Master Thesis

Improving the fitting of hearing aids by using prior information from community audiograms

Research Number: 592

Author:

Thomas Janssen

Supervisors:
Dr. T.M. (Tom) Heskes
Dr. P.E.M. (Paul) Ligthart
T.M.H. (Tjeerd) Dijkstra

Colophon

Author:	**Thomas Janssen**
Student Number:	**0213896**
Email:	**tjanssen@sci.ru.nl**
Research number:	**592**
Date:	**August 2008**
University:	**Radboud University Nijmegen**
Address:	**Toernooiveld 1**
	6525 ED Nijmegen
Research Institute:	**Institute for Computing & Information Sciences (ICIS)**
Section:	**Intelligent Systems**
External Company:	**GN ReSound, Algorithm R&D**
Address:	**Horsten 1**
	5612 AX Eindhoven
Computer Science Supervisor:	**Dr. T.M. (Tom) Heskes**
Email:	**t.heskes@science.ru.nl**
Management Supervisor:	**Dr. P.E.M. (Paul) Ligthart**
Email:	**p.ligthart@fm.ru.nl**
External Supervisor:	**T.M.H. (Tjeerd) Dijkstra**
Email:	**tdijkstra@gnresound.com**

Abstract

This thesis consist of two parts. In the first part we study a large database containing more than 100.000 audiograms, the hearing loss as a function of frequency. We fit a quadratic curve to each of the audiograms and study the distribution of regression coefficients. We conclude that there is no evidence for different categories of audiograms as other authors have claimed. Rather, the distribution of regression coefficients follows a multivariate normal distribution. To test the quality of the quadratic model, we used leave-one-out cross validation and show that the quadratic model does reasonably well, but that the cross-validation error is about 3 times as large as the test-retest variability. The ultimate goal of the analysis is an automated procedure for measuring an audiogram.

The second part consists of the problem of adoption of the new measuring technology by audiologists. Qualitative market research among five audiologists/hearing aid specialists has given us insight in factors which influence the adoption rate of the new measuring technique. We conclude that most audiologists are satisfied with the current staircase technique, do not think the procedure can be performed much faster and perceive audiometric equipment as expensive. The performance risk is the dominant factor in the decision process, while the financial risk is perceived differently by independent and chain store audiologists. To lower the switching costs, the new technology should be designed towards currently available software, for example a plugin. We also conclude that the biggest improvements can be achieved with young children or people with a mental handicap. This is a suitable target market for the new technology.

Acknowledgements

This thesis is part of the Management & Technology master track of the Radboud University Nijmegen and incorporates both computer science as well as management aspects. This document is the result of a year long lasting research project at the Radboud University and GN Resound, an external company specialized in hearing aids and located at Eindhoven.

There are several people who have contributed to this thesis, who I would like to thank. First of all, I would like to thank my supervisors: Tjeerd Dijkstra of GN Resound for his explanation of hearing aids, supervision, guidance and feedback on my results, Tom Heskes (Computer Science) and Paul Ligthart (Management) for their supervision and help during the project.

I would like to thank GN Resound for providing a place to work to complete my research and making me feel at home. Especially, I would like to thank Judith Verberne of GN Resound for her support with the measurement of an audiogram. Special thanks goes out to the participants in the empirical study, who have been very cooperative in helping me in my empirical research.

I would also like to thank my parents and family for their support during my study and my friends for providing moral support and some time off when I was stuck during this thesis.

And I could not have done this without you, my dear. For your undying support, thank you Corina.

Contents

1 Introduction **15**
- 1.1 Hearing impaired problems . . . 15
- 1.2 User satisfaction . . . 16
- 1.3 Research Objectives . . . 16
- 1.4 Past & current research . . . 18
  - 1.4.1 MarkeTrak . . . 18
  - 1.4.2 HearClip . . . 19
- 1.5 Structure of this thesis . . . 20

2 Audiometry **21**
- 2.1 Psychophysics . . . 21
  - 2.1.1 General psychophysical methods . . . 22
    - 2.1.1.1 Method of constant stimuli . . . 22
    - 2.1.1.2 Method of limits . . . 23
    - 2.1.1.3 Method of adjustment . . . 24
- 2.2 Methods for Manual Pure-Tone Threshold Audiometry . . . 25
  - 2.2.1 Reference equivalent threshold sound pressure level . . . 26
  - 2.2.2 Determination of the threshold . . . 27
  - 2.2.3 Adjustments to the ANSI standard . . . 28
- 2.3 Most Comfortable Loudness & Loudness Discomfort Level . . . 28
  - 2.3.1 Most Comfortable Loudness . . . 28
  - 2.3.2 Loudness Discomfort Level . . . 29
- 2.4 Audiogram . . . 30
- 2.5 Test-Retest Reliability . . . 30
  - 2.5.1 Determining the Test-Retest Reliability . . . 31
- 2.6 Improvements and automation . . . 32

3 Data Mining **33**
- 3.1 Introduction to Data Mining . . . 33
- 3.2 Preprocessing . . . 33
- 3.3 Different types of models . . . 34
  - 3.3.1 Regression . . . 34
    - 3.3.1.1 Linear regression . . . 34

3.3.1.2 Goodness of fit 36
3.3.1.3 Orthogonal Design Matrix 36
3.3.2 Cross Validation 38
3.3.3 K-nearest neighbors 39
3.3.3.1 K-Nearest Neighbors as classification 40
3.4 Improvements and automation 40

4 Dataset **43**
4.1 Introduction on Dataset 43
4.2 Datasets 43
4.2.1 Removing incorrect values 43
4.2.2 Beltone Connect 44
4.2.2.1 Preparation of the data 44
4.2.3 GN Resound Copenhagen 44
4.2.3.1 Selection Criteria 46
4.2.3.2 Preprocessing the data 46
4.2.3.3 Presentation of the data 47

5 Analysis of the datasets **49**
5.1 Pittman Criteria 49
5.2 Most Comfortable Loudness vs Loudness Discomfort Level 51
5.3 Curve Fitting 53
5.3.1 Alternative criteria 54
5.3.2 Analysis of the curve fitting coefficients 57
5.4 Analysis of personal information 59
5.5 Cross Validation 61
5.6 Summary of conclusions 63

6 Introduction to management thesis **65**
6.1 User satisfaction 65
6.2 Problem statement 67
6.3 Research plan 68

7 Market adoption **69**
7.1 Understanding the customer 69
7.2 Competition in the same market 70
7.3 Switching Costs 71
7.4 Perceived Risk 72
7.4.1 Performance Risk 72
7.4.2 Financial Risk 73
7.4.3 Time Risk 74
7.5 Adoption categories 74
7.5.1 Innovators 74

7.5.2 Early Adopters . . . 75
7.5.3 Early Majority . . . 75
7.5.4 Late Majority . . . 76
7.5.5 Laggards . . . 76
7.6 Propositions . . . 77

8 Empirical Study 79
8.1 Methodology . . . 79
8.2 Participants . . . 80
8.3 Research Topics . . . 80
8.4 Empirical Study Questions . . . 81

9 Empirical Results 83
9.1 Results of the empirical study . . . 83
9.2 Conclusions . . . 91

10 Outlook 93

A Interview Research Audiologist A 99

B Interview Independent Hearing Aid Professional B 103

C Interview Chain Hearing Specialist C 107

D Interview Chain Hearing Aid Professional D 111

E Interview Clinical Audiologist E 115

List of Figures

1.1 Kochkin active vs. non-owners 18
1.2 Kochkin user satisfaction 19

2.1 Typical Psychometric function 22
2.2 Example of method of limits 23
2.3 Example of staircase method 24
2.4 Reference equivalent threshold sound pressure levels for supra-aural earphones 26
2.5 Example of two audiograms 30

3.1 Examples of K-Nearest Neighbors 40

4.1 Age histogram of the Copenhagen dataset and Kochkin data 48

5.1 Categories of Pittman Criteria in percentages 50
5.2 Average audiograms of Pitmann Criteria 52
5.3 MCL vs LDL Beltone 53
5.4 r^2 values of curve fitting Beltone Pittman criteria 54
5.5 Categories of alternative criteria in percentages 55
5.6 Average of alternative criteria 56
5.7 r^2 values of curve fitting Beltone alternative criteria 57
5.8 Regression coefficient matrix Beltone 58
5.9 Regression coefficient matrix Copenhagen 60
5.10 Cross validation Beltone 61
5.11 Cross validation Copenhagen 62

6.1 Kochkin active vs. non-owners 66
6.2 Kochkin Satisfaction instruments 67

7.1 Understanding customers 70
7.2 Five Different categories of adopters 75

List of Tables

2.1 Positive & negative aspects psychophysical methods 25

4.1 Summary Beltone Connect dataset 45
4.2 Simple statistics Beltone Connect dataset 45
4.3 Summary GN Resound Copenhagen dataset 47
4.4 Simple statistics GN Resound Copenhagen dataset 47

5.1 Alternative criteria 55
5.2 Statistics coefficients by age and gender 59

7.1 The seven Perceived risk facets 73

9.1 List of relevant quotes 86
9.2 Short list of interviews-factors 87
9.3 General thoughts based on different topics 89

Chapter 1

Introduction

Today as well as in the past, many people suffer from hearing loss. A hearing impaired person cannot hear soft sounds and needs amplification of the sound in order to restore audibility. The amplification is provided by a hearing aid. This hearing aid needs to be fitted for the particular hearing loss of the patient. A standard part of every fitting procedure is the measurement of hearing loss as a function of sound frequency, the so-called audiogram.

This thesis consist of two parts. An automated and improved procedure for the measurement of an audiogram is the ultimate goal of the first part of this thesis. The second part consists of the problem of adoption of the new technology by audiologists. Health care is known for its somewhat conservative attitude towards adoption of new technology. This thesis will introduce factors which are involved in the adoption of a new technology. A qualitative market research study will be used to find which factors are involved with the adoption of the automated and improved fitting technology.

1.1 Hearing impaired problems

Hearing aids are designed and fitted to reduce the problems faced by hearing impaired. Hearing loss involves the loss of hearing ability, which can be described by decreased audibility, decreased dynamic range and signal to noise ratio loss [11].

Decreased audibility Hearing impaired people do not hear soft sounds. People with severe hearing loss may not hear any speech sounds, people with a mild loss are likely to hear some sounds, but not others. To overcome this problem, a hearing aid has to provide more amplification for frequencies where hearing loss is the greatest.

Decreased dynamic range The dynamic range is defined as the difference between discomfort level and the threshold of hearing. Soft sounds can be made audible by amplifying them. But it is not appropriate to just amplify everything. Hearing loss increases the threshold of hearing much more than it increases the level of loudness discomfort. Therefore, the dynamic range of hearing impaired is smaller than that of a normal hearing person. A hearing aid controls this problem by using an amplifier which automatically turns itself down as the sound gets stronger or turns itself up when the sound gets weaker, thus controlling the sound level to fit within the dynamic range. This is called compression.

Signal to noise ratio loss A third problem is the loss of understanding speech in noise, known as signal to noise ratio loss (SNR-loss). When sounds of different frequencies enter the ear, separate signals are send to the brain. The brain can then partly ignore sounds that are considered noise, thus focussing on the important sounds like speech. A person with SNR-loss has lost some of this ability, making it more difficult to distinguish between the important sounds and background noise, affecting speech recognition. The lower the ratio, the more obtrusive the background noise is and the more difficult it is to recognize speech. Hearing aids today are not able to fully control this problem, but try to remove noise from the incoming sounds or emphasize sounds coming from a certain direction.

1.2 User satisfaction

Within the Netherlands as well as the United States, about 10 to 12 percent of the population has a hearing loss that affects his or her ability to understand normal speech. This percentage will further grow because of ageing effects. Roughly one out of five persons with hearing loss, i.e., 2 to 3 percent of the population owns a hearing aid. However, despite rapid advancements in technology, user satisfaction leaves much room for improvement: it has been estimated that one out of four buyers is very dissatisfied with the sound quality of his or her hearing aid [17]. An important factor is that the sophisticated algorithms running on today's hearing aids have not been tuned optimally to the impairment and preferences of the user.

The current problems are mostly related to user satisfaction. Over the past decade, signal processing in hearing aids has progressed from simple analog circuitry to sophisticated digital audio algorithms. Despite technical advances, satisfaction with hearing aids leaves much room for improvement [16, 17]. In the United States alone, there are about 30 million hearing impaired persons, while market penetration is only 6 million buyers. Of those 6 million, 1 out of 4 buyers is very dissatisfied with the sound quality of his hearing aid [17].

1.3 Research Objectives

The time spent on each patient for which a hearing aid needs to be fitted, is a problem audiologists and patients encounter. It is found to be quite time consuming to measure the hearing loss (accurately), because each frequency is measured independently. This could lead to loss of concentration by the patient, possibly resulting in inaccurate measurements. By reducing the time audiologists spend on the measurements, they can spend more time on counseling, an important aspect of the fitting procedure [17]. This could lead to preventing loss of concentration by the patient, possibly leading to more accurate results. Therefore, both the audiologist as the patient could benefit from an improved fitting procedure.

To try and overcome this problem, this thesis will focus on two aspects. First, it will focus on improving the fitting of hearing aids, by using community audiograms, i.e. a large set of audiograms from other patients. From these audiograms information can be “borrowed”. A bayesian approach, a commonly used technique in machine learning and statistics, is based on using prior information. Therefore, a prior needs to be found. This thesis will try to find this prior information from a large set of audiograms by fitting a quadratic curve to each individual audiogram and analyzing the three regression coefficients from each fit. The subquestions within this part of the thesis are:

1. What relationship does there exist between the three coefficients of different individuals?
2. Are the regression coefficients different for patients of different gender or age?
3. How well can individual frequencies be predicted based on a quadratic fit of hearing losses on other frequencies?

There are many other measurements involved in the fitting of hearing aids, like SNR-loss, Most Comfortable Loudness and Loudness Discomfort Level measurements. These measurements could also benefit from prior information, but they are not within the scope of this thesis and will therefore not be discussed.

The second aspect will be the marketing of a software tool which incorporates the new algorithm. The automated and improved technology will be developed to assist audiologists in measuring the hearing loss. The main question is: what problems does this new technology have when trying to conquer a market position? This problem can be divided in three subquestions:

1. What type of adopters are audiologists?
2. What factors influence their decision to use the new technology?
3. Is there a difference between audiologists of a large chain of stores and independent audiologists?

Health care is known for its somewhat conservative attitude towards new technologies [15], usually leading to a slower adoption rate [7]. The second part of this thesis will investigate some of the characteristics of the target market. Empirical study will be used to find which

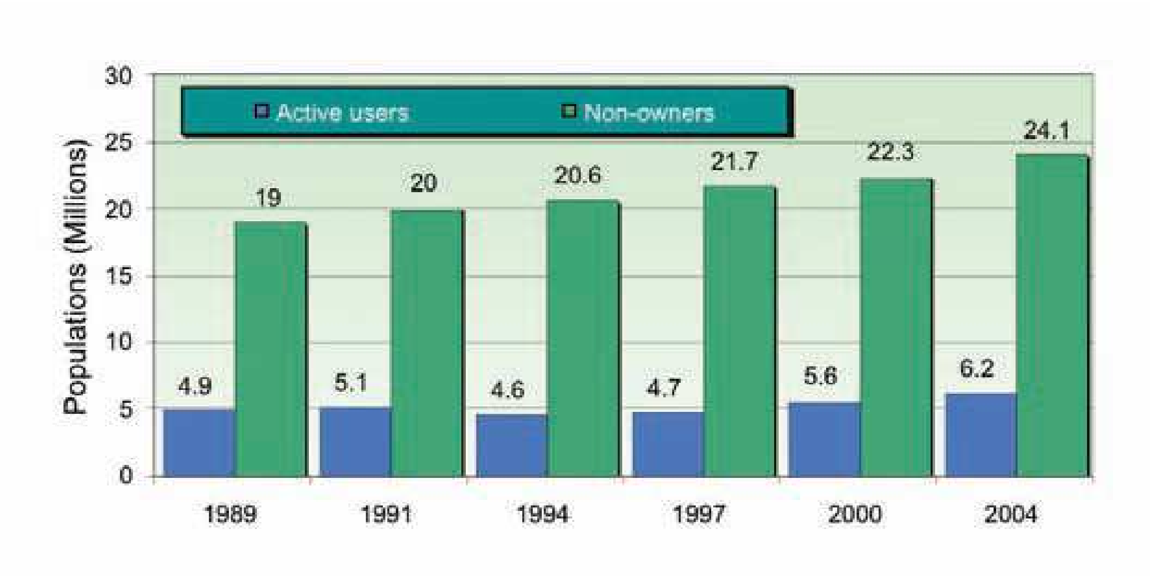

Figure 1.1: Kochkin active vs. non-owners. An active user is defined as a hearing instrument owner who uses their hearing instruments even if only once a year [17].

factors are important with the adoption of the automated and improved fitting technology. It is important to note that only the influence of the new algorithm will be investigated. The actual success of adoption will include more factors, like the usability and customer support of the tool. These factors are not within the scope of this thesis.

1.4 Past & current research

In the past, there has been much research based on improving different aspects of hearing aids. This section covers briefly some of this research.

1.4.1 MarkeTrak

For years, Sergei Kochkin (PhD, executive director of the Better Hearing Institute, Alexandria, Va.) reviews the current state of hearing aids in the US and publishes the results in MarkeTrak. In his publications, he covers significant trends in the hearing loss population, along with user (dis)satisfaction. Kochkin's latest MarkeTrak (MarkeTrak VII) [17] shows that a large number of hearing impaired do not use a hearing aid (see figure 1.1). Several factors contribute to this, with user satisfaction being an important factor. Of the population which does use hearing aids, user satisfaction leaves much room for improvement [16] (see figure 1.2).

Table 3. Satisfaction with Hearing Instruments

	1984	1989	1991	1994	1997
Satisfaction with Hearing Instruments		(n=1,632)	(n=2,323)	(n=2327)	(n=2,720)
Total owner population					
% Satisfied		58.5%	57.8%	53.4%	54.5%
% Neutral		22.6%	21.7%	26.6%	26.1%
% Dissatisfied		19.0%	20.3%	20.0%	19.4%
% hearing instruments in drawer (not used)	13.5%		12.0%	17.9%	16.2%
New hearing instruments (< 1 year)					
% Satisfied			66.3%	70.7%	63.1%
% Neutral			21.8%	22.7%	26.7%
% Dissatisfied			11.9%	6.6%	10.2%
% hearing instruments in drawer (not used)			3.0%	3.5%	4.6%
New hearing instruments (< 4 years)					
% Satisfied			60.7%	58.7%	59.3%
% Neutral			21.6%	25.0%	26.3%
% Dissatisfied			17.8%	17.3%	14.9%
% hearing instruments in drawer (not used)			7.7%	11.1%	8.8%

Figure 1.2: Kochkin user satisfaction with hearing instruments which are 4 years old or less [16].

1.4.2 HearClip

Currently, a research project called "HearClip" is performed by three departments at three universities: the department of Information and Knowledge Systems at the Radboud University of Nijmegen, the Signal Processing Systems group of the Department of Electrical Engineering at the TU Eindhoven, and the Clinical and Experimental Audiology department at the Academic Medical Centre (AMC) in Amsterdam. The goal of the project is to improve the hearing aid fitting procedure over the hundreds of parameters, by using Bayesian incremental utility elicitation [14]. It would be easy if it is possible to know what determines user satisfaction. Since it is not known, they model the uncertainty about user satisfaction through probability distributions. By asking questions of the type "which of the two audio samples sounds best?", they elicit responses from a hearing aid user. Combining a probabilistic response model with the probabilistic user satisfaction model in a Bayesian fashion, they take all responses into account during a fitting session.

1.5 Structure of this thesis

This thesis will be structured in the following way: Chapter 1 will be this introduction to the thesis, where the two research objectives of the thesis are presented. Chapter 2 will introduce the notion of audiometry and methods to measure the hearing loss. Chapter 3 will continue with the introduction on Data Mining and introduce some techniques which are used to improve the fitting of hearing aids. In chapter 4 the used datasets will be introduced to the reader. Chapter 5 covers the analysis of the datasets. The market adoption associated with the introduction of a tool based on this new algorithm will be covered in chapter 7. An empirical study will be performed to find out which factors are involved in the adoption process of a new measuring technology. An outlook will be presented in chapter 10, where topics for further research will be presented.

Chapter 2

Audiometry

Audiometry is the measurement of the hearing ability. Audiometric tests typically seek the hearing thresholds of a patient, but may also test the ability to distinguish speech from background noise or discriminate between sound intensities. The measurement of the hearing loss is the main task for fitting a hearing aid. This is done by presenting a pure-tone sound of some frequency to the patient and ask him if it is perceived. Because the natural behavior of perception has some fluctuation, it is possible that sometimes the response is positive and other times the response on the same tone will be negative. The threshold can therefore not be defined as the level below which detection never occurs and above which detection always occurs. The threshold is defined as the level where 50% of the trials is perceived (see figure 2.1).

This chapter will introduce the measurement procedure of the data on which this thesis is based. First an introduction into psychophysics will be presented, which incorporate three common methods to measure the hearing loss, in order to present some background information. Then a standard will be introduced to measure the hearing threshold, followed by an explanation of two other audiometric measurements. The last section will introduce the notion of an audiogram.

2.1 Psychophysics

In 1860, the German physicist G. T. Fechner introduced techniques for measuring sensations. These techniques were termed "psychophysics" [13]. Psychophysics refers to the notion that mental events (psycho) can be measured (physics). Within this stands the concept of the threshold. It is defined as the smallest amount of stimulus energy necessary to produce a sensation. A problem with the measurement of the threshold, is that the sensitivity to a stimulus tends to fluctuate from moment to moment. Therefore several measurements are averaged to get an accurate estimation of the threshold.

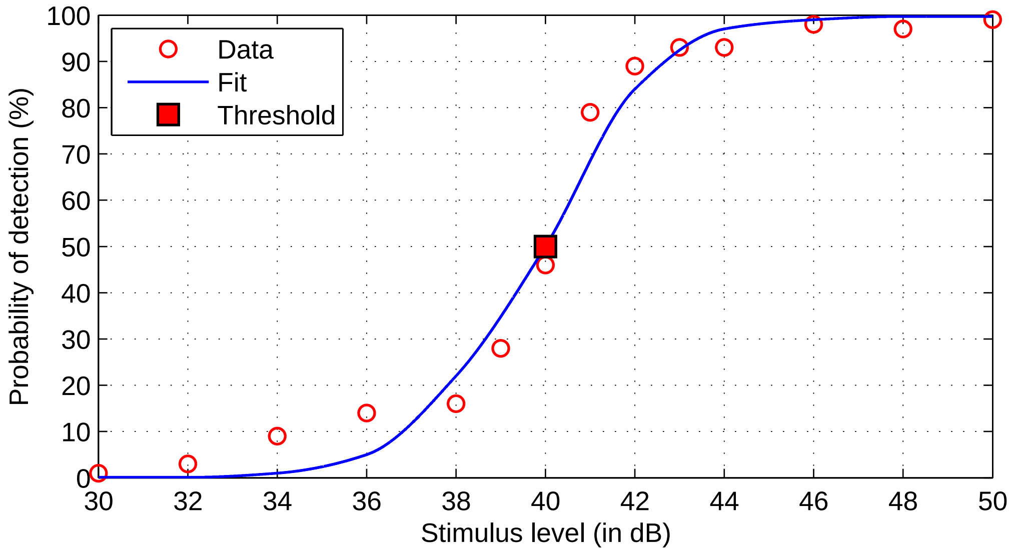

Figure 2.1: Typical Psychometric function. Threshold is at 40 dB (50% of detection).

2.1.1 General psychophysical methods

Fechner developed three methods to obtain thresholds [13]: the methods of constant stimuli, of limits and of adjustment. These methods are still in use today, although most have been modernized. In the following sections, the measuring of the thresholds will be presented. Each method will be illustrated with an example from pure-tone audiometry, where the task is to find the hearing threshold for a pure-tone of fixed frequency. For purposes of illustration, the hearing threshold will be 40 dB.

