Objective Prediction of Pure Tone Thresholds in Normal and Hearing-Impaired Ears with Distortion Product Otoacoustic Emissions and Artificial Neural Networks

by

Rouviere de Waal

November 2000

In Partial Fulfillment of the requirements for the Degree

DPhil: Communication Pathology,

Faculty of Humanities,

University of Pretoria, Pretoria

Acknowledgements

The author is especially grateful to:

Prof. S.R. Hugo and Mrs. M.E. Soer

for the guidance and supervision provided throughout the completion of this research project

and

Ben de Waal,

who spent tedious hours of his precious time to program all the neural networks and other software that were used throughout the study.

Abstract

Title : Objective Prediction of Pure Tone Thresholds in Normal and Hearing-Impaired Ears with Distortion Product Otoacoustic Emissions and Artificial Neural Networks
Name : Rouviere de Waal
Promoter : Prof S.R. Hugo
Co-promoter : Mrs. M.E. Soer
Department : Communication Pathology, University of Pretoria
Degree : DPhil: Communication Pathology

In the evaluation of special populations, such as neonates, infants and malingerers, audiologists have to rely heavily on objective measurements to assess hearing ability. Current objective audiological procedures such as tympanometry, the acoustic reflex, auditory brainstem response and transient evoked otoacoustic emissions, however, have certain limitations, contributing to the need of an objective, non-invasive, rapid, economic test of hearing that evaluate hearing ability in a wide range of frequencies. The purpose of this study was to investigate distortion product otoacoustic emissions (DPOAEs) as an objective test of hearing. The main aim was to improve prediction of pure tone thresholds at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz with DPOAEs and artificial neural networks (ANNs) in normal and hearing-impaired ears. Other studies that attempted to predict hearing ability with DPOAEs and conventional statistical methods were only able to distinguish between normal and impaired hearing.

Back propagation neural networks were trained with the pattern of all present and absent DPOAE responses of 11 DPOAE frequencies of eight DP Grams and pure tone thresholds at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz. The neural network used the learned correlation between these two data sets to predict hearing ability at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz. Hearing ability was not predicted as a decibel value, but into one of several categories spanning 10dB.

Results for prediction accuracy of normal hearing improved from 92% to 94% at 500 Hz, 87% to 88% at 1000 Hz, 84% to 88% at 2000 Hz and 91% to 93% at 4000 Hz from the De Waal (1998) study to the present study. The improvement of prediction of normal hearing can be attributed to extensive experimentation with neural network topology and manipulation of input data to present information to the network optimally. The prediction of hearing-impaired categories was less satisfactory, due to insufficient data for the ANNs to train on. A prediction versus ear count correlation strongly suggested that the inaccurate predictions of hearing-impaired categories is not a result of an inability of DPOAEs to predict pure tone thresholds in hearing impaired ears, but a result of insufficient data for the neural network to train on.

This research concluded that DPOAEs and ANNs can be used to accurately predict hearing ability within 10dB in normal and hearing-impaired ears from 500 Hz to 4000 Hz for hearing losses of up to 65dB HL.

Key words: otoacoustic emissions, distortion product otoacoustic emissions, artificial neural networks, prediction of hearing threshold, age and gender, objective hearing assessment.

Opsomming

Titel : Objektiewe voorspelling van Suiwertoondrempels in Normale en Gehoorgestremde ore met behulp van Distorsie Produk Otoakoestiese Emissies en Kunsmatige Neurale Netwerke
Naam : Rouviere de Waal
Promotor : Prof. S.R. Hugo
Medepromotor : Mev. M.E. Soer
Departement : Kommunikasie Patologie, Universiteit van Pretoria
Graad : DPhil: Kommunikasie Patologie

In die evaluasie van spesiale populasies, soos neonate, kleuters en persone wat gehoorverliese voorgee, moet oudioloë dikwels steun op objektiewe metings om gehoorvermoë te evalueer. Huidige objektiewe oudiologiese prosedures, soos timpanometrie, die akoestiese refleks, ouditiewe breinstam respons en transient-ontlokte otoakoestiese emissies, het egter soveel tekortkominge, dat daar steeds 'n behoefte bestaan vir 'n objektiewe, vinnige en ekonomiese toetsprosedure, wat suiwertone in 'n wye frekwensiegebied evalueer. Die doel van hierdie studie was, om distorsie produk otoakoestiese emissies (DPOAEs) te ondersoek as a moontlike nuwe objektiewe gehoortoets. Die hoofdoel van die studie was om suiwertoondrempel

voorspelling te verbeter by 500 Hz, 1000 Hz, 2000 Hz en 4000 Hz met DPOAEs en kunsmatige neurale netwerke in normale en gehoorgestremde ore. Ander studies wat gepoog het om gehoorvermoë te voorspel met DPOAEs en statistiese metodes, was slegs in staat om tussen normale en gehoorgstremde ore te onderskei.

Neurale netwerke is opgelei met die patroon van alle aanwesige en afwesige DPOAE response van 11 DPOAE frekwensies en agt DP Gramme, sowel as suiwertoondrempels by 500 Hz, 1000 Hz, 2000 Hz en 4000 Hz. Die neurale netwerk het die geleerde korrelasie tussen die twee data stelle toegepas om gehoorvermoë te voorspel by 500 Hz, 1000 Hz, 2000 Hz en 4000 Hz. Gehoorvermoë is nie as 'n desibel waarde voorspel nie, maar in 'n kategorie met 'n grootte van 10dB.

Resultate het bevind dat voorspellingsakkuraatheid van normale gehoor verbeter het van 92% tot 94% by 500 Hz, 87% tot 88% by 1000 Hz, 84% tot 88% by 2000 Hz en 91% tot 93% by 4000 Hz van die vorige studie (De Waal, 1998) tot die huidige studie. Die verbetering in voorspellingsakkuraatheid kan toegeskryf word aan uitgebreide eksperimentering met neurale netwerk topologie en manipulering van inset data om optimale voorstelling van inligting in neurale netwerk opleiding te bewerkstellig. Die voorspellings van kategorie waardes by gehoorgestremdheid was minder bevredigend weens onvoldoende data vir die opleiding van die neurale netwerk. Die voorspellingsakkuraatheid versus die hoeveelheid ore in elke kategorie is met mekaar gekorreleer. Hierdie bevinding dui daarop dat die onvermoë om kategorië met gehoorverliese te voorspel, nie 'n tekortkoming van DPOAEs as suiwertoonvoorspeller is nie, maar 'n gevolg is van die onvoldoende data wat die neurale netwerk gehad het in die opleidingsfase.

Die gevolgtrekking van hierdie studie dui daarop dat DPOAEs en neurale netwerke gebruik kan word om gehoorvermoë binne 10dB akkuraatheid te voorspel in normale en gehoorgestremde ore, van 500 Hz tot 4000 Hz vir gehoorverliese tot en met 65dB.

Sleutelwoorde: otoakoestiese emissies, distorsie produk otoakoestiese emissies, neurale netwerke, voorspelling van gehoordrempels, ouderdom en geslag, objektiewe meting van gehoorvermoë.

