KTH Speech, Music and Hearing

Subjectively preferred sound representation for different listening situations

Samuel Munkstedt

[samuelm@speech.kth.se]

Supervisor: Peter Nordqvist

Approved: 06-03-07

Examiner: Björn Granström

....

(signature)

Stockholm
8 mars 2006

Master Thesis in Hearing Technology
(Examenarbete i Hörselteknik)

Department of Speech, Music and Hearing
Royal Institute of Technology
S-100 44 Stockholm

KTH Tal, musik och hörsel

Examensarbete i Hörselteknik

Subjektivt vald ljudåtergivning för olika lyssningssituationer

Samuel Munkstedt
[samuelm@speech.kth.se]

Godkänt: 06-03-07	Examinator: Björn Granström	Handledare: Peter Nordqvist

Subjectively preferred sound representation for different listening situations

During the fitting process of a hearing aid the pure tone thresholds are used as input to a prescription formula. The output from the prescription formula is the gain-frequency response to be used in the hearing instrument. The goal with the prescription is to maximise speech audibility and speech comfort. However, the majority of the prescription formulas are only optimised for speech in quiet and other listening situations are not considered. Due to problems with the traditional hearing aid fitting, research today look for other means for a more individual based fitting for different listening situations. One must consider that the hearing aid users are unique individuals with different sound quality preferences.

The aim with this study is to let the user select the preferred frequency response for three different listening environments. The task for the subject is to adjust the frequency response so that the listening *comfort* is optimised and at the same time the speech is understandable. The results from this study show that the test subjects want different frequency response for different listening situations. The hearing-impaired appreciated to set their own amplification and the subjects' frequency response differed from the prescribed frequency response. In conclusion the hearing aid should be individual fitted with an emphasis on user-interaction.

Subjektivt vald ljudåtergivning för olika lyssningssituationer

Under utprovningsfasen av en hörapparat används brukarens tonaudiogram som ingångsdata till en preskriptionsformel. Resultatet från preskriptionsformeln är hörapparatens förstärkningskurva. Målet med preskriptionen är att maximera talhörbarheten och lyssningskomforten. Majoriteten av preskriptionsformlerna är optimerade för tal i tyst miljö och tyvärr är inte andra ljudmiljöer inkluderade. På grund av problem med traditionell hörapparatutprovning försöker dagens forskning undersöka andra möjligheter för en mer individuellt anpassad utprovning. Det är viktigt att inse att hörapparatanvändare är unika individer med olika preferenser vad det gäller ljudkvalitet.

Målet med den här studien är att låta användaren ställa in den ljudåtergivning som föredras för tre olika ljudmiljöer. Testpersonens uppgift är att förändra ljudåtergivningen så att lyssningskomforten optimeras samtidigt som talet är förståeligt. Resultatet visar att testpersonerna vill ha olika ljudåtergivning för de olika ljudmiljöerna. De hörselskadade uppskattade att ställa in sin egen förstärkning och det visade sig att deras ljudåtergivning skiljde sig från den föreskrivna ljudåtergivningen. Slutsatsen blir att hörappartutprovning bör vara individuellt anpassad med fokus på stor delaktighet av brukaren.

Acknowledgments

First of all, I would like to thank my supervisor Peter Nordqvist for proposing this interesting research topic to me. He has given me valuable comments on my work and interesting facts about hearing technology. His positive mindset about me, my work and this topic has given me the energy to do my very best. I also would like to thank Daniel Höglind and Jonas Nahlin for lending me their voices in the recording of sound material. In addition would I like to thank all the people giving me of their valuable time to participate in my study. Furthermore, I would like to thank Fredrik Hedberg for reviewing the final paper and the presentation. Finally, I would like to thank and send my love to Johanna Westlin for lending me her voice, for doing the illustration on page 20 and for always supporting me.

List of Figures

List of Tables

List of Abbreviations

ADC	Analogue to Digital Converter
AGC	Automatic Gain Control
AVC	Automatic Volume Control, slow AGC
BILL	Base Increases at Low Levels
BTE	Behind The Ear
CB	Critical-Band
CIC	Completely In the Canal
CR	Compression Ratio
DAC	Digital to Analogue Converter
DI	Directivity Index
DSL[i/o]	Desired Sensation Level - Prescriptive procedure
DSP	Digital Signal Processor
FFT	Fast Fourier Transform
FIG6	Data from Figure 6 in Killion's [1993] article - Prescriptive procedure
FPGA	Field Programmable Gate Array
GUI	Graphical User Interface
HA	Hearing Aid
HL	Hearing Level
HTL	Pure Tone Threshold
IG	Insertion Gain
ITC	In The Canal
ITE	In The Ear
LDL	Loudness Discomfort Level
LTI	Linear Time Invariant
MCL	Most Comfortable Level
NAL	National Acoustic Laboratories of Australia
NAL-NL1	Non-Linear v.1 - Prescriptive procedure
NAL-RP	Revised, Profound correction - Prescriptive procedure
OSPL90	Prescriptive procedure based on maximum output
PC	Peak Clipping
POGO	Prescription of Gain & Output - Prescriptive procedure
PSD	Power Spectral Density
PWM	Pulse-Width-Modulated
REAR	Real Ear Aided Response
REIG	Real Ear Insertion Gain
REUR	Real Ear Unaided Response
RMS	Root Mean-Square
SII	Speech Intelligibility Index
SL	Sensation Level
SNHL	SensoriNeural Hearing Loss
SNR	Speech-to-Noise Ration
SPL	Sound Pressure Level
TILL	Treble Increases at Low Levels
UCL	UnComfortable Loudness level

Contents

1 Introduction

More and more people around the world suffer from hearing losses. The increasing average age and the growing population are the main reasons for this. In Sweden it is estimated that 830 000 people, 9% of the Swedish population, would benefit from using a hearing aid [SBU, 2003]. However, studies made by Kochkin [2002] have shown that only about 50-60% are satisfied with their hearing instruments. This indicates that, even though the first electrical hearing instrument[1] was introduced 100 years ago, hearing technology even today is a prominent research area. The introduction of the digital hearing aid[2] in the middle of the 90th opened the door for a new world of signal-processing possibilities. Advanced algorithms as feedback suppression, noise reduction, listening environment classification and directionality could now be implemented in the hearing instrument [Nordqvist, 2004]. The hardware components are getting smaller and at the same time more powerful. More memory and more efficient, yet less power consuming, Digital-Signal-Processors (DSP) will counteract the hardware limitations in the future. The new algorithms have however not resulted in big improvement when it comes to speech intelligibility [Bentler and Duve, 2000]. More research on software development will be very important for future hearing instruments. In the future hardware components will probably not set the limitations any more and focus will shift more and more to software development.

1.1 Problem formulation

During the fitting process of a hearing aid (HA) the pure tone thresholds are used as input to a prescription formula. The output from the prescription formula is the gain-frequency response to be used in the hearing instrument. The goal with the prescription is to maximise speech audibility and/or speech comfort. However, the majority of the prescription formulas are only optimised for speech in quiet. Other listening situations such as noisy traffic environments or a noisy party with many speakers are not considered. Studies made by Keidser [1995] and Keidser [1996] indicate that the user would benefit from using different amplification schemes for different listening situations. Prescription formulas require an accurate loudness model of the hearing impairment. Due to problems with the traditional hearing aid fitting, research today look for other means for a more individual based fitting for different listening situations. One must consider that the hearing aid users are unique individuals with different sound quality preferences. It will therefore remain important with individual fitting and after adjustments of the hearing aid.

The aim with this work was to let the user select the preferred frequency response for three different listening environments. The task for the subject was to adjust the frequency response so that the listening *comfort* was optimised and at the same time the speech was understandable. The result was then compared against a number of prescription formulas.

[1] The first carbon hearing aid, consisting of a *carbone microphone*, a *battery* and a *magnetic receiver*, appeared in 1899. It was called Akoulallion and was big as a table. The sound level produced was 20-30 dB larger than without a microphone and were satisfactory for people with mild to moderate hearing losses. The fist wearable model called the Akouphone or Acousticon appeared in 1902 [Dillon, 2001]

[2] In 1996, the first fully digital Behind The Ear (BTE), In The Ear (ITE) and In The Canal (ITC) hearing aids were introduced on the commercial market. Research in digital hearing aids started in the 1960s, [Dillon, 2001], and today most hearing aids are digital

1.2 Method

When making this study of subjectively preferred sound representation some sort of testing equipment had to be developed. A computer program on a standard PC was chosen to be the best solution. The test program[3] was made as a Graphical-User-Interface in MatLab. MatLab is user friendly and easy to work with once you learned the syntax. It has a lot of built in functions in the included toolboxes, such as the signal processing toolbox. The final program is run on a standard PC with a pair of headphones connected. The sound is presented monaurally and the test subjects are presented graphical feedback on the screen. For the normal hearing persons the testing could have been made in stereo but since the perceived loudness is increased for binaural stimulation, the results would not have been comparable with the impaired hearing. No limitation in frequency bandwidth was made, i.e. a sampling frequency of 44.1k Hz was used throughout the recording and testing. This gives a bandwidth of 22k Hz which covers the hearing range of the human ears. In contrast, traditional hearing instruments usually have a bandwidth of 6k-7k Hz. The test program is divided into two parts. In the first part the subjectively based settings are made and in the second part they are evaluated. The evaluation is made with a double elimination tournament between three prescription formulas and the subjectively made settings.

1.3 Outline of the report

Chapter 2 of the report will introduce the reader to the theory on which the study is based on. It will include sections about the theory of sound, components of speech, basics in signal-processing and hearing aids. Several of the prescription methods that are used to calculate the insertion gain for the hearing aid are described. In chapter 3 the developed test program will be in focus where the various parts of the program will be thoroughly described. The test can be divided into two main parts, a fitting part and an evaluative part. Furthermore, the recording process for the sound material used in this study is described. In chapter 4 the results from the study is reported and a discussion of these results follows in chapter 5. Finally the conclusions are made in chapter 6 and an outlook for further research within this area is given. To help the reader find information about certain topics an *Index* has been made and it can be found before the *References* on page 46.

[3]Throughout the report, when I'm using the term *test program* I mean the program developed by me.

2 Theory

The theory chapter strives to give the reader all the information needed to understand the implementation and results of this study. It also tries to make the reader understand the techniques chosen when developing and conducting this study. The chapter covers areas in the theory of sound, speech, hearing aids and signal processing. The reader is expected to have some knowledge about algebra to be able to fully understand the signal-processing section.

2.1 Theory of sound

A brief introduction to the theory of sound will be given in this section. Traditional acoustics will be left out because this study mainly involves sound that already been transformed to signals, either in the ear as nerve impulses or as electrical signals. An overview of psychoacoustics, i.e. the study of subjective human perception of sounds, will be made and important features such as: loudness, phon and masking will be introduced.

2.1.1 Sound Pressure Level

The physical definition of sound is; *time-varying sound pressure*[4]. The unit of the sound pressure $p(t)$ is Pascal (Pa). The sound pressure can vary from the absolute threshold (10^{-5} Pa) to the threshold of pain (10^{2} Pa) [Zwicker and Fastl, 1999]. To make it easier to describe these big variations the term Sound Pressure Level (SPL) or denoted L is used. A reference pressure $p_o = 20\mu\text{Pa}$ divides the sound pressure p and the resulting quotient is converted to a logarithmic scale, see equation 1:

$$L = 20\log(p/p_0)\ \text{dB} \tag{1}$$

Now the variations in sound pressure are much easier to handle, from the absolute threshold at 0 dB SPL to the threshold of pain at 140 dB SPL. For each increment of 6 dB the sound pressure is doubled, i.e. a 6 dB increase corresponds to the double sound pressure.

2.1.2 Electrical Sound Representation

When making a digital recording of a sound it is sampled with a certain frequency. Each sample stores information about the sound pressure level measured at that specific time. To be able to reconstruct the sound correctly, the number of samples per second, called the sampling frequency (f_s), must be at least twice the highest frequency of the sound, see equation 2.

$$f_s = 2B \tag{2}$$

The Bandwidth (B) is the range of frequencies in a signal, in this case the highest frequency of the sound. The relation in the equation above (equation 2) was discovered by Nyquist and it is hence called the Nyquist sampling frequency [Kamen and Heck, 2000]. To find a measure of the sound pressure at the recording site the Root-Mean-Square (RMS) of the digital sound file (wav) can be calculated (see equation 3 on page 4). N is the number of

[4] or acoustic pressure

elements in the sound vector x. A scalar of value '3' (the last term in equation 3) is added to get the momentary peak SPL instead of the maximum continuous SPL.

$$RMS = 10 \cdot \log_{10}(\frac{1}{N} \cdot x' \cdot x) + 3 \tag{3}$$

The RMS value gives information about the power stored in the recorded signal. If the RMS power of a calibration sound of a known SPL is calculated, the SPL for any recorded sound can be calculated. For this to be true the same recording equipment with the same volume setting must be used when recording the calibration sound. In section 3.2.3 it is described how to calculate the SPL at the recording sites.

2.1.3 Loudness

Loudness is a subjective measure. It means that the perceived loudness for a certain sound intensity can vary between individuals [Zwicker and Fastl, 1999]. For the occurrence of hearing loss, the perception of loudness is altered. A hearing impaired does not perceive sounds as loud as a normal hearing person. Loudness can be affected by three parameters: sound intensity, frequency and duration. The alteration of loudness due to sound intensity is quite intuitive but maybe not the alteration due to frequency and duration. The human ears sensitivity to different frequencies can be seen in figure 1 on page 5 where curves of equal loudness are plotted. The loudness is altered by duration because the human ears use a *temporal integration* of the sound intensity within a short time frame (200ms). The perceived loudness will be an average of the intensity during this time, which means that a quiet sound of short duration is more difficult to hear than the same sound of longer duration [Zwicker and Fastl, 1999].

Traditional hearing aid fitting was basically to normalise loudness for the hearing-impaired, i.e. to make sounds as loud as they are perceived by a normal hearing. However, in hearing aid fitting loudness normalisation for hearing-impaired is not a target in itself because even normal hearing may benefit amplification in the treble frequencies [Smeds and Leijon, 2000]. Another reason not to focus too much on loudness normalisation is stated by Leijon [1989]; 'Because of the non-linear growth of loudness, the optimal frequency-response *shape* probably depends on the preferred overall *listening level*'. This indicates that the optimal frequency response for a hearing-impaired does not depend on the preferred loudness level of the normal hearing.

2.1.4 Loudness Level and Phon

In the 1920's Barkhausen introduced the measure of Loudness Level. Loudness level was needed to characterise the loudness *sensation* of any sound. Phon is the unit used for the scale based on perceived loudness [Zwicker and Fastl, 1999]. In figure 1 on page 5 are phon curves plotted to illustrate the human ears' sensitivity to different frequencies. Lines that connect points of equal loudness for each frequency are called equal-loudness contours. An equal-loudness contour with a certain phon value must go through the sound pressure level at 1k Hz that corresponds to the same value in dB SPL [Zwicker and Fastl, 1999]. E.g. a 40 phon sound of any frequency is perceived as loud as a 1k Hz pure tone with sound pressure 40 dB SPL. It is common to use dB HL, as in Hearing Level, when measuring a subject's loudness perception, i.e. measuring the tone audiogram.

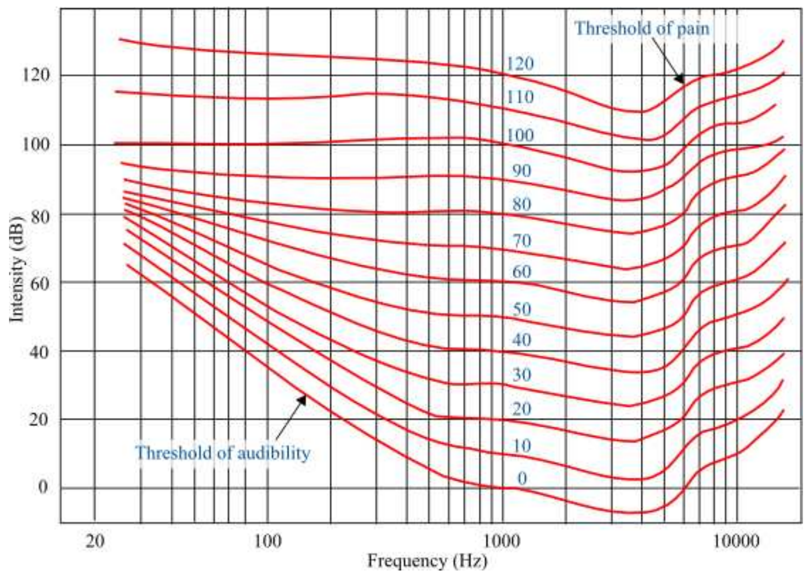

Figure 1: Fletcher-Munson equal-loudness contours. Shows the human ears' varying sensitivity to different frequencies.

2.1.5 Audiogram

The tone audiogram is a measure of hearing threshold for a set number of frequencies. The sound intensity levels are measured in dB HL, i.e. the varying sensitivity to different frequencies has been considered. These are the frequencies used to measure the audiogram in Sweden: 125, 250, 500, 750, 1k, 1.5k, 2k, 3k, 4k, 6k and 8k Hz. One must keep in mind that the audiogram only shows what the user cannot hear. At the point of threshold we know that the subject does hear, but only that he/she hears something. However, below the point of threshold we know for sure that the subject does not hear anything. The term Most Comfortable Level, MCL, is often used in hearing aid contexts. It is the hearing level that people experience as the most comfortable over time and it approximately bisects the area between the hearing threshold (HTL) and the discomfort level (DCL). In figure 2 on page 6 it is illustrated how the MCL bisects the HTL and the DCL.