2.1.1.1 Method of constant stimuli

To measure the threshold, the method of constant stimuli repeatedly presents a tone throughout the experiment. On each presentation the patient is asked whether he heard the tone. Usually there are between five and nine intensity levels used in random order, where the lower end should be a tone level that can almost never be perceived and the upper end should be detected almost every time. Increasing the intensity level of the tone should therefore increase the probability of perception. The 50% threshold is located somewhere in the range of the measured stimulus levels. The responses for each intensity level are kept and the percentage of yes answers are plotted in a graph called a psychometric function. A fit to the data follows a particular S-shape called an ogive. An example is presented in figure 2.1, where we have plotted data for 14 intensity levels, more than the usual five to nine intensity levels to illustrate the method. A red circle defines the percentage of perceived to the total amount of presented tones on the specific stimulus value. For example, at 34 dB, the percentage of perceived intensities was measured at 9%. The blue line is the fit to the measured percentages. The red square defines the 50% threshold which is located on

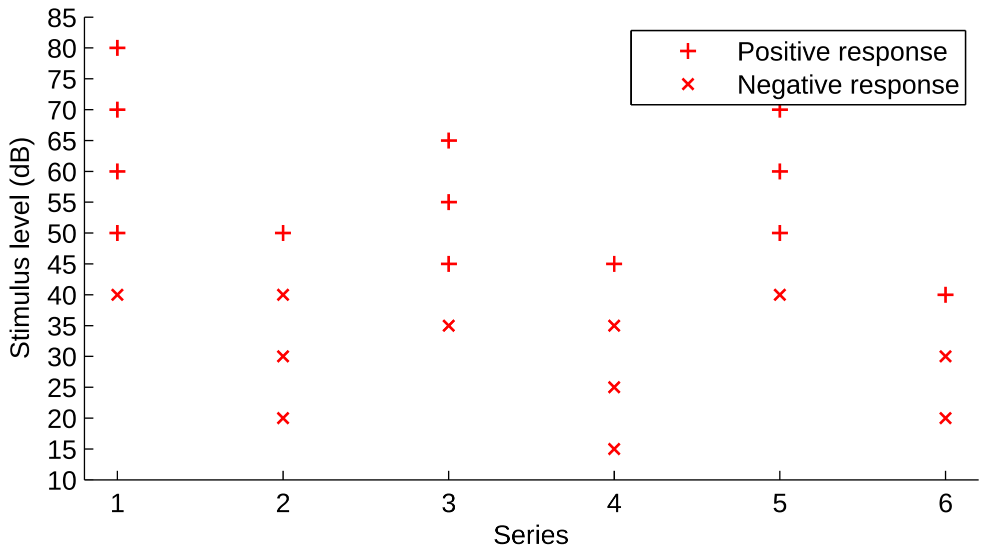

Figure 2.2: Example of method of limits with a combination of ascending and descending series. Threshold at 40 dB.

the fitted line, and is taken as the threshold. In the presented figure, the threshold is 40 dB.

High accuracy is an advantage of this method, because there is a little chance of perceptual learning by the subject by using the random order. A disadvantage is that it is very time consuming, because it relies on the measurement of many intensity levels.

2.1.1.2 Method of limits

Another measuring technique is the method of limits. The experimenter begins by presenting a tone well above or well below the threshold. On each successive presentation, the threshold is approached by changing the tone intensity by an amount, usually 10 dB (see section 2.2 for more details), in either an ascending or descending series, until the patient changes his response. In the case of an ascending series, the experimenter increases the tone intensity by an amount, until the patient reports the presence of the tone, called the transition point. When this happens, the series is terminated. In the case of a descending series, the intensity is decreased until the patient reports the disappearance of the tone. The threshold is then determined as the average of transition points over several series. Figure 2.2 presents an example of the use of the method of limits, combining ascending and descending series. It starts with a descending series at 80 dB. The patient responds positive until 40 dB is reached. This is the transition point. Then an ascending series is started at 15 dB, where the patient responds negative until 50 dB is reached. The average over 6 series is 40 dB, which is taken as the threshold.

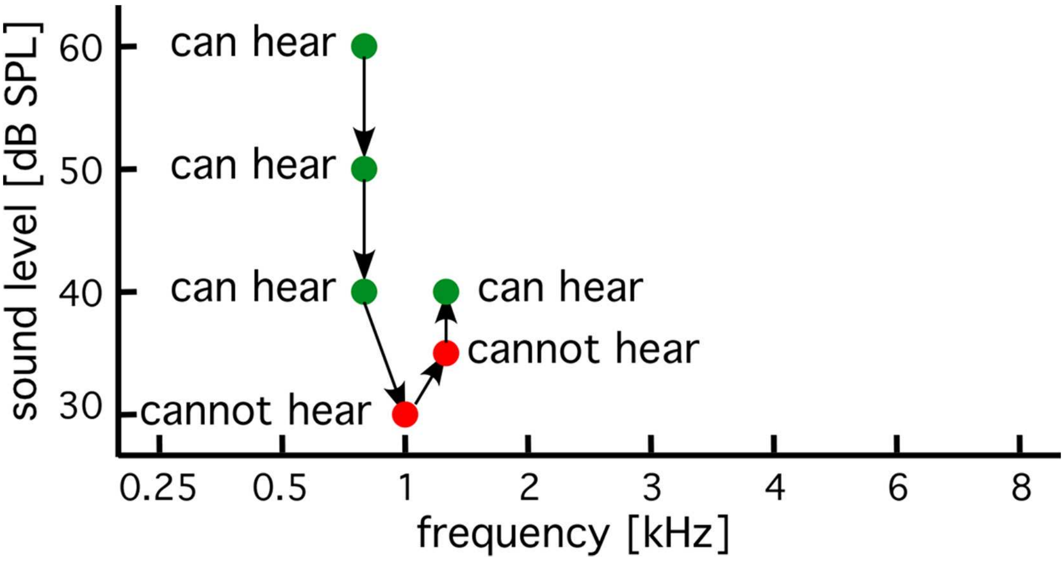

Figure 2.3: Example of the staircase method at 1 kHz. Threshold at 40 dB.

Staircase method

A variation of this method is the staircase method [10]. The experimenter begins the same as with the original method, by presenting a sequence of tones which progressively increase or decrease in intensity. When the patient's response changes, the tone intensity is recorded and the direction of the sequence is reversed. This procedure continues until a sufficient number of transition points have been recorded. The threshold is taken as the average of these points. Figure 2.3 shows an example of a trial at 1 kHz.

The method of limits (and especially the staircase method) is an efficient means of measuring thresholds and results are usually satisfactory. It is less time consuming than the method of constant stimuli, because there is a pattern in the presentation of the tone intensity. Those intensities which are well above or below the threshold are never presented again. The disadvantage of the latter, is that the pattern can lead to (perceptual) learning, giving the subject a chance of "predicting" the next response, which leads to a loss of precision.

This method is usually chosen in the process of determining the hearing loss of a patient. Section 2.2 outlines a standard for performing the hearing loss test.

2.1.1.3 Method of adjustment

The method of adjustment gives direct control over the sound intensity to the patient. The audiometer used with this method, is often called a "von Békésy" audiometer. It was invented in 1946 by the Hungarian-American physicist Georg von Békésy and is an instrument controlled by the patient. Modern Békésy audiometers control the frequency,

Measurement method	Advantage	Disadvantage
Method of constant stimuli	High accuracy	Very time consuming
Method of limits	Efficient, less time consuming	Less precise as other methods, because of (perceptual) learning
Method of adjustment	Active participation, prevention of loss of concentration/tiredness	Time consuming & depends on subjective threshold, because of loss of control by the audiologist.

Table 2.1: Positive & negative aspects of the different psychophysical methods.

repetition rate and rate of change in amplitude of the tone automatically, but the direction of change in tone level is controlled by the patient. The patient is presented an intensity level far below or above the threshold and is asked to increase or decrease the intensity until the perception changes. Generally a fairly large amount of ascending and descending series is to be presented and the mean of the results is taken as the threshold.

The advantage of this method is that the problem of loss of concentration or tiredness can be prevented, because of the active participation of the patient, possibly leading to maintain high performance. A disadvantage is the time consuming nature of this method. The patient does not have the experience the audiologist has in conducting the test, resulting in the fairly large amount of series. Another disadvantage is that the method depends more on the subjective threshold of the patient than the other methods. The patient controls the transition point and the audiologist has no control over it. In other methods, the audiologist can verify the transition points if he is in doubt over the responses. With the method of limits, there is no real option to do this, except for repeating the trial. The audiologist loses some the control over the procedure, possibly resulting in a loss of precision.

2.2 Methods for Manual Pure-Tone Threshold Audiometry

In 2004, the American National Standards Institute, Inc. (ANSI), approved the ANSI S3.21-2004 standard [3] as a revision of the ANSI S3.21-1978 Methods for Manual Pure-Tone Threshold Audiometry and the S3.6-2004 Specification for Audiometers standard [2]. The purpose of this standard is to present procedures for conducting manual pure-tone threshold audiometry whose use will minimize intertest differences based on test method. This section will introduce the methods for manual pure-tone threshold audiometry.

Pure-tone threshold audiometry is the procedure used in the assessment of an individual's

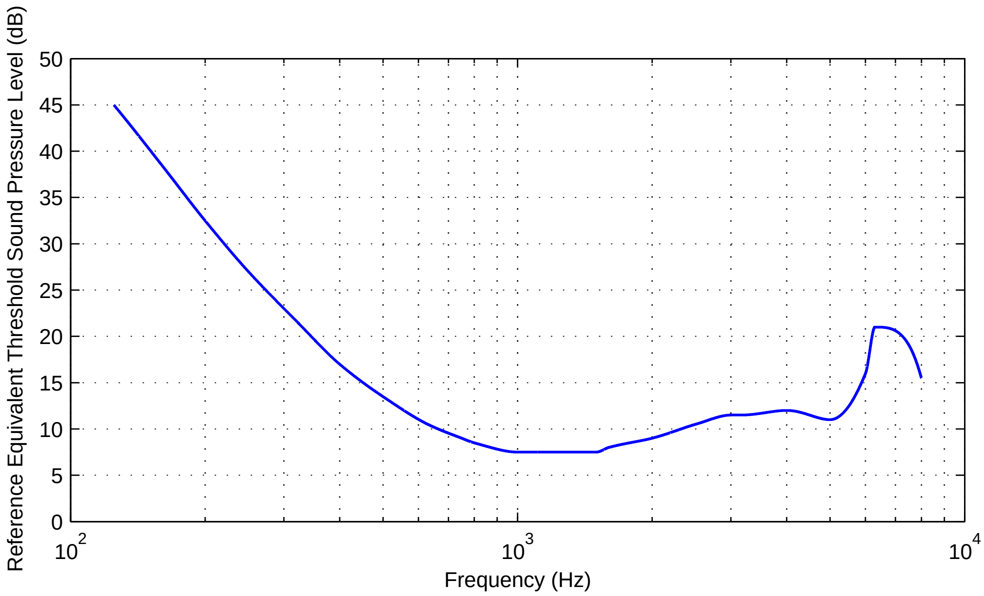

Figure 2.4: Reference equivalent threshold sound pressure levels for supra-aural earphones [2].

threshold of hearing for pure tones. It includes manual air-conduction measurements at octave intervals from 0.25 through 8 kHz and at intermediate frequencies as needed.

The standard defines the measurement of frequencies at octave intervals from 0.25-8 kHz. These are 0.25 kHz, 0.5 kHz, 1 kHz, 2 kHz, 4 kHz and 8 kHz. These are the most widely tested frequencies and done everytime. If the difference between two adjacent frequencies is more than 15 dB, a frequency in between is measured, respectively the 0.375 kHz, 0.75 kHz, 1.5 kHz, 3 kHz and 6 kHz frequencies.

2.2.1 Reference equivalent threshold sound pressure level

The measured thresholds are absolute levels in terms of sound pressure (dB SPL). To quantify the hearing loss relative to normal hearing, the sound pressure level which corresponds to normal hearing is found. These values are the Reference equivalent threshold sound pressure levels [2] (RETSPL). They are the standards which are used for defining normal hearing. It is based on the sound pressure level needed for an average person in the age of 18 to 30 years to detect a sound [2]. The RETSPLs are referred to as 0 dB hearing level (0 dB HL). If we speak about hearing loss in the following sections, it is always the hearing loss relative to 0 dB HL.

The ANSI specified for this purpose reference RETSPLs for several devices. For supra-aural earphones meeting the requirements described in IEC 60318-3, the RETSPLs are defined in figure 2.4. In this figure the minimal sound pressure level is presented which is needed for a normal hearing person to perceive a sound at a specific frequency. For example, a sound at a frequency of 0.25 kHz, the minimal SPL is 27 dB. For a sound at 1 kHz (10^3 Hz), the minimal SPL is 7.5 dB. Because the ear is much less sensitive for sound at lower frequencies than at higher frequencies, the minimal SPL at the lower frequencies is higher to be able to perceive the sound. The minimum lies between 1 and 5 kHz.

2.2.2 Determination of the threshold

First the patient needs to be familiarized with the listening task by a signal presented at 1 kHz at an estimated hearing level at which the patient has a clear response. There are two methods for familiarization:

- Present the tone continuously at a hearing level which will not be heard by the patient. Gradually increase the sound pressure level of the tone until a response occurs. Switch the tone for at least two seconds off and present it again at the same level. If there is a second response, proceed to the threshold measurement. Otherwise repeat the process.

- Present the tone at a hearing level of 30 dB. If a clear response occurs, proceed to the threshold measurement. Otherwise present the tone at 50 dB and at successive increment of 10 dB until a response occurs. The proceed to the threshold measurement.

The determination of the threshold occurs by presenting tones of 1-2 seconds duration, by using the method of limits (see section 2.1.1.2). The time interval between successive presentations is varied, but is never shorter than the test tone. The level of the first presentation of the tone for measurement is 10 dB below the level at which the patient responded during the familiarization process. Successive levels are presented, incremented in 5 dB steps until the first response occurs. Then the intensity is decreased by 10 dB and another ascending series is begun. The threshold is defined as the lowest hearing level at which responses occur in at least 50% of a series of ascending trials, with a minimum of two responses out of three required at a single level. The rate of frequency change is one octave per minute. The minimum recording period at each frequency is 30 seconds [2].

The better ear is tested first (if this information is available). The frequency of the first test tone is 1 kHz, followed by higher frequencies in ascending order. A retest of the 1 kHz is then done, to check for a possible learning effect. If the retest of the 1 kHz frequency differ from the first test by more than 5 dB, the lower of the two thresholds is accepted and at least one of the higher frequencies is retested. The procedure is continued by the assessment of the lower frequencies. The order of frequencies does not significantly influence the test results, but the above defined order will ensure consistency of approach to each test patient and minimize the risk of omissions.

The resulting hearing thresholds will be recorded in a tabular of graphical form (see figure 2.5). These values are relative to the Reference equivalent threshold sound pressure level (RETSPL, discussed in the next section), that is:

$$HL = SPL - RETSPL \tag{2.1}$$

with HL as the hearing loss and SPL the sound pressure level.

2.2.3 Adjustments to the ANSI standard

The ANSI standard as described above is not used everywhere. In the Netherlands, for example, the measuring method has some adjustments [20]. The frequencies and their order stays the same. With each frequency, the patient is presented a tone which is about 30 dB above the hearing threshold. A higher level is not recommended, because it could temporarily cause a threshold increase. The intensity of the tone decreases in steps of 10 dB, until the tone cannot be heard. Then the level is increased in steps of 5 dB, until a positive response occurs. In theory, this procedure is performed five times, but in practice it is done three times. The threshold is the lowest level on which a response occurs on more than half of the presented tones.

2.3 Most Comfortable Loudness & Loudness Discomfort Level

Apart from the hearing loss, two other measurements are commonly done, measurement of the Most Comfortable Loudness & Loudness Discomfort Level. These will be introduced in this section.

2.3.1 Most Comfortable Loudness

The Most Comfortable Loudness (MCL) is defined as the hearing level at which speech is most comfortably loud over a period of time. By several researchers it is regarded as the level at or near which a patient's speech-recognition ability is at its maximum [22].

There is not a real standard defined to measure the most comfortable loudness. Punch described in [23] a procedure to measure this level. The used stimulus consists of speech. The Independent Hearing Aids Fitting Forum's 7-point categorical scale is used to obtain the patient's judgement about speech loudness. The seven possible responses are:

- (7) Uncomfortably loud.
- (6) Loud, but OK.
- (5) Comfortable, but slightly loud.

- (4) Comfortable.
- (3) Comfortable, but slightly soft.
- (2) Soft.
- (1) Very soft.

The test is initiated at 30 dB HL. The intensity is then increased in steps of 5 dB and the responses are recorded. At the first occurrence of response "(5) Comfortable, but slightly loud" or higher, the intensity level is decreased in 10 dB steps, until a response of "(3) Comfortable, but slightly soft" or below is obtained. Then the intensity level is again increased with steps of 5 dB, until two ratings of "(4) Comfortable", out of three in succession, are obtained in the ascending series at the same intensity level. This intensity level is recorded as the MCL.

2.3.2 Loudness Discomfort Level

The Loudness Discomfort Level (LDL), also known as Uncomfortable Loudness (UCL), is defined as the hearing level at which the speech and pure tones become uncomfortably loud for the patient [22]. Together with the hearing threshold, LDL gives an indication of the dynamic range. When it is measured, the patient must feel confident that the intensity level will never be increased more then the patient allows for. Generally speaking, the Most Comfortable range is usually placed midways between the Hearing Threshold Level and the Loudness Discomfort Level.

The purpose is to measure the lowest intensity level which is uncomfortably loud to the patient. As with the measurement of the MCL, there is no real standard defined to measure the LDL. The British Society for Audiology propose the following procedure to measure the loudness discomfort level [6]:

The test is started at a pure-tone level which is comfortable for the subject and is predicted from the audiogram. Tones are presented for a duration of 1 second, with at least a 1 second quiet period. Ascend in steps of 5 dB until the subject indicates that the uncomfortable level is reached.

Another method is described in [23]. It uses the same stimulus and procedure, but with some changes. The procedure starts at 60 dB HL. The upper turning point is rated at "(7) Uncomfortably loud" and the lower turning point is rated at "(6) Loud, but OK" or lower. The UCL is recorded as the intensity level at which two ratings of (6), out of three in succession, are obtained in ascending series.

From research [22], it appears that there is no improvement in speech recognition above the LDL threshold.

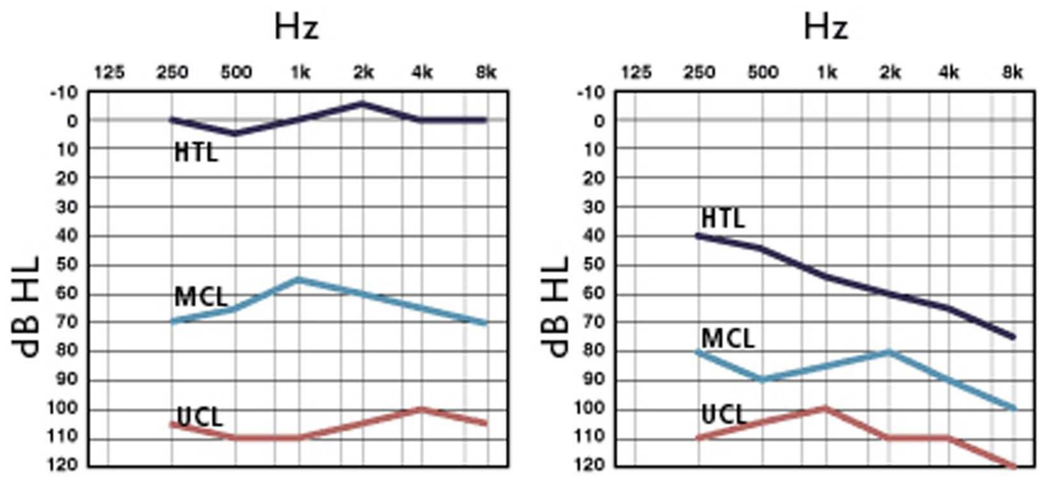

Figure 2.5: Example of two audiograms. The hearing loss is plotted downwards. Hearing loss is in dB HL.

2.4 Audiogram

The hearing loss is determined by using a measuring method (see section 2.2.2). The patient is presented different tones on different frequencies and has to respond if he or she has heard it, in order to find the hearing loss. The hearing threshold is defined as the minimum sound pressure level that is capable of evoking an auditory sensation in a specified fraction of series (see equation 2.1). The hearing threshold over all measured frequencies can be represented in an audiogram, which is a graphical representation of the hearing loss. Figure 2.5 presents two audiograms, with HTL as the hearing threshold and MCL and LDL (or UCL) as defined in 2.3. The left audiogram is a normal hearing person, the right audiogram is from a person with hearing loss. The slope of the right audiogram is characteristic to persons with hearing loss because of age, where the loss at the higher frequencies is higher than at the lower frequencies. The UCL is overall rather flat, with a a minor increase in dB HL at higher frequencies. The MCL is located in between these values. The horizontal scale represent the frequencies in Hz or kHz, while the vertical scale will be linear, labeled by hearing loss in dB. Conventionally, normal hearing is at the top of the graph and hearing loss is plotted downwards.

2.5 Test-Retest Reliability

The results of the testing of the hearing threshold depends on the time and place it was done and the instruments used. If a patient comes at a later time and has his hearing threshold retested, it is very likely that there are some (small) differences. The test-retest reliability quantifies how close retests are to the original test and are used to evaluate the consistency of measurements. The closer the two measurements are, the greater the test-retest reliability of the procedure is. The test-retest reliability is quantified as the percentage of retest variability within x dB as opposed to the original test (see section 2.5.1 for more details).

A drawback of this test is the practical effect. If a retest is done shortly after the test, patients tend to "learn" the responses and are able to influence the measurement. This leads to a higher correlation between the measurements. To try to avoid this problem, it is best perform the tests a week apart.

2.5.1 Determining the Test-Retest Reliability

In order to determine the Test-Retest Reliability, Schmuziger et al. [24] make use of two earphones, a Sennheiser HDA 200 circumaural and an Etymotic Research ER-2 insert earphone. They measure the hearing threshold four times (each earphone twice) for each test subject in the same session. After each measurement, the earphones were removed for a short period of time. Pure-tone measurements were done at the octave frequencies between 0.5-6 kHz and at 1.5, 3 and 6 kHz, and at 1/6 octave frequencies in the 8-16 kHz range. The subjects are normal hearing, with hearing loss $\leq$25 dB HL from 0.5-8 kHz (including frequencies of 1.5, 3 and 6 kHz).

From the tests Schmuziger et al. performed, they concluded that for the Sennheiser earphone the hearing losses measured at the retest were for 98% to 100% within 10 dB of the original test and decreased to 96% at 16 kHz. 25+ dB occurred only at one test subject at 12.5 kHz. For the ER-2 earphone, 89-99% was within 10 dB of the original test and decreased to 85% at 16 kHz. It can be concluded that the test-retest reliability differs between earphones and the difference in hearing losses increases particularly at high frequencies (at 16 kHz). For the purpose of measuring the hearing loss, the reliability overall is quite high.

Smith-Olinde et al. [26] & Burns et al. [9] have also measured the test-retest reliability and presented the mean differences between the tests and retests and the standard deviation. Smith-Olinde et al. have measured the 0.5, 1, 2 & 4 kHz from 43 healthy adults and the values range from -10 to 10 dB, with means of $[0.93, 0.23, 0.93, 0.69]$ dB for the respective frequencies and standard deviations of $[3.14, 1.86, 1.97, 4.02]$ dB. Note that the standard deviation at 0.5 and 4 kHz is higher than the frequencies in between. The step size in the measurement procedure is 5 dB. Burns et al. measured the 0.5, 1, 2, 3, 4 & 6 kHz frequencies from 20 healthy subjects with at least one week between two measurements. The means for the respective frequencies are $[1.0, 2.2, 1.5, 2.0, 1.4, -1.7]$ dB and the standard deviations are $[4.9, 4.2, 4.7, 4.7, 4.7, 7.6]$ dB. Note that at the 6 kHz frequency, the standard deviation is higher. The step size in the measurement procedure is 2 dB.

2.6 Improvements and automation

The above described procedures can take up a lot of time. For normal hearing, the average time is 5 minutes. But when the hearing loss is more apparent, the procedures may take up to 30 minutes. And this is one of the problems audiologists come across (see chapter 1). To reduce this problem, improvements for the measurement need to be found. This section tries to cover some possible improvements and ways to automate this.

The threshold at each frequency is measured independently. But looking at the audiograms in figure 2.5, there exists a relationship across the different frequencies: a threshold at some frequency gives information about thresholds at adjacent frequencies. The relationship between adjacent frequencies can be used to predict others. This can lead to measuring less frequencies and/or fewer trials for a particular frequency.

The next chapter introduces some techniques used to gather information from the audiograms. Chapter 4 introduces the data we will use to find the prior. This will be done in chapter 5.

Chapter 3

Data Mining

This chapter will introduce tools to extract information from a large dataset, called Data Mining. Several problems are associated with it, which will be explained later in this chapter.

3.1 Introduction to Data Mining

The amount of information in the world is growing rapidly. With it, the problem of handling and processing the information arises. For small sets of information it is possible to process information by hand. But if these datasets get to large (in the number of records or attributes), it is not feasible to recognize patterns by hand. Data mining is the science of trying to extract useful information and patterns from data. These can be used to predict the outcome of an event given a set of observations or find hidden relationships between several attributes, like predicting the behavior of the worlds climate during a period, based on observations in the past.

3.2 Preprocessing

There are several practical problems that are involved when using a dataset. Before any processing can be done, some preprocessing needs to be done, in order to remove or deal with common problems, like outliers and missing values.

Outliers are the problem of having attribute values which are outside the regular range in which the attributes fall. An example is an audiogram with hearing loss values $[11, 22, 33, 44, 55, 66]$. These values are incorrect, since the hearing loss is measured in 5 dB steps. Outliers can come from errors during measurement, testing or by exceptional cases. It is therefore important to try to eliminate or reduce the number of outliers in the data.

Missing values are another problem when you are faced with a dataset. Real world data can have records in which one or more attribute values are not specified. If having to many of these values, it can have a serious impact on the performance of the algorithm. Acuna & Rodriguez [1] note that rates of 1% or less are generally considered to be trivial, 1-5% is considered manageable, but 5-15% may be of serious impact on any kind of interpretation.

