoesel, RS. Tyler, (2003). "Speech perception, localization, and lateralization with bilateral cochlear implants," *Journal of Acoustic Society of America,* vol. 113, no. 3, 1617-30

[44] KD. Brown, TJ. Balkany, (2007). "Benefits of bilateral cochlear implantation: a review," *Current opinion in otolaryngology & head and neck surgery,* vol. 15, no. 5**,** 315-318

[45] S. Das, CA. Buchman, (2005). "Bilateral cochlear implantation: Current concepts," *Current Opinion in Otolaryngology & Head & Neck Surgery*, vol. 13, 290-293

[46] J. Murphy, G. O'Donoghue, (2007). "Bilateral cochlear implantation: An evidence-based evaluation," *Laryngoscope,* vol. 117, 1412-1418

[47] W. Noble, R. Tyler, C. Dunn, S. Witt, (2005). "Binaural hearing has advantages for cochlear implant users also," *Hearing Journal,* vol. 58, 56-64

[48] BC. Papsin, KA. Gordon, (2008). "Bilateral cochlear implants should be the standard for children with bilateral sensorineural deafness," *Current Opinion in Otolaryngology & Head & Neck Surgery*, vol. 16, 69-74

[49] RJM. Van Hoesel, (2004). "Exploring the benefits of bilateral cochlear implants," *Audiology & Neuro-Otology*, vol. 9, 234-246

[50] J. C. Middlebroks and D. M. Green, (1991). "Sound localization by human listeners," *Annual Review of Psychology*, vol. 42, 135-159

[51] S.V Robert, A. D. Doug, F. L. Roger, P. L. Robert, and G. Michael, (1990). "A computer interface for psychophysical and speech research with the nucleus cochlear implant," *Journal of the Acoustical Society of America*, vol. 87, no. 2, 905–907

[52] G. Michael, I. Colin, and S. A. Brett, (2002). "Cochlear implant communicator," *International Patent No. WO 02/045991 A1*

[53] G. Luc and W. Jan, (2000). "A concept for a research tool for experiments with cochlear implant users," *Journal of the Acoustical Society of America*, vol. 108, no. 6, 2949–2956

[54] L. Johan, B. Bart, M. Marc, W. Astrid van, and W. Jan, (2005). "A flexible auditory research platform using acoustic or electric stimuli for adults and young children," *Journal of Neuroscience Methods*, vol. 142, 131–136

[55] Mark Harrison, (2003). "*SPEAR3 Hardware Technical Manual*, CRC for Cochlear Implant and Hearing Aid Innovation," HearWorks Pty. Ltd., 384-388

[56] CRC & HearWorks, (2003). "Spear3 speech processing system," Online, Product Brief. Pty Ltd. HearWorks, *SPEAR3 Software Development Library and Programmers Guide*, Cochlear Implant and Hearing Aid Innovation, and HearWorks Pty. Ltd., Albert Street East Melbourne Vic, 3002 Australia, 1.03 edition, 384-388

[57] G. Luc and W. Jan, (2000). "A concept for a research tool for experiments with cochlear implant users," *Journal of the Acoustical Society of America,* vol. 108, no. 6, 2949–2956

[58] S. S. Joshua, T. S. Chandra, C. M. Leslie, (2008). "Developing a Flexible SPEAR3-Based Psychophysical Research Platform for Testing Cochlear Implant Users," *Technical Report, Duke University*

[59] Microsoft Corp, (1997). "Windows Sockets 2Application Programming Interface", Revision 2.2.2

[60] N. R. Gunupudi, A. Lobo, D. Kim, R. Ramachandran, V. Gopalakrishna, N. Kehtarnavaz and P. C. Loizou**,** (2009). "Open Architecture Research Interface for Cochlear Implants," *12th Quarterly Progress Report, NIH-NO1-DC-6-0002*

[61] Craig C. Wier, Walt Jesteadt, and David M. Green, (1976). "A comparison of method-of-adjustment and forced-choice procedures in frequency discrimination," *Perception and Psychophysics,* vol. 19, no. 1, 75–79

[62] KD. Brown, TJ. Balkany, (2007). "Benefits of bilateral cochlear implantation: a review," *Curr Opin Otolaryngol Head Neck Surg,* vol. 15, 315-318

[63] J. Murphy, O'Donoghue G, (2007). "Bilateral cochlear implantation: An evidence-based evaluation," *Laryngoscope,* vol. 117, 1412-1418

[64] D. K. Eddington, B. B. Poon, H. S. Colburn, V. Noel, B. Herrmann, J. Tierney, M. Whearty, and C. Finley, (2003). "Speech Processors for Auditory Prostheses," *5th Quarterly Progress, NIH Contract N01- DC-2-1001. Cambridge.*

[65] R. G. Klumpp and H. R. Eady, (1956). "Some measurements of interaural time difference thresholds," *J Acoust Soc Am,* vol. 28, no. 5, 859-860

[66] B. B. Poon, R. Y. Litovsky, J. R. Madell, D. K. Eddington, and H. S. Colbumrn, (2001). "Sound localization and speech intelligibility with bilateral cochlear implants," *Conference of Implantable Auditory Prostheses, Pacific Grove, CA.*

[67] R. Van Hoesel, R. Ramsden, and M. Odriscoll, (2002). "Sound-direction identification, interaural time delay discrimination, and speech intelligibility advantages in noise for a bilateral cochlear implant user," *Ear Hear,* vol. 23, no. 2, 137-49

[68] N. I. Durlach and H. S. Colburn, (1978). "Binaural phenomena. Handbook of Perception," New York, Academic Press, vol. 4, 365-466

[69] M. F. Dorman, P. C. Loizou, and D. Rainey, (1997) "Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs," *J Acoust Soc Am,*, vol. 102, 2403–2411

[70] C.M. Rankovic, R. L. Freyman, and P.M. Zurek, (1992). "Potential benefits of adaptive frequency-gain characteristics for speech reception in noise," Journal of Acoustical. Society of America, vol. 91, 354-362

[71] B. J. Gantz, R. S. Tyler, J. T. Rubinstein, A.Wolaver, M. Lowder, P. Abbas, Brown, C., M. Hughes, and J.P. Preece, (2002). "Binaural cochlear implants placed during the same operation." *Otol Neurotol* vol. 23 no. 2. 169-180

VITA

Nageswara Rao Gunupudi was born in Mogallu, Andhra Pradesh, India. He obtained his diploma in Electronics and Communication Engineering with honors from GMR Polytechnic, Rajahmundry, India in 1999. Later he obtained his Bachelors in Technology (B.Tech) degree from National Institute of Technology, Warangal (NITW), India, in April 2003. After graduation, he joined Emuzed Inc (now Emuzed is known as Aricent Inc), a successful start-up company in India, where he worked for 4.5 years on various video signal processing algorithms and video codecs for mobile platforms. From January 2008, he has been a graduate student and the research assistant at the speech processing and cochlear implants lab in the University of Texas at Dallas (UTD) to earn his Master of Sciences degree in Electrical Engineering. His research interests include signal processing algorithms for video, image, speech, audio applications, and efficient implementation of complex multimedia and image processing algorithms on embedded platforms.