2.1.6 Masking

Masking of sound is a psychoacoustic phenomenon that describes the masking effect when a high intensity sound and a quieter sound is presented simultaneously to the ear. This phenomenon can occur both in the frequency domain as well as in the time domain. A high intensity sound of a certain frequency will make the ear less sensitive to sounds in the surrounding frequencies. Upward masking, i.e. the masking of higher frequency sounds, is a bit more prominent but also lower frequency sounds are masked. In figure 3 on page 6 one can see the masking effects of three 60 dB sounds of critical-band[5] wide noise at the frequencies 250, 1k and 4k Hz. The dashed line is the threshold in quiet, compare to the

[5]Critical-bands are the human ears' built in band-pass filters. Sounds within a frequency region/band are grouped together and presented to the brain as one sound. The bandwidth is approximately 1/3-octave, varying a little with frequency [Dillon, 2001].

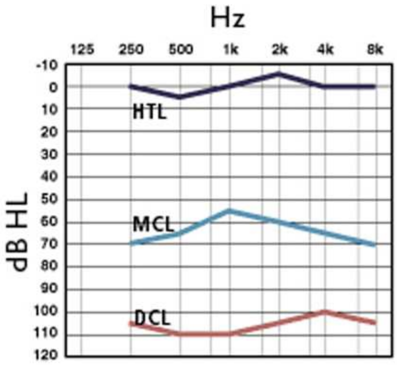

Figure 2: Shows that the Most Comfortable Level approximately bisects the area between the hearing threshold and the discomfort level, from GN-ReSound [2005]

phon curves in figure 1 on page 5. The solid line shows the hearing threshold when the noise of 60 dB has been applied. One can see that the noise is masking more of the higher frequencies than the lower frequencies.

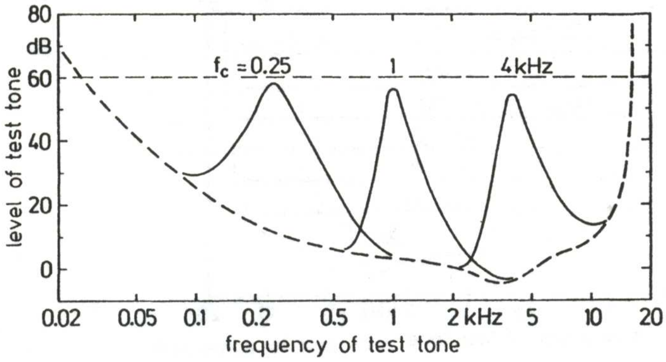

Figure 3: Thresholds of pure tones masked by critical-band wide noise of 60 dB located at 250, 1k and 4k Hz, from Zwicker and Fastl [1999]

Temporal masking mainly occurs when a faint sound follows a loud sound. The faint sound will then not be heard by the human ears. Temporal masking can also occur if a faint sound is preceding a louder sound, although the time window for this temporal effect is much smaller. Logically a faint sound will be masked if presented at the same time as the louder sound, a simultaneous-masking. This is described by figure 4 on page 7.

2.2 Components of Speech

Speech can be divided into two components: vowels and consonats. Vowels are speech sounds that are produced with a comparatively open vocal tract and with vibration of the vocal cords. Consonants on the other hand are produced with a closure at one or more points along the vocal tract, sufficient to cause audible turbulence. Vowels have most of their energy in the lower frequencies, from 250 to 1500 Hz. Consonants on the other hand are more spread out over the frequency band, from 250 to 8000 Hz [Norman et al.,

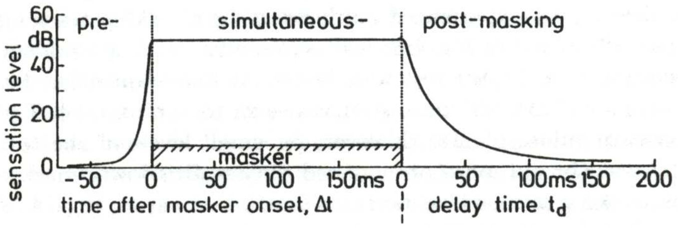

Figure 4: Describing the three time regions for temporal masking of a high intensity sound. One can see that the post-masking region is much longer than the pre-masking region, from Zwicker and Fastl [1999]

2006]. Some consonants do not produce any vibration of the vocal cords and they have frequencies from 4k to 8 kHz, such as 'f' and 's'. In figure 5 the Swedish speech sounds are described by frequency and intensity level. The *speech-banana* shows how vowels and consonants spoken with normal loudness are experienced at a distance of one meter from the speaker [Norman et al., 2006]. One can see that the vowels have a higher intensity level than consonants and are hence easier to hear. For a background noise that is equally spread over all frequencies and of 50 dB HL intensity, all consonants will be difficult to hear because they will vanish in the noise.

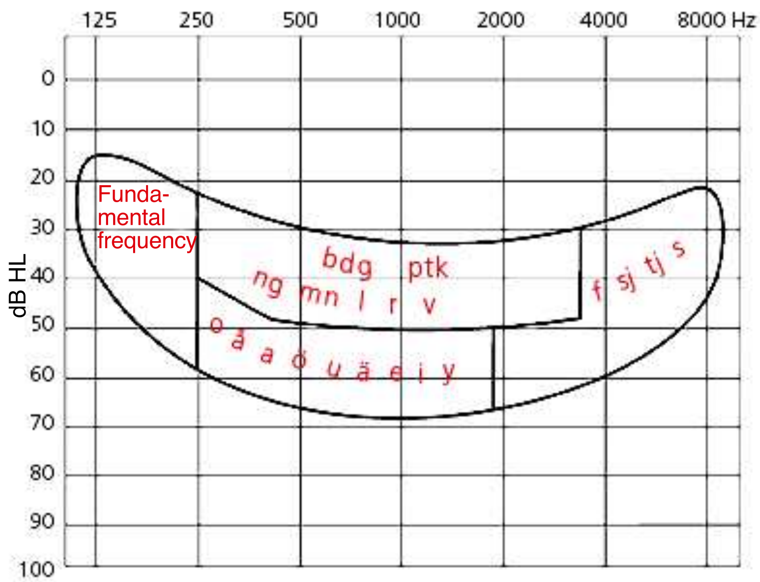

Figure 5: Swedish speech sounds related to hearing level, from Fant [1959] and Fant [1995]

2.3 Hearing Aids

In this section the reader will be informed about the technical aspects of hearing aids, such as the electrical components and what they do. Inside a modern Hearing Aid (HA) we find many advanced electronic components. A small microphone transforms the sound pressure variations to an analogue[6] electrical signal. On some hearing aids two or more microphones are used to enhance sounds coming from certain directions, preferably ahead of you. The enhanced directivity[7] can increase the signal-to-noise ratio with 2-3 dB, [Smeds and Leijon, 2000]. The analogue signal from the microphone is converted to a digital signal by an Analogue to Digital Converter (ADC). The digital signal is processed and algorithms are applied in the Digital Signal Processor (DSP). The DSP may need to read from a working memory, a non-volatile memory that keeps its content when power is removed. Some hearing aids have different insertion gain settings stored in memory to be used in different listening situations. The processed signal is finally sent to the amplifier where the signal is amplified to the desired level. A common amplifier to use is a Class D type with a Pulse-Width-Modulated[8] (PWM) output signal. When the signal is passed through a passive low-pass filter, a continuous signal representing the average of the PWM-signal is formed. This signal finally reaches the loudspeaker where it is transformed to sound waves (Texas Instruments [2004]). Additional components in the HA may be a *telecoil* to be used in locations equipped with induction-loop or a *FM-receiver* for radio systems. In figure 6 is the the block diagram of a digital hearing aid.

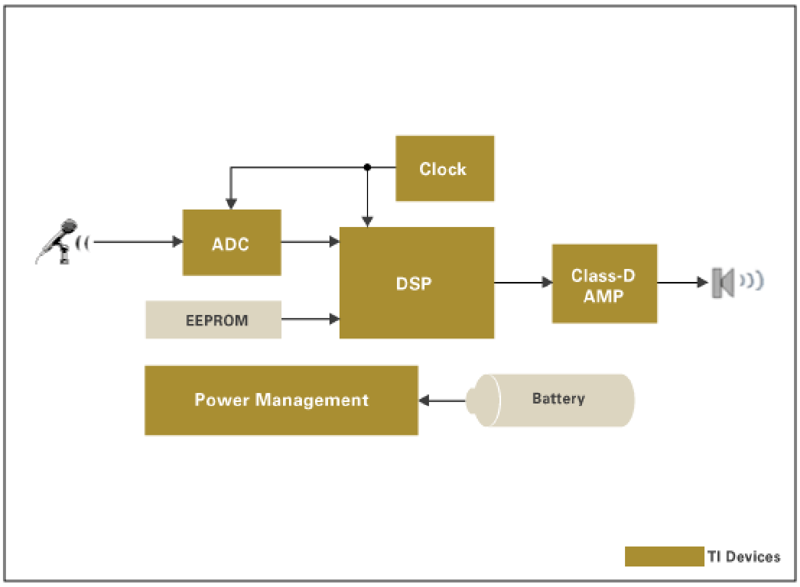

Figure 6: Block diagram of a Digital Hearing Aid, from Texas Instrument

There are many different types of hearing aids on the market today. Four models with different physical shape are listed below in table 1 on page 9. Features that we probably will see in future hearing aids are: automatic sound classification for automatic determination of alarm signals, listening environment classification to configure the HA for the included listening environments, ability to control the HA by giving speech commands, binaural signal-processing for enhanced directivity.

---

[6] continuously variable

[7] Directivity is a quantitative measure of the focusing of acoustic energy

[8] PWM - the information is represented by alterations in the width of the signal pulse, i.e. the time that the signal is high or low.

BTE	-	Behind The Ear. The most common model on the market and it is often equipped with a telecoil.
ITE	-	In The Ear. It fills the cavity of the outer ear.
ITC	-	In The Canal. It takes advantage of the amplifying shape of the outer ear. The shape of the outer ear can increase the sound pressure level at eardrum for sound coming from certain angles with up to 10 dB, [Shaw and Vaillancourt, 1985].
CIC	-	Completely In the Canal. It is as small as a cigarette stub and is fitted in the ear canal. CIC also takes advantage of the amplifying shape of the outer ear.

Table 1: Four HA models on the market today with different physical shape

2.4 Prescription Method

A prescription method is used by the HA-fitting institutes to calculate the targeted insertion gain needed[9] for the HA. It is a mathematical formula with hearing threshold as input and insertion gain for specified frequencies as output. Insertion Gain (IG) or the Real Ear Insertion Gain (REIG) is the frequency dependent gain needed for the hearing-impaired to normalise loudness, which most prescription methods strive to do. The definition of IG is the Real Ear Aided Response (REAR) minus the Real Ear Unaided Response (REUR) for each frequency[10]: $IG = REAR - REUR$. Where REAR is the sound pressure level at the eardrum when a hearing instrument is used and REUG when not used. In other words it tells us how much higher intensity a sound is presented with at the eardrum when a hearing aid is inserted in the ear.

During the 1980th, Swedish Hearing Aid fitting-institutes started to calculate the IG purely based from the measured tone audiogram [Smeds and Leijon, 2000]. Traditional hearing aid fitting and fine-tuning was a very straightforward procedure with minimum of user interaction [Elbering, 1999]. The newest technology automatically became the best and traditional subjective evaluation became less important. A common misunderstanding was that the prescription formula was a targeted method for each patient and not a starting point for individual fitting based on an average individual [Smeds and Leijon, 2000]. This simplified viewpoint led to dissatisfied HA users and once again the importance of individual testing, fitting and evaluation became clear. The importance of *perception* in HA fitting is also discussed by Schweitzer et al. [1999] where they state: '...fitting not by prescription but by perception'. Individual fitting after the prescription procedure is important for several reasons: big variations of the individuals ear, ear-canal and external ear and individual preferences of sound quality and loudness. It is impossible to find an optimal setting of the hearing aid for all listening situations. The final setting will always be a compromise of speech intelligibility, sound quality and localisation [Elbering, 1999]. It is therefore questionable if traditional fitting can give an adequate setting that meets the user's needs for all situations. Elbering [1999] recommends that the user should be

[9]How much gain that is targeted depends on the formula used and the theories on which the prescription method is based.

[10]The relation is true for a constant sound pressure level at free space.

integrated directly into the fitting and fine-tuning process.

Prescription methods and HAs can be divided into two main types: linear and non-linear. The following sections (2.4.1 to 2.4.8) will explain and give examples of the two.

2.4.1 Linear Hearing Instruments

A Linear Time-Invariant (LTI) system or simply *linear system* must undergo three criteria: homogenous, additive and time invariant [Kamen and Heck, 2000]. In a homogenous system a change of the input level will produce a proportional change on the output. It can easily be described by a equation: $y(t) = x(t) \cdot k$, where $y(t)$ is the output, $x(t)$ is the input and k is some scalar gain-constant. Any change in $x(t)$ will result in a proportional change (k) in $y(t)$. The system is additive if the response $y(t)$ to the sum of inputs $(x_1(t) + x_2(t))$ is equal to the sum of the responses $(y_1(t) + y_2(t))$ to the inputs $x_1(t)$ and $x_2(t)$ respectively [Kamen and Heck, 2000]. Time-invariance means that no matter when in time the input signal reaches the system, it will be processed congruently. When plotting a linear system in an input/output-graph it will always give a straight line. A hearing instrument have however physical limitations which constrain the dynamic range and cause non-linearity for large inputs. Two gain settings for a HA are depicted on a log-log input/output graph in figure 7: $y_{dB} = x_{dB} + log(k)$. In the first case (the lowest of the two lines) $log(k)$ is set to 20 dB, hence a linear amplification of 20 dB. For each input level the output will be 20 dB higher. The upper line of the two has an amplification of 30 dB. The reader may notice that for input signals bigger than 90 dB respectively 80 dB the hearing aid's limitation level is reached and the output will be 110 dB SPL. This is called clipping and leads to a distortion of the sound. In the following sections (2.4.2 and 2.4.3) two linear prescription methods are described.

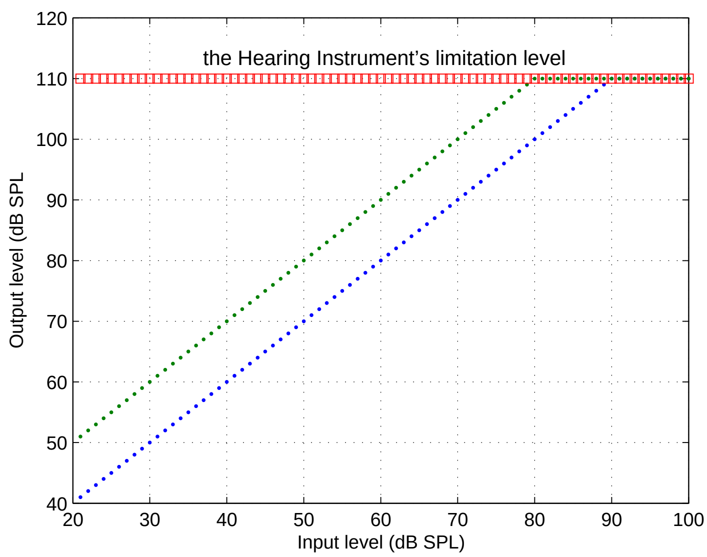

Figure 7: Input/Output graph of a Linear Hearing Instrument. At an input level of 110 dB SPL the HA's physical limitation level is reached and the output sound is distorted

2.4.2 NAL-RP

The NAL formula is developed by the National Acoustics Laboratories of Australia with the intent to maximise *speech intelligibility* at a listening level preferred by the aid wearer [Dillon, 2001]. The first NAL formula was published in 1976 and has since then been revised several times. The latest revised version is the NAL-NL1 (described in section 2.4.6), a Non-Linear prescription formula. NAL-RP (Revised, Profound) is an updated version from NAL-R (Revised) to meet the needs from users with severe hearing losses [Smeds and Leijon, 2000]. The idea behind the formula is that the maximum speech intelligibility is reached when all critical-bands gives equal contribution to a sound's total energy [Smeds and Leijon, 2000]. In the equations below the NAL-RP prescription formula is described. H_i is the hearing threshold for frequency i. Equation 5 shows H_{3FA} which is the sum of the thresholds at 500, 1k and 2k Hz and it is used in the calculation of the *insertion gain* (equation 6). In table 2 are the values for the constant called PC that is used in equation 6. Hearing threshold at 2k Hz is used as input to select the PC constant [Dillon, 2001].

$$H_{3FA} = (H_{500} + H_{1k} + H_{2k})/3 \tag{4}$$

$$X = \begin{cases} 0.15 \cdot H_{3FA} & \text{for } H_{3FA} \leq 60 \\ 0.15 \cdot H_{3FA} + 0.2 \cdot (H_{3FA} - 60) & \text{for } H_{3FA} > 60 \end{cases} \tag{5}$$

$$IG_i = X + 0.31 \cdot H_i + k_i + PC \tag{6}$$

PC values in dB, Frequency in Hz							
$H_{2\,\text{kHz}}$	250	500	1k	2k	3k	4k	6k
≥ 90	0	0	0	0	0	0	0
95	4	3	0	-2	-2	-2	-2
100	6	4	0	-3	-3	-3	-3
105	8	5	0	-5	-5	-5	-5
110	11	7	0	-6	-6	-6	-6
115	13	8	0	-8	-8	-8	-8

Table 2: PC values used for equation 6 of the NAL-RP prescription formula

2.4.3 POGO

The POGO formula, short for Prescription of Gain & Output, is based on the half-gain rule with an added low cut in the low frequencies up to 1k Hz. The half-gain rule is based on the observation made by S.F. Lybarger in 1944; the amount of insertion gain wanted is approximately half the amount of the loss in hearing threshold [Dillon, 2001]. The low cut in the POGO formula was added to decrease the masking of high frequency sounds. The information in low frequencies consists of high intensity vowels and uninteresting ambient noise. A cut in this frequency region is justified by that the vowels don't need the same amplification as the more quiet consonants in the higher frequencies. The insertion gain is set to be half the hearing loss plus a constant, the low cut (see table 3 on page 12). In the test program the revised POGO procedure, POGOII is used. It was developed/revised by

Schwartz, Lyregard and Lundh in 1988 to provide additional gain for people with severe hearing losses [Dillon, 2001]. POGOII formula is described below in table 3.

$$IG_i = 0.5 \cdot H_i + k_i, \quad \text{for } H_i < 65$$
$$IG_i = 0.5 \cdot H_i + k_i + 0.5 \cdot (H_i - 65), \quad \text{for } H_i \geq 65$$

Frequency (Hz)	250	500	1k	2k	4k
k_i (dB)	-10	-5	0	0	0

Table 3: The POGO II formula. The low-cut is added with the k_i vector.