There are many algorithms proposed to deal with these missing values [29]. Among them are the straightforward most frequent element, replacing the missing value with a value which is out of range, the Mean Imputation or Median Imputation and just deleting the rows which contain missing values, also known as Complete Case Analysis. Another method is K-nearest neighbor (KNN) Imputation [29]. With this method it tries to find the K nearest neighbors and use those to replace the missing values. This technique will be discussed in section 3.3.3.

3.3 Different types of models

There are several types of algorithms that build models [28]. Among them there are models which capture the data into several parameters, like regression, while other models cluster the data based on a non-parametric algorithm, like k-nearest neighbors. Regression is used in this thesis to model the audiograms. K-nearest neighbors is chosen for calculating missing values, because it is one of the simplest algorithms, it is fast and it finds those hearing loss curves which are most similar to the curve which has missing values. Regression and k-nearest neighbors will be discussed in the following sections.

3.3.1 Regression

Regression is a predictive modeling technique where the target variable to be estimated is continuous [28]. It is the goal of learning a function that maps a set of observations into a response output. For our application to audiograms, the observations are the sound frequency and the response data the hearing loss.

3.3.1.1 Linear regression

We will use the following quadratic model for hearing loss as a function of frequency:

$$HL_f = \beta_0 + \beta_1 \log_2 (\frac{f}{f_{ref}}) + \beta_2 (\log_2 (\frac{f}{f_{ref}}))^2 \tag{3.1}$$

with f the frequency, f_{ref} the reference frequency of 1.4349 kHz (see section 3.3.1.3 for an explanation) and HL_f the hearing loss at frequency f. We use $\log_2(\frac{f}{f_{ref}})$ as predictor, because the standard audiometric frequencies are separated by a factor 2 (an octave). We divide by $f_{ref} = 1.4349kHz$ to facilitate in the interpretation of β_0. For $f = 1.4349kHz$ we get:

$$HL_{1kHz} = \beta_0 + \beta_1 \log_2(\frac{1.4349}{f_{ref}}) + \beta_2(\log_2(\frac{1.4349}{f_{ref}}))^2 = \beta_0 \tag{3.2}$$

We will use the more general notation $HL_{f_i} = y_i$ and $\log_2(\frac{f_i}{f_{ref}}) = x_i$ and we assume that the residuals are normally distributed with variance σ^2, that is:

$$f(y_i|x_i) \sim N(\beta_0 + \beta_1 x_i + \beta_2 x_i^2, \sigma^2) \tag{3.3}$$

The conditional likelihood is given by:

$$\mathcal{L} \equiv \prod_{i=1}^{n} f_{y|x}(y_i|x_i) \propto \sigma^{-n} \exp\{-\frac{1}{2\sigma^2} \sum_{i}^{n} (y_i - (\beta_0 + \beta_1 x_i + \beta_2 x_i^2)^2)\} \tag{3.4}$$

Maximizing over $\mathcal{L}(\beta_0, \beta_1, \beta_2, \sigma)$, known as Maximum Likelihood Estimator (MLE), gives the parameters β_0, β_1, β_2 & σ.

The log of $\mathcal{L}$, known as the log-likelihood, is given by:

$$\begin{aligned} \log \mathcal{L} &= -n \log \sigma - \frac{\sum_i^n (y_i - (\beta_0 + \beta_1 x_i + \beta_2 x_i^2)^2)}{2\sigma^2} \\ &= -\frac{SSE}{2\sigma^2} - n \log \sigma \end{aligned} \tag{3.5}$$

Since the maximum of $\log \mathcal{L}$ does not depend on the addition of $-n \log \sigma$ or the scaling by $2\sigma^2$, the maximum of $\log \mathcal{L}$ is equal to the minimum of the sum of squared errors (SSE):

$$SSE = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i + \beta_2 x_i^2))^2 \tag{3.6}$$

The SSE can be written in matrix form as:

$$SSE = (y - X\hat{\beta})(y - X\hat{\beta})^T \tag{3.7}$$

with $\hat{\beta} = (\beta_0, \beta_1, \beta_2)$, $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ and X as the design matrix

$$X = \begin{pmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{pmatrix}$$

with the columns representing the frequencies and an extra first column of 1's to calculate the intercept. $\hat{\beta}$ follows from minimizing the SSE and is given by:

$$\hat{\beta} = (X^T X)^{-1} X^T y \tag{3.8}$$

The Maximum Likelihood estimate of σ^2 is given by:

$$\sigma^2 = \frac{1}{n-p} \sum_{i=1}^{n} (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2))^2 \tag{3.9}$$

with n the number of audiometric frequencies, in our case 6, p as the number of parameters, in our case 3, and y_i as the observed hearing loss response value.

3.3.1.2 Goodness of fit

The goodness of fit of a model describes how well it fits the set of observations. The closer the fit is to the observed data, the better the model fits. Even if the regression coefficients maximize the likelihood, it does not mean the model is a good fit for the data. For example, if an audiogram for some frequencies shows a large hearing loss (due to damage on specific frequencies), it will be hard to model it using a quadratic model. Therefore, it is not only necessary to construct a model, but also to check if the model fits the observed data. The correlation between two variables [28] describes the strength of the linear relationship. Pearson's correlation coefficient between two variables x and y is defined by:

$$r(x,y) = \frac{Cov(x,y)}{\sigma_x \cdot \sigma_y} \tag{3.10}$$

where the covariance $Cov(x,y)$ and standard deviation σ_x are defined by:

$$Cov(x,y) = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \bar{x})(y_k - \bar{y}) \tag{3.11}$$

$$\sigma_x = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}(x_k - \bar{x})} \tag{3.12}$$

with $\bar{x}$ the mean of x.

The r^2 statistic, also known as the coefficient of determination, is an example of a goodness of fit parameter. Along with the SSE, the r^2 statistic uses two other types of errors, the total sum of squares (SST) and the regression sum of squares (SSM). They are defined as:

$$SST = \sum_i (y_i - \bar{y})^2 \tag{3.13}$$

$$SSM = \sum_i (f(x_i) - \bar{y})^2 \tag{3.14}$$

with y_i as the given response, $\bar{y}$ as the mean value of y_i and $f(x_i)$ as the predicted response value. r^2 is then given by:

$$r^2 = \frac{SSM}{SST} = 1 - \frac{SSE}{SST} \tag{3.15}$$

The r^2 for a regression model ranges between 0 and 1. The closer the value is to 1, the better the data can be explained by the model.

3.3.1.3 Orthogonal Design Matrix

The values of the regression coefficients will depend on the parametrization of the predictor x. Concomitantly the correlations between regression coefficients of different patients depend on the parametrization. We solve this problem by using a parametrization such that the columns of the design matrix are orthogonal:

$$\sum_{i=1}^{n}(X_{i,1}^{T}X_{i,2}) = \sum_{i=1}^{n}X_{i,2} = 0 \tag{3.16}$$

$$\sum_{i=1}^{n}(X_{i,1}^{T}X_{i,3}) = \sum_{i=1}^{n}X_{i,3} = 0 \tag{3.17}$$

$$\sum_{i=1}^{n}(X_{2,i}^{T}\cdot X_{i,3}) = 0 \tag{3.18}$$

with $X_{.,j}$ as column j of the design matrix and n the number of frequencies. QR decomposition is a technique to decompose a matrix into an orthogonal (Q) and an upper-triangular (R) matrix. The orthogonal matrix satisfies the equations defined in 3.16-3.18.

We create the design matrix X

$$X = \begin{pmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix} \tag{3.19}$$

by taking the $\log_2$ of the standard audiometric frequencies described in section 2.2. From this matrix we create the orthogonal design matrix X_{orth}:

$$X_{orth} = \begin{pmatrix} 1 & -2.6049 & 5.8503 \\ 1 & -1.5630 & -1.1701 \\ 1 & -0.5210 & -4.6803 \\ 1 & 0.5210 & -4.6803 \\ 1 & 1.5630 & -1.1701 \\ 1 & 2.6049 & 5.8503 \end{pmatrix} \tag{3.20}$$

As we can see, the representation of the original 1 kHz is now -0.52 (element (3,2) in the matrix). The reference frequency is thereby:

$$\log_2(1) - \log_2(f_{ref}) = -0.5210 \tag{3.21}$$

$$0 - \log_2(f_{ref}) = -0.5210 \tag{3.22}$$

$$\log_2(f_{ref}) = 0.5210 \tag{3.23}$$

$$f_{ref} = 2^{0.5210} = 1.4349 \tag{3.24}$$

Therefore β_0 represents the hearing loss at 1.4349 kHz.

3.3.2 Cross Validation

Cross validation is a model evaluation technique [28, 30] and is the process of using a training set to train the model and a test set (or validation set) to evaluate the performance of the model on validation data. There are three common types of cross validation: holdout validation, k-fold cross validation & leave-one-out cross validation.

Holdout validation

Holdout validation [28] is the most basic kind of cross validation. The data is separated into two sets, a training set and a test set. The model is trained only on the training set. Using parameter values fitted from the training data, predictions are made for the data in the test set. The difference between predicted and observed test set value are accumulated to compute the performance of the model. A common metric is the mean absolute error defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|f_i - y_i| \tag{3.25}$$

with f_i the predicted value and y_i the observed value from the test set. This metric is preferred over the mean squared error, because with the latter metric, outliers can influence the result disproportionally. This is undesirable in the evaluation of a regression model of an audiogram. A downside of the holdout validation technique is that it may depend heavily on the division of data, which could influence the evaluation and you are not using all your data to fit the model.

K-fold cross validation

An improvement over the holdout validation technique is the k-fold cross validation [30]. The data is divided into k subsets and the holdout validation technique is performed k times. Each time, one of the k subsets is used as the test set and the other $k-1$ subsets form the training set. The error is the average error across the k runs. The advantage is that this technique relies less on the division of the data, because each data point is once in the test set and $k-1$ times in the training set. The disadvantage is that the algorithm runs k times, before an evaluation can be made.

Leave-one-out cross validation

Leave-one-out cross validation (LOOCV) [30] is a special case of the k-fold cross validation, where k equals the number of frequencies in the dataset. The model is trained k times on all frequencies, except for one. Each training leaves a different frequency out, such that each frequency is left out once. Each time the model is trained on the remaining frequencies. Based on that model, a prediction is made for the frequency that was left out. The absolute difference between the predicted value and the true value is taken as the absolute error. The average over the k errors is taken as the mean absolute error to evaluate the model. A disadvantage is that it is expensive to run. Leave-one-out cross validation will be used to evaluate the prediction of an audiogram based on a regression model.

Parametric Bootstrapping

The pattern of MAEs for each audiometric frequency can have two sources: they can represent the characteristics of the dataset or they can represent the characteristics of the quadratic model. To investigate this, we use parametric bootstrap. The idea is to simulate audiograms from the original dataset and use these simulated audiograms as the input for cross-validation. If the results on these simulated sets are similar to the results of the cross-validation we done before, the results are largely due to the model, otherwise due to characteristics of the datasets.

To bootstrap an audiogram, we calculate its regression coefficients as in equation 3.8. Then we simulate the hearing loss with:

$$HL_{sim} = \beta_0 + \beta_1 x + \beta_2 x^2 + N(0, \sigma) \tag{3.26}$$

with σ as the standard deviation, calculated by taking the square root of the Maximum likelihood estimation of σ^2, given by equation 3.9.

3.3.3 K-nearest neighbors

K-nearest neighbors (KNN) [28] is an algorithm which tries to find the k nearest neighbors of a data point and use only these data points as input for the rest of the process. It does this by calculating the proximity between the given data point and all the other points present. The k-nearest neighbors of a point in space are those k points which have the closest proximity (or highest similarity) to the given data point. Commonly used distance measures are the Euclidean Distance, Manhattan Distance and Cosine Similarity [28]. Figure 3.1 shows an example of KNN and algorithm 1 offers an overview of KNN.

Algorithm 1 K-Nearest Neighbors [28]

Input: a dataset D, k the amount of neighbors to be found, x the example.
Output: k-nearest neighbors.

for each data point $d \in D$
    Calculate $distance(x, d)$.
end
Return D_x as the k closest data point to x.

Important with this technique is the value of k. Choosing a value which is too small may result in overfitting the model, because the dataset could contain noise. If k is chosen too large, it may include points in the neighborhood which are far away from the given point. So choosing the right k is an important factor.

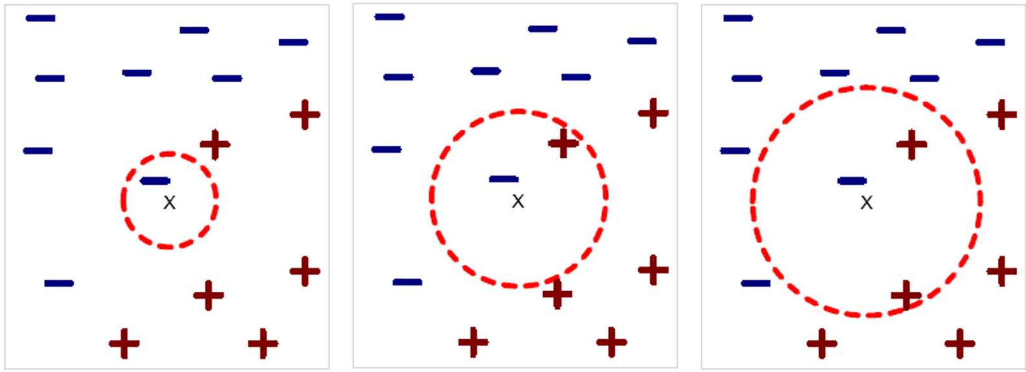

Figure 3.1: Example of K Nearest Neighbors [28]

3.3.3.1 K-Nearest Neighbors as classification

The KNN Impute algorithm can be used as a classification algorithm [29] to replace the missing values. When the list of closest data points D_x to x is obtained, several techniques can be deployed to classify x. One commonly used method is using the majority class of its nearest neighbors [28]:

$$MajorityVoting : y' = \arg\max_{v} \sum_{(z_i, y_i) \in D_x} I(v = y_i) \tag{3.27}$$

where v is the class label, z_i are the neighbors, y_i the class label for one of the nearest neighbors and $I(\cdot)$ as an indicator function that returns the value 1 if its argument is true and 0 otherwise.

3.4 Improvements and automation

In the previous chapter about audiograms, it was discussed that there is a relationship between the hearing loss at adjacent frequencies, making it possible to improve and automate the procedure of measuring the hearing loss. The improvement consists of using the knowledge from audiograms of other patients as a prior in a bayesian framework.

We plan to use the Bayesian framework for the estimation of $\beta_0 \ldots \beta_2$. In that framework, we model the prior over β_0, β_1 & β_2 as a multivariate normal. We can define it as:

$$Pr(\beta) \sim N(\mu, \Sigma) \tag{3.28}$$

with β as the vector containing β_0, β_1 & β_2, μ as the mean values of β and Σ as the covariance. Since it is difficult to visualize a three-dimensional density, we plot the following marginal densities:

$$\Pr(\beta_0) = \int\int \Pr(\beta_1, \beta_2)\, d\beta_1\, d\beta_2 \tag{3.29}$$

$$\Pr(\beta_1) = \int\int \Pr(\beta_0, \beta_2)\, d\beta_0\, d\beta_2 \tag{3.30}$$

$$\Pr(\beta_2) = \int\int \Pr(\beta_0, \beta_1)\, d\beta_0\, d\beta_1 \tag{3.31}$$

$$\Pr(\beta_0, \beta_1) = \int \Pr(\beta_2)\, d\beta_2 \tag{3.32}$$

$$\Pr(\beta_0, \beta_2) = \int \Pr(\beta_1)\, d\beta_1 \tag{3.33}$$

$$\Pr(\beta_1, \beta_2) = \int \Pr(\beta_0)\, d\beta_0 \tag{3.34}$$

Because the Beltone and Copenhagen datasets contain discrete values (multiple of 5 dB) and the marginal density plots, which are binned for evaluation purposes, depend on the choice of the origin of the bin grid, we use the averaged shifted histogram (ASH) technique [25] to remove this dependence. The averaged shifted histograms technique does not depend on the origin of the bin grid, because it computes histograms using the same bin width, but with different origins and averages over the different histograms. As the number of histograms in the average increases, ASH converts from a step function to a continuous function. The obtained average histogram is used in the marginal density plots.

This thesis will focus on finding an estimate of the prior, by finding the marginal densities, in order to investigate the feasibility of a bayesian framework possible.

Chapter 4

Dataset

This chapter will outline the datasets that will be used within this thesis. The attributes will be presented, along with the processing done on the sets.

4.1 Introduction on Dataset

This thesis will focus on improving the fitting of hearing aids by using prior information from community audiograms. We will be using real data from real measurements of patients. The next section will introduce the data that will be used. Furthermore, it will explain what preprocessing has been done to make fully use of the data.

4.2 Datasets

One of the main components of this thesis is data. Without data, it will not be possible to test the assumptions made and refine the algorithm. This section will discuss the datasets that will be used. The data that will be used in this thesis will come from two different sources, GN Resound Copenhagen and Beltone Connect. Both sets will be explained more in the next sections.

4.2.1 Removing incorrect values

Before a dataset can be used in the analysis, it is necessary to filter out values which are incorrect. To do so, we remove every audiogram which contains incorrect data. The occurrence of values not conforming to a divisor of 5 dB (see section 2.2.2 for this value), is a possible indication of corruption or misprint. There can also be values which appear correct at first, but are incomplete or fake when examined more closely, An example is a record where all the values are the same. They can be legitimate records, but it is also possible that they are test records. In order to be on the safe side, we remove these records from the dataset. Audiograms with large leaps among adjacent frequencies are also suspicious. Many can be explained by the assignment of a zero value, which most probably means the frequency has not been measured.

4.2.2 Beltone Connect

The first dataset was made available by Beltone Connect. This set has been created over many years and contains 101.906 audiograms. The audiograms are taken from hearing aid stores in the United States, which measured the hearing loss of patients needing a hearing aid. These records have the following 15 attributes:

- (8) The hearing loss for 0.25 kHz, 0.5 kHz, 1 kHz, 2 kHz, 3 kHz, 4 kHz, 6 kHz & 8 kHz frequencies (in dB HL).
- (1) Most Comfortable Level for speech (in dB)
- (1) Loudness Discomfort Level for speech (in dB)
- (5) Loudness Discomfort Level for 0.25 kHz, 0.5 kHz, 1 kHz, 2 kHz & 4 kHz band (in dB).

This set contains the hearing loss on frequencies as defined in 2.2 as well as 3 and 6 kHz. We have no way to know if all records are unique, because there is no information available which indicate a unique relationship to a patient.

4.2.2.1 Preparation of the data

The preparation of the Beltone dataset starts with the removal of incorrect values, as described in section 4.2.1. The Beltone dataset contains 101.906 audiograms. In this dataset, missing values are denoted by "-1". Of these records, 22.011 records contain one missing value at the standard audiometric frequencies and 28 records contain two or more (respectively 21.6% and 0.0% of the dataset), leaving 79.867 complete records. The remaining set contains 309 records (0.4%) with one or more values not conforming to a 5 dB divisor, which are removed from the dataset. There appeared to be 205 audiograms (0.3%) in the resulting set, which have identical values across all frequencies. The removal of giant leaps (60 dB or more) resulted in the removal of 74 audiograms (0.1%), leaving 79.279 audiograms to use for the analysis. We have removed 22.627 audiograms (22.2%) from the original set. For a summary of this dataset, please see table 4.1. Because we are left with more than 79.000 audiograms, we decided not to process audiograms with missing values. We have summarized some simple statistical information (mode, median, mean and standard deviation for each frequency) about this remaining set in table 4.2.

4.2.3 GN Resound Copenhagen

The second dataset is from GN Resound Copenhagen and will be used to compare the results with those obtained by the Beltone set. This set contains 688 audiograms of patients

Processing	Number of removed audiograms
One missing value	22.011
Two or more missing values	28
Not conforming to 5 dB divisor	309
Identical values	205
Leaps ≥60 dB across adjacent frequencies	74
Total number of removed audiograms	22.627
Audiograms for analysis	79.279

Table 4.1: Summary of the Beltone Connect dataset.

Frequency (kHz)	Mode (dB)	Median (dB)	Mean (dB)	Std (dB)
0.25	30	35	38.9	17.0
0.5	40	40	40.3	16.8
1	45	45	45.4	16.4
2	55	55	54.6	15.0
4	65	65	67.2	15.8
8	80	80	79.2	17.3

Table 4.2: Simple statistics of the Beltone Connect dataset.

who participated in a trial to test a hearing aid prototype. The dataset consists of the following 37 attributes:

- (1) Patient's unique ID.
- (1) Patient's first name.
- (1) Patient's last name.
- (1) Patient's birthdate.
- (1) Patient's gender.
- (2) Patient's type of hearing loss (Sensorineural, Conductive & Mixed Sensorineural/Conductive), for each ear one attribute.
- (2) Year of measurement of the hearing loss, for each ear one attribute.
- (2) The brand name of the hearing device, for each ear one attribute.
- (2) The device family name given the brand name, for each ear one attribute.
- (2) The form factor of the device (BTE, Open Fitting BTE, PBTE, SPBTE, ITE, ITC, CROS, CIC, Spectacles, None), for each ear one attribute.

- (22) The hearing loss on 0.125 kHz, 0.25 kHz, 0.5 kHz, 0.75 kHz, 1 kHz, 1.5 kHz, 2 kHz, 3 kHz, 4 kHz, 6 kHz & 8 kHz frequencies, for each ear one attribute.

This set contains the most widely used as well as the 0.125, 0.75, 1.5, 3 and 6 kHz frequencies. Each patient has a unique patient ID and it is possible for one patient to have multiple records, by having recorded test-retest measurements.

4.2.3.1 Selection Criteria

There are some attributes in the Copenhagen dataset which we will use to identify special relationships within this data, like age and gender. We need to set up some extra selection criteria, before we can use this information, in order to remove possible biased results. These criteria are:

- The age of the patient (30 years and older).
- Use only one audiogram per patient (Uniqueness).

We will only use patients with a minimum age of 30 years, so children are not taken into account. Children can have some specific characteristics, which can bias the results for adults and elderly. We also remove any duplicate information, taking only one measurement of a patient. If we would keep track of all measurements, the audiograms could lie too close together and do not represent a group of unique patients. These attributes will be used to research a relationship between them and the hearing loss.

4.2.3.2 Preprocessing the data

The complete dataset consists of 688 records. We make use of the widely used frequencies as described in section 2.2, along with the 6 kHz frequency. We first remove all children and patient audiograms with unknown age, leading to 538 audiograms. There are no patients left without a specification of gender. Then we take of all patients the audiogram with the least amount of missing values, leading to 221 audiograms of unique patients.

Of the remaining audiograms, there are 59 audiograms with one missing value at the standard audiometric frequencies and 7 with two or more (respectively 26.7% and 3.2% of the remaining set). Of the 155 complete audiograms, there was only one with a leap $\geq$60 dB, which is removed. No other incorrect values were found. Because there are relatively many audiograms with one missing value, these are filled in. KNN Impute (see section 3.3.3) is used for this purpose (with k=10), to compute the nearest neighbors. Of these neighbors, the mode is chosen as the missing value (majority of vote, section 3.3.3.1). This method is used, because audiograms which lie close together are expected to have roughly the same hearing loss. The filled audiograms had no incorrect values, ending the preparation with 213 audiograms. Please see table 4.3 for a summary. We have not included the amount of audiograms with one missing value, because we have processed them already. We have summarized some simple statistical information (mode, median, mean and standard deviation for each frequency) about this remaining set in table 4.4.

Processing	Number of removed audiograms
Unknown age or younger than 30 years	150
Unknown gender	0
Duplicate audiograms of patients	317
Two or more missing values	7
Not conforming to 5 dB divisor	0
Identical values	0
Leaps $\geq$60 dB across adjacent frequencies	1
Total number of removed audiograms	475
Audiograms for analysis	213

Table 4.3: Summary of the GN Resound Copenhagen dataset.

Frequency (kHz)	Mode (dB)	Median (dB)	Mean (dB)	Std (dB)
0.25	15	30	33.1	21.2
0.5	35	35	37.7	22.4
1	45	45	45.2	21.8
2	55	55	55.6	19.1
4	75	70	67.4	19.1
6	75	75	75.2	19.7
8	80	75	75.6	21.1

Table 4.4: Simple statistics of the GN Resound Copenhagen dataset.

4.2.3.3 Presentation of the data

This part of the thesis will cover the preprocessed Copenhagen data in a graphical way. In figure 4.1 the age distribution of the Copenhagen dataset is presented, along with the age distribution among hearing impaired in the United States in 1997 by Kochkin [16]. The data is divided into six age categories (18-34, 35-44, 45-54, 55-64, 65-74, 75+) and each category presents the percentage of patients in that category opposed to the total number of patients. As we can see, most patients in the Copenhagen dataset are between 45 en 65 years old. Overall, the patients in the Copenhagen dataset are somewhat younger compared to the Kochkin patients, specifically visible at the 45-54, 55-64 and 75+ category, while at the other categories, the distribution is comparable.