Table of Contents:

Chapter 3: Parameters that Influence Pure Tone Threshold Prediction Accuracy with Distortion Product Otoacoustic Emissions and Artificial Neural Networks 63

List of tables:

List of Figures:

List of Appendices:

Abbreviations Used in this Study:

2f1-f2	:	The cubic distortion frequency in DPOAE testing
ABLB	:	Alternate Binaural Loudness Balance
ABR	:	Auditory Brainstem Response
AEPs	:	Auditory Evoked Potentials
ANNs	:	Artificial Neural Networks
ANS	:	Artificial Neural System
ART	:	Acoustic Reflex Threshold
BBN	:	Broadband Noise
dB	:	decibel
DP	:	Distortion Product
DP Gram	:	Distortion Product Audiogram
DPOAEs	:	Distortion Product Otoacoustic Emissions
EcochG	:	Electrocochleography
EEG	:	Electroencephalogram
EOAEs	:	Evoked Otoacoustic Emissions
f1 or f2	:	The primary frequencies in DPOAE testing
GM	:	Geometric Mean
HL	:	Hearing Level
Hz	:	Hertz
I/O Function	:	Input/ Output Function
L1 or L2	:	Loudness levels of the primary frequencies in DPOAE testing
LLR	:	Long Latency Response
MLR	:	Middle Latency Response
NPR	:	Negative Predictive Value

OAEs	:	Otoacoustic Emissions
OHC	:	Outer Hair Cells
PPV	:	Positive Predictive Value
PTA	:	Pure Tone Average
PTTs	:	Pure Tone Thresholds
SFOAEs	:	Stimulus Frequency Emissions
SLRs	:	Short Latency Responses
SOAEs	:	Spontaneous Otoacoustic Emissions
SPAR	:	Sensitivity Prediction with the Acoustic Reflex
SPL	:	Sound Pressure Level
TBOAEs	:	Tone Burst Otoacoustic Emissions
TEOAEs	:	Transient Evoked Otoacoustic Emissions

Chapter 1: Orientation and Statement of the Problem

1.1 Introduction

The quest for the development of an optimal objective diagnostic procedure to aid in the assessment of persons regarded as *difficult-to-test*, such as neonates and infants, has kept many researchers intensely occupied in the last two decades (Kemp, 1979; Tanaka, O-Uchi, Arai & Suzuki, 1987; Bonfils & Uziel, 1989; Lonsbury-Martin, 1994; Stover, Gorga & Neely & Montoya, 1996a; Koivunen, Uhari, Laitakari, Alho & Luotonen, 2000). Despite phenomenal advances in the ability to record electrical potentials generated at various levels of the nervous system and discoveries of active biological mechanisms in the cochlea, audiologists are still spending large amounts of time (Lee, Kimberley & Brown 1993; Vohr, White, Maxon & Johnson, 1993; Quinonez & Crawford, 1997) and money (Mauk & Behrens 1993; Weber, 1994) to attempt to evaluate difficult-to-test populations with equipment that allows only a limited frequency range of evaluation (Kemp & Ryan, 1993).

The purpose of this chapter is to present a brief overview of the ongoing struggle for the development of an optimal objective diagnostic audiologic procedure to aid in the assessment of difficult-to-test populations. This overview clearly indicates the need for a simple, cost effective, non-invasive yet accurate and objective method and states the reasons for difficulties experienced in this seemingly impossible quest. Furthermore, this chapter will present the purpose of this study, and plot a brief course of how the main objectives would be obtained. Lastly, this chapter will outline the objectives of following chapters to provide a more detailed description of the scope and objectives of the study.

1.2 The Origin of Objective Procedures

For many decades, diagnostic audiology relied on behavioral testing procedures in which hearing thresholds were determined by studying the listener's motor responses (Yantis, 1994). The first behavioral audiology test battery was developed in 1920 when bone conductors and speech channels became a standard feature included in an audiometer's capabilities (Brunt, 1994). For three decades, audiological tests were developed with only these basic features. Tests that were developed included the ABLB test for loudness growth to indicate cochlear pathology in the 1930s, the tone decay test to indicate retrocochlear pathology in the 1940s, and the SISI (Short Increment Sensitivity Index) for cochlear pathology in the 1950s (Brunt, 1994).

Prior to the 1960's, the evaluation of *difficult-to-test* populations, such as neonates, infants and very young children typically consisted of the observation of behavioral responses to an array of noise makers, such as bells, whistles, rattles, rustling paper and a spoon stirring in a cup (Ewing & Ewing, 1944). Hardy (1962) also used phonemes such as "S,S,S" presented behind the child to observe head turning responses but concluded that if the child does not respond, it can be suspected that something is wrong, although not necessarily that a hearing loss existed. Responses observed in neonates usually involved the moro reflex and for older infants localization responses and conditioned orientation reflexes (Suzuki & Ogiba, 1961). Before the development of objective diagnostic procedures in the 1970's, hearing tests to evaluate neonates and infants were not geared to evaluate pure tone thresholds objectively and accurately (Martin, 1984).

The first objective physiological procedure was developed in the 1970s. Progress in technology enabled audiologists to measure minimal changes in air pressure in the external meatus, which resulted in a completely new diagnostic tool, tympanometry. This allowed audiologists to obtain information not only about middle ear pressure and tympanic membrane movement but also about the stapedius reflex. Immittance testing enabled the Audiologist to obtain a variety of objective diagnostic functions, such as an indication of middle ear pathology, cochlear pathology when loudness recruitment is present, and retrocochlear pathology when reflex decay occurs. What is more important is that immittance procedures allowed audiologists to verify results obtained with behavioral audiometry objectively, for the first time. The prediction or evaluation of pure tone thresholds in difficult-to-test populations was still problematic. One application of the acoustic reflex, the SPAR-test (sensitivity prediction with the acoustic reflex), was developed by Jerger in the 1970s to predict hearing ability (Jerger, 1974). SPAR predicts hearing ability as normal, moderately impaired, or severely impaired. However, SPAR was influenced by a number of variables, such as chronological age (children between 0 and 10 are most accurately predicted), minor middle ear abnormalities, and audiometric configuration and still did not predict specific decibel levels for pure tones. Even though prediction of moderate hearing levels was only slightly better than chance, it offered a rapid estimate of hearing sensitivity useful for screening purposes (Northern & Gabbard, 1994).

The second development toward objective audiological measurements also occurred in the 1970s when audiologists began to measure the electric potentials of the nervous system with surface electrodes. Auditory evoked potentials (AEPs) occur in different time intervals after stimulation and provide information about the cochlea (EcoghG), auditory nerve (auditory brainstem response (ABR)), and brainstem (middle latency response (MLR) and late latency response (LLR)). AEPs measure the integrity of the auditory pathway at certain sites and are not a test of hearing such as pure tone measurement, which evaluates the entire auditory system (Cope, 1995). However, AEPs not only enabled audiologists to confirm behavioral test results but also to measure auditory status at certain sites in difficult-to-test populations. (Robinette, 1994). ABR is still currently the preferred objective diagnostic method in the evaluation of neonates and very young infants (Weber, 1994) despite limitations such as a limited frequency area that can be evaluated (Kemp & Ryan, 1993), lengthy test times, expense and expertise required for diagnostic applications and the unfortunate possibility of sedation in the case of very young children (Hall III & Mueller III, 1997). The quest for an optimal procedure to determine pure tone thresholds in difficult-to-test populations rapidly, economically, accurately and across a wide range of frequencies continued.

At the end of the 1970s, another objective way to evaluate hearing ability was discovered by David Kemp (1978), namely otoacoustic emissions—that is, energy generated by a normal cochlea either spontaneously or in the presence of acoustic stimulation. Kemp's (1978) original reports were greeted rather skeptically, and much early research only replicated his study to confirm the presence of otoacoustic emissions. After two decades of intensive research, however, there is currently much excitement among researchers, since certain types of otoacoustic emissions such as transient otoacoustic emission (TEOAE) prove to be highly applicable in the areas of hearing screening and distortion product otoacoustic emissions (DPOAEs) in screening and even diagnostic audiology (Kummer, Janssen, & Arnold, 1998; Martin, Probst, & Lonsbury-Martin, 1990b; Stach, Wolf, & Bland, 1993; Stover, et al. 1996a).