2.4.4 Non-linear Hearing Instruments

A Non-Linear Time-Invariant system lacks one or more of the criteria stated for the LTI system defined in section 2.4.1 [Smeds and Leijon, 2000]. Non-LTI systems are often said to be non-linear and for most cases it means that the homogeneity criteria is broken. For example a 10 dB change on the input will not give a 10 dB change on the output. Instead a non-linear hearing instrument will adapt and change the amplification according to the input level. This is called Automatic Gain Control (AGC) or Compression. There are many types of AGCs developed to serve different purposes and in table 4 some of them are described:

AGC-O	-	*Output*-controlled compression which restrict the signal to fit the HA's physical limitations.
AGC-I	-	*Input*-controlled compression to fit the signal into the hearing impaired user's dynamic range.
TILL	-	*Treble Increases at Low Levels*, high frequency sounds are more amplified for low-level input signals.
BILL	-	*Bass Increases at Low Levels*, high frequencies are amplified to the same high level for all inputs but low frequencies are only amplified for low input signals.
Multichannel	-	Some HA's treat low and high frequencies in separate channels equipped with different AGCs.
Adaptive Compression	-	After a high intensity sound the time to adapt the AGC can be set automatically. After a long duration of high intensity sound the HA will have a longer adaptation time than after a short duration sound.
Syllable Compression	-	An AGC-I with short attack[a] and adaptation time to make it possible to adapt to each syllable.

[a]Attack time: the time it takes for the HA to cut the amplification for a sudden high intensity sound

Table 4: Various types of Automatic Gain Control or Compression, from K-AMP [1995]

In figure 8 on page 13 the input/output graph for a HA with AGC-O is depicted. This type of AGC has a varying compression ratio depending on the input level. For levels up to 60 dB SPL the compression is 1:1, the same as for the linear HA above. When the input level is increased above 60 dB SPL the compression is increased to 2:1, i.e. for a

6 dB change of the input the output changes 3 dB. The compression ratio is increased even more for higher input levels to avoid reaching the HA's physical limitations. Varying compression ratio for the entire frequency band is uncommon in modern hearing aids, instead a multichannel AGC is used with many frequency channels [K-AMP, 1995]. Some hearing aids even have a channel for each critical-band with independently controlled compression ratios.

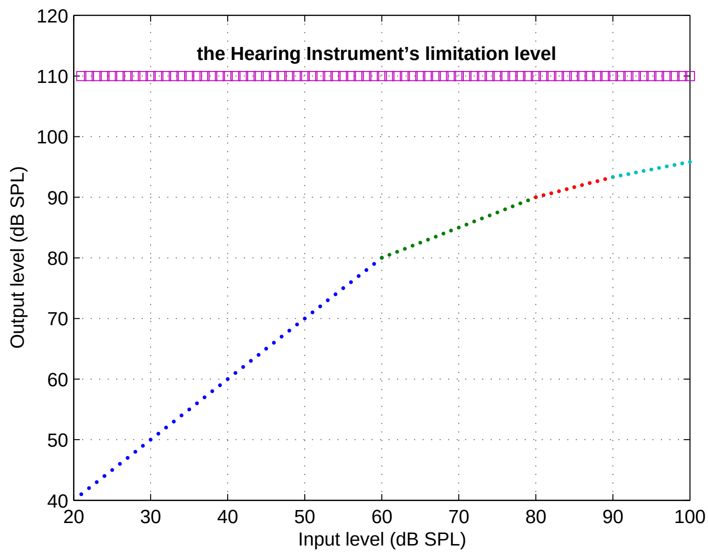

Figure 8: Input/Output graph for a non-linear Hearing Instrument. The compression ratio is successively increased.

2.4.5 FIG6

The FIG6 prescription method, named after the data from Figure 6 in Killion's [1993] article - 'Prescriptive procedure', specifies how much gain that is required to normalise loudness for different input signals. It is not based on individual loudness perception but an average of loudness perception for a large group of people [Dillon, 2001]. To calculate the prescribed insertion gain, hearing thresholds (H_i) for frequencies (i) and the input signal level are necessary. The input signals are divided into three groups based on their intensity level. Input signals for levels up to 40 dB SPL, 65 dB SPL and 95 dB SPL are separated and treated differently (see table 5 on page 14). FIG6 is used in the test program as the initial condition for the user. As can be seen the insertion gain is zero for a normal hearing person with hearing thresholds less than 20 dB HL, hence a normal hearing user will start the test program with a flat gain-frequency response.

2.4.6 NAL-NL1

NAL-NL1 is as NAL-RP (sec 2.4.2) developed by the National Acoustics Laboratories of Australia. NL1 is short for non-linear version 1. Its target is not to restore normal loudness but to maximise speech intelligibility. The overall loudness perceived by the

40 dB SPL input levels		
$IG_i = 0$	for	$H_i < 20\ dB\ HL$
$IG_i = H_i - 20$	for	$20 \leq H_i \leq 60\ dB\ HL$
$IG_i = 0.5H_i + 10$	for	$H_i > 60\ dB\ HL$
65 dB SPL input levels		
$IG_i = 0$	for	$H_i < 20\ dB\ HL$
$IG_i = 0.6(H_i - 20)$	for	$20 \leq H_i \leq 60\ dB\ HL$
$IG_i = 0.8H_i - 23$	for	$H_i > 60\ dB\ HL$
95 dB SPL input levels		
$IG_i = 0$	for	$H_i < 40\ dB\ HL$
$IG_i = 0.1(H_i - 40)^{1.4}$	for	$H_i \geq 40\ dB\ HL$

Table 5: FIG6 prescription formula

hearing impaired is set to be the same or less than the perceived loudness of a normal hearing person. The reason for this is that people are less able to recover information from sound if they are forced to listen to it at high sound levels [Dillon, 2001]. Unfortunately only Speech in Quiet was considered during the development of the NAL-NL1, as for most of the present prescription methods.

2.4.7 DSL[i/o]

The idea behind the Desired Sensation Level prescription formula is to map the dynamic range of a normal hearing person into the dynamic range of a hearing-impaired [Smeds and Leijon, 2000]. It has *hearing thresholds* as input and the *real-ear aided gain* as output. The formula is commonly used on children because its fitting methods are easily applied on children [Dillon, 2001].

2.4.8 Prescription method summary

In section 2.4 the basics of hearing aid prescription were described. A few examples of both linear (NAL-RP and POGO) and non-linear (FIG6, NAL-NL1 and DSL[i/o]) prescription methods were given and the difference between a linear and non-linear HA was declared. One important issue that Smeds and Leijon [2000] points out is that there is a connection between the formulas developed for linear hearing aids and those developed for non-linear hearing aids. As long as we are considering speech of normal loudness they should prescribe similar results.

2.5 Evaluation of Hearing Aid fitting

To be able to measure the quality of a HA fitting some sort of quantitative measure of user satisfaction can be used. To evaluate psychoacoustic tests one can preferably use an adaptive selection strategy. The fundamental of adaptive testing is that the response of the participant determines the subsequent test conditions. In the test program an adaptive procedure called Double Elimination Tournament (DET) is used. DET is a simple adaptive procedure where it is not required to compare all conditions. Other adaptive strategies available are *Iterative round robin* and *Modified simplex procedure* but DET is chosen because it is relatively simple to implement and it has a fast runtime.

2.5.1 Double Elimination Tournament

The Double Elimination Tournament is described by Neuman et al. [1987] and it is a fast evaluation tool for HA fitting. The tournament works like an elimination soccer tournament but with an extra safeguard against occasional poor judgement. DET is used in the test program to compare *gain-frequency responses* based on prescription and based on individual settings. Three prescribed settings and one individual setting are the conditions in the tournament. DET is as most efficient when the number of conditions is a power of two, hence a group of four is a good number. The conditions are grouped into two pairs at random and the user is presented to the first group of sounds. The participant can listen and compare the two sounds until he/she has settled for the one that sounds most pleasing. The winner is placed in the *winners' group* and the loser in the *losers' group*. The next two sounds are compared and a winner is determined. The conditions in the winners' group are paired and the conditions in the losers' group are paired. The comparison continues with pairs randomly chosen from the winners' or losers' group. If a condition from the winners' group wins a comparison, it will stay in the winners' group. If it loses it will be moved to the losers' group. A condition from the losers' group that loses a comparison has *lost twice* and is removed from the tournament. Finally when there is only one condition left in the winners' and losers' group, they are compared until one of them has two defeats. The remaining one is the final winner. As one notices this procedure is quite complicated to implement because a running tally of winners and losers is necessary to determine how the tournament is to proceed. However, if it is correctly implemented on a computer it will run fast and give non-erroneous results.

2.6 Signal-Processing

In the test program developed in MatLab many signal-processing techniques are used. This chapter strives to make the reader understand the development of the test program and the choices made considering these techniques. To be able to process the recorded sounds, the digital sound files are imported into MatLab and stored in a vector. Base functions are created and mapped to the desired frequency scale. A continuous insertion gain curve is created by interpolating the gain-frequency points from the prescription formula. The base functions and the insertion gain curve are added and normalised and finally used as a filter-curve. The curve is used by a Finite Impulse Response-filter or FIR-filter to filter the sound vector. The FIR-filtering is the heart of the program and therefore an introduction to FIR and basic signal-processing follows.

2.6.1 Finite Impulse Response

Filtering is a form of convolution. Convolution is algebraically to multiply two polynomials whose coefficients are the elements in the vectors b and x, see equation 7.

$$y_c(n) = \sum_j b(j)x(k+1-j) \tag{7}$$

For a filter y_c is the output signal, b contain the filter coefficients and x is the input signal. When a linear phase is desired for a digital filter, FIR is often used as opposed to Infinite Impulse Response (IIR). The difference between FIR and IIR is that a FIR-filter has an impulse response[11] $h[n]$ that is zero for all samples n≥N, where N is the window size. The IIR-filter has an impulse response of infinite length and has often a non-linear phase which may cause some distortion in the time signal [Kamen and Heck, 2000]. However, when designing the FIR filter there are also some things to be considered. A FIR filter of linear phase will cause a time delay of the signal being processed, i.e. the signal cannot be filtered in real-time [Kamen and Heck, 2000]. Also it generates a ripple in magnitude plot if the filter coefficients are truncated. The ripple was first discovered and explained by Josiah Willard Gibbs (1839-1903) and is referred to as Gibbs phenomenon [Kamen and Heck, 2000]. He discovered that for a discontinuity of a signal $x(t)$ the Fourier series representation (the frequency content) of that signal will have an overshoot. This overshoot will tend to be off by 9% as the number of Fourier terms goes to infinity. Hence when the filter coefficients are truncated the output will have a ripple in the magnitude plot. To truncate the coefficients a *window* of size N is used, see equation 8.

$$W[n] = \begin{cases} 1, & 0 \leq n \leq N-1 \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

The FIR filter is constructed by multiplying the window $W[n]$ by the impulse response of the IIR filter $h[n]$, see equation: 9.

$$h_d[n] = W[n] \cdot h[n] \tag{9}$$

To filter the sound in the test program, the built in MatLab function FIR2 is used. FIR2 filter uses a Hamming window to truncate the infinite-length impulse response h[n]. Instead of using a rectangle window a Hamming window will reduce the ripple in frequency response of the FIR filter otherwise caused by the abrupt truncation [Kamen and Heck, 2000]. A Hamming window is smoother and tapers off gradually and is described by the function below.

$$w[n] = 0.54 - 0.46 \cdot \cos(\frac{2\pi \cdot n}{N-1}), \text{ for } 0 \leq n \leq N-1 \tag{10}$$

The window will be smoother if the integer N is increased. By choosing an even order (N-1) of the FIR-filter the returning filter coefficients will be an odd number (N). Hence the length of the filter is set to be $N = 2m+1$ where m is an integer or an integer divided by 2. Now the filter undergoes the linear phase criteria: $h_d[n] = h_d[2m-n]$, i.e. it is symmetric around $m = (N+1)/2$. This is according to the criteria that a filter $H_d(\omega)$ has linear phase if its impulse response has even symmetry [Kamen and Heck, 2000]. The time delay with a linear phase filter stated above is not a problem in the test program because the filtering process is not made in real-time anyway.

[11] The impulse response of a system is the output of the system when its input is the unit impulse. The unit impulse $\delta(t)$ is a signal equal to zero for all $t \neq 0$ and $\int_{-\varepsilon}^{\varepsilon} \delta(t)dt = 1$ for any $\varepsilon > 0$ [Kamen and Heck, 2000]

2.6.2 Signal-processing in MatLab

The MatLab function *filter* is a *direct form II transposed* implementation of the standard difference equation (see figure 9). The input is x, the output is y and the filter coefficients are stored in b. Z^{-1} is a delay-tap to shift the data one step.

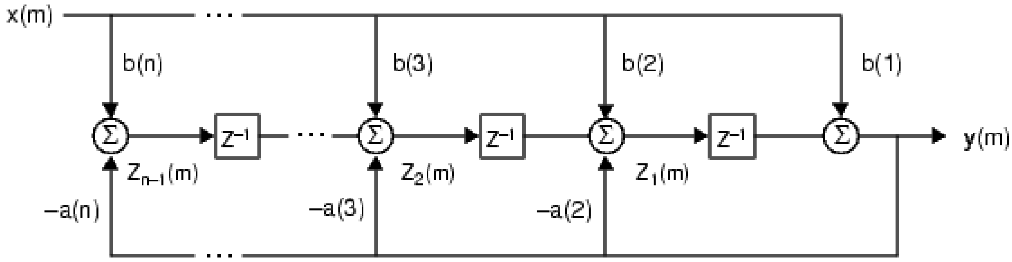

Figure 9: The structure for the *filter* is the Direct form II transposed. The input is x, the output is y and the filter coefficients are stored in b.

The structure can also be expressed as an equation for an element n, see equation 11:

$$\begin{aligned} y(n) = b(1) \cdot x(n) \quad &+ \quad b(2) \cdot x(n-1) + \ldots + b(nb+1) \cdot x(n-nb) \\ &- \quad a(2) \cdot y(n-1) - \ldots - a(na+1) \cdot y(n-na) \end{aligned} \tag{11}$$

It can also be expressed as a rational transfer function by describing the input/output function in the z-domain, see equation 12:

$$Y(z) = \frac{b(1) + b(2)z^{-1} + \ldots + b(nb+1)z^{-nb}}{1 + a(2)z^{-1} + \ldots + a(na+1)z^{-na}} \cdot X(z) \tag{12}$$

A FIR filter does not have any *poles*, i.e. it only has b-components and all the a-components are zero. This means that there is no feedback in the *direct form II transposed* figure (see figure 9 on page 17). It makes sense that it has no poles because the FIR-filter coefficients are of finite length. After the b-components have been removed from equation 12 the final input/output equation for the FIR-filter becomes:

$$Y(z) \quad = \quad b(1) + b(2)z^{-1} + \ldots + b(nb+1)z^{-nb} \cdot X(z) \tag{13}$$

What the function *FIR2* does in MatLab is to calculate the filter coefficients b that are used by the filter, see equation 11.

3 Implementation

In this chapter the test program will be in focus. First some feasible solutions are presented and their pros and cons are evaluated. In section 3.2 the recording process of the sound material is presented and technical details about the recording equipment are explained. The last section gives an overview of the various parts of the main test program. The test can be divided into two main parts, a fitting part and an evaluative part.

3.1 Feasible solutions

A few possible solutions to construct the listening test in this study were considered. Two different hardware solutions were chosen from, either to use a standard computer or to use programmable electronics such as a Field Programmable Gate Array (FPGA). One advantage with the FPGA was that the program would run fast and a real-time solution would probably be easy to implement. A negative issue with the FPGA was the need for external hardware such as an amplifier and controller buttons. It would also be a bit more difficult to have a detailed graphical feedback to show the progress of the test. The programming language used for a FPGA is VHDL and it requires much more developing time than a high order programming language. On these bases a Windows PC was decided to be the most efficient platform to work on. The numerical computing environment *MatLab* or the industry standard object-oriented language *C++* were the two programming environment to choose from. MatLab was chosen because of its many built in signal-processing functions that would come in handy in this project.

3.2 Recording of sound material

One of the aims with the study was to simulate a dialogue in a real listening situation as good as possible. To get the sound material needed for this simulation, one could have used a number of phonemic balanced word lists. Another solution was to use phonemic balanced sentences with little or no redundancy and add noise from different sound environments to reach the desired Signal-to-Noise Ratio (SNR). However, the easiest way to get an authentic listening situation was to record it on spot, with real people and real noise. The problem with a real recorded situation was however the deficiency to control the quality of the recordings on forehand. It was difficult to set a predefined SNR and to get a noise that does not fluctuate over time. However, with vast amount of recorded material, sections that suited the desired criteria was localised and extracted. The recorded sound material was 45 minutes in total and the extracted material was approximately 2 minutes. The targeted SNRs for the recordings in noisy environments were 0 dB, and it was met for Speech in Babble (see table 9 on page 22 for data on the recordings). If a listening situation has a high SNR it is easier to hear the speaker than for a low SNR. Due psychoacoustic and acoustic phenomena a normal hearing person can understand speech in an environment of approximately -2 dB SNR, i.e. the speech has a lower intensity level than the noise.

3.2.1 Recording equipment

The Recording equipment was an iRiver iHP-120 mp3/wav recorder. The sound was recorded in stereo, but with only one channel in use, in the uncompressed wav-format. Details of the recording equipment can be found in table 6 on page 19.