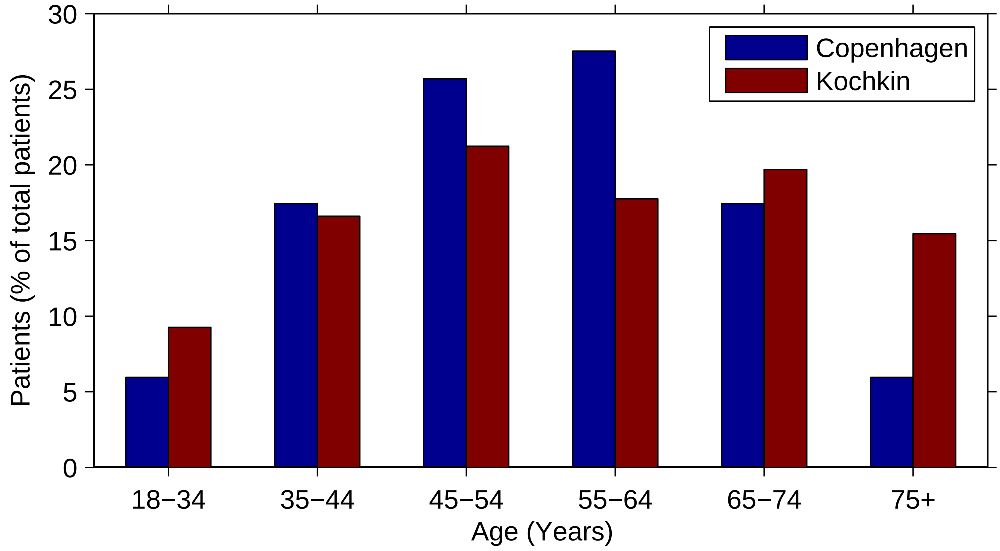

Figure 4.1: Age histogram of the Copenhagen dataset (blue) and Kochkin data (red) [16].

Chapter 5

Analysis of the datasets

In the previous chapter the two datasets were introduced. This chapter will cover the analysis of the Beltone and Copenhagen datasets. First we compare the datasets with results reported by Pittman. The purpose of this comparison is to ensure that our datasets are typical of the population of hearing impaired. Criteria will be used to divide the dataset based on its shape. Then we will investigate if we can fit a curve into the data to represent the audiogram and determine the goodness of fit, as described in chapter 3.

5.1 Pittman Criteria

Analysis starts by dividing the dataset in similar audiograms by using the shape. This section will introduce the criteria, which we will call the Pittman Criteria from now on, on which this division will be based. Pittman describes in [21] the objective to characterize the sensorineural hearing losses of a group of children and adults along three parameters important to the hearing instrument fitting process:

- Audiometric configuration.
- Asymmetry of loss between ears.
- Progression of loss over several years.

To define the audiometric configuration, Pittman describes six categories of audiograms. The criteria to define each category are [21]:

- 1. *Sloping:* thresholds occurred at equal or successively higher levels from 0.25 to 8 kHz. The difference between thresholds at 0.25 and 8 kHz was always $>$ 20 dB. To accommodate severe-to-profound hearing losses with only one or two thresholds in the low-frequency region (corner audiograms), thresholds indicating no response at the limit of the audiometer were not considered in the algorithm.
- 2. *Rising:* thresholds occurred at equal or successively lower levels from 0.25 to 8 kHz. The difference between thresholds at 0.25 and 8 kHz was always $>$ 20 dB.

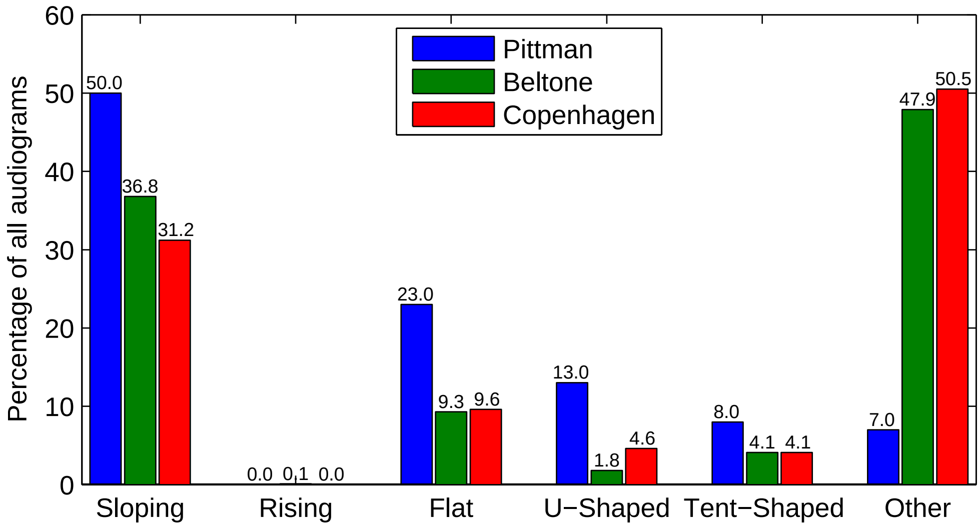

Figure 5.1: Categories of Pittman Criteria in percentages of the Pittman, Beltone & Copenhagen data.

- 3. *Flat:* thresholds across frequencies did not vary more than 20 dB from each other.
- 4. *U-Shaped:* one or more adjacent thresholds between 0.5 and 4 kHz were $\geq$ 20 dB relative to the poorer threshold at 0.25 or 8 kHz.
- 5. *Tent-Shaped:* one or more adjacent thresholds between 0.5 and 4 kHz were $\leq$ 20 dB relative to the better threshold at 0.25 or 8 kHz.
- 6. *Other:* configurations that did not meet the criteria for inclusion in any other category.

We use these criteria for a first analysis of the datasets. We compare them to the results of Pittman, in order to show that our sets are similar in characteristics to Pittman's data and represent the typical population of hearing impaired. The original figure in [21] is replotted in our own figure to make the comparison easier. We have removed the analysis of children, because in our analysis we omit children (see section 4.2.3.1) or have no knowledge about them. In figure 5.1 for each category the percentage of audiograms of the full dataset is presented. We see that the sloping category is the largest category within each dataset, if we do not consider the "other" category (on which we will come back soon). The u-shaped category is the second largest for all datasets. The tent-shaped and flat category seem to be relatively small for our own datasets, where Pittman has a larger proportion. A difference between the datasets is that the Beltone Connect set has a very small amount of rising audiograms, while both other sets have none. But the real difference between our datasets and Pittman's data is in the "other" category. Our data seem to have many audiograms (about 50%) which met that category, contrary to the Pittman data, which has only a small portion in that category, the smallest over all categories. We expect that the categories by Pittman are too specific for her data, so many of our audiograms fall in the “other” category. We will investigate this further later on. We conclude that our two datasets are quite similar with respect to the division in categories, with the exception of the “other” category.

Next we compare our data to the Pittman data by taking the mean of the audiograms in a category and plot them. This is done in figure 5.2. On the left side the Pittman data is plotted, the Beltone set in the middle and on the right the Copenhagen set. At each category, the mean of the audiogram is plotted, with 1 standard deviation for the errorbars (as described in [21]). For the sloping, flat and u-shaped categories, the audiograms are quite similar. For the u-shaped category, all three sets show the highest loss at 4 kHz, following the same kind of slope. The Pittman analysis provided for the tent-shaped and “other” category four examples instead of the average to show that their set contained some specific characteristics, which would be lost if the mean was taken. In our analysis, we have omitted the examples, because the tent-shaped set did not contain any specific audiograms which where far off from the mean. The “other” category, on the other side, seemed to contain many audiograms which could be part of any of the other categories, but did not meet the criteria. We conclude that the audiograms in each of the categories are similar in shape.

5.2 Most Comfortable Loudness vs Loudness Discomfort Level

In order to compare our datasets to other sets and show our sets are not very different from others, we make a comparison between the most comfortable loudness and loudness discomfort level of the Beltone dataset and data from Schwartz et al. & Pascoe, presented in [11]. We can only compare the Beltone set, because the Copenhagen set does not contain information about most comfortable loudness and loudness discomfort level. The MCL and LDL are plotted against the average hearing threshold in the $[0.5 \dots 4]$ kHz range.

In figure 5.3 we see that there is a relationship between the average hearing loss and the most comfortable loudness and the loudness discomfort level. The MCL and LDL get higher as the average hearing loss increases. If the most comfortable loudness gets higher, the loudness discomfort level gets higher as well, although at a slower rate. The figure also shows that the LDL in the Beltone set is comparable to the others sets, while the MCL is not as steep as the other sets, resulting in a higher MCL at the lower average losses, while at the higher average losses the MCL is significantly lower than the other sets.

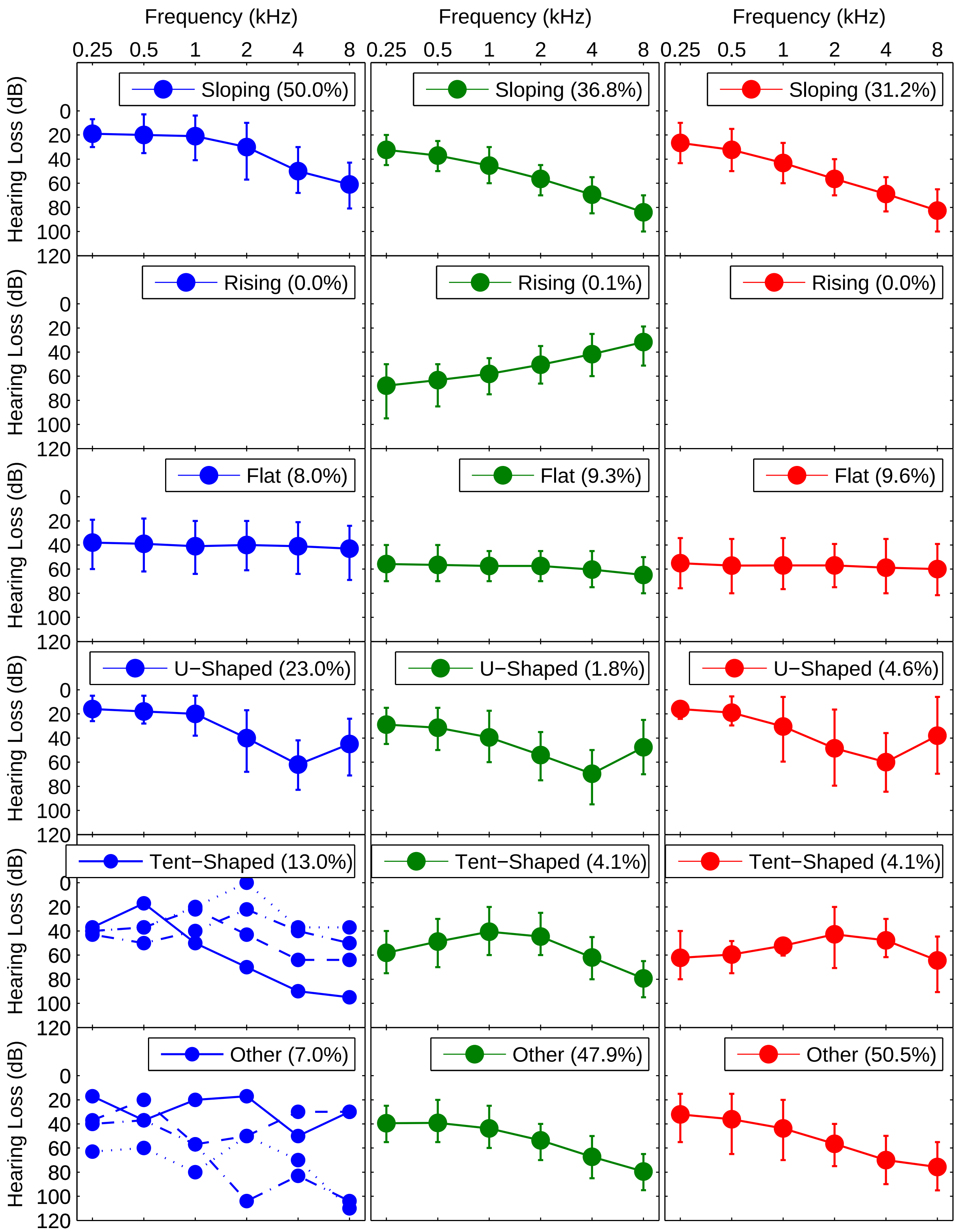

Figure 5.2: Average ±1 std audiograms of the Pittman (left), Beltone (middle) & Copenhagen (right) data, split for the different Pittman categories.

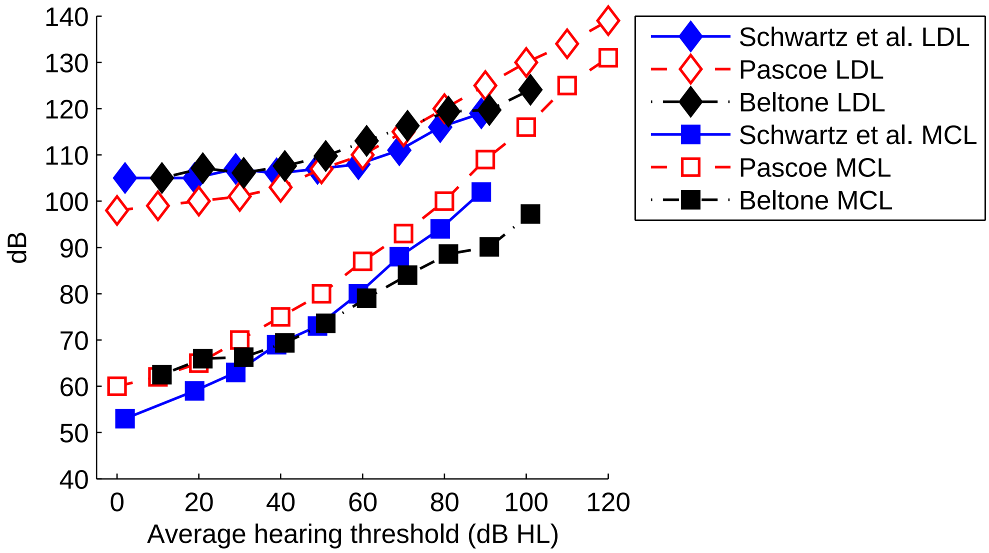

Figure 5.3: Most comfortable loudness vs Loudness discomfort level of the Beltone Connect set and data from Schwartz et al. & Pascoe presented in [11].

5.3 Curve Fitting

Now that we have established that our data represent the typical hearing impaired population, we can start the analysis of our datasets. We fit a curve to each individual audiogram using linear regression, as described in 3.3.1.1, leading to three coefficients (β_0, β_1 & β_2). For each audiogram we calculate the goodness of fit. For each category based on the Pittman criteria, we have plotted the r^2 value in figure 5.4. The majority of audiograms in the sloping and tent-shaped category are fitted quite well. The flat category is somewhat more equally divided. This can be explained by the definition of this category. The difference between the minimum and maximum value of the audiogram is within 20 dB. Within this audiogram, the hearing loss can bounce from the minimum to maximum and back, giving a golf-shaped audiogram. This is difficult to fit with a quadratic curve, leading to a fit which looks like a straight line. This results in a low r^2 value. The u-shaped category contains audiograms where a fit results most of the time in a value between 0.6 and 0.9.

One surprising result is the other category, which we already mentioned in 5.1. From the definition in 5.1 we find that the other category should contain audiograms which do not follow a shape with some kind of relationship between adjacent frequencies. This is displayed by Pittman's data in figure 5.2. However, if we take a look at the r^2 values of the other category in figure 5.4, we see that many audiograms are being fitted quite nicely, which contradicts the definition. Because the percentage of audiograms in the other category in the Beltone dataset is much higher than in the Pittman data, we think the definitions of the Pittman criteria are too strict and based too much on the dataset that Pittman had. We took a look at some of the respective audiograms and concluded that

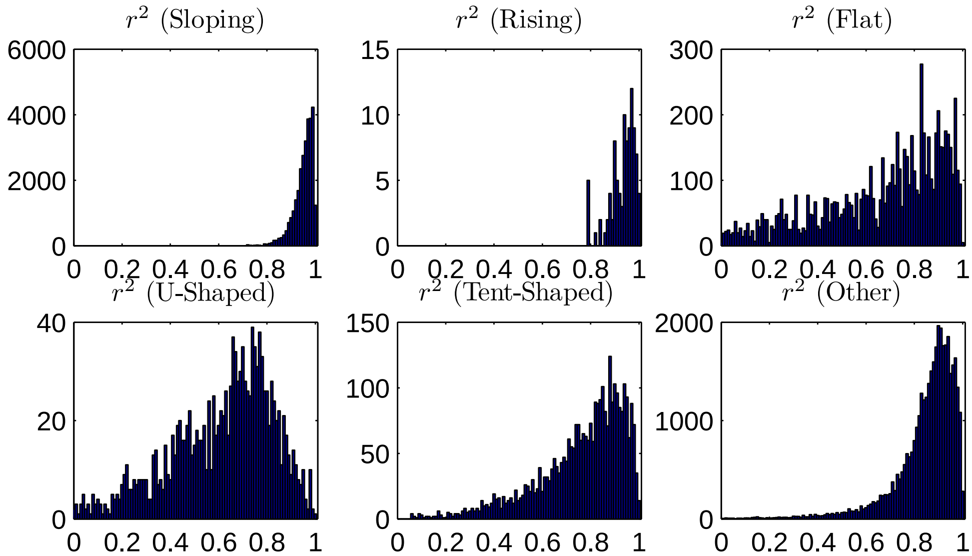

Figure 5.4: r^2 values (goodness of fit) of the regression curve fitting of the Beltone Connect dataset based on the Pittman criteria.

there are many audiograms in the other category which could fit in with another category, but was not because it did not fully meet the requirements. An example is an audiogram which was sloping except between 0.25 and 0.5 kHz, where it was rising 5 dB. We will describe alternative criteria to determine the type of shape an audiogram has.

5.3.1 Alternative criteria

Because many audiograms in the Beltone and Copenhagen dataset did not meet the criteria Pittman described and fell into the other category, we will describe alternative criteria on which the type of audiogram can be determined. Many audiograms in the other category had a sloping shape, but had a decrease in hearing loss of 5-10 dB in the lower or higher frequencies and did not meet the sloping criteria. This 5-10 dB difference is of the same order of magnitude as the test-retest reliability, as described in 2.5. We feel that such a decrease should not exclude the audiogram from the sloping category and find the criteria too strict to determine the type of shape of an audiogram. We also found that many u-shaped and tent-shaped audiograms had a great increase in the hearing loss at a specific frequency, while the lower frequencies are rather flat. Fitting a curve to this does not result in a u-shaped or tent-shaped curve, but rather in a more sloping or rising curve. We will determine the type of shape based on the fitting of a curve. A u-shaped audiogram should have a u-shaped curve fitted to it and a sloping audiogram has a sloping curve fitted to it. The criteria we determined are described in table 5.1. The values for β_1 and β_2 are determined by taking the Pittman categories and taking the 15 and 85 percentile of the coefficients of the audiograms. We have capped the β_1 values of the u-shaped and tent-shaped categories to 10 $\frac{dB}{octave}$. If this value gets too large, the fit does not follow a

Type of shape	σ (dB)	r^2	β_1 $(\frac{dB}{octave})$ & β_2 $(\frac{dB}{octave^2})$
Sloping	$\sigma \geq 11$	$r^2 \geq 0.6$	$(\beta_1 \geq 0$ & $\lvert\beta_2\rvert < 2)$ or $\beta_1 \geq 10$
Rising	$\sigma \geq 11$	$r^2 \geq 0.6$	$(\beta_1 < 0$ & $\lvert\beta_2\rvert < 2)$ or $\beta_1 < -10$
Flat	$\sigma < 11$		
U-Shaped	$\sigma \geq 11$	$r^2 \geq 0.6$	$\lvert\beta_1\rvert < 10$ & $\beta_2 < -2$
Tent-Shaped	$\sigma \geq 11$	$r^2 \geq 0.6$	$\lvert\beta_1\rvert < 10$ & $\beta_2 \geq 2$
Other	$\sigma \geq 11$	$r^2 < 0.6$	

Table 5.1: Alternative criteria to determine the type of shape of an audiogram.

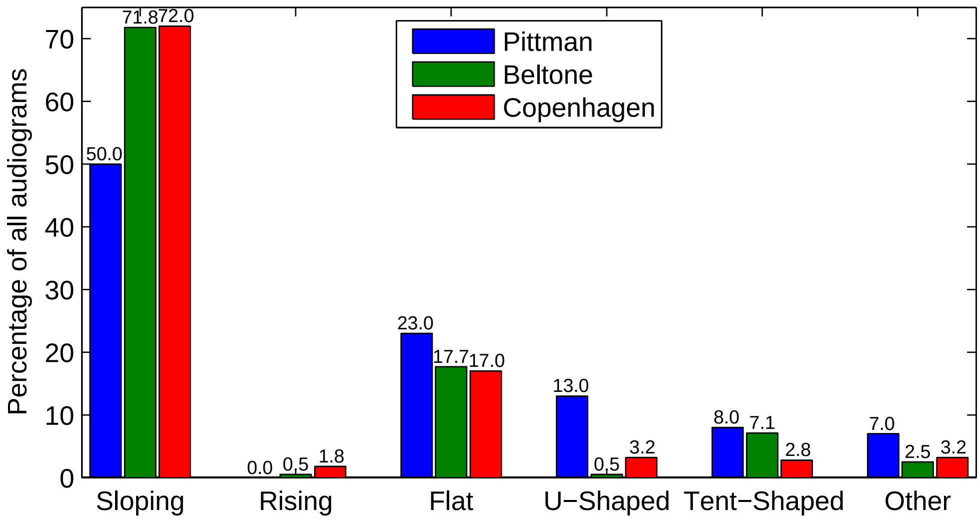

Figure 5.5: Categories of alternative criteria in percentages of the Pittman, Beltone & Copenhagen data.

u-shape or tent-shape anymore. The linear coefficient dominates the curve, giving a more sloping or rising type of curve. We therefore added an extra optional criterium to the sloping category, stating that any audiogram with a linear component larger than 10 $\frac{dB}{octave}$ is sloping. While there are a few cases where the fitted curve still shows a tent-shape or u-shape, they represent less than 1% of sloping audiograms. The same goes for the rising category. The flat category has audiograms where the standard deviation is smaller than 11 dB. The Pittman criteria state that an audiogram is flat when the difference between the min and max is smaller than 20 dB. Taking the most extreme audiogram which meets this criteria, the standard deviation is a little bit smaller than 11 dB. This is our criterium value.

We have replotted the percentages and shaped in figures 5.5 and 5.6. Because we did not have the raw data Pittman used, we have not recategorized the Pittman data. Comparing figure 5.5 to figure 5.1, our datasets have more sloping audiograms and more flat audiograms, while the other category decreased to a percentage Pittman had. The u-shaped

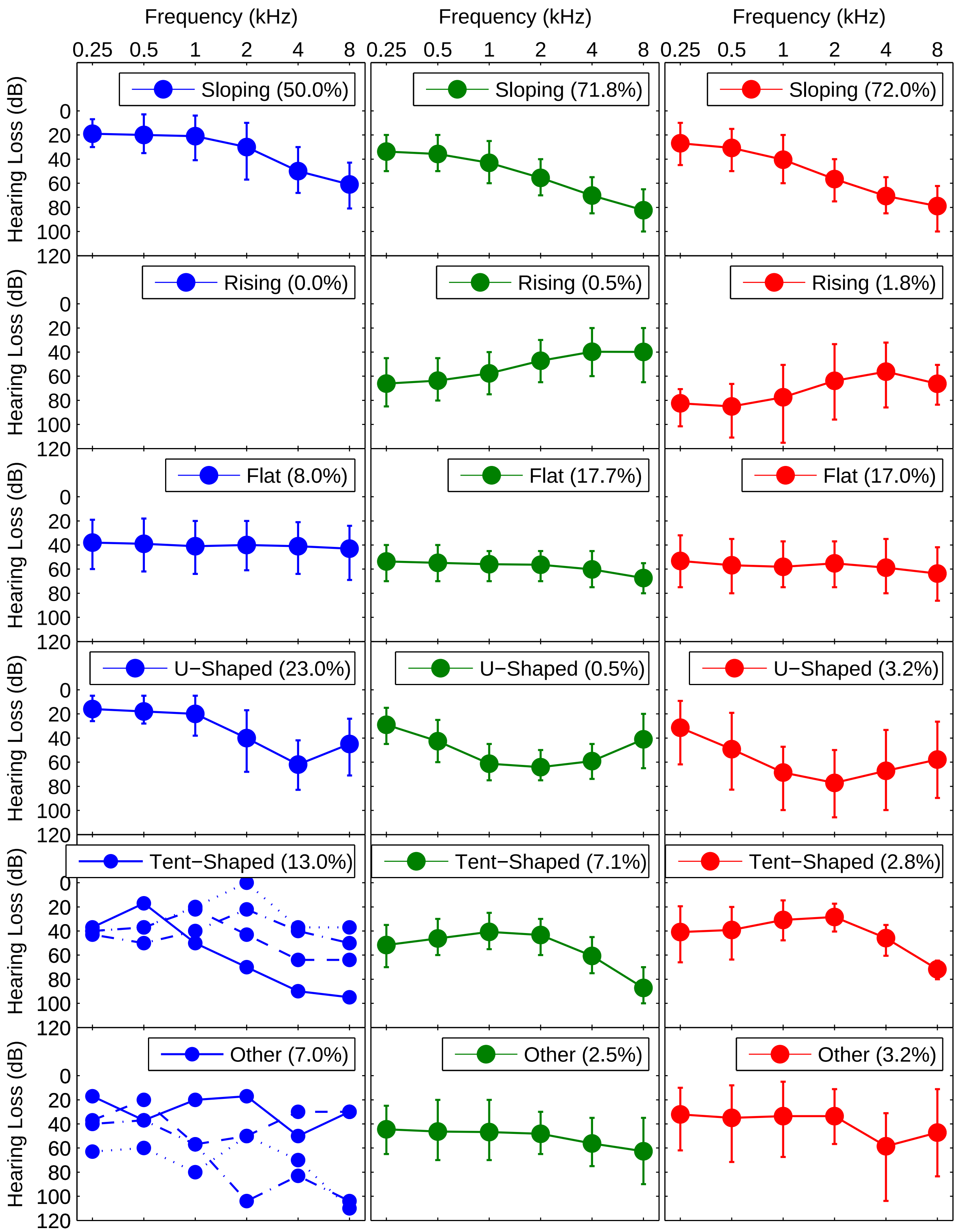

Figure 5.6: Average ±1 std audiograms of the Pittman (left), Beltone (middle) & Copenhagen (right) data, split for the different alternative categories.