Otoacoustic measurement will certainly never replace pure tone audiometry, immittance, or ABR, but OAEs reveal diagnostic information regarding the auditory system that is not available from any other test and offers the possibility of objective evaluation of pure tone thresholds in special populations. Many researchers hoped that this relatively new field in audiology would prove to be the long-awaited objective, rapid, and accurate test of auditory function (Kemp, 1990; Lee, et al. 1993; Kimberley Hernadi, Lee & Brown 1994b; Danhauer, 1997).

The hope that OAEs are the optimal new objective test of hearing arose despite facts that suggest that this prospect may be impossible. OAEs measure the functioning of the outer hair cells (OHC), which apparently is only involved with the amplification of sounds and fine-tuning of the cochlea to specific frequencies (Kemp, 1978). It is the inner hair cells that receive 95% of the afferent auditory nerve fibers to carry auditory signals to the brain (Dallos, 1997). Furthermore, pure tone threshold estimation involves measurement of the entire auditory system, outer-, middle- and inner ear as well as the auditory nerve, brainstem and cortex. OAEs evaluate OHC functioning only. Despite these fundamental differences in the two procedures, many researchers found a significant correlation between one type of emission, the distortion product otoacoustic emission (DPOAE) and pure tone thresholds (PTTs) (Probst & Hauser, 1990; Lonsbury-Martin, Harris, Stagner, Hawkins & Martin, 1990; Avan, Elbez & Bonfils, 1997; Nieschalk, Hustert & Stoll, 1998). Good correlation between DPOAEs and PTTs was also found for degree and configuration of hearing loss (Martin, Ohlms, Franklin, Harris, Lonsbury-Martin, 1990). These findings fueled expectations that DPOAEs could possibly predict pure tone thresholds objectively (Kemp, 1997).

Many researchers studied the correlation between DPOAEs and PTTs (Martin et al. 1990a; Lonsbury-Martin et al. 1990; Probst & Hausser, 1990; Gorga, Neely, Bergman & Beauchaine, 1993) and some attempted to predict PTTs as normal or hearing-impaired with DPOAEs (Kimberley, Kimberley & Roth 1994a; Kimberey et al 1994b; Moulin et al 1994; De Waal, 1998). Several difficulties were experienced in the development of DPOAEs as a diagnostic procedure: Firstly, the distortion product has numerous stimulus parameters that should be carefully chosen to ensure optimal measurement and many researchers investigated optimal levels and ratios for input stimuli (Gaskill & Brown, 1990; Avan & Bonfils, 1993; Stover et al. 1996a; Mills, 1997). Secondly, the measured responses are often "noisy", incomplete, (Probst & Hauser, 1990; Stover et al. 1996a) idiosyncratic (Lonsbury-Martin, Harris, Stagner, Hawkins, & Martin, 1990) and complex which makes the determination of the non-linear correlation between DPOAEs and PTTs even more difficult (Ruggero, 1993; Kemp, 1997; Nakajima, Mountain & Hubbard, 1998). Thirdly, the non-linear correlation between the two data sets of which one is noisy and incomplete as well as the numerous variables that influence that correlation require a special data processing technique that is capable of predictions of pure tone thresholds based on the determined correlation between the data sets (Kimberley et al. 1994a). Many researchers used conventional statistical methods such as multivariate discriminant analysis (Kimberley et al. 1994b; Vinck, Cauwenberge, Corthals & De Vel, 1998) and struggled with the noisy incomplete data, non-linear correlation and number of variables involved. Fourthly, prediction of normal hearing at frequencies lower than 1000 Hz proved to be problematic for most researchers due to the rising noise floor caused by subject artifacts such as breathing and swallowing (Gorga et al. 1993) and limitations in measuring equipment (Kemp, 1997). Lastly, the research needed to fully understand the fine structure of the distortion product and the significance of other distortions as the cubic distortion tone (2f1-f2) in pure tone prediction is only scraping the surface (Kemp, 1997). The quest to predict pure tone thresholds with DPOAEs objectively, rapidly, economically and accurately is apparently still limited to categorization of hearing ability as normal or impaired at frequencies above 1000 Hz (Gorga et al. 1993; Kimberley et al. 1994a; Kimberley et al. 1994b; Stover et al. 1996a; Kummer et al. 1998a; Zhao & Stephens, 1998).

A new form of information processing, called artificial neural networks (ANNs), was applied to this problem (De Waal, 1998; Kimberley et al. 1994a) that excel in dealing with noisy, complex or incomplete data sets (Nelson & Illingworth, 1991), non-linear correlations (Rao & Rao, 1995) and predictions based on learned correlations (Blum, 1992).

Kemp (1994a) found the neural network approach to be more effective in the classification of normal and impaired hearing than multivariate statistical methods. De Waal (1998) predicted normal and impaired hearing at frequencies ranging from 500 Hz to 4000 Hz and predicted normal hearing as low as 500 Hz objectively with 92% accuracy for the first time. The fore mentioned study also indicated that it is probable that pure tone thresholds could be predicted for impaired hearing within 10dB for hearing loss up to 65dB HL if the neural network has enough data to train on. The application of the neural network approach to this field of diagnostic audiology revealed that distortion product otoacoustic emissions are suitable as a diagnostic test of hearing to evaluate pure tone thresholds in normal and hearing-impaired ears objectively, accurately, rapidly, economically across a wide range of frequencies.

However, certain difficulties were encountered in the De Waal (1998) study that could have influenced prediction accuracy negatively. These included:

- The search for optimal neural network topology such as desired number of middle level neurons, input data manipulation and error tolerance levels.
- Variables included in neural network training such as DPOAE amplitude that made the neural network so complex that no convergence was possible and the question of specificity of which the age variable can be presented to the neural network.
- Optimal definition of DPOAE threshold, to ensure that all present responses are valid yet no valid responses are discarded.
- Which aspect of the distortion product to best correlate to pure tone thresholds- the f1 frequency, f2 frequency, 2f1-f2 frequency or a combination of frequencies.

1.3 Purpose of This Study

Against this background it seems possible to develop DPOAEs as an objective pure tone predictor with the use of artificial neural networks to aid in the assessment of difficult-to-test populations.

The purpose of this study is therefore to further develop DPOAEs as an objective diagnostic test of hearing by addressing the problems experienced in the De Waal (1998) study to improve prediction accuracy of pure tone thresholds from 500 Hz to 4000 Hz with DPOAEs and artificial neural networks. Optimal parameters for DPOAE measurement will be identified with an in-depth literature study and applied to measure DPOAE responses in subjects with normal hearing, various degrees of sensorineural hearing loss and of various ages. The correlation between DPOAEs and pure tone thresholds (PTTs) will be studied with the use of artificial neural networks. There will be extensively experimented with optimal neural network topologies, error tolerance levels, manipulation of input data and the inclusion of DPOAE amplitude for neural network training. The effect of DPOAE threshold defined as 1, 2 or 3dB above the noise floor, the effect of the inclusion or omission of noisy low frequency data on prediction accuracy and specificity of which the age variable should be presented to the neural network will be investigated. The main purpose of this study is to improve prediction accuracy of PTTs at 500, 1000, 2000 and 4000 Hz with DPOAEs and ANNs

1.4 Outline of the Thesis

Chapter One: Orientation and Statement of the problem

This chapter provides a brief overview of the development of objective tests in audiology, formulate the need for an accurate, cost effective procedure and delineate the purpose of this study, which is to further investigate DPOAEs as a diagnostic test of hearing by improving prediction accuracy of PTTs with DPOAEs and artificial neural networks (ANNs).