Sampling frequency = 44.1kHz
Bitrate = 1411k (44.1kHz*16 bit*2 channels)
Frequency Range = 20 Hz to 20k Hz
SNR = 90 dB
Automatic Gain Control (AGC) is turned off
External microphone volume = 13 (13 out of 20)

Table 6: Specifications of the iRiver iHP recording equipment

The microphone used for the recordings was an AV-Jefe TCM110 omnidirectional microphone with 1k ohm impedance. The microphone had a flat frequency characteristic between 20 Hz and 7k Hz, i.e. it gave a correct loudness representation for the most important frequencies containing speech information. See figure 10 for the frequency characteristics of the microphone.

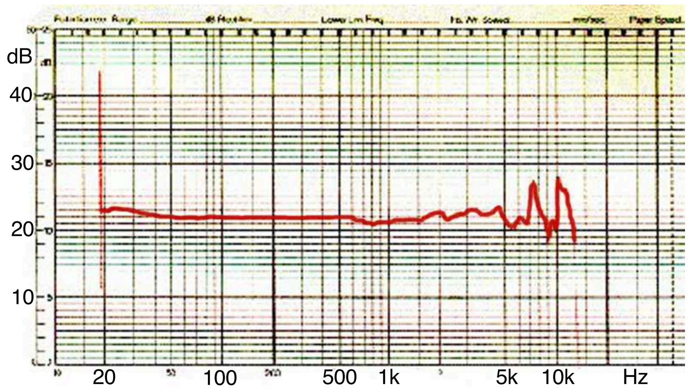

Figure 10: Frequency characteristics for the microphone TCM110. One can see that the microphone has a reasonable flat response up to 7k Hz.

3.2.2 Listening Situations

In traditional prescription methods, conducted by the NAL-institute for instance, speech in quiet has been the only listening situation of concern. A hearing impaired person often manage quite well in quiet situations but has problems when faced to a situation with background noises. In this study, speech in three listening situations has been selected. The situations are: speech with no background noise, speech in traffic noise and speech in a room with many background speakers. These environments are in the thesis called: Speech in Quiet (SIQ), Speech in Traffic (SIT) and Speech in Babble (SIB). The three situations have different Signal-to-Noise Ratio (SNR), where speech in babble has the worst SNR (see table 7 on page 21). Both female and male voices have been used in the recordings. In figure 11 on page 20 the frequency content, FFT, for a male voice in SIQ is plotted. All pauses and silent sections have been removed from the sound file before the FFT calculation. The curve is declining with local maxima at 600, 1k, 2.5k and 3k where more prominent speech sounds are located. One can also see that the vowels, located between 200 and 1k Hz, have a higher intensity level than the consonants.

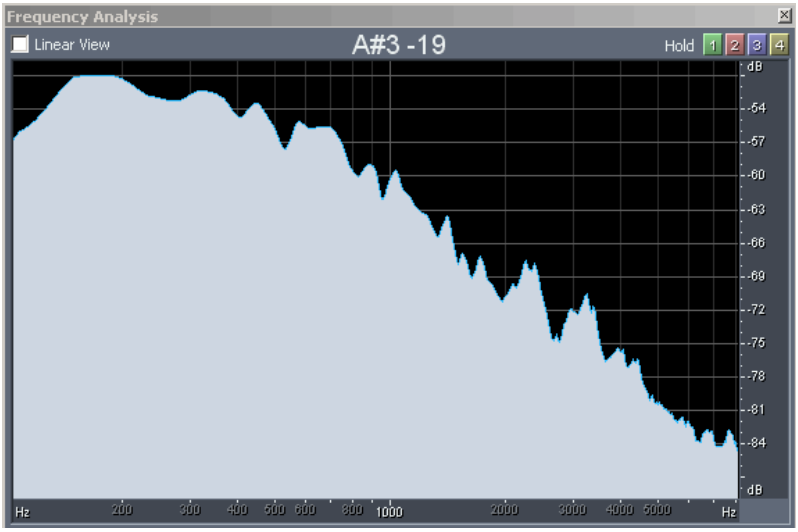

Figure 11: The FFT for a male voice recorded with no background noise

When the dialogues were recorded a small microphone was placed behind the outer ear to simulate the position of a BTE hearing aid (see figure 12). This placement conveys that the resonance effects of head and torso were being considered. Dialogues with up to 10 minutes duration were recorded. Sections with a steady sound pressure level of background noise were later on selected and extracted. The final sound clips have 5 seconds duration and it only consists of the opponent's speech and the background noise.

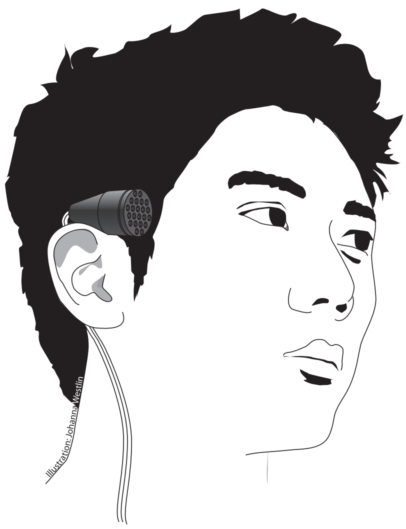

Figure 12: The microphone placement simulating the position of a BTE hearing aid, illustration J. Westlin

3.2.3 Calculation of the SPL at the recording sites

A calibration tone of 1000 Hz and 93.8 dB SPL was recorded with the recording equipment, with exactly the same settings as for the listening environment recordings. The average RMS power (see section 2.1.1) of the signal was calculated in Adobe Audition 1.5. The Adobe Audition calculation showed that 93.8 dB SPL was the equivalent of -13.65 average RMS power. The RMS values for the listening environments were then calculated and finally the dB SPL at the recording sites could be calculated. To calculate the SPL for speech and noise at the recording sites, the known RMS and SPL values for the calibration tone were used (see equation 14).

$$\mathrm{SPL}_{at\ site} = \mathrm{SPL}_{calibration} - (\mathrm{RMS}_{calibration} - \mathrm{RMS}_{at\ site}) \tag{14}$$

The original recorded material was listened to and speech and noise were separated for each environment. To do this the sound file was opened in Adobe Audition 1.5 and sections with only speech was extracted and pasted into a separate track. When the whole sound file had been processed the outcome was one track with only speech and one track with only noise. The tracks with only speech did inevitable also have noise in the background but because the speech had a higher intensity than the noise in most of the recordings this was not a problem. However, it implies that the worst SNR value that can be detected was 0 dB, i.e. negative SNR was not possible. The RMS value for each track was calculated and by comparing the speech-track with the noise-track, it was possible to calculate the SNR (see equation 15).

$$\mathrm{SNR} = \text{Signal level} - \text{Noise level (dB)} \tag{15}$$

In the tables below all the RMS and SPL data are presented (see table 7, 8 and 9). In table 8 the sound files have been filtered through a A-weighted filter to get SPL values closer to how the human ears would experience the sound. In table 9 on page 22 the SNR values for the three listening environments are displayed. The reader may notice that the SNR for *speech in quiet* and *traffic* differs between the flat and the A-weighted. The reason for this is that SIQ and SIT have noise in the bass and treble frequencies that are masking the speech. This noise are being suppressed for the A-filtered sound.

	Avg. RMS Power	SPL dB(flat)
Calibration tone 1kHz	-13.65	93.8
Speech in Quiet		
Only speech:	-42.17	65.28
Only silence:	-63.18	44.27
Speech in Traffic		
Only speech:	-28.78	78.67
Only traffic:	-30.65	76.80
Speech in Babble		
Only speech	-27.17	80.28
Only babble	-27.10	80.35

Table 7: Sound Pressure Levels at Recording sites for the speech and the noise. The calibration tone is used to calculate the SPLs at the recording site from the average RMS Power.

	Avg. RMS Power	SPL dB(A)
Calibration tone, 1kHz	-13.70	93.8
Speech in Quiet		
Only speech:	-45.44	62.06
Only silence:	-68.36	39.14
Speech in Traffic		
Only speech:	-32.19	75.31
Only traffic:	-35.74	71.76
Speech in Babble		
Only speech	-29.26	78.24
Only babble	-29.28	78.22

Table 8: A-weighted Sound Pressure Levels at Recording sites for the speech and the noise. The calibration tone is used to calculate the SPLs at the recording site from the average RMS Power.

	SNR, dB(flat)	SNR, dB(A)
Speech in Quiet	21.01	22.92
Speech in Traffic	1.87	3.55
Speech in Babble	-0.07	0.02

Table 9: Signal to Noise Ratio for the recorded sounds

3.3 Equipment for the Test program

The test was implemented on a PC with standard hardware and with MatLab installed. Additional equipment was a pair of full-sized headphones that were connected to the sound card. Details about the headphones can be found in table 10.

Half-Sealed Full-Size Headphones: Philips SBC HP890.
Frequency range: 5 - 30000 Hz
Impedance: 32 ohm
Sensitivity: 106 dB SPL for 1mW input

Table 10: Details of the headphones as supplied by the manufacture

At first a laptop was used to make it possible to test on people not situated in Stockholm. However, after some testing it was discovered that the performance of the laptop was too poor. When too much data were sent to the sound card the computer froze and the test had to be restarted. A desktop computer with better performance therefore replaced the laptop. To make the results fully comparable both computers had to be calibrated and their differences had to be taken account of. In the next section the calibration of the test equipment is described.

3.4 Calibration of the test equipment

A calibration of the test equipment was needed to be able to calculate the frequency response for the entire system. A small microphone was placed in the outer ear and the headphones were put on. White noise was played in the headphones and the noise was recorded. A Fourier Transform analysis was made in Adobe Audition 1.5 and the resulting frequency response can be seen in figure 13. The figure shows that the system has a reasonable flat response up to 7k Hz (see marker) with a smaller knob at 3k to 6k Hz. From 7k Hz and above the response is fluctuating considerably. The big fluctuation is probably due to the occurrence of standing waves in the ear cavity. The declining trend however, is probably due to the system's poor response to very high frequencies.

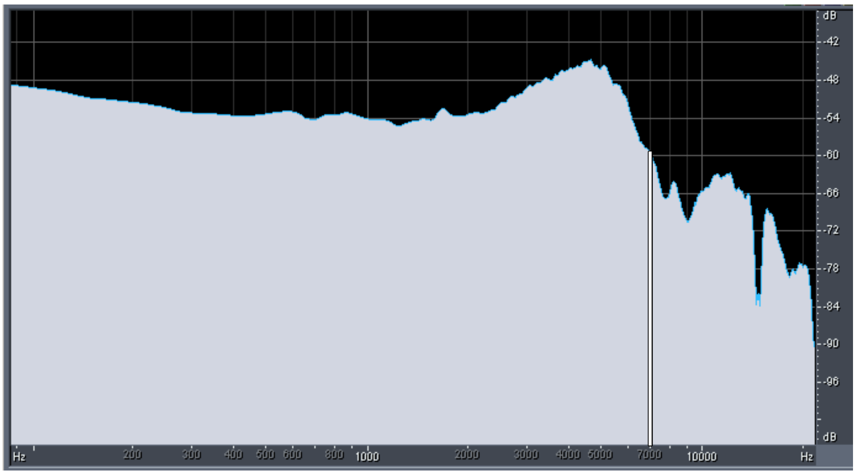

Figure 13: The FFT of the testing system for white noise. A line at 7k Hz marks where the response is starting to fluctuate more significantly.

A calibration of the output volume for the two systems was also made. First a calibration tone of a 500 to 1k Hz noise was played and the actual dB SPL was measured with a Sound Level Meter. A microphone also recorded the sound and the RMS of the recorded sound was measured. Now it was known how much an electronic value (RMS) corresponded to a certain sound pressure value (SPL). To measure the sound pressure level for the system's volume control a microphone was placed in the outer ear cavity between the headphone and the ear. A sound was played for a set number of volume settings and the response was recorded. The sound played during the measurements was the same sound that the participants listened to when setting the volume in the study. It was a speech in babble sound of 20 seconds duration that was normalised to -25 dB RMS. The RMS levels were calculated in Adobe Audition 1.5 and then used to calculate the corresponding SPLs. See table 11 for the SPLs of the laptop and desktop computer.

Volume	10%	20%	30%	40%	50%	60%	70%	80%	90%
dB SPL (desktop)	67.1	71.8	76.0	77.5	80.6	82.1	82.1	83.4	85.0
dB SPL (laptop)	47.6	51.4	55.4	56.6	59.6	61.1	62.3	62.6	63.9

Table 11: Calibration of output volume for the two systems, measured at the outer ear cavity.

Two linear approximations were made from the data set, one for 10% to 50% volume and a second for 50% to 90% volume. All the participants' volume settings were within these regions. The linear approximation was later on used to calculate the exact SPL for any volume setting made by the subject. In figure 14 the approximations are plotted as lines and the sample values are marked as 'x'. The bottom line represent the laptop volume and it be seen that is has an output that is approximately 20 dB lower than the desktop computer.

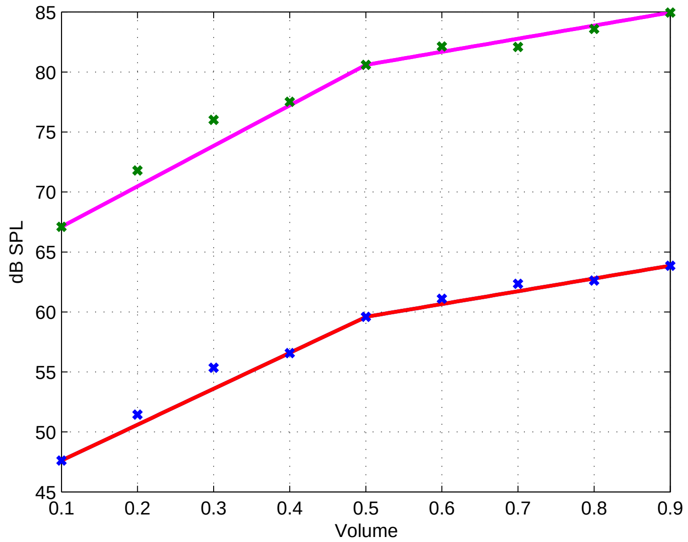

Figure 14: Sound volume for the two computers. The bottom line illustrate the laptop and the top line illustrate the desktop computer

3.5 Test program

In this section the various steps in the test program are described. A flowchart has been made to guide the reader through the steps and it can be seen in figure 15 on page 25. The code written for the test is in total 1300 lines were 500 lines control the functionality and 800 lines control the graphics. The program consists of several functions (called m-files in MatLab) where each function performs a small task or controls the progress of the program and the graphics.

3.5.1 The first part of the Test program

In the beginning of the test the participant is greeted with a start-up screen with brief details about the research and three input fields. The participant can here enter personal details as name, age and audiogram (see Appendix A). The audiogram data are used to calculate the insertion gain using the FIG6 prescription method, described in the theory chapter (see section 2.4.5 on page 13). The next step is to set the desired sound volume and to do this a new GUI is loaded. Here the user can change the volume with a slider and choose the side with impaired ear that will be used during the test (see Appendix B). To be able to change the Windows XP volume from MatLab a command-line tool is

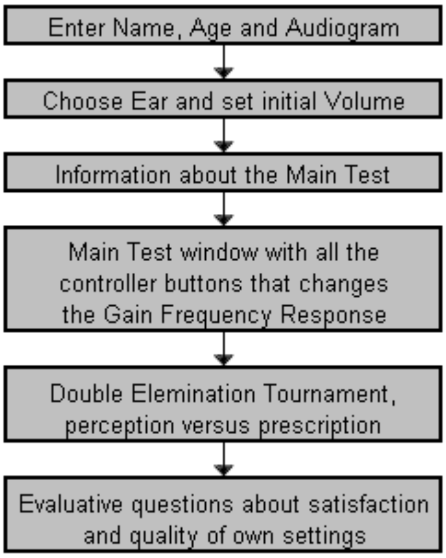

Figure 15: Flowchart for the test program

installed. It is called NirCmd and is developed by NirSoft [2005] and it makes it possible to change the volume from the command prompt. To change the volume the following command are sent from MatLab to the command prompt:

```
dos('C:...\nircmd.exe setsysvolume VOL')
```

To set the minimum volume *VOL* should be set to 0 and for maximum it should be set to 65535.

When the *test-sound* button is pressed a sound of 20 seconds duration is played in the headphones. During this time the user can set the preferred volume with a slide-control. The sound that is played is filtered with the FIG6 filter-curve and normalised to -25 dB RMS. This RMS level is kept through all filtering of sound in this program, which gives enough dynamic space for eventual fluctuations in the signal without clipping. The dynamic space is independent of the initial volume setting and it gives all the test subjects equal testing conditions. Once the user has settled for a comfortable level, hence the *Most Comfortable Level*, an information screen is loaded. In the information screen the user is informed about the main test and the tasks to be performed by the test subject. The user is asked to find a most comfortable sound environment for each of the three listening situations. To find a comfortable sound the participant has four groups of controller buttons and by manipulating these the the sound representation will change. It is up to the user to decide which dimension of sound quality (intelligibility, loudness or clearness) to strive for. This study does not focus on speech intelligibility but on a most comfortable or pleasing sound environment. This has been deliberately chosen to add more aspects than just speech intelligibility. The information screen can be seen in Appendix C.

3.5.2 Main Graphical-User-Interface

In the main part of the test the user is asked to set the desired frequency response for three different listening situations (see Appendix D for the main GUI screen). For each situation the user can control the frequency response of the sound independently. The subject can also go back and change settings for the other sound environments without losing any previous settings. The main GUI has four groups of controller buttons visible for the user. Information on their exact effect on the frequency response is not displayed. When not giving visual information about what the user is doing, the subject has to fully rely on other senses, in this case the sense of hearing. The user clicks on the controllers and listens to the filtered sound file. This process of clicking and listening will continue until the user is satisfied with the result and has reached the most comfortable sound representation. To make the comparison between settings easier the user has two modes, mode A and B, where different settings can be stored. The user will end up with one setting for each of the three listening situations.