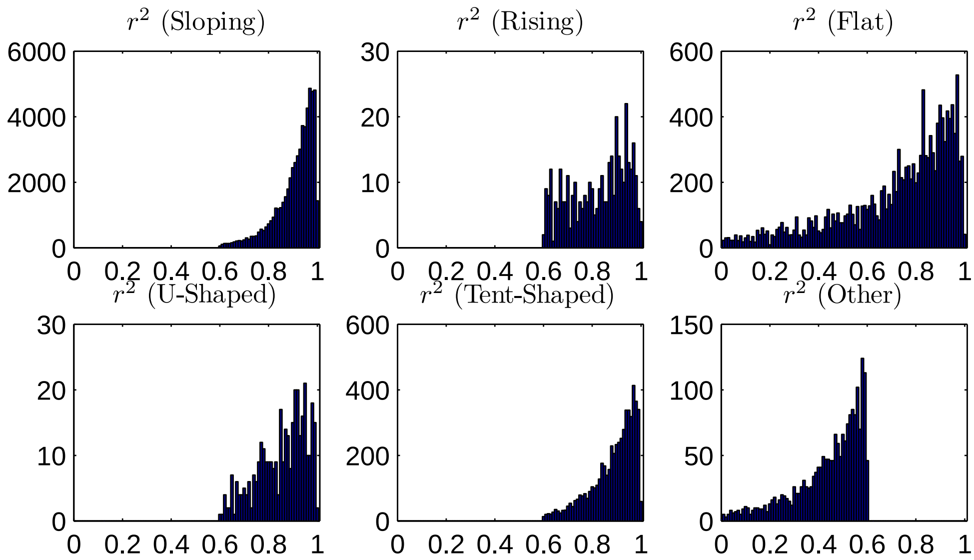

Figure 5.7: r^2 values (goodness of fit) of the regression curve fitting of the Beltone Connect dataset based on the alternative criteria.

and tent-shaped categories have decreased also. This can be contributed to the removal of audiograms which Pittman defined as u-shaped or tent-shaped while the fitted curve did not follow such a shape. Comparing figure 5.6 to figure 5.2, our sets have shifted a little at the u-shaped and tent-shaped category. The resulting r^2 values are presented in figure 5.7. We see that the other category does only contain values below 0.6, the flat category contains values between 0 and 1, because of the criterium of the standard deviation, and the remaining categories contain values of 0.6 and above.

5.3.2 Analysis of the curve fitting coefficients

We will use the Beltone Connect dataset for the analysis of the curve fitting coefficients. The Copenhagen data will be used to support our conclusions. First we take a look at the fitting of a curve to each audiogram. Based on the r^2 values, we continue our analysis with audiograms of all categories except for the "other" category. We have chosen this selection to remove outliers from the analysis, but still use audiograms which do not fit perfectly. This amounts to taking an r^2 threshold of 0.6. One can view the results of the regression as a probability density of regression coefficients $\Pr(\beta_0, \beta_1, \beta_2)$. To gain insight in this 3D density, we plot the marginal densities $\Pr(\beta_0)$, $\Pr(\beta_1)$, $\Pr(\beta_2)$ and $\Pr(\beta_0, \beta_1)$, $\Pr(\beta_0, \beta_2)$ and $\Pr(\beta_1, \beta_2)$ in figure 5.8. The marginal densities $\Pr(\beta_0)$, $\Pr(\beta_1)$, $\Pr(\beta_2)$ are plotted on the diagonal of figure 5.8. All three are roughly Gaussian. The β_0 coefficient (hearing loss at 1.4349 kHz) has a mean of 54.4 dB and an std of 12.2 dB. The β_1 coefficient has a mean of 8.1 dB and an std of 4.4 dB. The β_2 coefficient has a mean of 0.8 dB and an std of 1.1 dB.

Next we look at the off-diagonal plots $\Pr(\beta_0, \beta_1)$, $\Pr(\beta_0, \beta_2)$ and $\Pr(\beta_1, \beta_2)$, which again

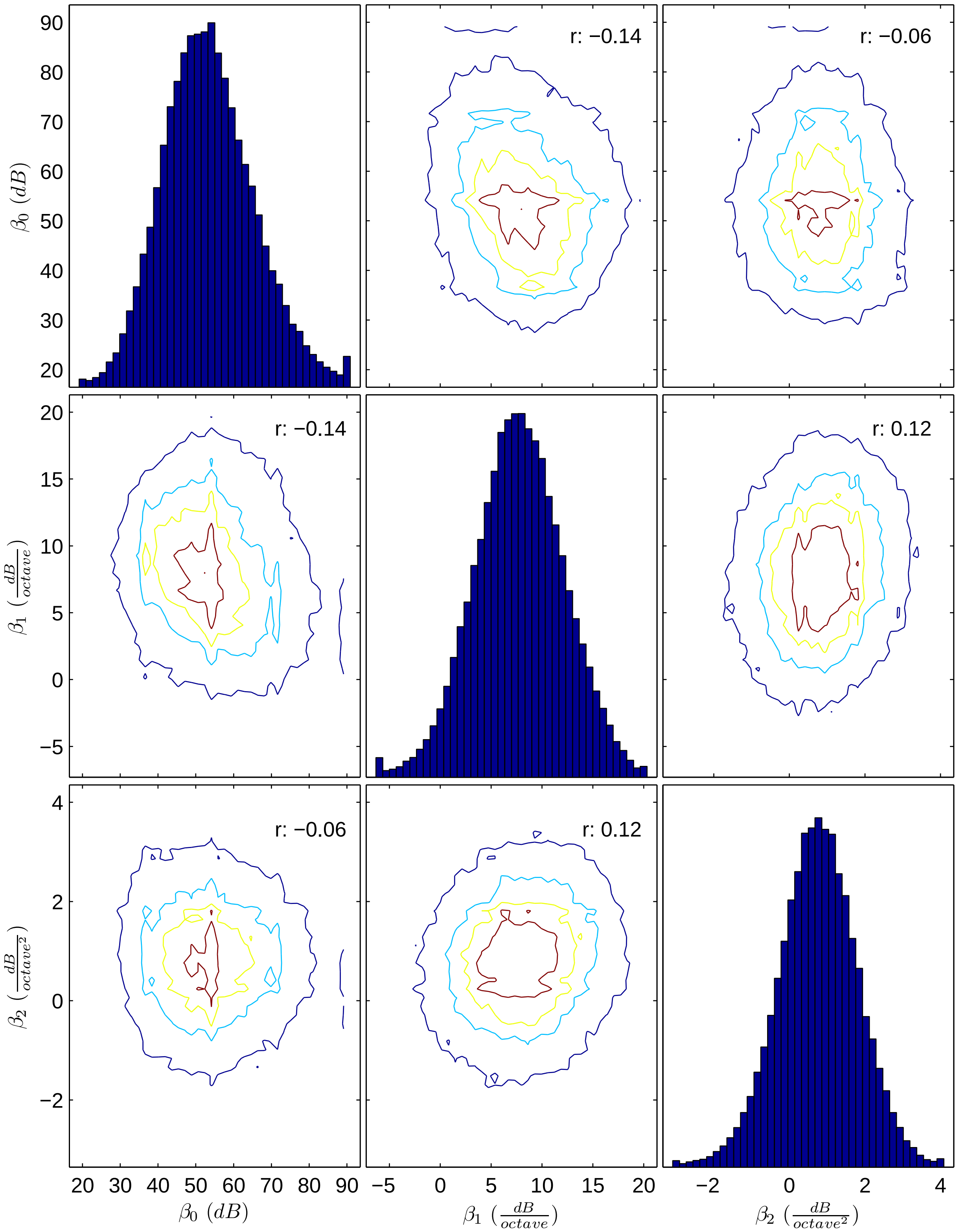

Figure 5.8: Regression coefficient matrix of the Beltone Connect dataset, with β_0 vs β_1 vs β_2. Each non-diagonal plot contains the correlation coefficient r between the two regression coefficients.

Age category	Size	$\bar{\beta}_0$ (dB)	$\bar{\beta}_1$ $(\frac{dB}{octave})$	$\bar{\beta}_2$ $(\frac{dB}{octave^2})$	r(β_0, β_1)	r(β_0, β_2)	r(β_1, β_2)
$age < 45$	51	60.1±19.8	5.8±5.6	0.0±1.4	-0.17	0.01	0.17
$45 \leq age < 65$	116	49.5±15.2	8.4±5.4	0.3±1.2	-0.01	-0.09	0.00
$age \geq 65$	46	49.8±12.5	10.8±4.5	0.8±1.0	-0.30	-0.37	0.17
Men	143	51.2±15.1	9.5±5.1	0.4±1.2	-0.10	-0.07	0.24
Women	75	53.4±18.5	6.2±5.6	0.2±1.4	-0.23	-0.21	-0.03

Table 5.2: Statistics of the β_0, β_1 & β_2 coefficients by age and gender.

look roughly Gaussian. In each plot we report the Pearson correlation coefficient for that particular pair of regression coefficients. The correlation between β_0 and β_1 is -0.14, which means there is a weak tendency for anti-correlation between the hearing loss at 1.4349 kHz and the slope. With a correlation coefficient of -0.06, the β_0-β_2 plot indicates that there is a very small negative relationship between the hearing loss at 1.4349 kHz and the curvature. The β_1-β_2 plot has a small positive correlation of 0.12, indicating a correlation between the slope and the curvature. Based on these results, we see that there is no evidence of different categories in the densities. Thus, in contrast to the analysis of Pittman, our results from a regression analysis do not show multimodal distribution of regression coefficients. Rather, the population of regression coefficients shows a smooth unimodal distribution. For the purpose of the rest of our analysis, we will omit the use of different categories. An analysis of the Copenhagen set (figure 5.9) shows the same relationships. This figure contains scatter plots instead of contour plots, because the size of the dataset is small enough to see relationships with this type of plot. The β_0 coefficient (hearing loss at 1.4349 kHz) has a mean of 52.3 dB and an std of 15.9 dB. The β_1 coefficient has a mean of 8.5 dB and an std of 5.5 dB. The β_2 coefficient has a mean of 0.3 dB and an std of 1.2 dB.

5.4 Analysis of personal information

The previous analysis is based solely on the hearing loss. The Copenhagen data contains information about the gender and age of a patient. Does this provide any more information for the type and shape of hearing loss? We have split up the Copenhagen dataset into three subsets based on the age, $age < 45$, $45 \leq age < 65$ and $age \geq 65$, and into two subsets based on gender, to draw conclusions about characteristics based on the gender of the patient. For each subset, we have calculated the mean of β_0, β_1, β_2 and the correlations between these coefficients and presented them in table 5.2. The average hearing loss at 1.4349 kHz ($\bar{\beta}_0$) and the standard deviation is higher when the patient is younger. The linear (β_1) and quadratic (β_2) components are smaller for this group of patients. Older patients have a relatively large negative correlation between β_0 and β_2. On average, the difference between men and women is not significant.

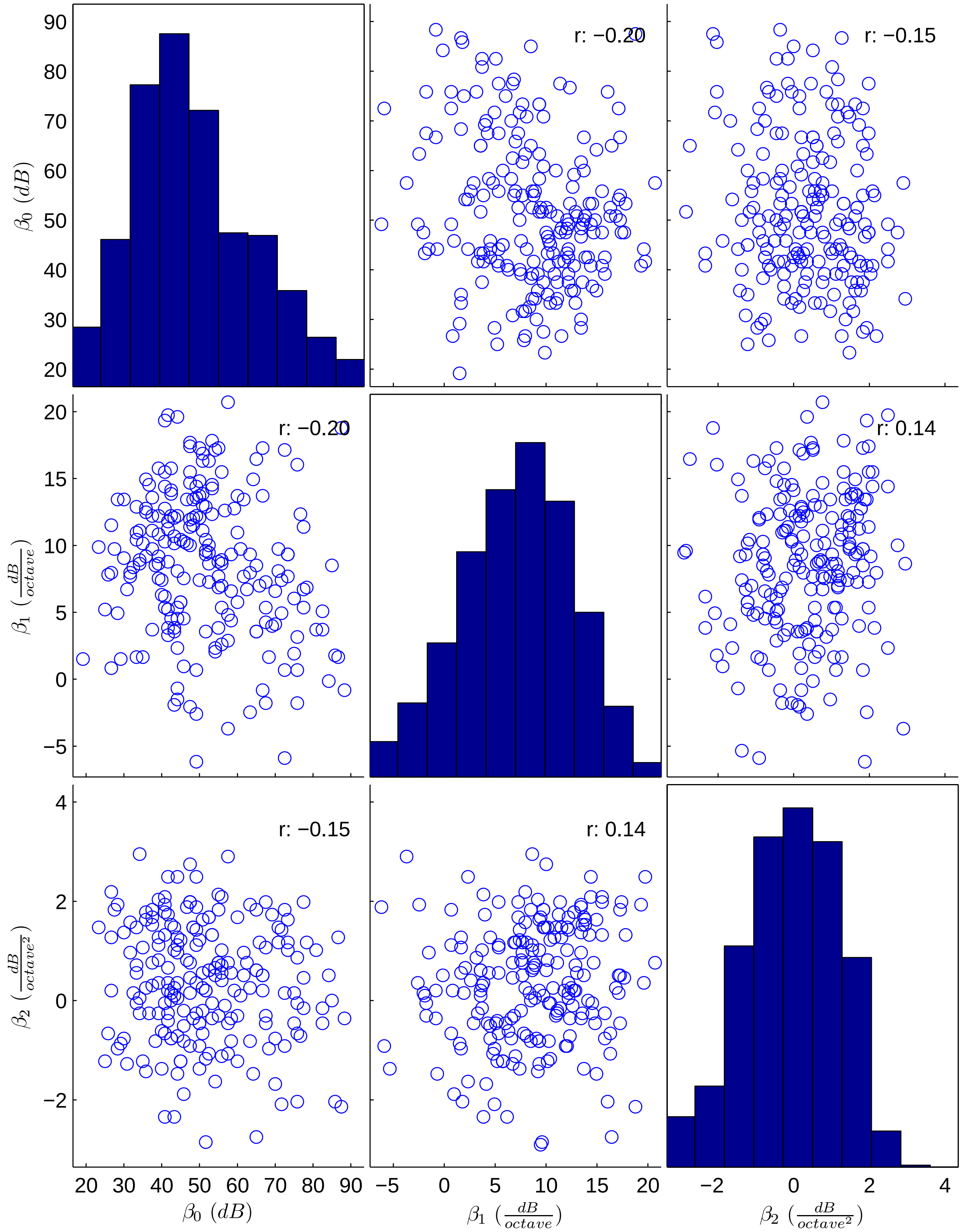

Figure 5.9: Regression coefficient matrix of the Copenhagen dataset, with β_0 vs β_1 vs β_2. Each non-diagonal plot contains the correlation coefficient r between the two regression coefficients.

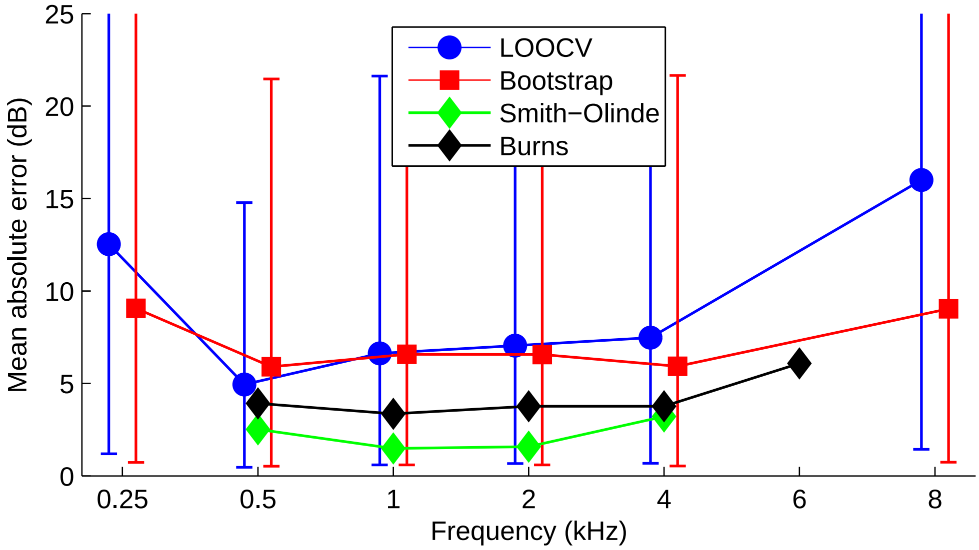

Figure 5.10: Cross validation of the Beltone (blue curve) and sampled (red curve) datasets with the mean absolute error and [5% 95%] interval errorbars for each frequency. Data points are shifted slightly for better visibility. Also provided are the test-retest reliability results by Smith-Olinde et al. [26] & Burns et al. [9]

5.5 Cross Validation

The r^2 model evaluation technique captures the overall fitting quality of the hearing loss. Leave-one-out cross validation is another model evaluation technique, which captures the fit quality for each audiometric frequency. We use this technique to evaluate how well the curve fitting technique can predict the hearing loss at the different frequencies. For each audiogram, we fit a curve based on all frequencies except one, predict the hearing loss at the missing frequency and evaluate the difference between predicted and observed value. This is done for each frequency. We will use the mean absolute error as evaluation metric as described in equation 3.25 to measure the average error for each frequency. In figure 5.10 we have plotted the mean absolute error with [5% 95%] interval errorbars for each frequency. We see that the error of the prediction at 8 kHz is larger than at the other frequencies, on average 1.5-2 times as large. The 0.25 kHz frequency is also a little harder to predict than the higher frequencies (with the exception of the 8 kHz frequency). The 0.5, 1, 2 & 4 kHz frequencies can be predicted with an average error between the 5 and 10 dB, well within the range of the test-retest reliability of 10 dB (see section 2.5). The Copenhagen dataset in figure 5.11 shows the same pattern.

The results from the cross-validation can have two sources: they can represent the characteristics of the dataset or they can represent the characteristics of the quadratic model. To investigate this, we use the parametric bootstrap (see section 3.3.2). In order to show if the results from cross-validation are characteristic to the datasets or to the model, we

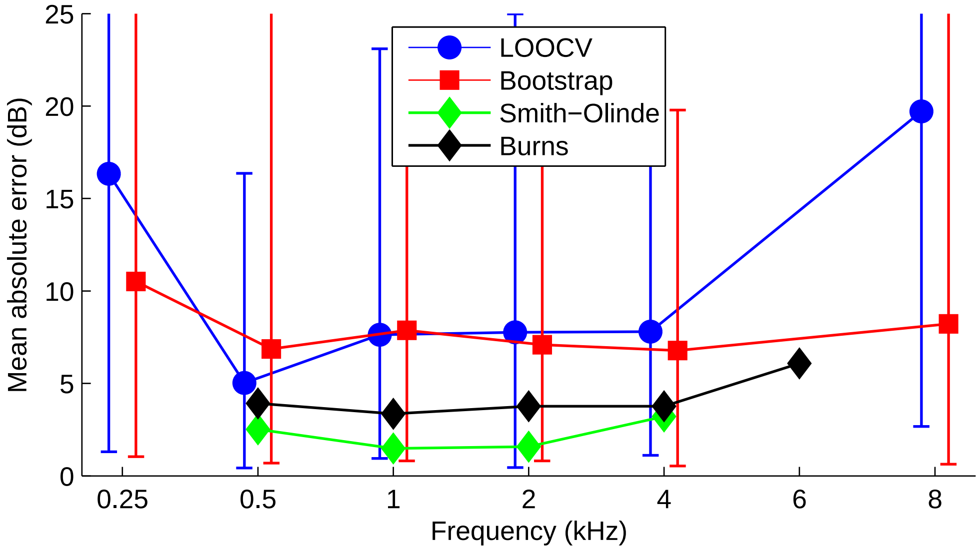

Figure 5.11: Cross validation of the Copenhagen and sampled datasets with the mean absolute error and [5% 95%] interval errorbars for each frequency and shape category. Also provided are the test-retest reliability results by Smith-Olinde et al. [26] & Burns et al. [9]

sample from each audiogram in the datasets a new audiogram, using equation 3.26. Then we use the cross-validation technique on these simulated datasets. If the results on these simulated sets are similar to the results of the cross-validation, the results are largely due to the model, otherwise due to characteristics of the datasets. For the Beltone set, the results of cross-validation on the simulated audiograms are presented in figure 5.10. The red curve is the cross validation performed on the sampled dataset. We see that at 0.5, 1, 2 & 4 kHz, the MAEs are similar to the cross validation performed on the Beltone dataset. At 0.25 kHz, the red curve shows a little lower MAE, while at the 8 kHz frequency, the difference is more than 5 dB. This shows that the unpredictability of the 8 kHz frequency is caused by the characteristics of dataset. The results for simulated audiograms from the Copenhagen dataset are presented in figure 5.11.

We have also plotted the MAE of the test-retest data presented by Smith-Olinde et al. and Burns et al. (see section 2.5.1). We see that the MAEs of the cross-validation of the Beltone dataset is higher than the data reported by Smith-Olinde and Burns. This indicates that there is room left to improve the performance of the quadratic model. But the results provided by Smith-Olinde and Burns indicate that the outer edges of the hearing range is less predictable, in line with our results. However, we must keep in mind that the test-retest results provided by Smith-Olinde et al. and Burns et al. are for healthy people, without a serious hearing loss and who are motivated to participate in the test. This could be of an influence on the results. Further tests can be done with test-retest reliability tests with people with a hearing disorder in order to support or contradict these test-retest results.

5.6 Summary of conclusions

Based on our analysis, we can conclude the following:

1. From figures 5.6 and 5.3, we conclude that the Beltone and Copenhagen data represent typical hearing impaired population.
2. All histograms of the regression coefficients are normally distributed, consistent in the assumption we made in equation 3.28. There is no evidence for discrete categories.
3. The three regression coefficients are weakly correlated, supporting the three-parameter model. If we had found strong correlations between regression coefficients, we would have support for a model with fewer parameters.
4. The correlations β_0-β_1 and β_0-β_2 are negative. This indicates the higher the hearing loss at 1.4349 kHz, the smaller the slope, respectively the curvature.
5. The correlation between β_1-β_2 is positive. This indicates the higher the slope, the higher the curvature. Is is expected that the age of patients is the cause. We cannot test this, because the Beltone dataset has no age and the Copenhagen dataset is too small.
6. There is no significant difference between men and women.
7. Younger people have on average a higher hearing loss at 1.4349 kHz (β_0). With the increasing age, β_1 and β_2 increase also. The patients in the Copenhagen dataset participated in a trial, so each patient has a known hearing problem. The younger patients in the dataset have a rather flat hearing loss, resulting in a small slope, while the elderly have a low hearing loss at the lower frequency, while it is rather high at the higher frequencies. This results in a larger slope.
8. LOOCV: the 8 kHz frequency is the least predictable hearing loss from a quadratic regression on the remaining HL data. Hence, 8 kHz should always be tested on any patient.
9. The quadratic model leaves room for improvement when comparing to the results of the test-retest reliability.

Chapter 6

Introduction to management thesis

This chapter will introduce the management part of this thesis. It will introduce the problem of adoption which arises with the introduction of the new method of fitting hearing aids. The new method is developed by the research department of GN Resound. In order to introduce this technique to the market, they need to market the product strategically. The target market are the audiologists, who will use this method to measure the hearing loss better and faster, so they can concentrate more on consulting the patient. Audiologists have some needs and expectations. The marketing department will try to market their product as the next best thing, while the audiologists do not see their needs fulfilled. In order to successfully market the product, the difference between the needs of the audiologists and the features provided by the product needs to be as small as possible. The marketing department needs to define their strategy based on the target market and the adoption rate of the market. This problem is relevant, because there is a significant part of the hearing impaired population which is not satisfied with their hearing aid device [16]. A product which can improve the measurement of the hearing loss and make room for better consulting, can lead to more satisfied customers.