Chapter Two: The Quest for Optimal Pure Tone Threshold Prediction

The second chapter will focus on current objective diagnostic procedures for pure tone evaluation purposes, motivate their need and discuss their limitations in the evaluation of certain populations. This chapter will also concentrate on specific prerequisites that objective tests should have as well as the data processing techniques used to develop these tests, to be considered effective and efficient. Special attention will be given to the history of the use of distortion product otoacoustic emissions (DPOAEs) to predict pure tone thresholds (PTTs), the limitations of previous studies and reasons for their struggle.

Chapter Three: Parameters that Influence Pure Tone Threshold Prediction Accuracy with Distortion Product Otoacoustic Emissions and Artificial Neural Networks

The third chapter will be a discussion of all the parameters that influence prediction accuracy of PTTs with DPOAEs and ANNs. The discussion will be divided into two segments, first, parameters that influence the distortion product in the recording, analysis and interpretation of measurements and second, parameters that influence the neural network in the designing-, training-, prediction- and analysis phase of operation.

Chapter Four: Research Methodology

Chapter four will be a discussion of the methodology for data collection, preparation and analysis, apparatus, subjects, the research design and the procedures chosen for optimal neural network functioning.

Chapter Five: Results

Chapter five will present results of all experiments: Pure tone threshold prediction accuracy at 500, 1000, 2000 and 4000 Hz, and the effect of subject-, DPOAE-, and ANN variables experimented with to determine effects on prediction accuracy.

Chapter Six: Discussion and Interpretation of Results

Chapter six will discuss and interpret all findings in terms of significance as well as readiness for broad clinical use. A few interesting case studies will also be discussed.

Chapter Seven: Summary, Evaluation of the Study and Conclusion

The last chapter will evaluate this study in terms of validity, reliability and limitations and make recommendations for future research.

References

Only references mentioned in this thesis appear alphabetically in the reference list.

1.5 Conclusion

> "It should be clear that in spite of the technical advancements made in detecting [DPOAEs], the relation between DPOAEs and sensorineural hearing loss is not fully understood. There is no physiological basis for assuming that DPOAE measures ought to perfectly correlate with pure-tone threshold, even in the case of a purely cochlear hearing loss. Currently available empirical evidence, however, suggests that the general relationships between pure-tone threshold and DPOAE will serve as an important tool for the audiologist" (Kimberley, Brown & Allen, 1997:201).

Chapter 2: The Quest for Optimal Objective Pure Tone Threshold Prediction

2.1 Introduction

David Kemp (1978) first described otoacoustic emissions (OAE) from the human ear and ignited a tremendous interest in these measurements to develop another objective diagnostic test of hearing. These relatively easy measurable active responses from the cochlea to sound stimulation, due to the basilar membrane's natural ability to amplify sound and tune in to specific frequencies, have kept many researchers occupied in the last two decades (Elberling, Parbo, Johnsen & Bagi, 1985; Bonfils, Avan, Francois, Marie, Trotoux & Narcy, 1990; Brass & Kemp, 1994; Kossl & Boyan, 1998). The main interests of most studies were to attempt to categorize pure tone sensitivity with these measurements as normal or impaired (Kimberley, et al. 1994b; Hurley & Musiek, 1994; Kimberley, et al. 1994a) or to gain more information regarding the site-of-lesion in diagnostic audiology (Tanaka, et al. 1987; Ohlms, Lonsbury-Martin & Martin, 1990; Robinette, 1992; Moulin, Bera & Collet, 1994). Most researchers, however, found it extremely difficult or even impossible to predict impaired pure tone thresholds (PTTs) or to categorize hearing ability at low frequencies as normal or impaired with distortion product otoacoustic emissions (DPOAEs)(Gorga, et al. 1993; Kimberley et al. 1994b; Stover, et al. 1996a; Zhao & Stephens, 1998). This unsatisfactory prediction of PTTs with DPOAEs is probably due to the large number of DPOAE stimulus parameters that influence optimal measurement (Bonfils, Avan, Londero, Trotoux & Narcy, 1991; Gorga, et al. 1993), the complex nonlinear nature of the measured responses (Nakajima, et al. 1998; Kummer, et al. 1998) and the inability of conventional statistics to address this problem sufficiently (Kimberley et al. 1994a).

This seemingly impossible quest to predict pure tone thresholds (PTTs) accurately with DPOAEs arises not from the need to replace existing conventional behavioral evaluation procedures, but to aid in the assessment of pure tone sensitivity in difficult-to-test populations such as neonates, infants, malingerers and the crucially ill (Balfour, Pillion, & Gaskin, 1998).

The purpose of this chapter is to evaluate current objective diagnostic procedures available in the assessment of pure tone thresholds in difficult-to-test populations, to identify their need and limitations and formulate the requirements for an optimal objective diagnostic procedure. The history of the development of DPOAEs as a pure tone predictor will be reviewed extensively, and evaluated to identify limitations and postulate reasons for their struggle. Prerequisites for optimal data processing techniques in the development of an objective diagnostic procedure will also be formulated.

The following section will discuss objective procedures in audiology to better understand the need for an optimal non-invasive, rapid, accurate, simple and cost-effective test to aid in the assessment of difficult-to-test populations.

2.2 Objective Diagnostic Procedures

What is meant with an *objective* test, why does audiologists have to rely heavily on these measures when assessing *difficult-to-test populations* and who is regarded as difficult-to-test?

2.2.1 Overview and Definition of Terms

An *objective test* according to Cope (1995) requires no voluntary response from the patient that an auditory stimulus was perceived. Even though the patient may still influence the results by interfering with the procedure, the subjectivity is transferred to the clinician who interprets the results. Objective tests are not a measurement of hearing as such, but evaluate the integrity of the auditory pathway at various levels, never in its entirety. The value of objective tests becomes apparent in the evaluation of subjects who cannot participate in conventional behavioral audiometry to respond to auditory stimuli voluntarily. Subjects who are too young, critically ill, subconscious, mentally incapable of providing cooperation or subjects who refuse to cooperate for whatever reasons, are all considered to be *difficult-to-test* and it is in the evaluation of these populations that audiologists have to rely heavily on objective evaluation procedures (Balfour, et al. 1998).

2.2.2 Current Objective Procedures Available to Predict Pure Tone Thresholds

As was stated in the definition of an objective test in the previous paragraph, objective tests measure the integrity of the auditory pathway at various levels and is not a measurement of hearing such as pure tone audiometry. Objective tests such as acoustic immittance measurements evaluate the mobility of the middle ear when air pressure is varied with tympanometry, or the ease of flow of energy through the middle ear in static acoustic immittance, or the lowest intensity needed to elicit a muscle contraction in the middle ear with acoustic reflex measurements (Block & Wiley, 1994). These measurements are valuable in site of lesion testing in diagnostic audiology and provide useful information regarding certain components of the auditory system but do not provide information about hearing sensitivity as such (Cope, 1995). Certain objective measurements however, have been adapted in an attempt to predict pure tone thresholds, such as the SPAR, ABR and some types of OAE measurements. Since the scope of this chapter is to discuss the quest for optimal objective pure tone threshold prediction, only those objective procedures that attempt to predict pure tone thresholds will be discussed in more detail.

2.2.2.1 Sensitivity Prediction with the Acoustic Reflex (SPAR)

SPAR was developed by Jerger in 1974 (Jerger, Burney, Mauldin & Crump, 1974) to predict hearing ability into categories as normal, moderately impaired, or severely impaired. This technique uses the difference between the thresholds of pure tone acoustic reflexes at 500, 1000 and 2000 Hz and broadband noise reflexes to predict hearing sensitivity. Normal hearing ability was accurately predicted 100% of the time when the 1000 Hz acoustic reflex threshold (ART) is 95 dB SPL or less and the noise-tone reflex difference is more than 20 dB. Severe hearing loss was accurately predicted 85% of the time when the broadband noise (BBN) threshold is more than 95 dB SPL. Moderate hearing losses were predicted accurately 54% of the time when the noise-tone ART difference is less than 20 and the BBN reflex threshold is 95 dB SPL or less. However, SPAR is influenced by a number of variables, such as chronological age (children between 0 and 10 are most accurately predicted), minor middle ear abnormalities, and audiometric configuration. Even though prediction of moderate hearing levels is only slightly better than chance, in difficult-to-test populations SPAR can often offer a rapid, economical, and objective estimate of hearing sensitivity and is also useful in screening (Northern & Gabbard, 1994).