The graphical display of the main GUI can be seen in Appendix D. The four base functions are considered most important so for this reason the GUI is graphically formed to put these at the centre of attention. The modes have no set order between them hence they are called mode A and B and not mode 1 and 2. However, the tuning switch have a given order where the user should start with coarse, hence the numbering. A colour coding for the base functions is added to make the visual feedback clearer. For example when pressing on the buttons for the green base functions, the green hand of time changes. The hands of time make it much easier to remember previous settings and to compare different settings. More important and frequently used buttons are made bold and bigger to immediately attract the test subjects' attention. All this graphical information strives to make the GUI as user-friendly as possible. In the next section the base functions will be described.

3.5.3 Base functions

The four groups of buttons control base functions of increasing degree. The base functions are used to calculate filter-curves that control the amplification for each frequency. The base functions used in this program are cosine functions with variable amplitude. The first base function is a cosine function from 0 to π, hence a tilt function where the controller sets the tilt amount. This means that the controller input is the slope k on the cosine function. Equation 16 shows how the base function is calculated where k is the slope, b is the base-number from 1 to 4 and x is a vector containing the frequency bins from 0 to 10. The next base function is a cosine from 0 to 2π, i.e. b is set to 2. The third from 0 to 3π and fourth from 0 to 4π.

$$y = k \cdot \cos(b \cdot \pi \cdot x/10) \tag{16}$$

The user can change between coarse- and fine-tuning with radio buttons in the interface. With the fine-tuning set, each click makes a 2 dB difference on the filtered output and with coarse-tuning each click makes a 8 dB difference. The four base functions can be seen in figure 16 on page 27. When the base functions are added to one curve they can form almost any arbitrary shape. However, the curve will be continuous and fairly smooth. The base functions give the users the possibility to find their preferred frequency response with only four controllers.

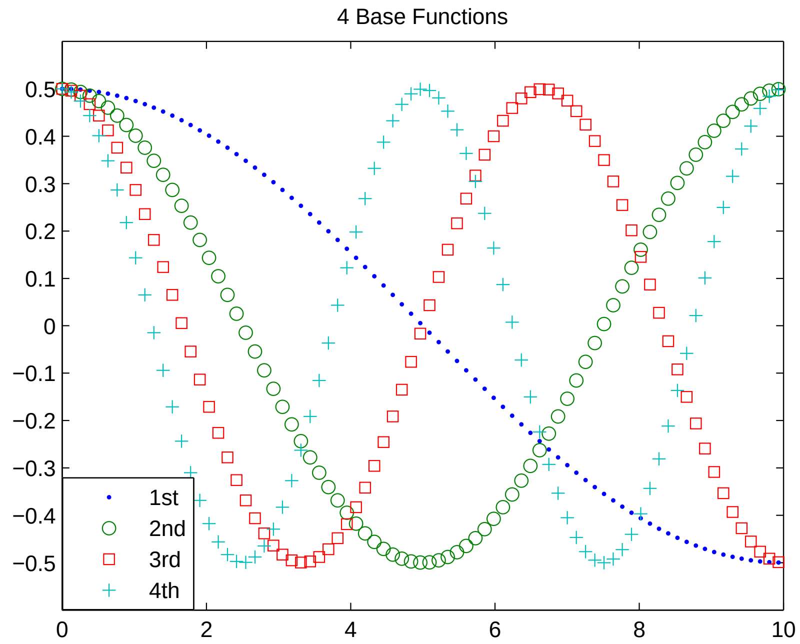

Figure 16: The four base functions used in the main test. They are cosine functions, from: 0 to π, 0 to 2π, 0 to 3π and 0 to 4π, mapped to the same frequency scale from 0 to 10. The four groups of buttons control the tilt/gain amount k for the base functions.

The first version of the Main GUI had a 5th button group, which shifted the centre frequency of the base functions on the frequency scale. This was left out due to too much complexity for the user to handle. Instead the centre frequency is set to 3k Hz for all base functions. To pin pole the centre at the right frequency a basic mapping function is used called *map.m*, more details are given in section 3.5.4.

3.5.4 Mapping

The human ears are as most sensitive to sounds of frequencies around 2k to 5k Hz (see figure 1 on page 5) and therefore it is desired to have the centre frequency in that region. To be able to shift the centre frequency to 3k Hz the base functions must be mapped to a different scale. The functions are mapped to a natural logarithmic frequency scale with a fixed centre frequency. As seen in the top plot of figure 17 on page 28 the incoming frequencies are mapped to new frequencies. The centre, 5 on the incoming x-scale, is mapped to 8 on the output scale called xm (x-mapped). To make the function continuous and with a smooth transition a 3rd degree *polyfit*-function is used as a transformation function. An example with the second base function can be seen in the bottom plot of figure 17. *Circles* illustrate the base function from the beginning and '+' mark the base function after mapping. As can be seen the centre of the function is moved from 5 to 8 on the frequency axis. 8 on the natural logarithmic scale corresponds to $e^8 \simeq 3000$Hz and 10 corresponds to $e^{10} \simeq 22050$Hz.

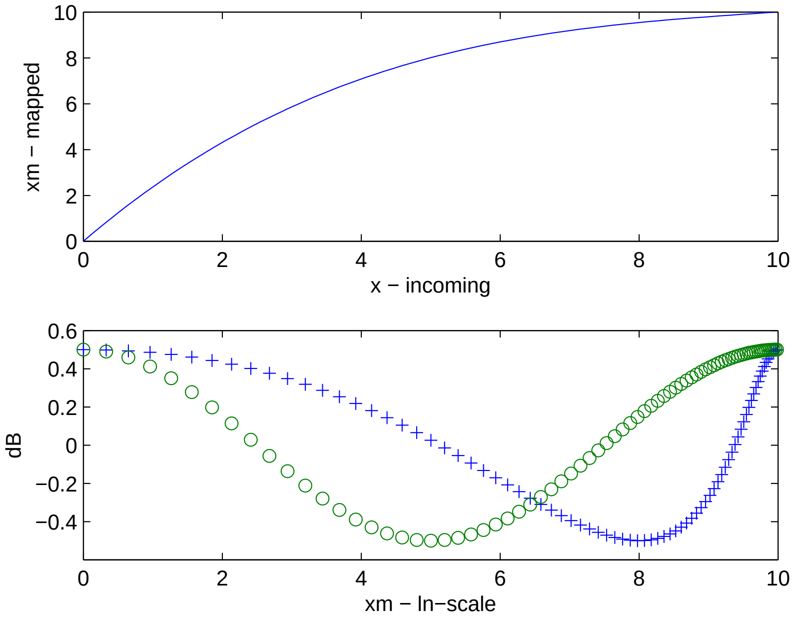

Figure 17: Illustrating the Map function for shifting the centre frequency. In the top plot the incoming values are mapped by a third degree function to new values to shift the centre frequency. In the bottom plot an example with the second base function is illustrated. The centre frequency is shifted from 5 to 8 on the natural logarithmic scale.

All contributions from the base functions and the FIG6 coefficients are added and transformed to a linear scale, 0-10 becomes 0-22050Hz. Then a normalisation of the filter curve is made to make the highest amplitude equal to one. This means that the filter will not amplify any frequencies but decrease the frequencies with amplitudes less than one in the filter-curve. The curve has a frequency span of half the sampling frequency, i.e. the last value corresponds to 22050 Hz. The filter-curve is sent to the MatLab function *FIR2* to calculate the filter coefficients. The *FIR2* function takes any continuous filter-curve with corresponding frequency bins and returns the filter coefficients needed for the *filter* function. The *filter* takes the filter coefficients and the sound vector as inputs and returns the filtered sound in a vector. Once the sound is filtered it is normalised to -25 dB RMS and saved as a wav-file. The saved file is then played in the user's headphones. The reason for saving is to save time when the user wants to listen to a sound environment without any changes of the filter-curve being made. In that case the saved file is played directly without doing the filtering process again.

When the user have found satisfying settings for the three listening environments the test continues to the evaluative part. In a pop-up window, the user can check that the right mode has been selected as the best setting for each listening situation. If not correct, the subject can choose to return to the main test window (see Appendix E).

3.5.5 Evaluative part of the study

An evaluation of the test is added to check how satisfied the participant is with the developed frequency responses. The evaluation is divided into two parts where the first part is the aural adaptive procedure called Double Elimination Tournament (see Appendix F). The second part is a questionnaire (see Appendix G) with questions about various means of user satisfaction. In the evaluative part of the test program, one non-linear and two linear prescription formulas are used: NAL-RP (section 2.4.2), POGOII (section 2.4.3) and FIG6 (section 2.4.5). The test sound files are filtered with the prescription formulas and the user's own setting for each listening situation. The filtered sound files are then the conditions for the DET; see section 2.5.1 to read about the rules for the tournament. When a winner has been determined in speech in quiet, the procedure is repeated for speech in traffic and speech in babble. After the DET the participant enters the last part of the test, which includes four questions about user satisfaction and quality of own settings. The questions can be answered on a five graded scale, e.g. from *not at all* to *very good*. The results from the questionnaire are presented in section 4.2.1.

3.6 Participants

Seventeen test subjects participated in the study, three with hearing impairment of varying severity. The subjects consisted of five females and twelve males ranging from 24 to 86 years old; see the age and sex distribution in figure 18. All participants have Swedish as mother tongue.

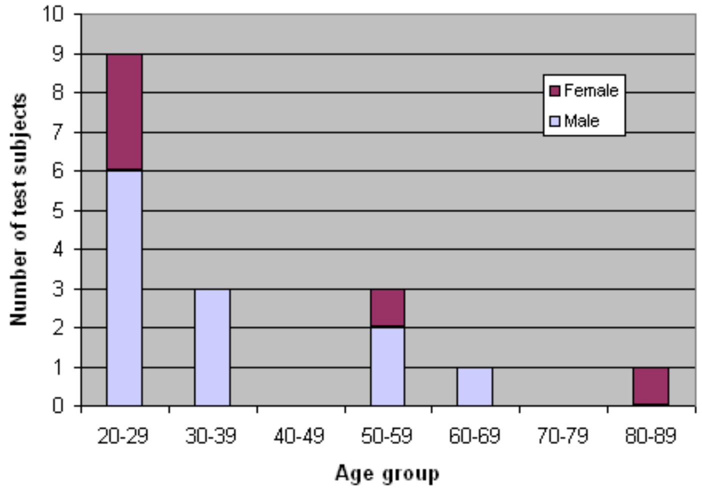

Figure 18: Age and sex distribution of the test subjects

For the hearing-impaired subjects the pure tone audiogram was measured. The hearing-impaired have impairments of varying severity and their audiogram can be seen in figure 19. The hearing-impaired are named HI-1, HI-2 and HI-3 in the figure below and in the following chapter where the results are presented. Points where audiogram data were missing were linearly interpolated with the closest known values. For HI-2 it was not possible to measure an audiogram and instead an average for Presbycusis hearing loss was used.

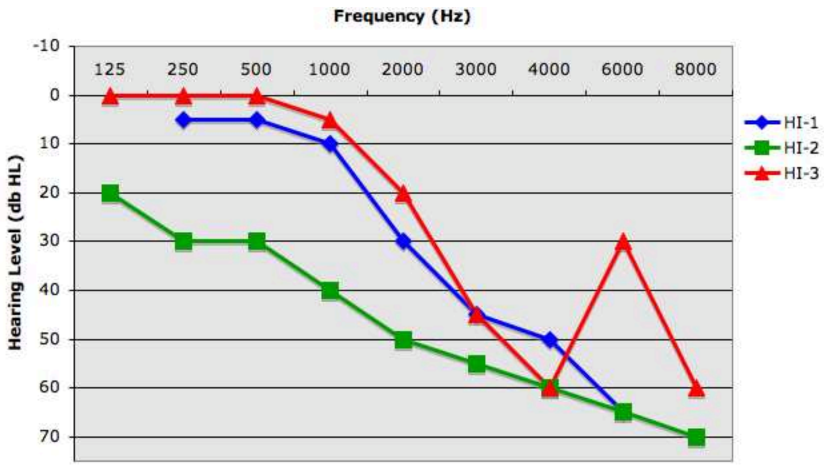

Figure 19: Pure tone audiogram for the three hearing-impaired subjects

4 Results

The results from this study are presented in two subsections: objective (4.1) and subjective (4.2) results. In the first part of the objective results, data from the normal hearing are presented. In the second part a case study of the three hearing-impaired is conducted. In the subjective results section discussions about the data and the observations are made. Some problems encountered during the study are also presented in section 4.3. A final overall discussion of the study and the results are made in chapter 5.

4.1 Objective results

The gain-frequency response curves from the normal hearing were collected. The curves from each listening situation were added and the average of the sum was calculated. The average gain-frequency responses are plotted in figure 20 on page 31. Because of the mean-value calculation, the amplitude scale is a relative dB scale and it does not correspond to any actual SPL. One can see that *Speech in Quiet* (circles) has a flat response up to 1k Hz and is then more gradually dropping. The drop in the high frequency is probably due to a wanted reduction in the high frequency noise from the original recording. In *Speech in Traffic* a definite peak at 2.5k Hz can be seen. Traffic noise in the low and high frequencies have been suppressed. For *Speech in Babble* one can see a local maximum at 3k Hz, which is within the human ears' most sensitive region to sound (see figure 1 on page 5). Smeds and Leijon [2000] state that even normal hearing may benefit from amplification in the treble frequencies and this can clearly be seen in SIT and SIB.

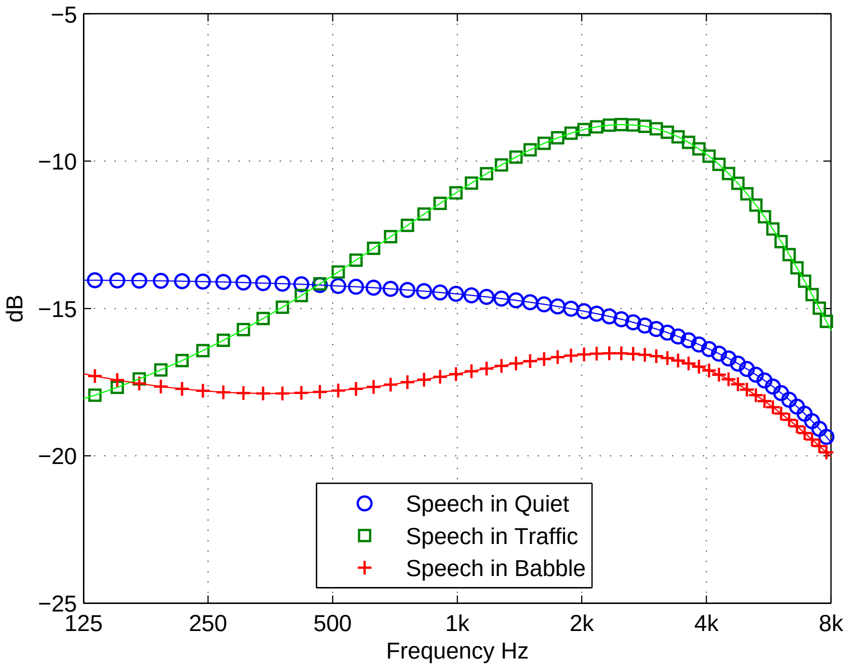

Figure 20: The average Gain-Frequency Response for the normal hearing subjects.

The standard deviation of the gain from the average frequency response of the normal hearing can be found in table 12.

(Hz))	125	250	500	1k	2k	3k	4k	6k	10k	16k
SIQ (dB)	14	16	18	18	17	16	15	13	15	27
SIT (dB)	11	11	10	8	8	8	9	10	12	15
SIB (dB)	16	18	19	20	19	18	16	15	17	19

Table 12: Standard deviation for the Gain-Frequency Response curves of the normal hearing subjects. The standard deviation indicates that there are big variations between the subjects' preferred insertion gain.

When reviewing the standard deviations for the frequency responses it is noticed that the deviations are relatively big. The main reason for this is that some participants did not manage to find a suitable setting at all. These participants had an urge to find out about the underlying techniques behind the base functions. To be able to learn about the techniques they experimented too much with the base functions and often failed to find a pleasing setting. For this reason the double elimination tournament was added to check if the participants really were satisfied with their own settings. The results from the DET were checked and the user's settings that had not finished first or second were removed from the data set. The number of settings remaining in the corrected data set was 86% in SIQ, 64% in SIT and 57% in SIB. A new average was calculated with the corrected data set and the result can be seen in figure 21.

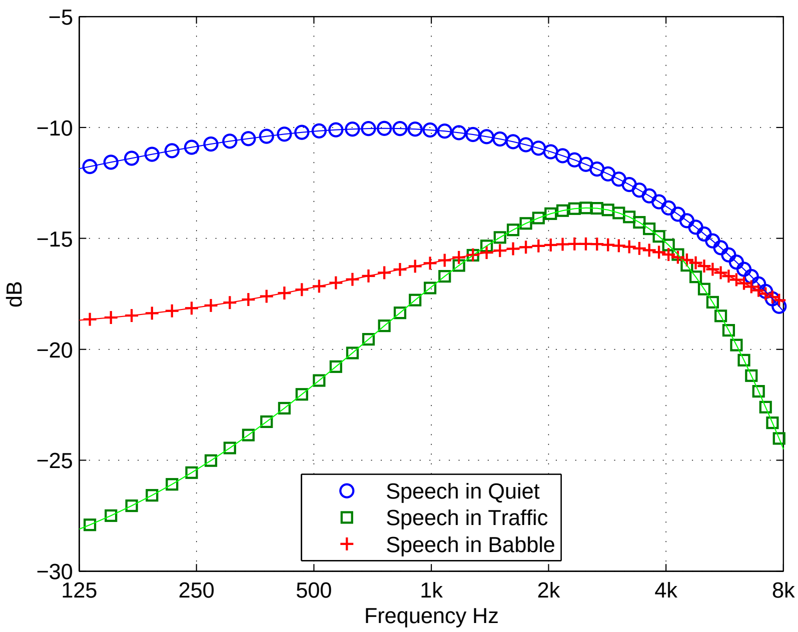

Figure 21: The average Gain-Frequency Response for the normal hearing subjects that were satisfied with their settings.