First an introduction will be given on the problem of user satisfaction with hearing aids. While the population of hearing impaired is large, a significant part of them is not satisfied with the device. Counseling can lead to better usage of a hearing aid.

6.1 User satisfaction

For years, Sergei Kochkin (PhD, executive director of the Better Hearing Institute, Alexandria, Va.) reviews the current state of hearing aids in the US and publishes the results in MarkeTrak. In his publications, he covers significant trends in the hearing loss population, along with user (dis)satisfaction. Kochkin's latest MarkeTrak (MarkeTrak VII) [17] shows that a large number of hearing impaired do not use a hearing aid (see figure 6.1) , even though they are in need of one, with the numbers growing over the years. Several factors contribute to this, with user satisfaction being an important factor. Of the population

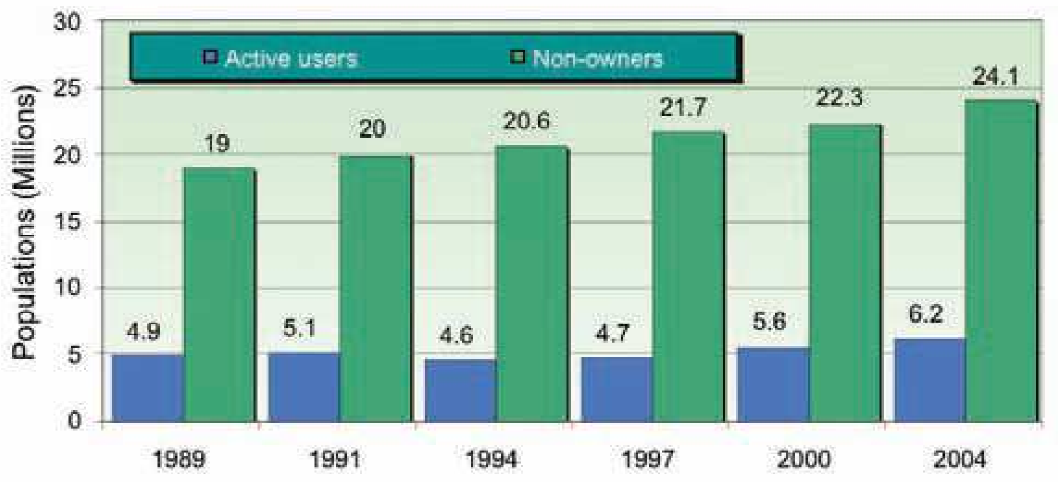

Figure 6.1: Kochkin active vs. non-owners. An active user is defined as a hearing instrument owner who uses their hearing instruments even if only once a year [17].

which does use hearing aids, user satisfaction leaves much room for improvement [16] (see figure 6.2). In 1997, the amount of dissatisfied hearing instrument owners was 19.4%, while the population that was neutral was 26.1%. Together, it sums up to more than 45% of the hearing impaired population with a hearing aid, who are not really satisfied with their device. While the percentage of dissatisfied population decreased over the years, most of it went to the neutral category. The percentage satisfied stayed steady over the years. But the percentage of hearing aids younger than 1 year old which are not used, has risen over the years. Considering the percentage of dissatisfied users, it is not surprising that 16.2% of the hearing impaired population does not use its hearing aid. Based on this analysis, there is a large group of hearing impaired who are not really satisfied with their hearing aid. There is a lot of room for improvement in the devices and the usage of them. Counseling can contribute to better usage of the hearing aid. But there should be time for that. The current measurement of hearing loss takes up time and could lead to loss of concentration by the patient. By reducing the time spent on the measurement, audiologists can spend more time on providing information, benefitting both the patient as the audiologist. The new measuring technique, which reduces the time spent on the pure-tone measurement and could prevent the loss of concentration during the process, can create more available time for the counseling of the patient, so patients know how to use hearing aids and what problems they can face and how to overcome them. This could lead to more satisfaction with the hearing aid.

Table 3. Satisfaction with Hearing Instruments

	1984	1989	1991	1994	1997
Satisfaction with Hearing Instruments		*(n=1,632)*	*(n=2,323)*	*(n=2327)*	*(n=2,720)*
Total owner population					
% Satisfied		58.5%	57.8%	53.4%	54.5%
% Neutral		22.6%	21.7%	26.6%	26.1%
% Dissatisfied		19.0%	20.3%	20.0%	19.4%
% hearing instruments in drawer (not used)	13.5%		12.0%	17.9%	16.2%
New hearing instruments (< 1 year)					
% Satisfied			66.3%	70.7%	63.1%
% Neutral			21.8%	22.7%	26.7%
% Dissatisfied			11.9%	6.6%	10.2%
% hearing instruments in drawer (not used)			3.0%	3.5%	4.6%
New hearing instruments (< 4 years)					
% Satisfied			60.7%	58.7%	59.3%
% Neutral			21.6%	25.0%	26.3%
% Dissatisfied			17.8%	17.3%	14.9%
% hearing instruments in drawer (not used)			7.7%	11.1%	8.8%

Figure 6.2: Kochkin satisfaction with hearing instruments [16].

6.2 Problem statement

The problem on which the management part of this thesis will be based, is the problem of market acceptance of a new hearing loss measurement technology. There are different factors involved in the adoption of technology. Each target market behaves in a different way. Health care is known for its somewhat conservative attitude towards new technology [15], usually leading to a slower adoption rate [7]. It needs to be researched how the market of audiologists respond to the new technology. The following research questions need to be answered:

1. What type of adopters are audiologists?
2. What factors influence their decision to use the new technology?
3. Is there a difference between audiologists in (a large chain of) stores and independent audiologists?

This problem exists, because over some decades, the techniques used to measure the hearing loss have not changed, while the equipment has progressed technologically. It seems that the equipment is getting faster and better, but the basic means to measure the hearing loss is at the same level as thirty years ago. To be able to market a new technology to the target market, we need to understand what type of adopters audiologists are and what attracts or retracts them from technological improvements. On one hand, we believe the satisfaction among hearing aid users can be improved with better consulting, because a relatively large part of the hearing aid users is not satisfied with the device and uses the device rarely [16]. On the other hand, current used methods won't allow many time to be spent on the consulting, because of the intensive task the patient has in the measurement procedure. Even though there are several (semi-)automated procedures developed in psychophysical literature over many years, like the von Békésy method (see section 2.1.1.3), they are not used widely. The question is why this has not happened. If we want to market this new technology, we need to understand the customers, the audiologists.

In this thesis, I will take a look at general factors (which are common in general markets) and will take a look at the way they influence the acceptance of the new fitting process. Which factors are involved in the market adoption of the new product? This will be done by an empirical study.

6.3 Research plan

In order to find answers to the research questions defined earlier, it is necessary to outline the research plan. This research plan introduces the theory and the empirical study to combine those two to form the answers. The management part of this thesis will be structured as follows: chapter 6 is this introduction to the adoption problem. The problems associated with the adoption of a new technology, like different categories of adopters and risks involved, are presented in chapter 7. There, three propositions are introduced on with the empirical study will be based. We will study which of these propositions hold and which of them fail, so we can gain insight in the adoption of new measuring technologies within the market of audiologists. The empirical study will be explained in chapter 8. Chapter 9 presents the results following the empirical study.

Chapter 7

Market adoption

A problem that arises with the introduction of a new technology or software, is the problem of market adoption. Market adoption is the theory of how and when people will start to use a new technology and which problems can arise with it [19]. There are several factors involved that have an effect on the decision of an audiologist to adopt the technology. This section will explain some of those factors.

There are several factors involved in the decision of adopting the new fitting product. It is important to understand what type of product you have and what the target market is. Adoption will only be a success if you know the characteristics of the customer. This involves understanding you customer, knowing what their needs are and how they think, knowing which type of adopter he is and which problems arise with competitors, like current software as NOAH or new types of automated fitting software and the costs involved for the customer when switching products. Understanding these factors make it easier to plan your marketing strategy and be able to adjust to certain market situations.

7.1 Understanding the customer

One of the key factors is understanding what the customer in the target market wants and how he thinks [19]. There are three critical issues that marketing must deal with in order to reach out to the customer. These issues are:

- What motivates a customer to buy the technology? What affects the customer to make a decision about the purchase?
- What affects the timing of customers' purchase decision? Are they planning on postponing the decisions in order to wait for new technologies in the near future? Are they postponing decisions until they must purchase a new technology?
- Are there different types of categories of buyers noticeable?

These issues are conceptually depicted in figure 7.1 ,which are specified towards high-tech

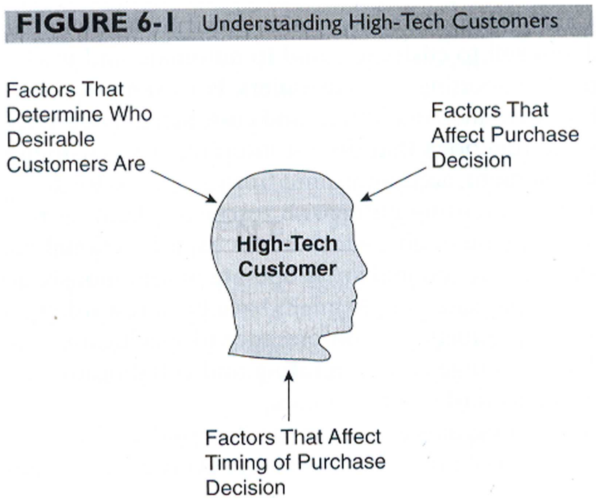

Figure 7.1: Three critical issues to understand customers.

customers, but are also applicable to understanding the market of audiologists. It is critical to understand these issues in order to speed up the adoption process of the technology. If a company knows what its target market will be and understands how the customers in it think, the whole marketing process can be fitted to their needs, pitfalls can be avoided and better services can be delivered. In the next sections, we will relate the presented factors to these issues.

7.2 Competition in the same market

An important factor which can influence the adoption of a technology is the presence of competitors [19]. Important is the size of the market which competitors control. If a particular company controls a large part of the potential market, then it could prove quite hard to compete with this. If all competing companies only control a small part of the market, it takes a different strategy in trying to gain market share. The competition in the market influences the factors that affect purchase decisions and the timing of purchase decisions, because if there are more competitors, customers will look around which company provides them the most value for their money, by product features and/or services. They also tend to postpone the decision to purchase based on the behavior of the market the companies operate in.

A software product which is used by many audiologists is NOAH[1]. This database is widely used and supported by many leading hearing aids companies, like GN Resound. Audiologists can use this product to measure and record the hearing loss of patients and passing the results to a common database, to use any fitting system provided for the NOAH soft-

[1] http://www.himsa.com

ware system. There are several versions of the software available, from the most basic fitting version to versions with more functions as billing and scheduling. As of January 2007, over 30.000 units have been sold worldwide.

One product which tries to improve the fitting of hearing aids is “Amplifit”, developed by Amplifon [5]. Hearing aid products in the market today have many variables which need to be set. Many of those are based on a algorithm which is found to be the best for that specific hearing aid device in field tests. The ideal would be that the algorithm is based on each individual patient. Current algorithms are not fit for an individual patient, but only for a larger group of patients, with large deviation between the individuals. Amplifit promises to fit a hearing aid based on the subjective input of the patient. With this product, the patient is more active in the process and therefore the product is based more towards the patient than to the product.

When introducing the new measuring software to the market, old software is likely to be replaced. With this comes the problem of switching costs, which will speed up the process of adoption when the costs are low or slow it down when the costs are high.

7.3 Switching Costs

An important factor which relates to the adoption of a technology is the switching costs [19, 8]. These costs are defined by the costs made from switching from one technology or brand to another. These costs can be split up in procedural switching costs, financial switching costs and relational switching costs.

Financial switching costs [8] consist of benefit-loss and financial-loss costs. This represents the loss of financial resources that comes with the purchase of the new technology.

Procedural switching costs consist of economic risks, costs from learning, evaluating and setting up the technology. Before a decision will be made to buy the technology, it has to be evaluated to see if it fits the requirements set by the customer. This procedure takes time, and time is money. A new technology, especially when it is radically new, needs time to be learned and needs time to be incorporated into the lives or company of the customer. If the usage of the technology involves many people, they need to be trained to control it and be able to use it. These all are costs on which a customer will make a decision.

Relational switching costs consist of loss of personal relationship and brand relationship costs. Emotional and psychological discomfort are involved when switching from technology or brand. This can prevent people to switch, because they like the service level or the brand name it carries. An unknown brand name can cause some discomfort, perceiving the switching costs higher than switch to the same technology, but from the better known brand.

The switching costs relate to the factors that affect purchase decisions. When the costs are perceived as too high, focus will be switched to competitors offering a same kind of technology, decisions can be postponed or even cancelled. This also involves the timing of the purchase. Postponing the decision to wait for better deals or lower switching costs (as the technology has time to prove itself) effects the timing decision, whereas focussing on competitors does not influence the timing decision right away.

The new product for measuring the hearing loss comes across this problem. The current software does its job well and is around for quite some time. Many audiologists use the NOAH software. The new measuring product needs to compete with this product and switching costs need to be relatively small in order to gain a market share. It also needs to compete with the relatively new product Amplifit. This product is set to improve the measuring of hearing aids and is developed by Amplifon, of which Beter Horen is part of. The Beter Horen stores will incorporate that product, so switching costs for those stores are expected to be higher, especially relational switching costs.

7.4 Perceived Risk

Perceived risk is defined as “a combination of uncertainty plus seriousness of outcome involved” [4]. This is involved in the adoption decision when the customer perceives some feeling of uncertainty, discomfort and/or anxiety, conflict aroused in the customer, concern, psychological discomfort, making the costumer feel uncertain, pain due to anxiety and cognitive dissonance. Perceived risk can be divided into seven types of risk facets. Table 7.1 (by Featherman [12]) will give an overview of these facets. In the following subsections, the risks that are the most important within this thesis will be further explained and will be related to the issues of understanding customers presented in section 7.1.

7.4.1 Performance Risk

A risk customers are always aware of, is the risk of a product failing to meet its expectations. One of the most important factors in considering a purchase, is that it must perform at least on par with the old technology and most of the time it should perform better. That means it should be able to have better performance, better service and/or better ease of use. A product will be purchased under certain expectations and the operations it will be used for are thus expected to be improved. New technology is expected to make the company perform on the same or better level than competitors, on a performance and/or cost efficiency. Failing to meet these expectations, or perceiving the performance risk as too high, affects the purchase decisions. Customers who perceive the performance risk tend to postpone the decision or even cancel the decision, which affects the timing of the purchase. Some consumers are more willing to take a performance risk, or do not perceive the risk as much as others, which will affect the decision and timing of the purchase.

Perceived Risk Facet	Description
1. Performance risk	"The possibility of the product malfunctioning and not performing as it was designed and advertised and therefore failing to deliver the desired benefits".
2. Financial risk	"The potential monetary outlay associated with the initial purchase price as well as the subsequent maintenance cost of the product". The current financial services research context expands this facet to include the recurring potential for financial loss due to fraud.
3. Time risk	Consumers may lose time when making a bad purchasing decision by wasting time researching and making the purchase, learning how to use a product or service only to have to replace it if it does not perform to expectations.
4. Psychological risk	The risk that the selection or performance of the producer will have a negative effect on the consumer's peace of mind or self-perception. Potential loss of self-esteem (ego loss) from the frustration of not achieving a buying goal.
5. Social risk	Potential loss of status in one's social group as a result of adopting a product or service, looking foolish or untrendy.
6. Privacy risk	Potential loss of control over personal information, such as when information about you is used without your knowledge or permission. The extreme case is where a customer is "spoofed" meaning a criminal uses their identity to perform fraudulent transactions.
7. Overall risk	A general measure of perceived risk when all criteria are evaluated together.

Table 7.1: The seven Perceived risk facets [12]

7.4.2 Financial Risk

Closely related to the performance risk is the financial risk. Most new technologies cost money to acquire. All technologies have some monetary costs to maintain them in operation. Customers are aware of the risk that comes with the purchase of a new technology. If it performs as expected, the financial risk is likely to be accepted. But what if a product cannot meet its expectations? Extra costs need to be made to maintain the operations on the level it performed before, like fixing costs or purchase of another technology. Customers have these kind of questions in mind when they are orienting on new technologies to acquire. Switching costs can contribute in perceiving this risk. If the expected win in performance is too little compared to the financial costs, the financial risk will be perceived more. This risk affects the purchase decision in that a perceived risk can contribute to the decision to purchase from a competitor of postpone or cancel the purchase. Postponing the decision also affects the timing of the purchase decision, as consumers which perceive the risk more tend to delay the decision more or wait for a situation when the decision has to be made.

7.4.3 Time Risk

Time risk is an aspect that involves the risk of wasting time when evaluating, buying and learning a product that does not meet the expectations or need by the customer. The current operations in a company can be interrupted and again time needs to be spent on evaluating, buying and learning the new technology. This risk can withhold customers to make a decision, because they don't want to make the wrong one. Procedural switching costs are involved with the time risk. Evaluating, installing and learning a new technology costs time. If the expected time is perceived as too much compared to the expected performance and costs, the time risk is perceived more. As with the performance and financial risks, the time risk affects both the decision to purchase and the timing of the purchase. If the time risk is perceived, the decision can be postponed or cancelled.

7.5 Adoption categories

One of the critical issues is the categories of adopters. These categories define the behavior of customers when it comes to adoption a new technology. It is important to understand which categories there are and how each category behaves, in order to know what can be expected of the target market customers. These categories can be roughly divided as innovators, early adopters, early majority, late majority and laggards (see figure 7.2). The timing of the purchase is depicted along the x-axis. Each of these categories will be explained in the following sections and will be related to the different factors affecting the adoption rate. These factors determine who desirable customers are, as seen in figure 7.1.

7.5.1 Innovators

Innovators are customers who buy a new technology, even though it is new and can contain some early problems. They are motivated by the fact that they can make a change in the adoption of the technology and love technology for its technology and thrive on the fact of being an innovator in the market. Because they are very early in the process of market entrance of the technology, they can come across some glitches or bigger problems, but they are willing to help the company to solve those problems. These type of adopters do not really perceive any of the described risks. Although they do see that there are risks

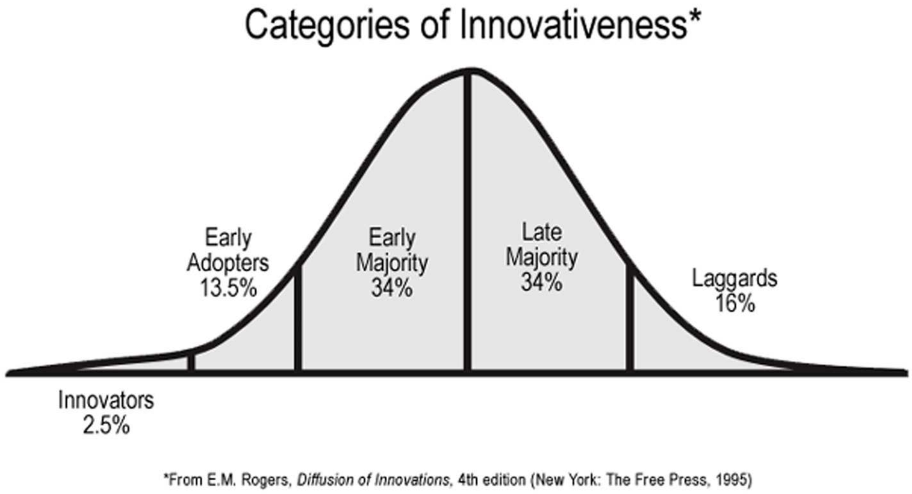

Figure 7.2: Five Different categories of adopters.

involved, they are willing to take them. If performance fails, they are more willing to accept it by helping the company to address the issues. New products are in the beginning often a bit higher prices, but this does not withhold innovators to decide to purchase the product (if the price is still reasonable). The timing of purchase is close to the introduction of the product. Failing to convince this category, may result in the failure of the product.

7.5.2 Early Adopters

Another category is the early adopters. They are looking for new technology to gain a dramatic competitive advantage over competing companies. They are looking for a high performance by looking at the performance risk, but do not really consider the financial risk. Because of the need of competitive advantage over competitors, the timing of the purchase is relatively quick. To get the advantage, the company needs to purchase it before other companies do it. Therefore, they are more willing to take the risks and their timing is right behind the innovators. Because of that, they are not very price sensitive and therefore typically demand a high service level for their needs and problems. But this category will bring real money to the table. For a company, it will be profitable to understand them.

7.5.3 Early Majority

Early majority, also known as pragmatists, is the next category which defines the market. This group is motivated by enhancements in the production cycle of the customers. They want technology that has proven itself and does not contain (many) problematic bugs, otherwise it could disrupt their normal operations, affecting the perceived performance risk. The performance risk affects the decision process more than the previous categories and decisions will be postponed or focused on other technologies if the risk is perceived as too high.

In order to reduce the risks involved with the adoption of a new technology, the early majority follows three principles:

- "When it is time to move, let us all move together". Because of this, adoption increases very fast at this stage. The performance, financial en time risks are reduced, as everyone moves almost together.
- "When we pick the vendor to lead us to the new paradigm, let us all pick the same one". A market leader will generally become from this principle. Performance and financial risks are reduced, since more customers decide on purchasing the product.
- "Once the transition starts, the sooner we get it over with, the better". This is why this stage occurs very quickly. The time risk is hereby reduced, as spending too much time on the transition decreases the chances of the product to be purchased.

It could prove quite difficult to convince this type of category. They are not likely to buy a new technology without an opinion from a trusted person. Because these two are mostly from the same type of category, it is necessary to think about how to convince one pragmatist, so that others can get references from him. This is important, because they form the bulk of the market.

7.5.4 Late Majority

The late majority are risk averse and technology shy. They base their opinion on pricing, as they are very sensitive to that, and on the maturity of the technology. The financial risk they perceive is interpreted higher than with the foregoing categories. If it shows some flaws, the odds are in favor of not acquiring the new technology. Performance must be almost perfect and any performance risk is considered as a decision to focus on competitors, postpone or cancel decisions. These people are motivated just to stay even with the competition and often rely on one trusted person to make their decisions. These factors lead to that the timing of purchases is relatively far from the introduction of the technology.

7.5.5 Laggards

The last category which can be distinguished, is the group of laggards. These people are skeptic of new technology and will only buy it they believe that all other options are worse and that the cost justification is absolute solid. Any performance, financial and time risk is considered as being a (major) disadvantage over the current used technology. Because of this behavior, this group is mostly left aside, because the profit will be low compared to the expenses that are made to convince these people.

As we can see, there are several types of adopters. Some of them tend to use new technology when it is available and do not experience many risks, while others are rather technology shy and base their decisions on the perceived risks. But the largest population is in the early majority and late majority category. In order to reach those categories with the new measuring technology, it is necessary to analyze the needs of these groups of audiologists and market the product towards these needs. The new measuring technology should be marketed towards the improved measurement of hearing loss and the reduced time for the measuring procedure, where focus is on the different risks which affect the adoption process. The success of adoption also comes with the competition in the target market.

7.6 Propositions

The performance, financial and time risks are some of the most important factors in the process of technology adoption. The remainder of this management thesis will consist of an empirical study of how these factors hold in the market of audiologists based on the new hearing aid measuring technology. To make this empirical study useful, we set up three propositions which we will answer to get an overview of the target market of the new measuring technology. The propositions are:

1. Why have the measuring methods remained the same over decades, while the equipment has advanced?
2. What is the relative importance of the risks when making a decision about a new measuring method?
3. Which risks will be most relevant in the adoption process of a new measuring method?

Chapter 8

Empirical Study

The management part of this thesis consists of an empirical study of the adoption rate of improved and automated measuring methods. The following questions are part of this study:

1. What type of adopters are audiologists?
2. What factors influence their decision to use the new technology?
3. Is there a difference between audiologists of a large chain of stores and independent audiologists?

This chapter will introduce the methodology which will be used to gather answers for these questions. For the sake of simplicity, audiologists and hearing aid professionals ("audicien" in Dutch) will both be referred to as audiologists, even though there is a difference between the two.

8.1 Methodology

There are two types of market research, qualitative and quantitative [18]. Qualitative research is an unstructured, primarily exploratory design based on small samples, intended to provide insight and understanding, while quantitative research seek to quantify data and, typically, apply some form of statistical analysis. This often requires a large group of participants. Qualitative research does not focus on generalization of the population of audiologists, like quantitative research does, but focuses on the validity and reliability of the different aspects of the decision making process by those audiologists. These aspects will not be generalized to the entire population, but provide insight in the different factors which influence the adoption of improved and automated measuring methods.

Because we are interested in some insight in the market of audiologists, we will apply qualitative research in order to answer the questions. This means that we will have a small sample (5 participants) of the population of audiologists, who will provide some insight in the factors. Note that it is not our intention to generalize over the population of audiologists, but to try to uncover different aspects which are involved in the decision process of purchasing the new measuring technology.