2.2.2.2 Sensitivity Prediction with Auditory Evoked Potentials

The measurement of auditory evoked potentials (AEPs) involve the extraction of tiny electrical amplitudes of the auditory system from larger signals such as electroencephalographic (EEG) activity and other general muscular activity by using surface electrodes and signal averaging techniques.

Auditory evoked potentials (AEPs) occur in different time intervals after stimulation and provide information about the cochlea, auditory nerve, and brainstem. AEPs are usually classified by their "latency epoch," the time domain within which the response occurs after stimulus onset (Ferraro & Durrant, 1994). AEPs occurring in the first 10–15 milliseconds are known as short latency responses (SLRs). SLRs include the auditory brainstem response (ABR) as well as components preceding the ABR that are recorded via cochleography (ECoghG). ECoghG can be used in pure tone sensitivity prediction to enhance wave I in ABR testing, when test conditions are less than optimal or when a hearing loss is present. SLRs arise from the periphery and brainstem (Ruth, 1994).

Middle latency responses (MLRs), which refer to components in the latency epoch of 10–50 milliseconds, are generated in structures beyond the inferior colliculus (Kraus, Kileny, & McGee, 1994). MLRs are clinically used to objectively determine hearing ability in the lower frequencies. They are also used to assess the cochlear implant function and to localize auditory pathway lesions (Kraus et al. 1994). However, MLRs are affected by sleep and cannot be detected in certain phases of sleep. It is possible to monitor sleep phases with EEG measurements and to conduct MLR testing only in favorable sleep periods, but this requires much more expertise and expensive equipment. The fact that MLRs are affected by the subject's level of consciousness has limited their popularity as an objective diagnostic procedure (Ferraro & Durrant, 1994).

Components generated beyond 50–80 milliseconds post-stimulus onset are long latency responses (LLRs) and are cortically generated (Kraus et al. 1994). An example of an LLR measurement is the N_1-P_2 Complex, which was successfully used as an indicator of hearing sensitivity in difficult-to-test populations and also to detect lesions in the central auditory pathway (Ferraro & Durrant, 1994). Just like the MLR, the N_1-P_2 Complex is sensitive to the subject's state of consciousness. Another example of LLR is the $P_{300,}$ whose most common uses include studies of aging, dementia, and attention disorders (Ferraro & Durrant, 1994).

The ABR dominated clinical attention to AEPs for about a decade and is still a very popular test of auditory function for difficult-to-test populations (Ferraro & Durrant, 1994). Behavioral evaluation of very young infants relies on spontaneous responses such as eye and head movements. Even with some kind of reinforcement, these responses cannot be elicited near the threshold value. Presentation of stimuli is via loudspeakers, which does not provide information about hearing ability in separate ears. All these limitations of behavioral hearing testing made ABR the preferred objective audiologic technique for infants younger than 6 months (Weber, 1994). With ABR, stimuli are presented via earphones, making it possible to test the hearing status of the individual ears. ABR enables the audiologist to obtain responses to low stimulus intensity levels from sleeping infants. As Robinette (1994) stated, the ABR is popular in the evaluation of hearing when traditional behavioral tests are precluded or their results are equivocal.

The ABR is, however, not without its own shortcomings. First, the frequency range in which hearing ability can be determined with ABR is limited. ABR testing with click stimuli provides only a one-point audiogram in the 2000–4000 Hz region. This is due to the type of stimuli needed to elicit an ABR—namely, abrupt onset acoustic clicks. The more abrupt the stimulus onset, the more neural fibers will respond in synchrony and the more clearly defined the ABR (Weber, 1994). The acoustic click has its greatest energy around 3000 Hz, therefore creating the stimulus range from 2000 to 4000 Hz.

This aspect implicates another limitation: If hearing in the 2000 to 4000 Hz region is normal, a hearing loss in lower frequencies can be overlooked. According to Kemp and Ryan (1993), a passable ABR response can be obtained during screening from a limited region of normal high frequency hearing although medium and low frequency hearing may be seriously impaired.

Attempts to gain information about the low frequencies created a new set of problems. The use of low frequency tone bursts with abrupt stimulus onset resulted in high frequency contamination. (An abrupt stimulus onset might stimulate broad areas of the basilar membrane.) Investigators have used several alternative techniques in an attempt to gain reliable and frequency-specific low frequency information such as masking techniques and filtering. A study by Balfour et al. (1998) conducted in clinical settings in the United States, indicated that the availability of filters and attenuators in ABR equipment in most clinical settings is very limited and that if tone burst stimuli are utilized, that they are unmasked for most facilities. Therefore, the quest continues for a sensitive electrophysiologic measure of low frequency hearing status that can be used with difficult-to-test populations (Weber, 1994).

The second shortcoming of ABR testing is the amount of time the test requires. It can take more than 30 minutes to obtain a single ABR threshold for each ear (Weber, 1994).

The third weakness is the possibility of sedation. When testing hearing ability close to the threshold, it can be affected by movement artifacts. The child should therefore be as still as possible, preferably asleep. When infants younger than 6 months are tested, it can be assumed that there will be periods of sleep long enough for ABR testing. For older infants, it is unfortunately often necessary to ensure adequate test conditions by giving the child some form of sedative, usually administered orally (Weber, 1994).

Lastly, ABR performed for diagnostic purposes requires highly trained personnel and is a relatively expensive procedure (Musiek, Berenstein, Hall III, & Schwaber, 1994).

At the end of the 1970s, David Kemp discovered a feature of normal cochleae that led to another objective way to evaluate hearing ability. He discovered otoacoustic emissions, which is tiny amounts of energy released by the outer hair cells. The energy is generated from a normal cochlea either spontaneously or in the presence of acoustic stimulation. It appeared that normal cochleae emitted these responses, whereas ears with a hearing loss >35 dB HL did not.

Kemp's (1978) original reports were greeted rather skeptically, and much early research only replicated his study to confirm the presence of otoacoustic emissions. After two decades of intensive research, this method is now generating much excitement among researchers, since certain types of otoacoustic emissions are proving to be highly applicable in the areas of hearing screening and even diagnostic audiology (Kummer et al. 1998; Martin, et al. 1990b; Stach, et al. 1993; Stover, et al. 1996a). Many researchers hope that this relatively new field in audiology will prove to be the long-awaited objective, rapid, and accurate test of auditory function to aid in the assessment of difficult-to-test populations.

Otoacoustic measurement will certainly never replace pure tone audiometry, immittance, or ABR, but OAEs offer diagnostic information regarding the auditory system that is not available from any other test. This new objective procedure deserves to be evaluated.

2.2.2.3 Otoacoustic Emissions

Otoacoustic emissions are low intensity acoustic signals generated by the outer hair cells (OHC) in the organ of Corti on the basilar membrane either spontaneously or in the presence of acoustic stimulation. Brownell (1990) describes the outer hair cell motility as a lengthening or shortening of the outer hair cells in response to acoustic stimulation. This active biological mechanism in the outer hair cells causes a vibration of the basilar membrane in an attempt to enhance the ear's sharpness and sensitivity (Attias, Furst, Furman, Haran, Horowitz, & Breslof, 1995) by providing the appropriate stimulus to the inner hair cell receptors (Kummer et al. 1998). This vibration, called an otoacoustic emission, can be recorded using a very sensitive microphone placed in the ear canal. Otoacoustic emissions are therefore not, themselves necessary for hearing but reflect processes in the cochlea necessary for hearing (Norton, 1993).