When comparing figure 20 and the corrected figure 21 the main difference is the shape of the SIQ curve. SIQ in the corrected figure shows a maximum at 700 Hz and a decrement in the low and high frequencies. All the curves for the three listening situations have a maximum but the slope to the peak is varying in steepness and the peak position is varying in frequency. For SIQ the gain difference between the maximum and minimum is 8 dB, for SIT it is 15 dB and for SIB it is 4 dB. The conclusion can be drawn that the normal hearing want fairly big variations of the sound representation for the three listening situations.

4.1.1 Results from the Double Elimination Tournament

Table 13 shows the results from the Double Elimination Tournament. The results have been separated between the normal hearing and the hearing impaired but a total has also been calculated. The table shows how many users that have placed their setting on 1st place, 2nd place or below. The sum for each listening situations on each row should add up to 100%. For example to calculate the number of users that has place their setting for SIT on 1st or 2nd place the values on the first row in the columns SIT are added: $3 + 6 = 9$, i.e. 64 % (9/14). In general the test subjects seem satisfied with their settings and only one user did not place any of the settings in 1st or 2nd place. Two subjects placed all their settings in the 1st place and five subjects placed two of their settings in the 1st place. In total 88%, 71% and 59% of the test subjects were satisfied with their settings for SIQ, SIT and SIB respectively.

	1st place			2nd place			Out		
	SIQ	SIT	SIB	SIQ	SIT	SIB	SIQ	SIT	SIB
NH	10	3	4	2	6	4	2	5	6
%	71.4	21.4	28.6	14.3	42.9	28.6	14.3	35.7	42.9
HI	2	2	1	1	1	1			1
%	66.7	66.7	33.3	33.3	33.3	33.3			33.3
Total	12	5	5	3	7	5	2	5	7
%	70.6	29.4	29.4	17.6	41.2	29.4	11.8	29.4	41.2

Table 13: Shows results from the DET for the 14 Normal Hearing (NH) and the 3 Hearing Impaired (HI). The '1st place'-column shows how many of the NH and HI that have placed their setting on first place in the DET for that particular listening situation.

4.1.2 Signal-to-Noise Ratio after average gain filtration

It was interesting to see if the signal-to-noise ratio had decreased or increased after the average gain had been applied to the sound representation. To investigate this the corrected average gain-frequency response as described above was used as a filter curve. The sound clips that was used for this calculation was the original sound files with speech and noise separated, i.e. the same files that was used for the first SNR calculations. The sound files with only speech and with only noise was filtered with the average filter curve. When the sound files been filtered they were opened in Adobe Audition 1.5 and the RMS power was calculated. From the RMS values the SNR was calculated as previously described in section 3.2.3. In table 14 the SNR values from before and after the filtration are shown. It can be seen that the SNR has increased after filtration for SIQ and SIT. For SIT the SNR has almost doubled with an increase of 81%. This means that the users have managed to increase the gain for the frequencies where most of the energy of the speech is located.

	Initial SNR, dB(flat)	Filtered SNR, dB	Difference, dB
Speech in Quiet	21.01	21.93	0.92
Speech in Traffic	1.87	3.38	1.51
Speech in Babble	-0.07	-0.04	0.02

Table 14: Signal to Noise Ratio after filtration by the users' average Gain-Frequency Response. The reader may notice that the SNR has increased for SIQ and SIT.

4.1.3 Power Spectral Density

In figure 22, 23 and 24 the Power Spectral Density (PSD) for the normal hearing are plotted. The PSD shows the distribution of signal power in the frequency domain. By definition the square root of the area under the PSD curve is the RMS of the signal. The sound files used in the test for SIQ, SIT and SIB are filtered with the user's setting for these environments. The volume level for each user has also been included to shift the curve to the right intensity level. This makes the PSD plot calibrated and it is now possible to calculate the SPL that the user had during the test. The curve is integrated over the frequency span to obtain the area. The area is a measure of the RMS level as stated above and it can now be transformed to the user's SPL during the test.

The main information that the PSD figures show is that even though all the subjects are normal hearing they want very different sound representation. For instance the difference in volume between the maximum and minimum volume setting is 33.3 dB, from 82.6 dB SPL to 49.25 dB SPL. These SPLs are the equivalent of a vacuum cleaner (70-80 dB SPL) and walking on a gravel path (50 dB SPL) respectively [Liljencrants, 2000].

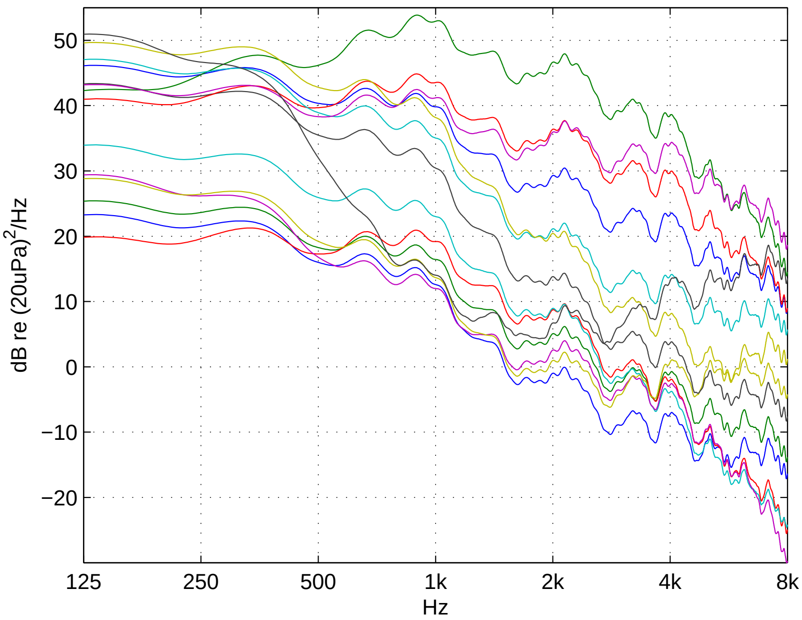

Figure 22: Power Spectral Density for Speech in *Quiet* for the preferred setting of the 14 normal hearing subjects

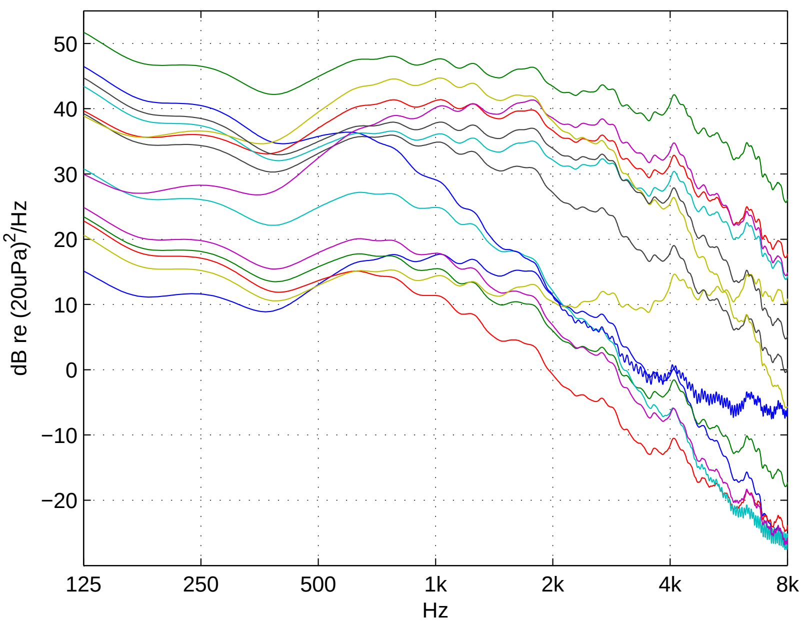

Figure 23: Power Spectral Density for Speech in *Traffic* for the preferred setting of the 14 normal hearing subjects

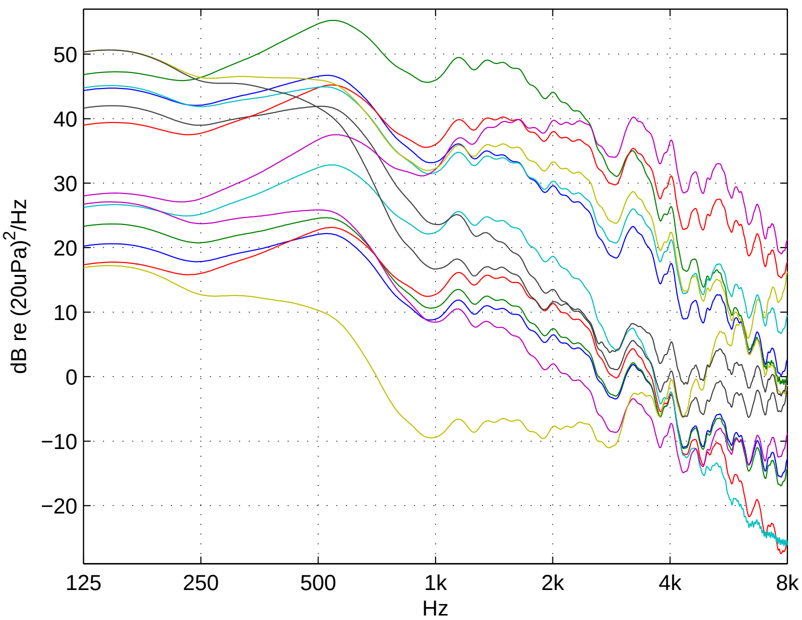

Figure 24: Power Spectral Density for Speech in *Babble* for the preferred setting of the 14 normal hearing subjects

4.1.4 Comparison between the prescription methods and the user's settings

In this section a case study of the hearing-impaired will be conducted. A comparison between the prescription methods and the hearing aid user's own settings will be made. The prescription methods are the same as in the DET (see section 2.5.1), i.e. FIG6, NAL-RP and POGO II. The hearing-impaired have impairments of varying severity and their tone audiogram can be seen in figure 19 on page 30.

In figure 25, 26 and 27 the gain-frequency response for the prescription methods and the user's own response for SIQ, SIT and SIB are displayed. The bright lines in the three bottom plots are the FIG6 response, the same as shown in upper left corner. The reason for including the FIG6 response is that when the user starts the test the FIG6 is the initial condition. This means that the difference between the two lines is the change in gain made by the user. The curves have been normalised to make the highest value 0 dB and on some graphs the 0 dB point is outside the plotted region (see bottom right plot of figure 25 on page 37 for an example). This makes the *shape* of the curves the main interest thus the average gain level is not possible to deduce.

The audiogram for hearing-impaired number one (HI-1) can be found in figure 19 on page 30. In figure 25 on page 37 the reader can see that the subject has chosen a moderately flat response for SIQ with a slow incline starting at 1k Hz and reaching the maximum at 6k Hz. The total difference in gain over this region is 10 dB to be compared with 28 dB for FIG6. For SIT and SIB the subject's gain-frequency response curve follows the shape of the FIG6-curve and it has the same break points. The gain difference for SIT and SIB is approximately the same as for FIG6. However, in SIT the gain in the bass-region is decreased as well as the peak at 6k Hz. In SIB the curve has a minimum at

1k Hz and then a steady incline in the higher frequencies. HI-1 was very satisfied with the SIQ and SIT settings hence they won the DET. For SIB the user placed FIG6 as winner and as the reader can see the FIG6 and the user setting for SIB are fairly alike.

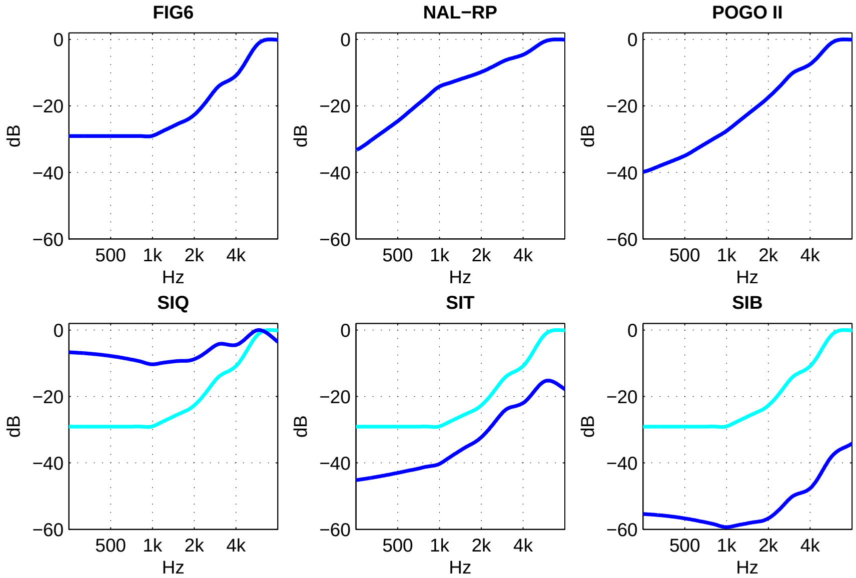

Figure 25: Gain-frequency response for Hearing-Impaired 1. The three top plots shows the prescribed gain-frequency response for FIG6, NAL-RP and POGO II. The three bottom plots shows the subject's preferred gain-frequency response for the three listening situations. The brighter line i these plots is the initial gain-frequency response from the FIG6 prescription formula.

The audiogram for hearing-impaired number two (HI-2) can be found in figure 19 on page 30. In figure 26 on page 38 the reader can see the subject's gain-frequency response in the three listening environments. The reader may notice that the curves roughly have the same shape for all three environments. The shape of the curve is in contrast to HI-1 not alike any of the prescription curves. One reason for this could be that an average audiogram for Presbycusis hearing loss was used for this subject. This indicates that the average audiogram and the subject's real audiogram might differ. HI-2 gain-frequency response has a peak at 1.5k Hz for all environments and the difference in gain is 10, 14 and 11 dB for SIQ, SIT and SIB respectively. In the DET the subject ranked her settings first or second in all listening environments. In conclusion it is interesting to note that this subject wants to have approximately the same gain-frequency response for all environments.

The audiogram for hearing-impaired number three (HI-3) found in figure 19 on page 30 indicates a noise hearing loss. The noise is located at 3k-4k Hz but the subject also has impaired hearing at 8k Hz. In figure 27 on page 38 the reader can see the gain-frequency response for HI-3. The shape of the curve is reasonably similar to the FIG6-curve for all listening environments. However in SIQ and SIB the curve has an overall declining trend, i.e. a little less gain is wanted in the high frequencies. HI-3 was very satisfied with the settings made and placed the SIQ and SIB settings as winners and the SIT setting on second place in the DET.

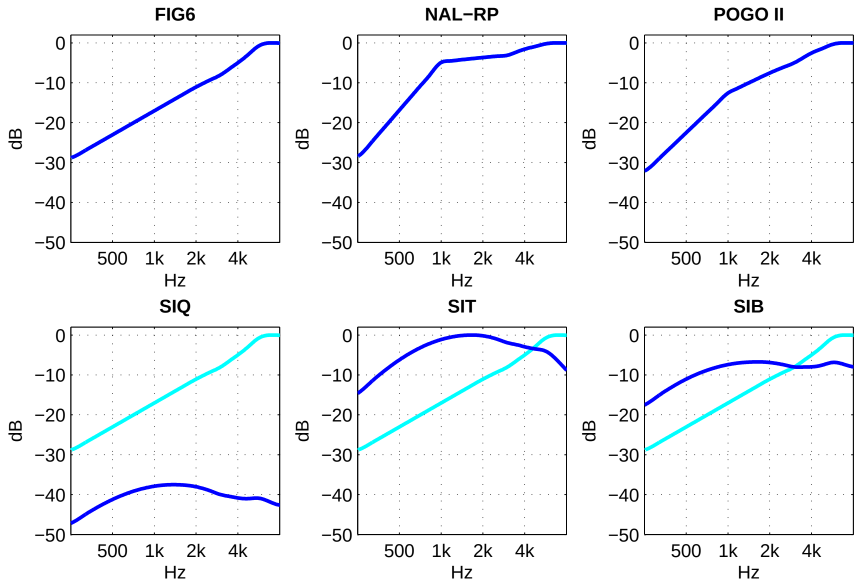

Figure 26: Gain-frequency response for Hearing-Impaired 2. The three top plots shows the prescribed gain-frequency response for FIG6, NAL-RP and POGO II. The three bottom plots shows the subject's preferred gain-frequency response for the three listening situations. The brighter line i these plots is the initial gain-frequency response from the FIG6 prescription formula.

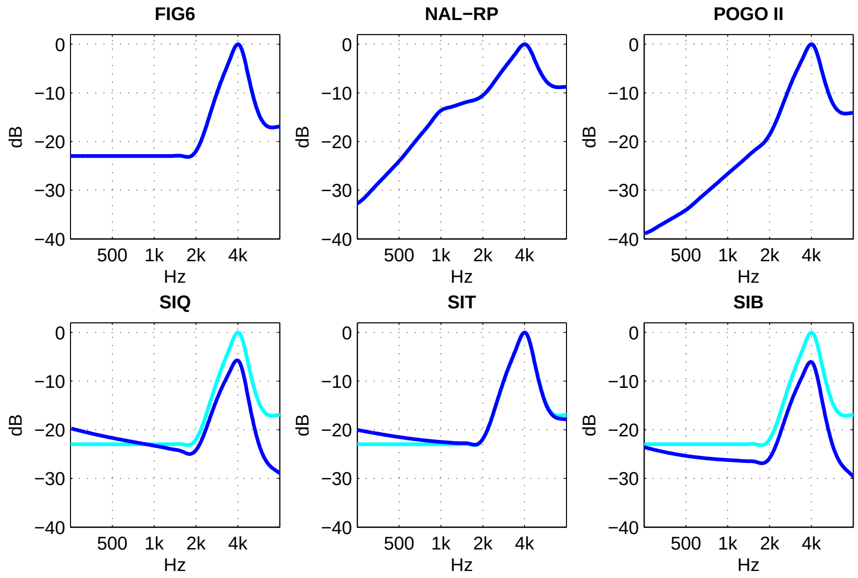

Figure 27: Gain-frequency response for Hearing-Impaired 3. The three top plots shows the prescribed gain-frequency response for FIG6, NAL-RP and POGO II. The three bottom plots shows the subject's preferred gain-frequency response for the three listening situations. The brighter line i these plots is the initial gain-frequency response from the FIG6 prescription formula.