8.2 Participants

As we have said before, the number of participants will be small (5). To be able to gather as much relevant information as possible, we will ensure the population is heterogenous. The participants consist of audiologists of stores (independent and chain stores) as well as participants who work as independent audiologists. We are only interested in the experience the audiologists have with certain measuring methods. Opinions can influence the outcome of the study, by biasing the responses and not presenting the current situation. The first participant is an independent research audiologist, found in appendix A. The second participant is a hearing aid professional at an independent store, found in appendix B. The third participant is a hearing aid specialist working at a store of a large chain in a small town, found in appendix C. The fourth participant is a hearing aid professional working at a store of a large chain in a city, found in appendix D. The fifth and last participant is a “klinisch fysicus” audiologist working at an “Audiologisch Centrum” and teaching at Fontys Hogeschool, found in appendix E.

8.3 Research Topics

Several topics are involved in order to answer the research questions. We will pick five of these topics, on which the empirical study questions are based. These five topics are:

1. The experience of the audiologist.
2. The patient.
3. The education of the audiologist.
4. The supplier of the measurement equipment.
5. The improved and automated measuring method.

Because we are interested in the experience the audiologist has with several measuring methods, this is the first topic to be addressed. The focus is on the experience of the audiologist, because that is what they are doing best. Opinions are useful when we want to research the needs of the audiologists, but can influence the results when we study the current situation and the factors that influence the adoption of improved and automated measuring methods by combining. The second topic is the patient. The type of patients can influence the decisions made on the usage of measuring methods. Certain patients respond better to certain methods, while other methods are not suitable for certain patients. The characteristics of the patients can influence the decision to use a certain method and not take others. The third topic is the education the audiologist had. If the education focuses on specific types of measuring methods, while discarding others, the education can have an influence on the decision making. The fourth topic is the supplier of the measurement equipment. The supplier can have great influence in the decision making if the supplied equipment implements a certain type of measuring method. The four previous topics are about gathering information about different factors influencing the decision making. The fifth and last topic is about the improved and automated measuring method, where conclusions are being drawn towards the adoption of the new technique based on the factors and risks from the four other topics.

8.4 Empirical Study Questions

To conduct the research, it is necessary to formulate questions which are based on the research topics. These questions should be targeted to answer the research questions introduced in this chapter. Each questions is formulated with one of the risks discussed in chapter 7 in mind.

First we ask two questions to gather information about who the participant is and how long he or she has experience in the working field. At the same time, these questions will make the participant feel more comfortable when the questions about the different topics are addressed. The questions are:

1. Are you an independent audiologist or working in a (chain) store?
2. How many years of experience do you have?

The first topic is about the experience of the audiologist. Here we try to uncover which measuring methods the audiologist has experience with and why they use a particular measuring method. The questions are:

3. What is the experience of the audiologist with different measuring methods?
4. Why does the audiologist use the current measuring method(s)? What are the experienced advantages and disadvantages?
5. Has the audiologist considered alternative measuring methods?

The second topic is about the patient of which the hearing loss is measured. We try to uncover what the experience of the audiologist is with patients, what characterizes them and therefore if some measuring method is preferred above another. The questions are:

6. What type of patients are measured and what are their characteristics?
7. Does each type of patient have its own type of measuring method?

8. What are the positive and negative effects for the patients regarding the measurement?

The third topic is about the education the audiologist had. Which measuring methods are educated and do they include methods which could automate the procedure? The questions are:

9. What education does the audiologist have?
10. Which measuring methods were introduced and why those?
11. Does the audiologist have had extra training based on new measuring techniques?

The fourth topic is about the suppliers of equipment for the measuring task. What is their offer for the audiologist? Do the suppliers support training and does that training include demonstrations and training of new measuring technologies? The questions are:

12. How do(es) the supplier(s) support the audiologist?
13. What type of measuring equipment does the supplier supply? Does that include equipment for several types of methods?
14. Which education or training does the supplier provide?

The fifth topic discusses the improved and automated procedure and what the audiologists would like to see in a new procedure. The questions are:

15. Would you consider a new technique which is more automated than the current measuring method?
16. If there would be a new technique developed, what would you like to see that it can do?

Based on these questions, which of the risks discovered in the five topics are most important? Are there some risks which influence the decision making more than others and in what decision does that result? What is the most influencing factor and how does it relate to the other risks described in chapter 7? The results of the empirical study are described in chapter 9.

Chapter 9

Empirical Results

This chapter covers the results of the empirical study we've conducted among 5 audiologists. For each factor introduced in chapter 7 we will evaluate what the experiences of the audiologist are and what influence the factor has on the adoption rate. First we will pick the most relevant quotes from the interviews and cluster them based on the same kind of quote, so called long list. From these clusters we will investigate what similarities and differences there are between the different participants, so called short list. Based on this short list we will investigate what factors are most relevant and what type of adopters audiologists are by answering the questions presented in chapter 8.

9.1 Results of the empirical study

The empirical study is focused on finding the aspects involved with the adoption rate of a new measuring technique. In chapter 8 we have introduced three questions which will help us in finding these aspects. For convenience of the reader, we will repeat the research questions here:

1. What type of adopters are audiologists?

2. What factors influence their decision to use the new technology?

3. Is there a difference between audiologists of a large chain of stores and independent audiologists?

The answers to these questions give us an insight which aspects play a role in the adoption process of a new measuring technique. To answer these questions, we have formulated three propositions, on which the questions of the interviews are based. The propositions are:

1. Why have the measuring methods remained the same over decades, while the equipment has advanced?

2. What is the relative importance of the risks when making a decision about a new measuring method?
3. Which risks will be most relevant in the adoption process of a new measuring method?

Answering these propositions give us insight in the type of adopters audiologists are, which factors play a role in the adoption process, what their relevance is and if their are differences between different type of audiologists. From these answers, we will answer the research questions.

From the conducted interviews, found in appendices A to E, we have picked the most relevant quotes, grouped by topic, which indicate a relationship towards a certain factor. Each quote will be referenced to a specific audiologist and are presented in table 9.1.

1.	I have only experience with the Hughson-Westlake technique (staircase method). And it suits my needs well (A).
2.	My main experience is with the staircase method, the Dutch variant. I have tried a technique where the patient controlled the intensity level, but I did not really like it. Gave up too much control (B).
3.	Because I only have experience with the staircase technique, I use that one. I am used to this method, so it goes naturally. There is no need for another technique (C).
4.	My only experience is with the staircase technique. This technique works fine for the task at hands (D).
5.	I have experience with the Hughson-Westlake technique (staircase method) [...]. I have seen other techniques, like von Békésy audiometry, but I do not really have any real experience with them. Hughson-Westlake is the standard in the world (E).
6.	The advantages of this technique is that it is fast and commonly used. The disadvantage is that this technique does not consider the risks of distraction of the patient and thereby I have to refocus the patient. But this is probably common for each technique (A).
7.	It covers what it should do, meaning obtaining the hearing loss. It does not consume more time than needed. Maybe the only [disadvantage] is that sometimes patients can forget the instructions.(B).
8.	The more automated procedure had the disadvantage that I did not have full control over the procedure (B).
9.	It is a natural way of measuring (B).
10.	It is fast, does what it has to do to accomplish the goal and is a simple test. I can control the procedure. The disadvantage that I experience is that a patient tends to be distracted if it takes too long. Elderly tend to be more exhausted at the end than younger people (loss of concentration) (C).
11.	Patients find it nice to know right away what their hearing loss problem is (C).

12. For me it works the way it should and I can perform it rather quickly. The advantage is that it is evaluated well. [Patients] know it is okay what is done. I do not experience any disadvantage in the staircase method for myself. A disadvantage [for the patient] is the loss of concentration by elderly (D).
13. It is easy to learn for someone who does not have experience with it. I do not think there is a method which performs better than this technique. The disadvantage for me is that it seems necessary to find the threshold at each frequency, so it could sometimes take a little longer. Sometimes [patients] get confused [...] (E).
14. The advantage of the von Békésy technique is that the patient can control the procedure by itself, so there does not have to be someone with him all the time, leading to more efficient use of time. The downside [of the von Békésy technique] is that it is only suitable for screening, not for diagnostic research. It is also more difficult when masking is used (E).
15. I have excellent experience with the staircase method, so I do not have a reason to consider alternatives (A).
16. The reason I stick with this method is that I have other priorities (A).
17. I am working in a store, which is commercial and time is money (B).
18. The current method is good enough for me and the software provides me everything I would like (C).
19. There are no possibilities at the moment to even consider [alternatives]. The equipment is not ready for it (D).
20. The Hughson-Westlake technique is the standard technique used, so there is not really any need to consider the other techniques (E).
21. Maybe they spoke about [several techniques] in literature [...]. The practical part was only about the staircase method, because it was the most commonly used method (A).
22. I do this too little to really spent time on extra training (A).
23. The real usage of the techniques is only performed at the working place and they basically just use the staircase method (B).
24. During these courses, there was no mentioning of new/other techniques on pure-tone audiometry (B).
25. In the literature they explained several techniques, but in practice you do not use any of them except for the standard staircase method (C).
26. I have had some extra training, but it was not specifically about new measuring techniques (C).
27. Van Boxtel provides the possibility to (re-)learn when you ask for some extra training [...] and these are not about new measuring techniques (C).
28. In the literature they spoke about the staircase method and I think a more automated procedure [...]. But in my working experience during my education we only used the staircase method (D).
29. I have had extra training outside of my education, [...], not about measuring techniques (D).

30.	Hughson-Westlake, von Békésy, variants of the Hughson-Westlake, like ascending or descending methods. But in practice, the Hughson-Westlake method is used, simply because there is no real alternative and is used widely (E).
31.	As a teacher here in Eindhoven, I introduce them all, but notice that the most reliable method is the Hughson-Westlake method (E).
32.	None of the courses and congresses I attend are about new techniques (E).
33.	Several suppliers do [provide courses], primarily based on how to perform audiometry [with current methods] (E).
34.	[...] there is no need for other equipment from both a financial as a time perspective (A).
35.	Most of the equipment is quite expensive (B).
36.	The headquarters of van Boxtel provides me the equipment and the decision comes from them (C).
37.	I do not have real influence on the decision making (D).
38.	If we obtain new equipment, reliability is the most important aspect. If it performs better, more reliable, then the financial part is of less importance (E).
39.	I would consider it. But I am not really convinced that it is possible, that the current method cannot get really faster or more accurate (A).
40.	[...] I would be interested. But I think it will be quite difficult to introduce a new measuring technique, because health insurance companies need to acknowledge the new procedure, the politics has its own interest (B).
41.	[The new technique] should be really fast and make a big improvement over the current method, would I be considering such a technique (C).
42.	It would be interesting, but I would first like to see it before I would believe it (D).
43.	I would be interested, but only if it provides me guarantees about reliability, a certain amount of time efficiency and maybe some other factors as well. It could provide useful if used with young children or people with a mental handicap [...] (E).

Table 9.1: List of relevant quotes by the different audiologists found in appendices A through E.

Based on the relevance of the quotes for the different factors, a short list is created. For each factor, there are three different grades: positive, neutral and negative towards the risk. Positive means that the risk is not really perceived as a barrier, while negative means that the factor plays an important part in the decision making. Neutral is that the risk is perceived and can be of an influence, but there are also opportunities. Each quote is interpreted within these grades and is presented in table 9.2. In this table, the experiences are clustered together in similar relationships towards a factor. The staircase technique (also known as the Hughson-Westlake technique), or a variant on that technique, is used by each participant in pure-tone audiometry. The participants experience the staircase technique as a technique which fulfills their needs, does it fast and is reliable.

Position towards risk	Performance Risk	Financial Risk	Time Risk
Positive		38	
Neutral	14, 31, 38, 42, 43	36, 37	39, 40, 41, 42, 43
Negative	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 18, 20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 39, 41	17, 34, 35	16, 17, 19, 21, 22, 25, 28, 30, 34

Table 9.2: Short list of the interviews involving the different factors.

The performance risk is therefore perceived negatively as they generally do not think a new method can be much faster for the task they need to perform and the current technique works fine. Although there are disadvantages experienced with the staircase method, like distraction or loss of concentration, it is considered as being a part of each measuring technique. Some prefer the staircase method above more automated procedures like the von Békésy technique, because they do not want to give up control over the procedure. Knowing what the next step is in the measurement procedure is considered as an important aspect. One participant noted that there is some need in better performance of the measurement of young children or people with a mental handicap, where keeping the focus of the patient for some time is more difficult. Current procedures do not fully fulfill the needs. Education is often focussed on the staircase technique, because it is the standard technique.

The literature does provide several techniques, but the practical part of the education (learning by working at a hearing aid store or AC) provides only the standard technique. Therefore, often the only experience is gained with the staircase technique. This is considered as a negative time risk, because the companies which provide the learning experience to the students cannot really provide the time to gain experience with several types of techniques. Because the main focus is on the staircase method, other techniques are not introduced during the practical part of the education. When considering the suppliers, none of them offer courses or congresses related to new measuring techniques. This could indicate that they do not see a better technique than the current staircase technique, indicating a negative performance risk.

Many suppliers offer equipment for several measuring techniques, but are generally quite expensive. This withholds many audiologists to consider other technologies or techniques when the current technique does fulfill the needs of the audiologist. Today, many techniques are used on a PC, where the software incorporates many modules to perform different types of measurement techniques and other techniques. These modules are relatively expensive, contributing to the negative financial risk. Two audiologists, both from a large chain of stores, do not really consider the financial risk, as the headquarters of the respected companies decide on the used equipment and the audiologists themselves do not have a real influence on this. Therefore, those are considered as a neutral financial risk. The participants are somewhat divided when asked if they would consider a new measuring technique. Some are not convinced that there is a big improvement in speed possible (negative performance risk), while others are interested and willing to spent time on the new technology, but only if there are big improvements (neutral performance risk and neutral time risk).

One participant (E) sees possibilities to improve the performance of the current methods, but only with young children and people with a metal handicap. Table 9.3 presents the general thoughts described here. Quotes representing the same experience are presented in one or two sentences. Each of the general thoughts is accompanied with the appendices of the participants supporting the general thought.

Based on these interviews, we see that the audiologists are satisfied with the current measuring technology, because it is fast and reliable. That is why the measurement techniques have not progressed over the last decades. The equipment has made improvements just to make the process of measuring easier, but they have not changed the method itself. To answer proposition 1 in section 7.6: that is why the equipment has been improved and the measuring technique itself has remained the same for the last several decades.

We see that the participants are cautious in trying and possibly purchasing a new measuring technique. The performance of the staircase method is perceived as very high and fulfills the needs of the audiologists. The performance risk is the factor which influences the adoption rate the most, considering the experiences of the participants. The financial risk is also perceived, but differs between the audiologists. Independent audiologists perceive the risk negatively, because the have more influence on the decision making process, while chain stores are controlled by the headquarters and the audiologists in the stores do not have much influence. Although audiologists are interested and would spent time on evaluating a new technique, financial risks can hold them back to decide to purchase the technology. The time risk is also perceived, relating to the education they have had and their interest in evaluating a new technique. These three risks seem to be the most relevant factors for the participants, considering their answers. This answers propositions 2 and 3. Based on these results, we can answer the questions presented in chapter 8.

What type of adopters are audiologists?

When considering the adoption of a new measuring technology, it seems that audiologists are not eager to spent financial resources on new measuring technology, because the current method works fine. There are many other measurements and fitting of the hearing aid to be done, that they do not really worry about the used technique. Quote 13 represents

Attitude towards risk	Performance Risk	Financial Risk	Time Risk
Positive		Education: Reliability above finance (E).	
Neutral	Improvement: Interesting, only with big improvement (D, E).	Supplier: Chain stores headquarters decide on equipment (C, D).	Improvement: Interesting, but likes to see first (A, B, D, E). Better performance with children/mental handicapped (E).
Negative	Audiologist/Patient: Hughson-Westlake works fast en reliable. Staircase has no real disadvantage compared to other techniques (A, B, C, D, E). Education: practical part uses only the most common technique. No other techniques introduced within courses (A, B, C, D, E). Improvement: Not convinced of improving drastically (A, C).	Supplier: Equipment (hard- and software) is expensive (A, B).	Audiologist: Other priorities (A) or independent store (B). Education: Techniques in literature, not in practice, because of learning at hearing aid stores or ACs (A, B, C, D, E).

Table 9.3: General thoughts of the participants, based on the different topics and factors.

the general thought of the audiologists, stating “The advantage of the Hughson-Westlake technique is that it is fast, reliable and reproducible.”. Some indicate that if the technique will prove itself, they will consider using the new technique. There seems to be a sceptical reaction towards a new technique, quoting audiologist A “I would consider it. But I am not really convinced that it is possible, that the current method cannot get really faster or more accurate.” (quote 39), because of the easy and fast way the staircase method works. They first need proof that the technique performs better, either by recommendation by another audiologist or by seeing the technique perform at least equal to or even better than the currently used technique. Because a relatively large part of the hearing aid stores are part of a chain of stores, they tend to follow the principle of “move together”, by using the technique in more stores at the same time. On the other side, because of the scale of a chain, the financial consequences are rather large. This indicates that the target market of this new technique is somewhere between the early (chain stores, ACs) and late majority (independent stores).

What factors influence their decision to use the new technology?

The most important factor seems to be the performance of the technique. The participants are satisfied with the current technique and are therefore not eager to use another technique. It fulfills the task at hands fast and reliable and it is the standard technique in the world. Many audiologists use it, so the technique has proven itself. Hearing aid stores and ACs use the staircase technique as the standard technique and that is where students gain their first full experience with pure-tone audiometry. Another factor is that the participants perceive new technology and equipment as expensive. The switching costs, financial as well as procedural, are perceived as too high compared to the benefits they perceive to gain. Audiologist B states this as "I am working in a store, which is commercial and time is money" (quote 17). Spending a lot of money on new technology of which they do not think could improve the current method drastically withholds them from purchasing the new technology. Furthermore, the time needed to spent on learning the new technique is, especially in stores which have a lot of patients coming in, too time consuming. If the current method works, why would they spent time on learning a new technique? But the performance of the current method has the most influence on the decision to use a new technology. Only if a new technique performs much better, then they would consider it, both from a financial and time point of view. Otherwise, the purchase of a new technique will not occur. This indicates that the three factors are related to some extent.

Is there a difference between audiologists of a large chain of stores and independent audiologists?

Based on the population of the participants, there is a difference between audiologists of large chain stores and independent audiologists in the perceived financial risk. Independent audiologists seem to be more aware of the costs of equipment than audiologists of large chain stores. This does not mean that the latter do not perceive the risk, only that the decisions come from the headquarters. They decide which equipment and technology is used in the stores, so they will perceive the financial risk more (no matter to what extent). Other than that, there are no real differences. However, there is a difference between an audiologist from a store and from an AC. The latter is confronted with young children and patients with a mental handicap. The full measurement procedure takes more time and improvement within these measurements is needed. Audiologists from stores are not confronted with young children (stated by the law). Improvements in the measuring procedure needs to be targeted at first towards audiologists of ACs, where the biggest improvements can be gained.

9.2 Conclusions

The empirical study has provided some insight in the target market of the new measuring technique. To summarize these insights, we can conclude the following:

1. The participants are a type of adopters between the early and late majority, depending on if they are independent or chain store audiologists. They need to be convinced that the new technique is an improvement over the current techniques, overcoming the performance risk of not being able to get any faster than the current methods.

2. The performance risk is dominant in the decision of a new measuring technique. The staircase technique is the standard technique in the world, works fast and reliable, and is often the only technique used in the practical part of the education. Less important are the financial and time risk. To some extent these three factors are related to each other.

3. The main reason that the pure-tone audiometry technique has remained the same over the years, is that it does exactly what it should do and does it fast and reliable. The disadvantages are perceived as being a part of each measuring technique.

4. Independent audiologists perceive the financial risk more compared to audiologists of a chain store. In chain stores, the headquarters of the respective company will make the decisions based on which equipment will be used.

5. Most of the participants like to known about a new measuring technique, but only consider purchasing the technique when it performs better than the current techniques. Some participants think this is hard to achieve.

6. The biggest improvements can be achieved at ACs, where young children and patients with a mental handicap are measured. These patients a more difficult to measure, because it is harder to keep their focus on the task.

The implication of the participants being between the early and late majority, is that they will consider the technique when it has proven itself and base their decision on the pricing of the technology. Their are differences in the perception of the risks, where the performance risk is the most relevant. Much of their decisions on a new measuring technology will be based on this perception. For the concept of perceived risks it means that they are not equally important and base their decision towards one or two of the risks, depending on the type of audiologist. Their decision is made on that one (or two), making other risks seem less important. The performance risk overrules the other risks, making it the most relevant risk to overcome. The difference in perception of the financial risk between audiologists indicates that there are different sub-markets in the market of audiologists itself. The target market consists therefore of special conditions, making it that the perception of risks is focused towards one or two risks, with the performance as the most relevant risk.

Because the overall feeling is that the performance cannot be improved much and they perceive the costs of new equipment as high, they will not be convinced easily to purchase the product. The performance and financial risks are perceived as high and these risks are important in the decision process. For a minor improvement in the performance and an expected high price for the technology, the switching costs are considered too high to justify a purchase. This implicates that it will be hard to convince the audiologists to purchase the new measuring technology. Influenced by the combination of performance, financial and time risk, the concept of switching costs is important in the adoption process.

Since an opinion of a trusted person is valuable in the decision of the early and late majority, it is necessary to prove the technology in a market which is suitable for gaining a big performance increase. This market is the measurement of young children and mental handicapped and those type of patients are measured at ACs. Convincing these audiologists could provide an entrance in the market of hearing aid specialists, because the audiologists at ACs are generally considered as trusted. Convincing chain stores also improves the adoption rate, where the principle of "When it is time to move, let us all move together" could lead to a faster adoption rate.

These insights can be used to define the target market of the measuring technique and market it successfully. There are several practical conclusions which can be drawn from these insights:

1. Audiologists are a type of adopters between the early and late majority, depending on if they are independent or chain store audiologists. Because they want technology which has proven itself, market the new technology towards usage with the more difficult patients, like children and mentally handicapped. These types of patients benefit the most from the new technique, because keeping the focus of these type of patients is more difficult than regular patients. It is therefore vital to convince this market and prove the technology provides a performance gain and create a trusted person to gain entrance to other markets.

2. Market the product as an enhancement over the current (staircase) technique, where the staircase method is used as a basis for the new technique. This should create a point of recognition for the audiologists, reducing the time risk, as they do not have to learn a whole new technique, but can rely on their understanding of the staircase technique. It also gives them the feeling they can retain control in the procedure.

3. Audiologists experience the costs of new software and equipment as expensive. Because many audiologists use a software package like NOAH, the new measuring technology should be designed for these software packages, for example in the form of a plugin. This lowers the switching costs (financial as well as procedural), making it more interesting to purchase the technique.

Chapter 10

Outlook

This thesis presented an analysis of audiograms, which is useful for the bayesian framework, and an empirical study to find which factors influence the adoption rate of the new measuring technology. This thesis is only the beginning of constructing the new measuring technique and providing it to the target market. Further research is needed to be able to produce this new technique, both on the development of the technique as on the research of the target market.

Further research must concentrate on developing the bayesian framework further presented in chapter 3 and developing the algorithm to automate the measurement of hearing loss. In order to provide the best possible algorithm, the following topics should be addressed first:

1. Perform the analysis of personal information (sex, age) on a larger set than the Copenhagen dataset, possibly on the Beltone dataset if this information is possible to retrieve. Other information could provide useful, like the profession of the patient. Certain jobs can cause a hearing loss at specific frequencies. This information could possibly provide useful in automating the procedure.

2. Test the test-retest reliability on people with a hearing disorder. Current research has mainly focussed on patients without a hearing disorder. Results between these two groups of patients could be different, but is not necessarily so.

3. Improve the quadratic model, so it performs closer to the test-retest reliability.

4. Perform cross-validation leaving more than one frequency at the same time out.

5. Analyze the data using other models and compare these to the quadratic model. Choose the best performing model to develop the algorithm.

6. Analyze if the procedure can be used with other measurements, like SNR-loss, Most Comfortable Loudness and Loudness Discomfort Level.

The results of the empirical study show that the performance of the technique is the most important factor influencing the adoption rate. The financial risk is perceived differently by different participants. This empirical study has been very small, with five different participants. In order to fully understand the target market and find evidence supporting or contradicting these results, further research should be performed. The following topics should be addressed:

1. Use a larger, more diverse, population of participants.
2. Why do educations not teach how to use different techniques? Why only at the practical part, where the staircase technique is the standard?
3. Are there more subtle differences between audiologists, which could indicate more differences between groups of audiologists, possibly based on more factors?

These topics should be addressed in further research, providing more insight in the development of the algorithm and the target market of the new measuring technique.