The primary value of otoacoustic emissions is that their presence indicates that the preneural cochlear mechanism (and middle ear as well) can respond to sound in a normal manner. A large area of the basilar membrane is stimulated, and the measured emissions are frequency specific and frequency selective, so it is possible to gain information about different areas of the cochlea simultaneously. "No other clinical test," wrote Kemp, Ryan, and Bray (1990), "specifically tests cochlear biomechanisms or combines the operational speed, non-invasivity, objectivity, sensitivity, frequency selectivity, and noise immunity of otoacoustic emission testing" (p. 94).

Kemp (1978) described two main classes of otoacoustic emissions: spontaneous otoacoustic emissions (SOAEs) and evoked otoacoustic emissions (EOAEs), which will be described below.

2.2.2.3.1 Spontaneous Otoacoustic Emissions (SOAEs)

SOAEs are tonal or narrowband low level signals that can be recorded in the absence of any auditory stimulation in only 50% of all persons with hearing levels <20 dB HL and in 60% of persons with hearing levels <30 dB HL (Lonsbury-Martin, 1994). Because of this low incidence of SOAEs, they are not viewed as a suitable clinical indicator of the mechanical activity of the cochlea (Lonsbury-Martin, 1994; Norton & Stover, 1994). After Kemp (1978) reported the existence of SOAEs, many clinicians hoped that they would be the objective basis for tinnitus. It has been proved, however, that most people are unaware of their spontaneous otoacoustic emissions, and only a very small percentage of people with tinnitus have recordable SOAEs that can be linked to their tinnitus (Norton, Schmidt, & Stover, 1990).

An interesting study by Kulawiec and Orlando (1995) investigated the effect of SOAEs on evoked OAEs and found that SOAEs contribute greatly to the level and shape of the frequency spectrum of TEOAEs. Present SOAEs increased the levels of the peak amplitudes at corresponding frequencies and as the number of SOAEs increased, the TEOAE levels increased. This phenomenon causes a large range of different levels found in TEOAE testing, which in turn is according to these authors, the primary reason why TEOAEs fail to predict actual hearing thresholds.

SOAEs can be used as a complementary technique for evoked otoacoustic emissions (Bonfils, et al. 1990) but due to the low incidence cannot be used as a screening or diagnostic procedure.

Several types of evoked OAEs exist, depending on the type of stimulus used during the measurement. Evoked emission types include stimulus frequency emissions, transient evoked otoacoustic emissions, and distortion product otoacoustic emissions.

2.2.2.3.2 Stimulus Frequency Otoacoustic Emissions (SFOAEs)

A stimulus frequency otoacoustic emission (SFOAE) is the most stimulus frequency specific of all emission types, but it is also probably the least clinically applicable (Norton & Stover, 1994). SFOAEs reflect the response of the cochlea at a certain pure tone, occurring simultaneously with and at the same frequency as the stimulus presented. When a tone is presented to the ear, the sound pressure measured in the ear canal is the sum of the sound pressure of the stimulus and the response. In the case of other evoked emission types, the stimulus sound pressure level is separated from the response either spectrally (as in the case of distortion product otoacoustic emissions) or temporally (as in the case of transient evoked otoacoustic emissions). Due to the lack of temporal or spectral separation techniques in measuring SFOAEs, more sophisticated equipment and processing of data are required, and therefore SFOAEs are not currently practical for clinical use (Lonsbury-Martin & Martin, 1990; Norton & Stover, 1994). Lonsbury-Martin, (1994) described this phenomenon quite effectively: "SFOAEs are technically difficult to measure, due to the complexities of separating the in-going acoustic stimulus from the out-going emitted response. Thus, to date, little information has accumulated concerning either their basic nature or their clinical utility" (p. 2). One possible way to overcome this problem is to use the nonlinearity of the cochlea by performing multilevel tests and with subtractions determine which part is due to the cochlea's nonlinear response (Kemp & Ryan, 1993). This can be time consuming and not a practical resolution for a screening test.

Another type of evoked otoacoustic emissions is transient otoacoustic emissions.

2.2.2.3.3 Transient Evoked Otoacoustic Emissions (TEOAEs)

TEOAEs are responses that follow a brief acoustic stimulus such as a click. TEOAEs can be recorded in nearly all persons with normal hearing (hearing levels < 20 dB from 500 Hz to 4000 Hz) and are absent in all ears with a hearing loss 30–40 dB HL. (Hearing loss > 40 dB HL according to Glattke, Pafitis, Cummiskey, & Herer, 1995; hearing loss > 35dB according to Robinette, 1992; or hearing loss > 30 dB according to Kemp et al. 1990.) The latest research indicated that when tone bursts are used to elicit TEOAEs instead of clicks, emissions can be evoked in ears with hearing losses at least up to 60 dB HL (Vinck, et al. 1998).

In measuring a TEOAE, a probe is inserted into the ear canal, containing a miniature sound source for delivering the stimulus and a very sensitive microphone for detecting the response. TEOAEs are obtained by using synchronous time-domain averaging techniques. Responses to several stimuli (e.g., 500–2000 clicks) are averaged to improve the signal-to-noise ratio and make the response distinguishable from the noise floor (Glattke et al. 1995). The ear canal sound pressure is amplified, filtered, and then digitized, and the first 2.5 seconds of the response are eliminated to remove the stimulus (Norton & Stover, 1994).

TEOAEs are frequency dispersive: high frequencies coded basally on the basilar membrane have a shorter latency (4 ms for 5000 Hz) than low frequencies, coded apically on the basilar membrane (20 ms for 500 Hz). According to Kemp et al.

(1990), this provides for temporal separation of the stimulus's and response's sound pressure level, both measured in the ear canal.

When it comes to the discussion of the frequency specificity of TEOAEs, the issue is often misunderstood because there are different areas of AOE measurement and analysis that involve frequency specificity. Kemp and Ryan (1993) distinguish between the three kinds:

- First, ***stimulus frequency specificity*** refers to the similarity of the stimulus to a pure tone. TEOAEs use broad band clicks and stimulate broad areas on the basilar membrane. However, when a pure tone is used for stimulation, it does not mean that only one point on the basilar membrane is stimulated as is often supposed. The traveling wave resulting from pure tone stimulation vibrates all the basilar membrane up to the place representing that frequency. Furthermore, excitation of the basilar membrane by a completely pure tone can involve up to a third octave range.
- Second, ***cochlear frequency specificity*** is the cochlea's ability to respond to the frequency of a stimulus in a certain place representing that frequency, and to generate responses for every frequency being stimulated independently. This is still not very specific: in the human cochlea only within a third octave specific.
- Third, ***response frequency specificity*** refers to the relationship between the response and the frequency area of the basilar membrane tested. In other words, can the frequency area on the basilar membrane being tested be determined by looking at the response?

It is evident that the use of frequency specific stimuli does not guarantee or create a frequency specific response.