The Power Spectral Density has also been plotted for the hearing impaired and it can be seen in figure 28, 29 and 30. The reader may notice the following: the peak at 4k Hz for HI-3 due to the Noise hearing loss, the flat response for HI-1 in SIT, the similarities of the subjects' curves in SIB from 500 to 2k Hz and the general differences between the subjects' settings for each listening situation.

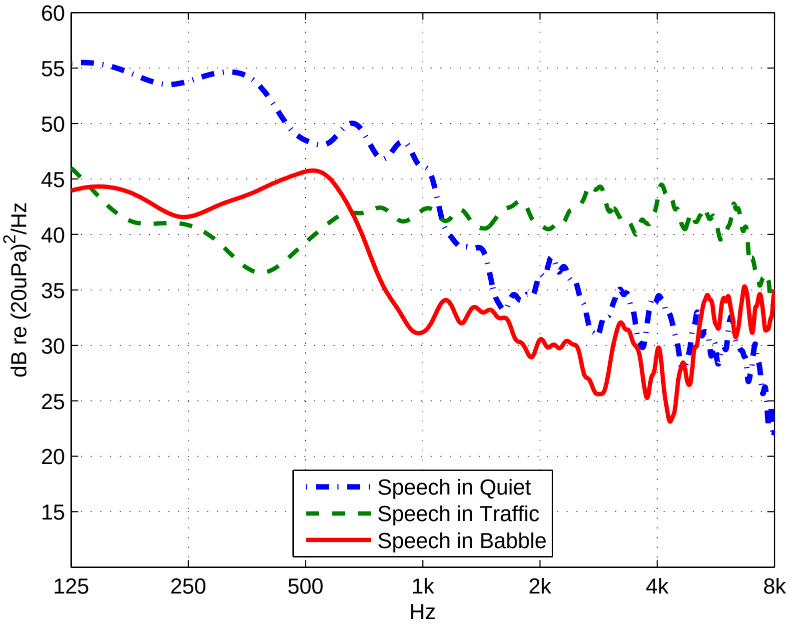

Figure 28: PSD for the preferred settings of Hearing-Impaired 1 for the three listening situations.

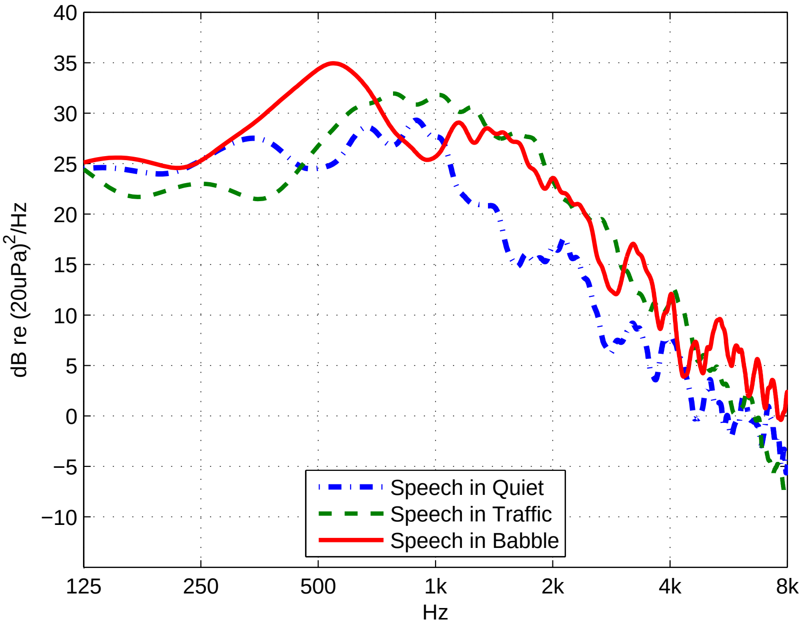

Figure 29: PSD for the preferred settings of Hearing-Impaired 2 for the three listening situations.

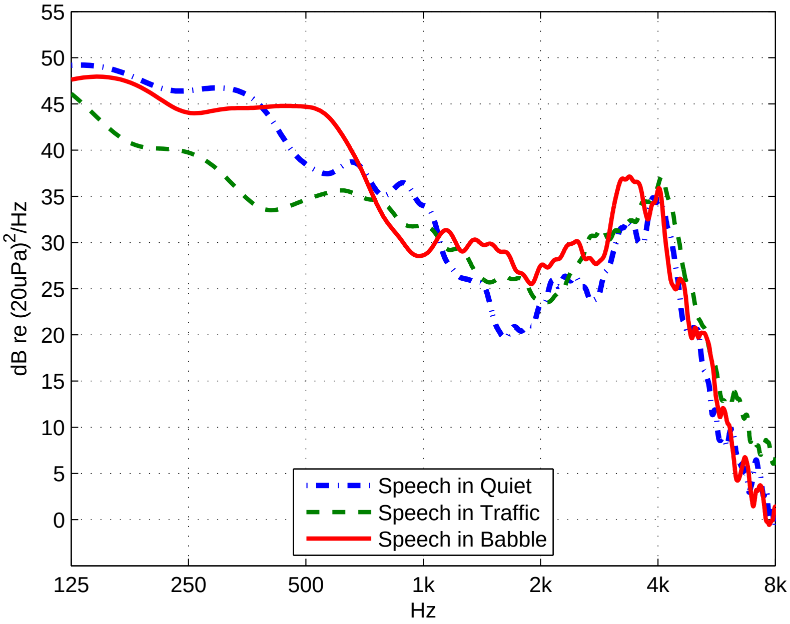

Figure 30: PSD for the preferred settings of Hearing-Impaired 3 for the three listening situations.

4.2 Subjective results

The following observations were made about the participants in this study:

- The participants do not really know what sound representation they want to have. It seems like they can prefer many different settings where each setting probably suits some of the wanted criteria, such as intelligibility or listening comfort.
- A common question among the participants was whether they should try to maximise speech intelligibility or not. They found it a bit difficult to set the sound representation without a certain target to meet. The 'most pleasant sound representation' was a too vague definition for them.
- Many participants found it quite difficult and time consuming to find the most pleasant setting. The trail-and-error progress of this program takes a lot of time and the fact that the user do not know what the controllers do can be frustrating.
- The hearing impaired appreciated the ability to set and control the sound representation.
- Some participants liked the idea behind the clock, which showed the user's settings. A round clock was a good way to illustrate positions without showing any endpoints. The hands of time also make it quite easy to remember previous settings and to compare different settings.
- The older test subjects found it in general more difficult to understand the speech and to find a comfortable frequency response. They were also not as experienced PC-users as the younger test subjects, which complicated the test.
- The participants spent most of the time in the fitting phase of the test, an estimate is 70-85% of the total time. The span in total time spent on the test between the users is from 20 minutes to 70 minutes where the average user spent 30 minutes.

4.2.1 Questionnaire

In figure 31 on page 42 the answers from the questionnaire are presented. The questions try to determine user satisfaction and the quality of their own settings. The four questions are shown as categories on the x-axis in the figure and the answers can be found in the legend. The height of each field of the bar represents the number of subjects with that answer, hence the total height is 17 (the number of participants). As one can see, 9 of the test subjects were really satisfied with their settings. However, few subjects thought that it was easy to find a suitable setting. 12 subjects said that it was a little easier to understand the speaker when the sound had been filtered with their own settings than from the initial sound. This implies that the perceived distance between speech and noise have increased, which also the SNR calculations indicate (see section 4.1.2). The final questions was added to check if the user had recognised his/hers settings in the DET. If the user was suspected of cheating[12] in the DET, the result from the DET and the answers from question 4 could be cross-examined.

[12]By cheating the author means that the subject has recognised his/hers own settings and selected them as winners even if they are not the best sounding

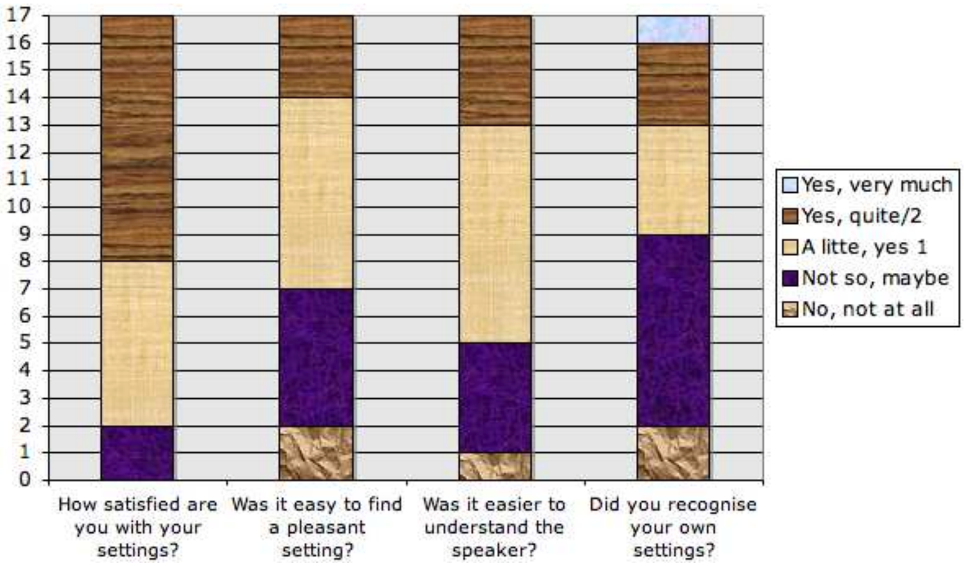

Figure 31: The answers from the questionnaire used in the test

4.2.2 Comments

One of the hearing-impaired subjects made this comment; 'With this setting I hear what they say but it will be too harsh and tiresome in the long run'. It seems like the preferred gain-frequency response will be a trade-off between speech intelligibility and listening comfort.

Another comment made by a number of users was that they would have appreciated a real-time system where the changes made are applied and noticed directly. One reason for this is that the 'sound memory' of some users is quite short and they have trouble to compare two sounds following each other and notice the differences. In general it can be difficult to compare sounds if you are not used to critical listening.

4.3 Problems

The laptop computer used from the beginning did not function as expected. When too much information was sent to the sound card Windows caused MatLab to shut down. A desktop computer replaced the laptop and because of this a calibration had to be done on both machines to make the results comparable.

5 Discussion

The main observation made from this study is that hearing aid wearers would appreciate different gain-frequency response for different listening situations. It is also clear that the test subjects' settings differ substantially from each other in both frequency response and sound pressure level. An interesting result was that the SNR had increased with the users' settings, although not a substantial amount for all the listening situations. However, it is not measured if the actual speech intelligibility has increased or not.

Another observation is that the participants in general do not really know what setting they would prefer. It seems like they can prefer many different settings. Each setting probably suits some of the wanted criteria such as intelligibility or listening comfort.

It is difficult to draw any general conclusions for the hearing impaired. For one subject the gain-frequency response coincided quite well with the ones from the prescription formulas. Another subject had very similar shape of the gain-frequency response for all listening situations. Nevertheless, due to the very big differences between all the test subjects the conclusion can once again be drawn that individual based fitting is very important. Is is not satisfactory enough to only use a standardised prescription formula based on an average human.

5.1 Limitations

When conducting this study some limitations and suppositions were made:

- The adaptation time for a hearing aid is at least 2 months according to Smeds and Leijon [2000]. Due to this adaptation time it is impossible to know if a user will be satisfied with the sound representation in the long run.

- In modern hearing aids compression is a key feature (see section 2.4.4 about non-linear HA:s). In the test program no compression is used, but it is not a big limitation because in a short time frame all hearing aids are linear. Smeds and Leijon [2000] state this fact and it is referred to in section 2.4.8 on page 14.

- The positive effects of binaural stimulation such as enhanced SNR and directivity are not considered.

- The sound files used in the test are normalised to a constant RMS-level. However, it is not certain that the perceived loudness level is the same for all the filtered sound files due to the human ears' varying sensitivity to different frequencies (see section 2.1.4). The RMS-normalisation makes on the other hand the loudness level reasonable constant for not too big changes of the gain frequency response.

- The difference between a HA and the headphones are considered small, i.e. the prescribed gain-frequency response is considered to have a correct representation.

- It is not investigated how the initial setting (FIG6) influences the final setting.

6 Conclusion

The results showed that all the normal hearing have chosen very different volume settings, hence it seems like the Most Comfortable Level varies substantially between individuals. It is expected that the normal hearing have the same hearing threshold but yet the MCL differ by 30 dB between them. This counteracts the claim that the MCL approximately bisects the area between the hearing threshold and the discomfort level (see section 2.1.5 on page 5).

In general the participants do not know what setting they would prefer and it seems like they can prefer many different settings. Further it is found that the participants have chosen different settings for each listening situation. This leads to the conclusion that it is impossible to find an ultimate setting that suits all conditions and listening situations. The final gain-frequency response will be a trade-off between speech intelligibility and listening comfort.

In conclusion it is clear that the individual differences in the preferred sound representation are very big. When considering hearing aid fitting this result implies that it is not satisfactory to only use a standardised prescription formula based on an average human. Nevertheless, the prescription formula can be used as a starting point for individual based fitting with an emphasis on user-interaction.

6.1 Outlook

- It might be a good idea to make a real time filtering program similar to the study conducted by Elbering [1999]. In a real time program it will be more difficult for the user to compare different settings but on the other hand every change on the controllers will generate a direct filtering of the output.
- The study should be conducted on more hearing-impaired test persons. Unfortunately is it quite difficult to be granted permission to test on hearing aid users.
- It would be interesting to conduct a similar test with subjectively controlled sound representation and investigate its effect on speech intelligibility.

Index

References

R.A. Bentler and M.R. Duve. Comparision of hearing aids over the 20th century. In *Ear Hear 21*, pages 625–639, 2000.

Harvey Dillon. *Hearing Aids*. Boomerang Press, Australia, 2001.

Claus Elbering. Hearing instruments - interaction with user preference. In *Auditory Models and Non-linear Hearing Instruments*, volume 18, pages 341–358. Oticon Research Centre, Danavox Symposium, 1999.

Gunnar Fant. *Acoustic Analysis and Synthesis of Speech with Applications to Swedish*. Number 1. LM Ericsson Technics, 1959.

Gunnar Fant. Speech related to pure tone audiograms. In G. Plant and K.E. Spens, editors, *Profound deafness and speech communication*, pages 299–305, London, 1995. Whurr Plub. Ltd.

GN-ReSound, 12 2005. URL `http://www.gnresound.dk/`.

Texas Instruments. Information for medical applications guide, 2004. URL `http://focus.ti.com/pdfs/vf/medeq/informationformedicalapplications.pdf`.

K-AMP. Agc - en liten guide, 1995.

Edward W. Kamen and Bonnie S. Heck. *Fundamentals of Signals and Systems using the Web and MATLAB*. Prentice-Hall, 2000.

G. Keidser. The relationship between listening conditions and alternative amplification schemes for multiple memory hearing aids. In *Ear Hearing*, volume 16, pages 575–586, 1995.

G. Keidser. Selecting different amplification for different listening conditions. In *Journal of the American Academy of Audiology*, volume 7, pages 92–104, 1996.

S. Kochkin. Marketrak vi: 10-year customer satisfaction trends in the u.s. hearing instrument market. In *Hearing Review 9*, pages 14–25, 46, 2002.

Arne Leijon. Optimization of hearing-aid gain and frequency response for cochlear hearing losses. Technical Report 189, School of Electrical and Computer Engineering, Chalmers University of Technology, Göteborg, Sweden, 1989.

Johan Liljencrants. *Elektroakustik*. Kungliga Tekniska Högskolan, 2000.

Arlene C. Neuman, Harry Levitt, Russell Mills, and Teresa Schwander. An evaluation of three adaptive hearing aid selection strategies. In *Journal of the Acoustical Society of America*, volume 82, pages 1967–1976, December 1987.

NirSoft. Nircmd, 2005. URL `http://www.nirsoft.net/utils/nircmd.html`.

Peter Nordqvist. *Sound Classificaton in Hearing Instruments*. PhD thesis, KTH, Stockholm Sweden, June 2004.

Carin Lövén Norman, Mats Nordén, Trinidad Rivera, Ingrid Adelsköld, Eva Volckerts, and Sören Birkhammar. Hörselboken, 02 2006. URL `http://www2.sit.se/orebro/horselboken/fakta/fb34.htm`.

SBU. The swedish council on technology assessment in health care, 2003. URL `http://www.sbu.se/www/Report.asp?ReportID=584`.

C. Schweitzer, M. Mortz, and N. Vaugan. Perhaps not by prescription - but by perception. In *Hearing Review Supplement*, volume 3, pages 58–62, 1999.

E. A. G. Shaw and M. M. Vaillancourt. Transformation of sound-pressure level from the free field to the eardrum presented in numerical from. In *Journal of the Acoustical Society of America*, volume 78, pages 1120–1123, September 1985.

Karolina. Smeds and Arne Leijon. *Hörapparatutprovning.* C A Tegnér AB, Bromma Sweden, 2000.

E. Zwicker and H. Fastl. *Psychoacoustics Facts and Models.* Springer, Munich Germany, 2nd updated edition, January 1999.