Bibliography

[1] Acuna, E. & Rodriguez, C., (2004), The treatment of missing values and its effect in the classifier accuracy. *Classification, Clustering and Data Mining Applications*, 639-648

[2] Accredited Standards Committee S3, Bioacoustics, (2004), American National Standards ANSI S3.6-2004 Specification for Audiometers

[3] Accredited Standards Committee S3, Bioacoustics, (2004), American National Standards ANSI S3.21-2004 Methods for Manual Pure-Tone Threshold Audiometry

[4] Bauer, R., (1967), Consumer behavior as risk taking, *Cox, D. (Ed.), Risk Taking and Information Handling in Consumer Behavior*

[5] Amplifon, Een nieuwe aanpasprocedure met Amplifit

[6] British Society of Audiology, (2004), British Society of Audiology: Recommended procedure, Pure tone air and bone conduction threshold audiometry with and without masking and determination of uncomfortable loudness levels

[7] Burke, D.E., Wang, B.B., Wan, T.T. & Diana, M.L., (2002), Exploring Hospitals' Adoption of Information Technology, *Journal of Medical Systems*, 26(4), 349-355

[8] Burnham, T.A., Frels, J.K. & Mahajan, V., (2003), Consumer Switching Costs: A Typology, Antecedents, and Consequences, *Academy of Marketing Science Journal*, 31(2), 109-126

[9] Burns, W. & Hinchcliffe, R., (1957), Comparison of the Auditory Threshold as Measured by Individual Pure Tone and by Békésy Audiometry, *The Journal of the Acoustical Society of America*, 29(12), 1274-1277

[10] Cornsweet, T. N., (1962), The staircase method in psychophysics, *American Journal of Psychology*, 75, 485-491

[11] Dillon, H., (2001), Hearing Aids, *Thieme Medical Publishers*

[12] Featherman, M.S. & Pavlou, P. A., (2003), Predicting e-services adoption: a perceived risk facets perspective, *International Journal of Human-Computer Studies*, 59, 451-474

[13] Gescheider, G. A., (1985), Psychophysics. Method, Theory, and Application (Second Edition)

[14] Heskes, T. & de Vries, B., (2005), Incremental utility elicitation for adaptive personalization, *The 17th Belgian-Dutch Conference on Artificial Intelligence*, 127-134

[15] Jaeger, J., F. & Monteiro, E., (2005), Realizing organizational benefits with ICT in healthcare: the challenge of integration, *Proc. Continuity Care*

[16] Kochkin, S., (1999), MarkeTrak V: Baby Boomers spur growth in potential market, but penetration rate declines, *The Hearing Journal*, 52, 33-48

[17] Kochkin, S., (2005), MarkeTrak VII: Hearing loss population tops 31 million people. *Hearing Review*, 12, 16-29

[18] Malhotra, N., K. & Birks, D., F., (2007), Marketing Research, An Applied Approach (Third European Edition)

[19] Mohr, J., Sengupta, S. & Slater, S., (2005), Marketing of High-Technology Products and Innovations (2nd edition). *Upp Saddle River NJ: Pearson Prentice Hall.*

[20] Nederlandse Vereniging voor Audiologie, (2000), Nederlands Leerboek Audiologie, *http://www.audiologieboek.nl/*

[21] Pittman, A.L. & Stelmachowicz, P.G., (2003), Hearing Loss in Children and Adults: Audiometric Configurations, Assymetry, and Progression, *Ear & Hearing*, 24(3), 198-205

[22] Punch, J., Joseph, A. & Rakerd, B., (2004), Most Comfortable and Uncomfortable Loudness Levels: Six Decades of Research, *American Journal of Audiology*, 13, 144-157

[23] Punch, J., Joseph, A. & Rakerd, B., (2004), Effects of Test Order on Most Comfortable and Uncomfortable Loudness Levels for Speech, *American Journal of Audiology*, 13, 158-163

[24] Schmuziger, N., Probst, R. & Smurzynski, J., (2004), Test-Retest Reliability of Pure-Tone Thresholds from 0.5 to 16 kHz using Sennheiser HDA 200 and Etymotic Research ER-2 Earphones, *Ear & Hearing*, 25, 127-132

[25] Scott, D. W., (1985), Averaged Shifted Histograms: Effective nonparametric density estimators in several dimensions, *The Annals of Statistics*, 13(3), 1024-1040

[26] Smith-Olinde, L., Nicholson, N., Chivers, C., Highley, P. & Williams, D. K., (2006), Test-Retest Reliability of In Situ Unaided Thresholds in Adults, *American Journal of Audiology*, 15, 75-80

[27] Willeboer, C. & Smoorenburg, G. F., (2006), Comparing Cochlear Implant Users Speech Performance with Processor Fittings Based on Conventionally Determined T and C Levels or on Compound Action Potential Thresholds and Live-Voice Speech in a Prospective Balanced Crossover Study, *Ear & Hearing*, 27(6), 789-798

[28] Tan, P-N., Steinbach, M. & Kumar, V., (2006), Introduction to Data Mining

[29] Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D. & Altman, R.B., (2001), Missing value estimation methods for DNA microarrays, *Bioinformatics*, 17(6), 520-525

[30] Wasserman, L., (2004), All of Statistics, A Consice Course in Statistical Inference, *Springer*

Appendix A

Interview Research Audiologist A

1. Are you an independent audiologist or working in a (chain) store?
   I am a research audiologist working at GN Resound Eindhoven.

2. How many years of experience do you have?
   My experience is 12 years, of which 7 years here at GN Resound Eindhoven.

3. What is your experience with different measuring methods, like staircase and method of constant stimuli?
   I have only experience with the Hughson-Westlake technique (staircase method), none other. And it suits my needs well. For my work here I do not measure the hearing loss frequently and this method is in my experience fast.

4. Why do you use this technique? What are its advantages? What are its disadvantages?
   My education only taught me this technique to use in practice, and it was the only method to be put into practice. Therefore I have made this technique my own. Over the years, I have done things my own way and have not followed the official procedure strict. The advantages of this technique is that it is fast and commonly used. The disadvantage is that this technique does not consider the risks of distraction of the patient and thereby I have to refocus the patient. But this is probably common for each technique.

5. Have you considered alternative techniques?
   I do not have experience with other measuring techniques. I have excellent experience with the staircase method, so I do not have a reason to consider alternatives. I do not perform this measuring technique all day, but only a few times each month. The reason I stick with this method is that I have other priorities, so basically I do not want to spent time on understanding and trying other techniques.

6. What type of patients do you measure and what are their characteristics?
   In the past, I have measured the somewhat “difficult” type of patients, who have

specific types of hearing losses and more time has to be spent on them. Here at Resound most of the time I come across patients with “standard” hearing losses.

7. Does each type of patient have its own approach of measurement?
   For the complex types of patients, extra work has to be done, like masking. But overall, I measure the patients the same way.

8. What are the advantages and disadvantages for the patient regarding the used methods?
   The advantage for the patient is that they have undergone the procedure before, like in an Audiologisch Centrum (AC). The procedure is known to them. Another advantage is that pure-tone audiometry is overall fast. For each ear the time taken to perform air conduction, in normal circumstances, is about three minutes, I think. If masking is used, the time for each ear will extend to about five to six minutes for each ear. A disadvantage is that patients sometimes can lose the concentration which is needed to perform the task well. They can get distracted, especially in rooms which are not fully soundproof. If I see that happen, I have to retest the current threshold to make sure the patient did not gave false responses. But the retesting is just around the just found threshold, so it does not take that much extra time.

9. What education did you have?
   My education is speech therapy, speech and language pathology and “akoepedie”.

10. Which measuring methods did they introduce and why those?
    From what I remember, they taught us only the staircase technique. Along pure-tone measurements, they also taught us other tests, but not other techniques for pure-tone audiometry. Maybe they spoke about it in literature, but I cannot remember that. The practical part was only about the staircase method, because it was the most commonly used method. Maybe now they teach extra techniques, I do not know that, but when I studied, staircase was the only technique.

11. Did you have extra training or education after your regular education and did that provide other measuring techniques?
    I have not had extra education. The measurement I perform now has become a habit for me. But I do this too little to really spent time on extra training. Maybe it is different at an AC, because they come across the measurement daily.

12. How does your supplier of hardware and software support you?
    I do not have any experience with support for the hardware I use, so I do not really know.

13. What types of measuring equipment does the supplier supply? Does that include equipment for different measurement techniques?
    The supplier does supply different types of equipment, including equipment for different techniques, but I do not make any use of that. The reason is that I do not

perform pure-tone audiometry much, so there is no need for other equipment from both a financial as a time perspective.

14. Does the supplier provide extra training and does this include introduction and training of new techniques?
    Our supplier does not provide any extra training, at least not that I know of. The NVA (Nederlandse Vereniging voor Audiologie) does provide courses on audiometry occasionally, but I do not take any of this. Again, the reason is the occasional measurements and therefore less time.

15. Would you consider a new technique which is more automated than the current measuring method?
    I would consider it. I first would like to know how it works and would like to try it out to see if it provides faster measurements. But I am not really convinced that it is possible, that the current method cannot get really faster or more accurate. So yes, I am curious, but sceptical.

16. If there would be a new technique developed, what would you like to see that it can do?
    The most important part is that it should retain the control of the audiologist over the procedure. I would like to keep the close relationship with the patient and would like to know what the technique does. But again, I do not think that there is much improvement in the measurement time.

Appendix B

Interview Independent Hearing Aid Professional B

1. Are you an independent audiologist or working in a (chain) store?
   I'm working as a hearing aid professional at an independent store (Kooij Optiek in Nijmegen), where we sell glasses and hearing aids.

2. How many years of experience do you have?
   Currently I have 15 years of experience, of which 8-10 years full time. I started out as a dispensing optician. Throughout the years, I specialized myself in audiometry. My fascination grew over the years and with it, my experience grew.

3. What is your experience with different measuring methods, like staircase and method of constant stimuli?
   My main experience is with the staircase method, the Dutch variant. My education taught me this technique, therefore I use this one. As I have gained experience, I use my experience to perform this measuring technique by doing it my way, not strictly following the official procedure. I do not really have experience with other techniques. I have tried a technique where the patient controlled the intensity level, but I did not really like it. Gave up too much control.

4. Why do you use this technique? What are its advantages? What are its disadvantages?
   The advantage of the staircase technique is that you can get an audiogram fast. It covers what it should do, meaning obtaining the hearing loss. It does not consume more time than needed. I do not experience any disadvantages of this method. Maybe the only thing is that sometimes patients can forget the instructions, but this is common among many techniques. The more automated procedure had the disadvantage that I did not have full control over the procedure. I am the expert on this procedure, the patient is not. I known what I am doing, the patient does not. So I want to know what actions to take in the next step of the procedure. The more automated procedure does not provide me that.

5. Have you considered alternative techniques?
   Besides the automated procedure, I have not. I am working in a store, which is commercial and time is money. I cannot spent more time on evaluating other techniques. And because of the advantages of the staircase method, I do not really have to spent this time, even tough I would personally like that.

6. What type of patients do you measure and what are their characteristics?
   I come across patients from 7 months to approximately 95 years old. Young children will only be measured at UMC Radboud, because they are difficult to measure, and they come here for a hearing aid. Measuring children takes more time and since we are a commercial company, we do not have the time, money and the need to really make the necessary investments. The largest group of patients is between 55 and 80 years old.

7. Does each type of patient have its own approach of measurement?
   No, each patient gets measured the same. There are no real differences and it saves some time.

8. What are the advantages and disadvantages for the patient regarding the used methods?
   I do not think there are real advantages or disadvantages for the patient. It is a natural way of measuring. The measurement is black or white. The only answers are that a patient can hear the tone or not.

9. What education did you have?
   I have done the SBBO education at the ROC. It is a learning/working experience, where you learn one day in the week and work the other days. I have completed this education 7 years ago.

10. Which measuring methods did they introduce and why those?
    There were some different techniques treated in the literature at school, and we had the chance of using these technique once at school. But the real usage of the techniques is only performed at the working place and they basically just use the staircase method. The education does not provide that.

11. Did you have extra training or education after your regular education and did that provide other measuring techniques?
    I have done some courses in audiometry. We can get training constantly, by the Stichting Audiciens Register (STAR). During these courses, there was no mentioning of new/other techniques on pure-tone audiometry.

12. How does your supplier of hardware and software support you?
    They supply the hardware and software. Once a year they perform the calibration of the equipment, which is stated by the law.

13. What types of measuring equipment does the supplier supply? Does that include equipment for different measurement techniques?
    Everything is possible, but they mostly focus on equipment based on the standard staircase method. The supplied software is an all-in-one tool, with most measuring techniques integrated. Most of the equipment is quite expensive. Probably because the market of the equipment is relatively small, so they can make a lot of money.

14. Does the supplier provide extra training and does this include introduction and training of new techniques?
    They provide courses for everyone in the company. These courses are pointed towards audiometry and are given by independent people. They do not introduce other techniques then the ones I know of. They all look the same, so I do not go to everyone of them.

15. Would you consider a new technique which is more automated than the current measuring method?
    If it is simple to comprehend by the patient and I can retain the control over the procedure, I would be interested. But I think it will be quite difficult to introduce a new measuring technique, because health insurance companies need to acknowledge the new procedure, the politics has its own interest, so I think there are some barriers to cross.

16. If there would be a new technique developed, what would you like to see that it can do?
    The new technique has to be easy to comprehend by the patient and should not take too long to perform. I can have the control over the procedure. I want to know how it works.

Appendix C

Interview Chain Hearing Specialist C

1. Are you an independent audiologist or working in a (chain) store?
   I am an hearing aid specialist of the van Boxtel chain, located in Druten.

2. How many years of experience do you have?
   I have 2 years of experience here at van Boxtel. No experience at other companies.

3. What is your experience with different measuring methods, like staircase and method of constant stimuli?
   I only have experience with the staircase method. If I notice that the patient does not respond to the presented tone, I will occasionally present it somewhat longer. But this depends on the time spent already.

4. Why do you use this technique? What are its advantages? What are its disadvantages?
   Because I only have experience with the staircase technique, I use that one. I am used to this method, so it goes naturally. There is no need for another technique. The advantage is that it is easy to find the hearing loss in a short time. It is fast, does what it has to do to accomplish the goal and is a simple test. I can control the procedure. The disadvantage that I experience is that a patient tends to be distracted if it takes too long.

5. Have you considered alternative techniques?
   I have not considered any alternative. The current method is good enough for me and the software provides me everything I would like.

6. What type of patients do you measure and what are their characteristics?
   The most common patients are elderly, but also younger people come here. What is sometimes difficult is the acceptance of the patient the fact that he has hearing problems. Some patients are "forced" by their families, but do not think they have a hearing problem. Those are the ones that are a little harder to help. But most of them find it nice that the store is specialized in hearing aids and then they cross a barrier.

7. Does each type of patient have its own approach of measurement?
   Very young children are measured at an AC and not here. From the age you can interact with them, I can measure them in a more playful manner. But the rest of the patients are measured quite the same. But it depends on the impression I get from them. The approach will differ between patients based on that.

8. What are the advantages and disadvantages for the patient regarding the used methods?
   The advantage I see when I observe the patients (or they tell me), is that this measuring method is fast and they find it nice to know right away what their hearing loss problem is. A disadvantage for the patient is that in some cases it can take some time. Especially with the elderly there can be a loss of concentration or distraction, which influences the results and the time spent on the test. Elderly tend to be more exhausted at the end than younger people. If that happens, then you need to slow the procedure down, taking more time for the measurement.

9. What education did you have?
   I have done an internal education at van Boxtel. It specializes purely on the profession, taking away some of the courses which regular education provides. That is why I can call myself a hearing aid specialist, rather than a hearing aid professional. I do not have the proper education for the latter. I have always been interested in sound and have been using it since I was young. I intend to follow an audiologist education in about a year.

10. Which measuring methods did they introduce and why those?
    In the literature they explained several techniques, but in practice you do not use any of them except for the standard staircase method. Stores only use that one, so it stays with that.

11. Did you have extra training or education after your regular education and did that provide other measuring techniques?
    I have had some extra training, but it was not specifically about new measuring techniques, only about new products.

12. How does your supplier of hardware and software support you?
    Van Boxtel supplies me with the equipment. The supplier of the equipment for van Boxtel calibrates the equipment each year, which they have to do by law.

13. What types of measuring equipment does the supplier supply? Does that include equipment for different measurement techniques?
    I do not know that specifically. The headquarters of van Boxtel provides me the equipment and the decision comes from them. So maybe there is, but then I do not know about it.

14. Does the supplier provide extra training and does this include introduction and training of new techniques?
Van Boxtel provides the possibility to (re-)learn when you ask for some extra training, but there are no standard courses for everyone to take, and these are not about new measuring techniques.

15. Would you consider a new technique which is more automated than the current measuring method?
Depends on the price and how it performs. It should be really fast and make a big improvement over the current method, would I be considering such a technique. I think that it could be interesting for hospitals, where they are time limited. Hearing aid stores have more time and you take the time for each patient.

16. If there would be a new technique developed, what would you like to see that it can do?
I do not really have any needs. The current method works fine for me. It is fast, does not take much time. I have the control over it.

Appendix D

Interview Chain Hearing Aid Professional D

1. Are you an independent audiologist or working in a (chain) store?
   I work at Beter Horen in Nijmegen and I get my degree in September 2008. Then I become a hearing aid professional.

2. How many years of experience do you have?
   I have over one year of experience and I have been educated by Beter Horen.

3. What is your experience with different measuring methods, like staircase and method of constant stimuli?
   My only experience is with the staircase technique. I have not had training in any other techniques.

4. Why do you use this technique? What are its advantages? What are its disadvantages?
   It is the standard technique in my education. This technique works fine for the task at hands. That is the advantage of it. I do not experience any disadvantage in the staircase method for myself. For me it works the way it should and I can perform it rather quickly.

5. Have you considered alternative techniques?
   No, I have not. There are no possibilities here at the moment to even consider it. The equipment is not ready for it.

6. What type of patients do you measure and what are their characteristics?
   Mostly elderly, with hearing problems because of age. Sometimes children, but they come here with a reference from an AC and we fit a hearing aid. All goes according to the NOAH-protocol, which defines how a patient with hearing loss should be treated and by who.

7. Does each type of patient have its own approach of measurement?
   Each patient gets treated the same. For some elderly we take it a bit slower during the explanation of the procedure, possibly with visual support, so they understand the procedure as good as possible.

8. What are the advantages and disadvantages for the patient regarding the used methods?
   The advantage is that it is evaluated well, so the patient can be explained that the measuring procedure works well. Then they know it is okay what is done. A disadvantage is the loss of concentration by elderly. Within 10 to 15 minutes we should be able to measure the hearing loss. If it takes longer, the patient gets really tired and we make a new appointment for measuring. But this happens almost never. But the longer it takes, the more some elderly experience loss of concentration.

9. What education did you have?
   My education was at Beter Horen, the audiologists academy. It is the same as the SBBO education by the ROC, and we get four months of theory and the rest working experience.

10. Which measuring methods did they introduce and why those?
    In the literature they spoke about the staircase method and I think a more automated procedure where the patient pushed the button when the perception of the tone changes, while the intensity changes automatically. But in my working experience during my education we only used the staircase method, which is quite common.

11. Did you have extra training or education after your regular education and did that provide other measuring techniques?
    I have had extra training outside of my education, but they where about the hearing aid devices itself, not about measuring techniques.

12. How does your supplier of hardware and software support you?
    They provide us with the calibration of the equipment, the rest of the support is done by the headquarters of Beter Horen.

13. What types of measuring equipment does the supplier supply? Does that include equipment for different measurement techniques?
    We only have the equipment the headquarter provides us. I do not have real influence on the decision making, so I do not known about that.

14. Does the supplier provide extra training and does this include introduction and training of new techniques?
    None, only from within the education. The headquarters can provide that, but I cannot control that.

15. Would you consider a new technique which is more automated than the current measuring method?

It would be interesting, but I would first like to see it before I would believe it. At first I would use it together with the current technique. I think there could be a market for this, but I do not know exactly what it should take. Especially in large chain stores, where there are several different stores of.

16. If there would be a new technique developed, what would you like to see that it can do?
    What I would really like is that bone conduction would because easier. Currently it is quite mechanically sensitive and if it falls, you need to recalibrate it. And it is a little hard, because of surrounding sounds. I would like that to be better.

Appendix E

Interview Clinical Audiologist E

1. Are you an independent audiologist or working in a (chain) store?
   I am a “Klinisch Fysicus” (KF) audiologist working at an Audiologisch Centrum in Venlo. I am also a teacher Audiologie at the Fontys Paramedische Hogeschool in Eindhoven.

2. How many years of experience do you have?
   Since 1989, when I started with my education. I have worked at an AC and got my education there. I am an audiologist for about 15 years now. I am teaching since April 2007 here at the Fontys Hogeschool.

3. What is your experience with different measuring methods, like staircase and method of constant stimuli?
   I have experience with the Hughson-Westlake technique (staircase method), using 5 up-10 down. I have seen other techniques, like von Békésy audiometry, but I do not really have any real experience with them. Hughson-Westlake is the standard in the world.

4. Why do you use this technique? What are its advantages? What are its disadvantages?
   I am using this technique, because it is the standard technique of pure-tone audiometry. It is practically the only technique used to measure the pure-tone hearing loss. The advantage of the Hughson-Westlake technique is that it is fast, reliable and reproducible. It is easy to learn for someone who does not have experience with it. The disadvantage for me is that it seems necessary to find the threshold at each frequency, so it could sometimes take a little longer. But I do not think there is a method which performs better than this technique. The advantage of the von Békésy technique is that the patient can control the procedure by itself, so there does not have to be someone with him all the time, leading to more efficient use of time. The downside is that it is only suitable for screening, not for diagnostic research. It is also more difficult when masking is used.

5. Have you considered alternative techniques?
   No, I have not. The Hughson-Westlake technique is the standard technique used, so there is not really any need to consider the other techniques.

6. What type of patients do you measure and what are their characteristics?
   I come across very young babies (a few weeks old) to eldery. An AC has its focus on both the very young as the elderly. Young children may only be measured at an AC, where the experience is to coop with the hearing loss of those children. Other patients, who are known by a KNO-doctor, can be redirected to a hearing aid professional.

7. Does each type of patient have its own approach of measurement?
   With very young children we measure based on the reaction of the child. I see it as an art and you should have several years of experience before you really know how to measure these children. You have to present a tone at the right moment in order to get a reaction which is reliable. The current methods only give a global indication of the hearing loss. It is necessary to have some objective tests, like measurements based on brain reaction, to really measure the hearing loss of young children reliable, not on subjective basis. For the other types of patients, the used approach is the same. For elderly, it is sometimes better to focus on speech, if they cannot concentrate for a long time. The most important task is explaining the task at hands, so they know what to do. Counseling is very important.

8. What are the advantages and disadvantages for the patient regarding the used methods?
   The patients do not really experience a disadvantage of the procedure. The most important task is explaining the procedure, giving instructions. They do not notice when they do not hear sound intensities. It can be performed fast. It is important to know if a patient still is concentrated, did not fall asleep. If so, then refocus/reinstructions are needed. Sometimes they get confused, especially when masking is used and report a sound to be heard, when it is in fact a sound resulting from the masking. Sometimes, they are dissatisfied when a test has to be redone (after some time), because they do not notice a difference in hearing, but that is the only negative experience I get from them.

9. What education did you have?
   I have studied physics in Utrecht and then "Klinisch Fysicus" at the Vrije Universiteit Amsterdam.

10. Which measuring methods did they introduce and why those?
    Hughson-Westlake, von Békésy, variants of the Hughson-Westlake, like ascending or descending methods. But in practice, the Hughson-Westlake method is used, simply because there is no real alternative and is used widely. As a teacher here in Eindhoven, I introduce them all, but notice that the most reliable method is the Hughson-Westlake method. I have tried some other techniques in my classes, but

the differences in results are very little. Hughson-Westlake and masking are the most prominent techniques introduced, the others are not used on regular basis.

11. Did you have extra training or education after your regular education and did that provide other measuring techniques?
    I have followed some courses and attended congresses, but they were not specifically about audiometry. As a KF audiologist, we attend congresses provided by the Nederlandse Vereniging voor Audiologie (NVA), for which we can obtain points, of which we must have a certain amount each year. Once a year I go to an international congress. None of the courses and congresses I attend are about new techniques.

12. How does your supplier of hardware and software support you?
    At Fontys, the calibration is done by the supplier. We also train our students in calibrating the equipment, which I think is very important. At the AC, an internal technician does the calibration. If I have some questions for the supplier, they respond fast.

13. What types of measuring equipment does the supplier supply? Does that include equipment for different measurement techniques?
    The PC is used at Fontys, where the software implements several modules. We currently do not have a module for the von Békésy technique, but I think they could have such a module, but I am not sure. If we obtain new equipment, reliability is the most important aspect. If it performs better, more reliable, then the financial part is of less importance.

14. Does the supplier provide extra training and does this include introduction and training of new techniques?
    Several suppliers do, primarily based on how to perform audiometry. They give tips on how to do the measurements and are primarily focussed on audiometrists. They only include the current methods.

15. Would you consider a new technique which is more automated than the current measuring method?
    I would be interested, but only if it provides me guarantees about reliability, a certain amount of time efficiency and maybe some other factors as well. If it takes less time and provides more time to be spent on other measurements, it is certainly interesting. It could provide useful if used with young children or people with a mental handicap, where it is harder to get the full hearing loss. If the automated procedure can be completed in less time, it could certainly help. But it depends on the reliability of the procedure.

16. If there would be a new technique developed, what would you like to see that it can do?
    I would like to see how the parametric shift and tilt technique (Guido Smoorenburg & Christina Willeboer [27]) could perform on pure-tone audiometry, in order to

measure less data points, if it is possible. The parametric shift and tilt method is a technique aimed to improve the fitting procedure of cochlear implant processors by making use of measurements of the electrically evoked compound actio potential (ECAP) and live-voice speech. Maybe such a method can also be applied to pure-tone audiometry.