The stimulus frequency specificity of TEOAEs is determined by the bandwidth of the stimuli being used to elicit a response. Emissions can be evoked at most frequencies in the normal cochlea. The broader the stimulus spectrum, the broader the emission spectrum (Norton & Stover, 1994). Broadband clicks are usually used for measuring TEOAEs, which allows for simultaneous multifrequency testing (Kemp & Ryan, 1993). TEOAEs provide simultaneous information regarding the functioning of the outer hair cells on the basilar membrane for a very broad region of frequencies (Kemp et al. 1990; Norton & Stover, 1994). Some cochlear frequency-specific information can be gained by analyzing the spectral distribution. Ueda, (1998) found that if certain frequencies are over stimulated in guinea pigs, the temporal shifts could be measured in the TEOAE spectrum and therefore proved that TEOAEs are frequency specific. Kemp et al. (1990) successfully used TEOAEs to identify frequency ranges of normal hearing in pathological ears. In a case with a high frequency hearing loss, they obtained emissions up to the frequency of the hearing loss and no emissions for the pathological frequencies. It should be noted, however, that no information regarding the thresholds of the pathological frequencies could be obtained. Many other researchers have also had difficulty in making comparisons between frequency-specific audiometric thresholds and frequency information provided by TEOAEs (Bonfils et al. 1990; Lee, et al. 1993; Lonsbury-Martin & Martin, 1990). The fact that no click emissions can be obtained when the hearing loss exceeds 30–40 dB HL has proven TEOAEs to be more applicable in the area of hearing screening than diagnostic audiology (Harris & Probst, 1991; Lee et al. 1993).

TEOAEs were the first method to be tried and recommended for neonatal hearing screening and are currently the most widely used OAE method for screening (Kemp & Ryan, 1993). TEOAEs can be measured very effectively in newborns. Both ears can be screened in a sleeping infant in about 10 minutes, compared to about 20 minutes with screening ABR (Norton & Stover, 1994). Another advantage of TEOAEs is that a broader frequency spectrum is being evaluated than with ABR, they do not require highly trained personnel, and they are objective and non-invasive (Lonsbury-Martin, McCoy, Whitehead & Martin, 1992; Stevens Webb, Hutschinson, Smith & Buffin, 1990).

TEOAEs do, however, have limitations. The first is that they are only recordable in normal and near-normal ears (30–35 dB HL) when clicks are used as stimuli. This implies that TEOAE data cannot be translated into "threshold data." An ear with a hearing loss of 65 dB will have the same absent response as an ear with a hearing threshold of 40 dB (Kemp et al. 1990). Although TEOAEs function as a wonderful screening procedure (Stevens et al. 1990), no information regarding hearing status can be obtained once the emission is absent, as in the case of mild and moderate hearing losses. A recent study by Harrison and Norton (1999) however, found that a broadband tone burst, which has all of its energy concentrated in a narrow bandwidth, could sometimes evoke and emission where a click could not. In a few isolated cases, they could elicit emissions with broadband tone bursts in ears with mild and moderate hearing losses. It is thus probable that with more research, it may be possible to distinguish between mild and moderate losses using broadband bursts. Vinck et al. (1998) also made a remarkable breakthrough in pure tone threshold prediction with tone burst otoacoustic emissions (TBOAE) and proved that TBOAEs could predict hearing sensitivity within 10dB for hearing losses up to 60 dB HL for all frequencies, 250 Hz to 8000 Hz. The authors considered a statistical artifact for these astounding findings but eventually attributed the results to several other explanations including the new use of TBOAEs, the use of multivariate statistical techniques, and the fact that pure tone thresholds are highly inter-related. Normal hearing at any given frequency might influence prediction accuracy of frequencies with absent emissions. In other words, prediction of PTTs at frequencies where there is a 60dB HL hearing impairment could have been influenced by inter-related frequencies where the hearing loss was less than 45 dB HL. The only limitation that they identified was that this process of emission measurement is still very time consuming and questionable for the use of small children. Advances in OAE software will hopefully enable audiologists to use tone bursts to predict PTTs more rapidly and easily in the future.

The fact that a more frequency specific stimulus has all of its energy concentrated in one area, instead of spreading it out over the basilar membrane, provides for more efficient basilar membrane stimulation and enables researchers to measure emissions in populations with hearing loss. This is the primary reason why TBOAEs and DPOAEs can be recorded in ears with a hearing loss and TEOAEs cannot.

Another weakness of TEOAEs is the variability in TEOAE spectrums for normal populations. Hurley and Musiek (1994) indicated that TEOAEs are affected by small changes in cochlear physiology that do not result in comparable changes in auditory threshold. In other words, they found considerable TEOAE variability among ears with similar hearing sensitivity. This aspect makes it almost impossible to predict hearing thresholds with TEOAEs. According to Kulawiec and Orlando (1995), the TEAOE variability is greatly influenced by the presence of SOAEs. A study by Avan, et al. (1997) indicated that ultra high frequency hearing status (8 to 16 kHz) also influences the TEOAE spectrum even though hearing in the normal frequency range (0.5 to 8 kHz) might be normal. This factor, together with age and probe tip placement is the reason why there is so much variability in TEOAE spectrums of normal hearing people. TEOAEs can therefore only classify hearing levels as normal (<20 dB HL) or abnormal (>20 dB HL) but fail to predict hearing sensitivity of specific frequencies.

Another aspect that should be kept in mind is the difference in the application of stimulus level in TEOAE recordings compared with its use in ABR recordings. According to Kemp and Ryan (1993) ABR measures a physiological threshold dependant on synaptic events and the ABR threshold is the dividing line between stimulus levels where one results in a response and the other doesn't. OAEs do not have such a threshold because they are presynaptic and for OAEs a detection threshold has to be determined which is the lowest point where a response can be distinguished from he noise floor. With TEOAE detection threshold determination, the exact threshold level can be a function of the equipment, background noise and subject cooperation and therefore TEOAE screening usually uses suprathreshold stimuli. It can be argued that the use of suprathreshold stimuli limits the sensitivity of a screening test for minor hearing impairments might be overlooked with high-level stimuli.

Finally, it seems that TEOAE amplitude and occurrence are negatively affected by increasing age. Norton and Widen (1990) reported a statistically significant decrease in TEOAE amplitude with increasing age even in a carefully screened sample. Kemp et al. (1990) also indicated stronger responses as well as responses at more frequencies for neonates than adults. It is still unclear whether the age-associated changes are due to normal developmental changes in the middle ear or to progressively impaired cochlear function.

All the emission types previously discussed—namely, SOAEs, SFOAEs, and TEOAEs—have one limitation in common. **None of these emission types can function as an objective test of hearing where pure tones can be predicted given only the otoacoustic emissions** (Bonfils et al. 1990; Hurley & Musiek, 1994; Lee et al. 1993; Lonsbury-Martin, 1994). The requirement for an emission type to be able to potentially predict pure tone thresholds given only the otoacoustic emissions is that the emission type should be present in normal and hearing-impaired ears (Kimberley, et al. 1994b). It should also be frequency-specific and easily compared to the frequencies of the behavioral thresholds (Lee et al. 1993). There is one emission type that might prove to be clinically applicable in the prediction of behavioral pure tone thresholds—namely, distortion product otoacoustic emissions (Lee et al. 1993; Lonsbury-Martin & Martin, 1990).

2.2.2.3.4 Distortion Product Otoacoustic Emissions (DPOAEs)

Distortion Product Otoacoustic emissions (DPOAEs) were identified by Kemp (1979), only one year after the initial identification of transient evoked otoacoustic emissions (TEOAEs) and spontaneous otoacoustic emissions (SOAEs). Distortion product otoacoustic emissions are different from the other emission types in a number of ways. Firstly, DPOAEs are elicited by the simultaneous presentation of two pure tones and the emission is an internal produced frequency different from the two stimuli, in frequency and amplitude. Secondly, in contrast to other emission types such as TEOAEs, SOAEs and SFOAEs, the distortion product can very easily be measured in many common vertebrae laboratory animals (Mills, 1997). Research on laboratory animals allows experimental control of certain factors that contribute to a better understanding of the characteristics of distortion product emissions and OAEs in general (Zhang & Abbas, 1997). DPOAEs have even been measured in the ear of a grasshopper with a complete different morphology. The hearing organ of a grasshopper does not have any sensory hair cells, but the dendrites of the ciliated receptor cells are responsible for generation of distortion (Kossl & Boyan, 1998). Distortion product otoacoustic emissions have therefore been proven useful in both clinical and research settings. Thirdly, DPOAEs can be measured in hearing impaired ears with elevated threshold levels of up to 65dBHL (Moulin, et al. 1994). This feature enables DPOAEs to provide more than just hearing screening information.