A Appendix - Startup screen from the test program

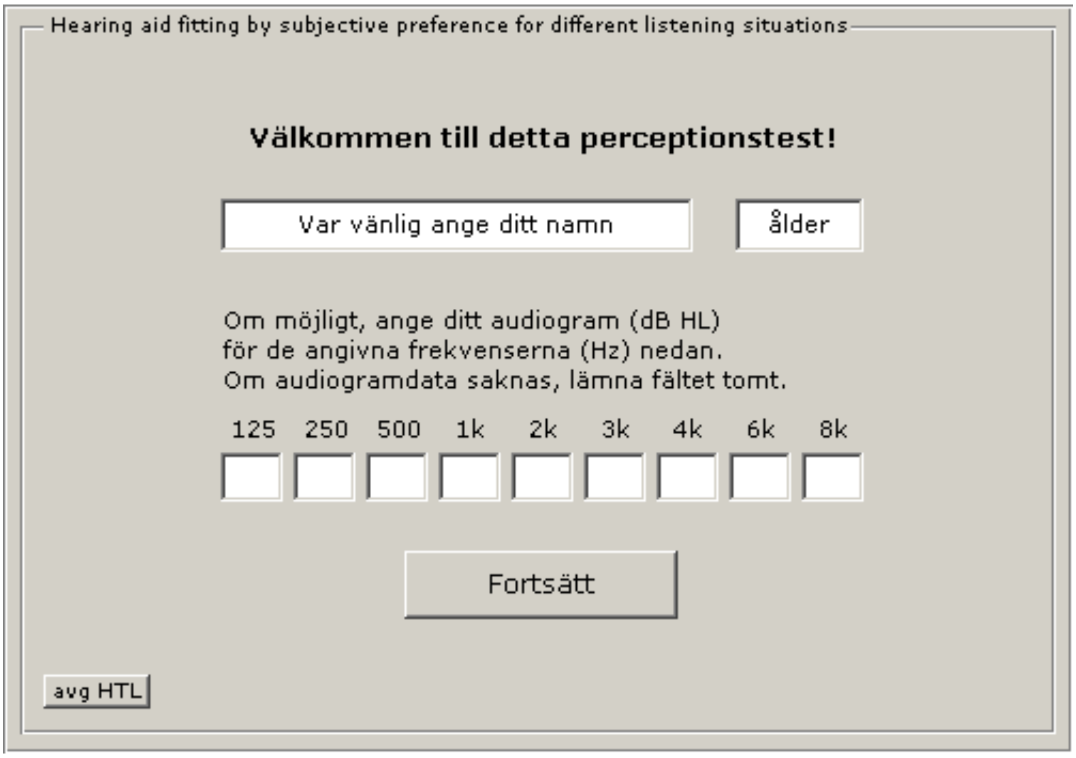

Figure 32: Startup screen where the subject enter name, age and audiogram

B Appendix - Choose ear and set the desired volume

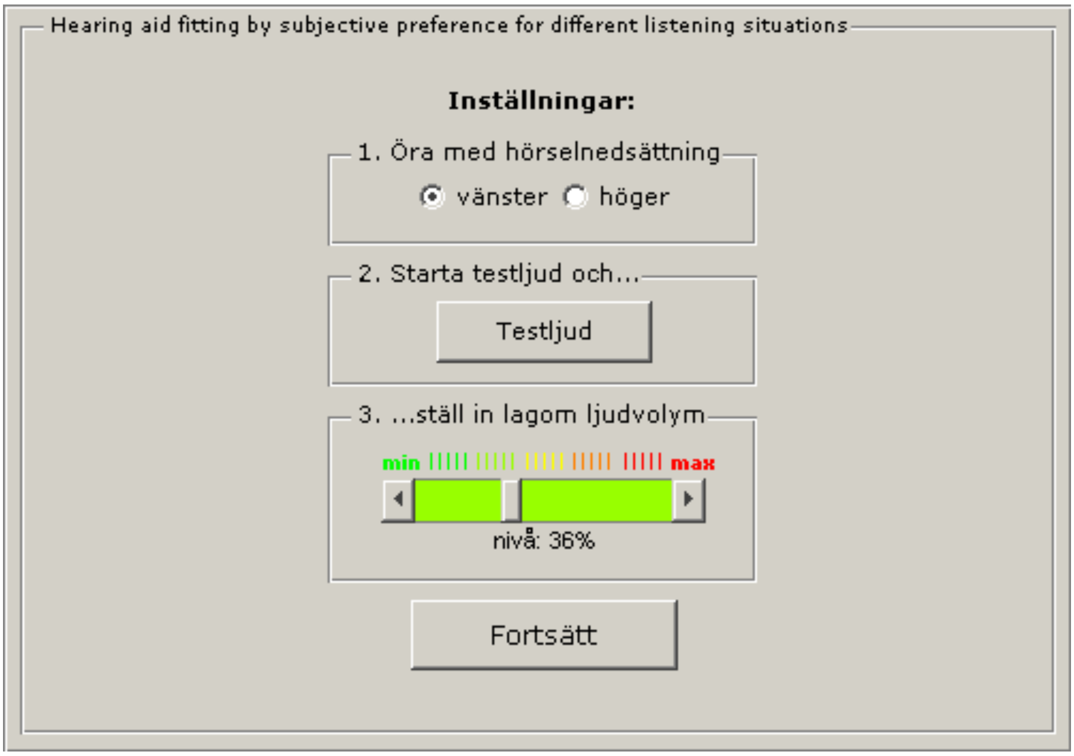

Figure 33: Screen where the subject change the volume and choose ear with hearing impairment

C Appendix - Information screen from test program

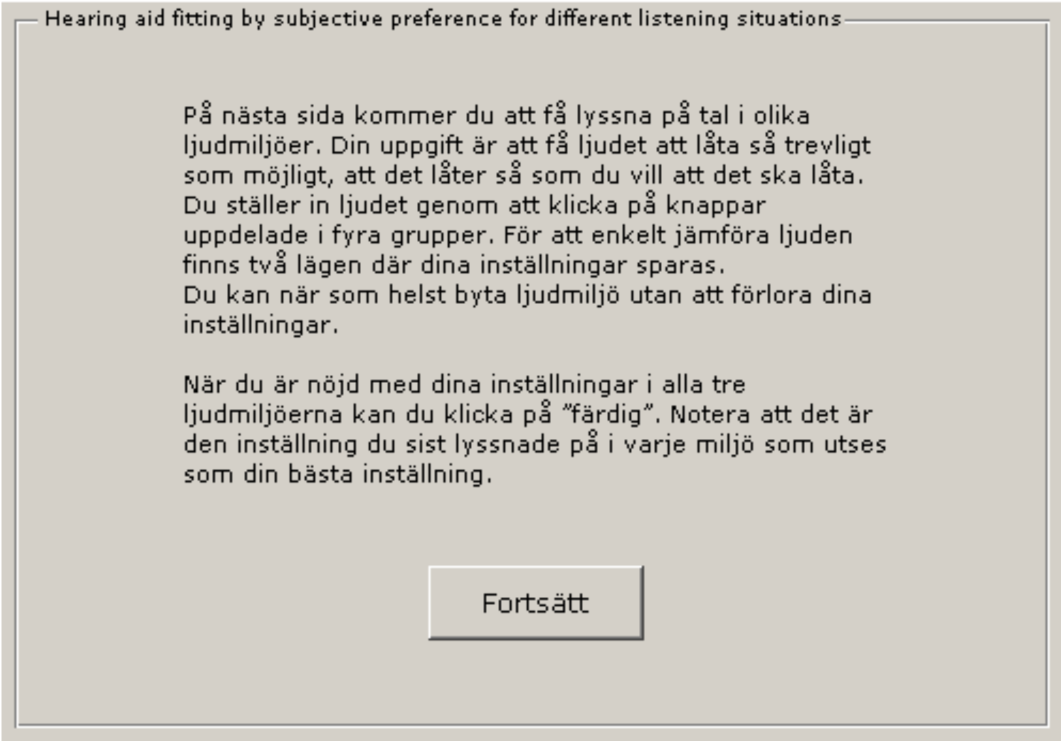

Figure 34: Information screen with instructions about the main test

D Appendix - The main GUI

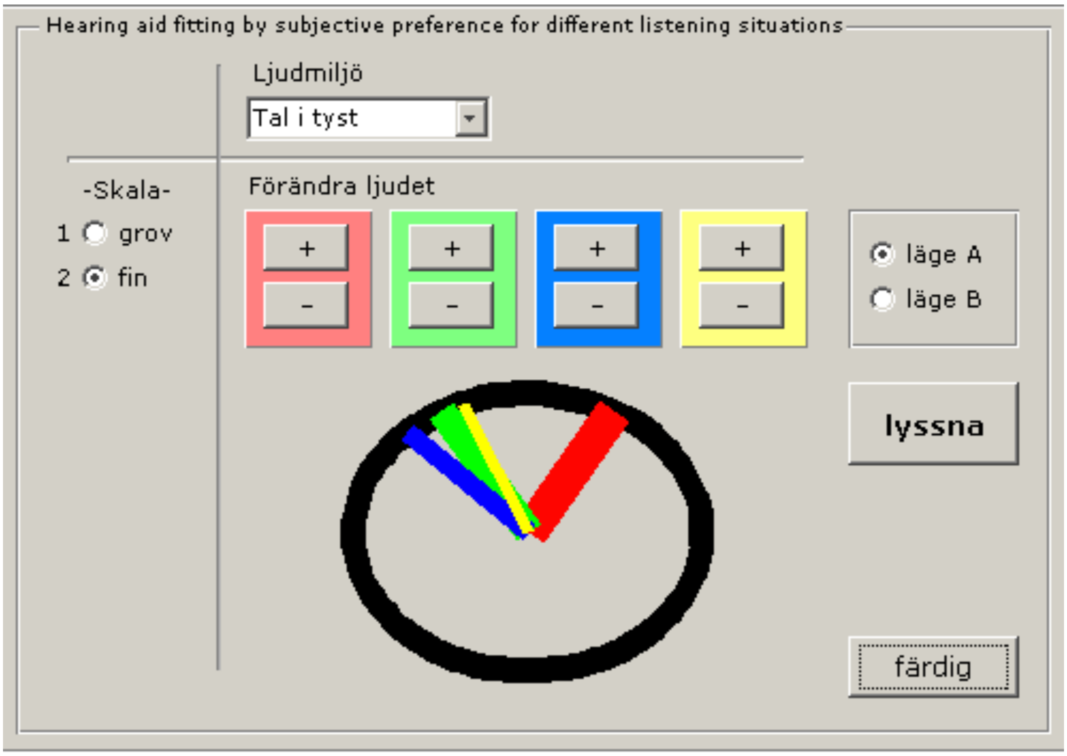

Figure 35: Main screen where the subject change the sound representation

E Appendix - Confirmation of the selected modes

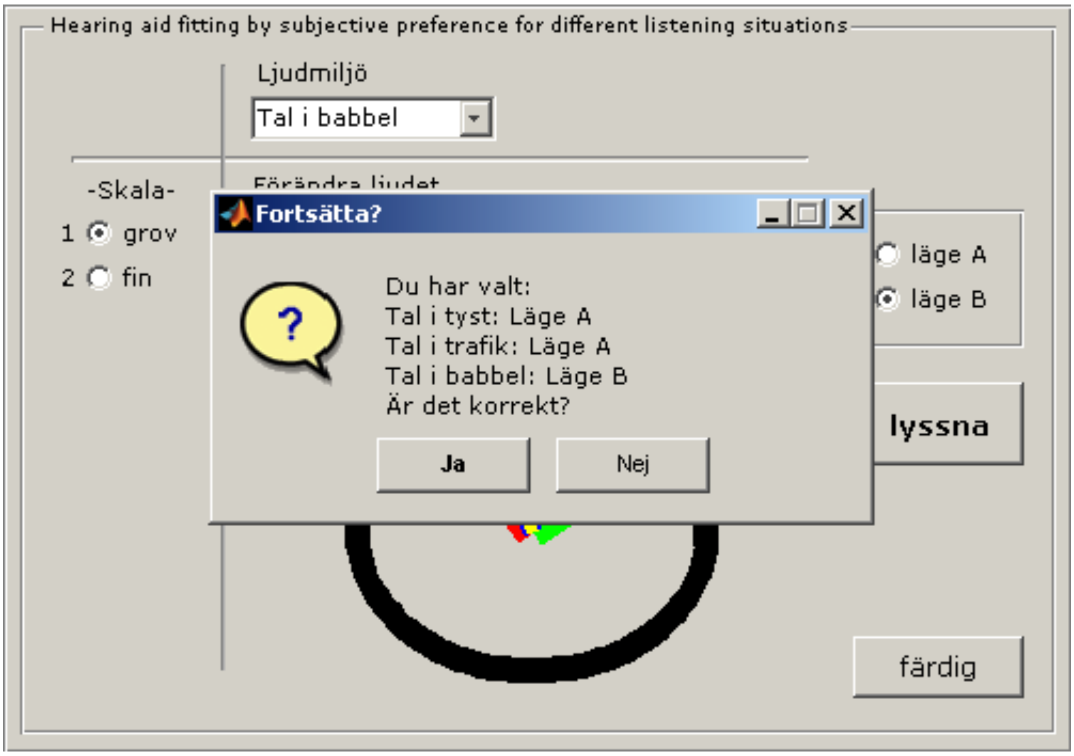

Figure 36: Screen where the subject is asked to confirm the settings made

F Appendix - Double Elimination Tournament

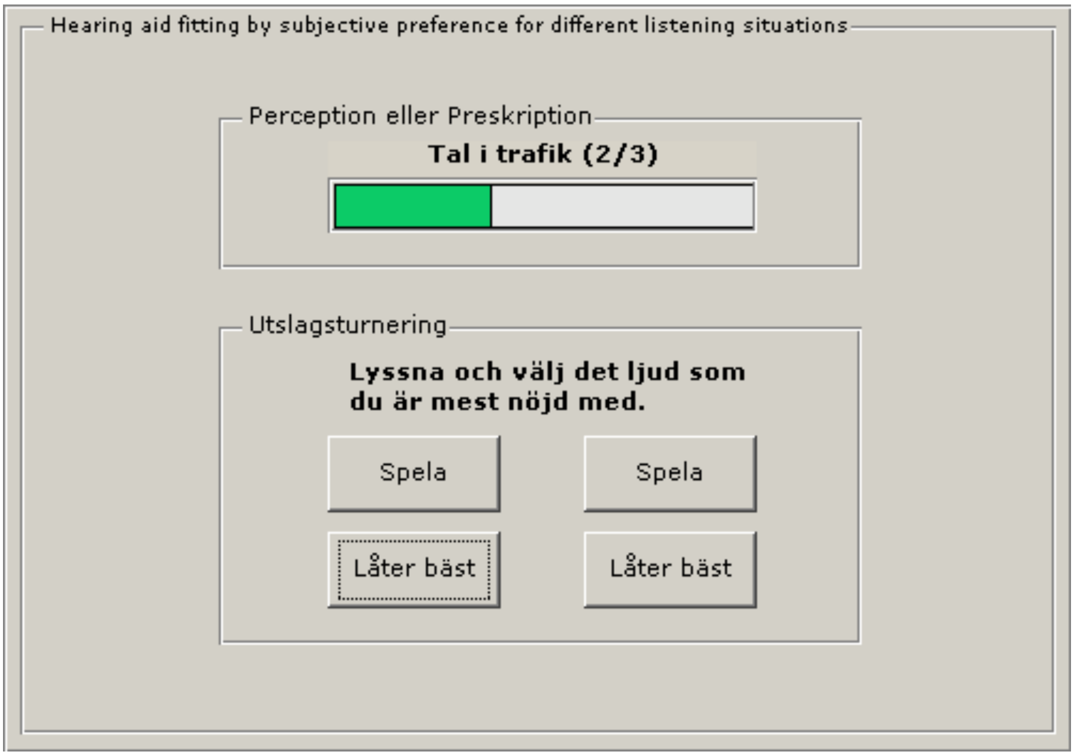

Figure 37: Screen for the evaluative tournament where the subject determines the best sound representation

G Appendix - Evaluative questions

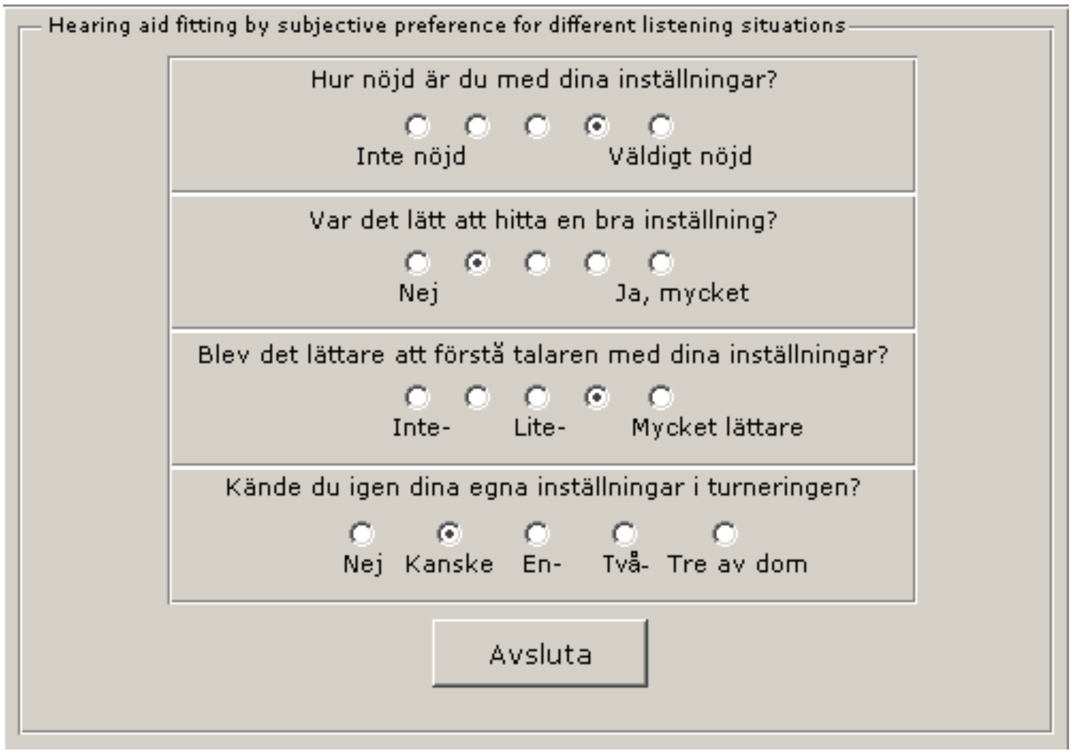

Figure 38: Screen with the questionnaire