These interesting differences between DPOAEs and other emission types led to an extensive investigation of DPOAEs to determine the clinical applicability of DPOAEs (Bonfils & Uziel, 1989). This clinical interest in DPOAEs is twofold. The first interest lies in the development of an objective test of auditory function. The second as a basis of a test uniquely sensitive to the functioning of the outer hair cells, and therefore a useful tool in differential diagnosis testing (Durrant, 1992). This research project focuses on the first interest: **To develop an objective, noninvasive test of auditory function with distortion product otoacoustic emissions.**

2.3 DPOAEs as an Objective, Non-invasive Technique

2.3.1 Definition

Distortion product otoacoustic emissions (DPOAEs) are elicited by the simultaneous presentation of two different pure tones, f1 and f2, where f1<f2. The distortion product response is a third tone of frequency, produced internally and in a frequency region different from the two primary frequencies. Responses can be expected at several different distortion product frequencies such as 2f1-f2, 3f1-2f2, 4f1-3f2, etcetera. Of all the distortion products, the cubic distortion product is the most prominent in humans and occurs at 2f1-f2 (Nielsen, Popelka, Rasmussen & Osterhammel, 1993).

The normal cubic distortion product is typically 60dB lower than the overall level of the primaries (Nielsen, et al. 1993). The relationship of the distortion product (2f1-f2) and the two primary frequencies (f1 and f2) can very clearly be seen in the spectrum of the ear canal sound pressure of normal hearing subjects undergoing DPOAE testing. This relationship is illustrated in Figure 2.1.

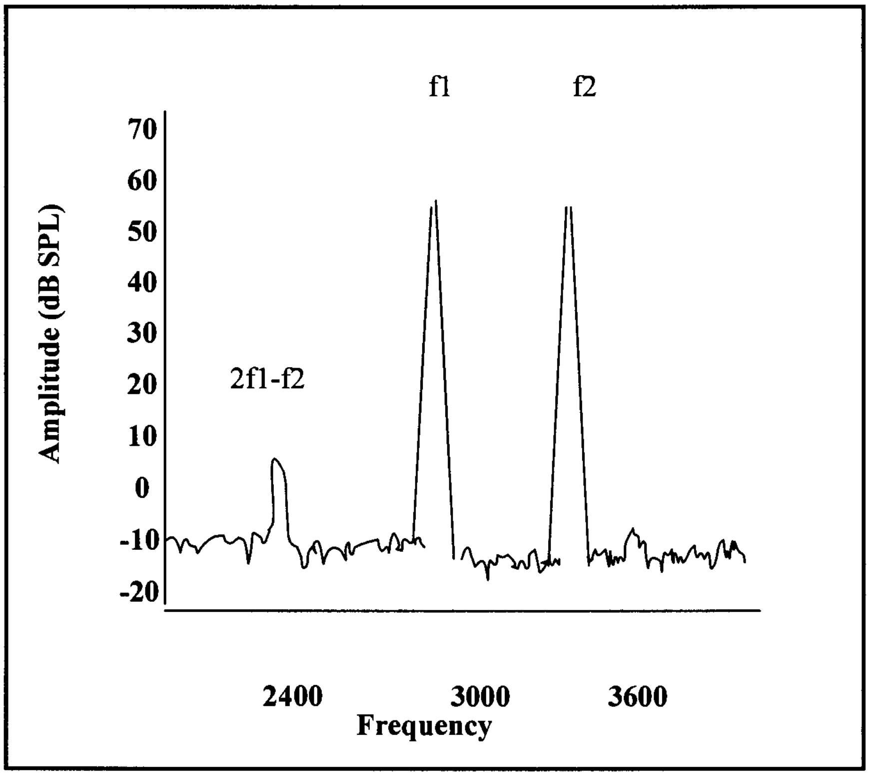


Figure 2.1: The spectrum of the ear canal sound pressure of a normal hearing adult undergoing DPOAE testing. f1 and f2 are the stimuli and 2f1-f2 is the response (from Norton & Stover, 1994:457).

The next section will discuss the measurement procedures and instrumentation necessary to elicit a distortion product otoacoustic emission.

2.3.2 Measurement Procedures and Instrumentation for DPOAEs.

For the measurement of a distortion product otoacoustic emission, two separate channels for stimulus generation and attenuation are necessary. These two channels should be electrically isolated to prevent distortion. The signals are presented to the ear canal via a probe microphone assembly with two delivery ports. Probe microphone systems for DPOAEs consist of a miniature sound source and a very sensitive microphone built into a unit small enough to fit snugly into a human ear canal (Siegel, 1995).

After the two signals are presented to the ear canal, the ear canal sound pressure is averaged to reduce the noise floor and spectrally analyzed for the levels of the primaries (fl and f2) and the response (2fl-f2). Figure 2.1 shows the spectrum of the sound pressure measured in the ear canal, depicting the two primary stimuli fl and f2, as well as the response, 2f1-f2.

A complete DPOAE system is presented schematically in Figure 2.2.

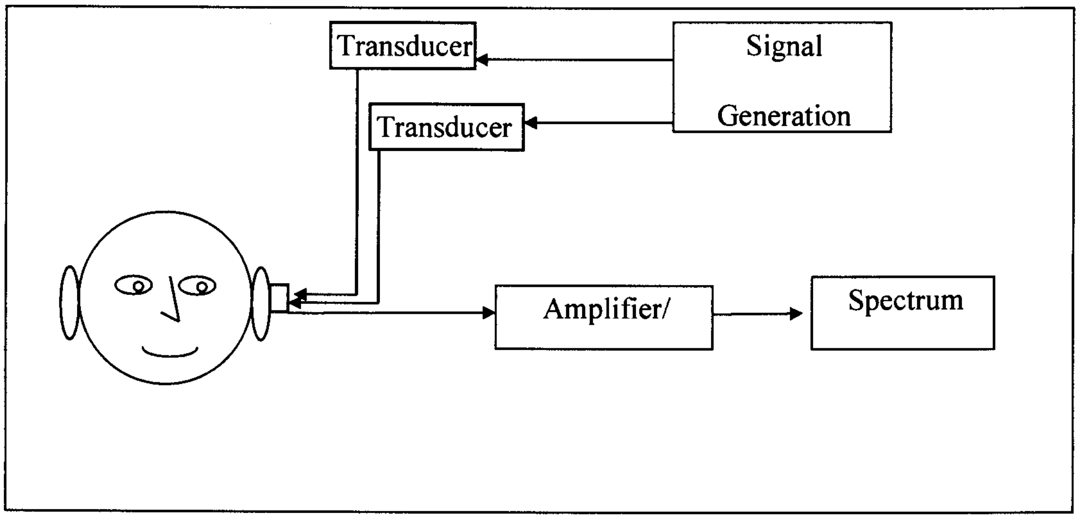


Figure 2.2: Schematic representation of a system used to measure distortion product otoacoustic emissions (Norton & Stover, 1994:456).

DPOAEs can be recorded in virtually all normal hearing ears (100% according to Lee et al. 1993; and 95% according to Kimberley et al. 1994b). DPOAE measurement uses frequency specific pure tone stimuli. The facts that specific input frequencies can be selected and that responses are measured at certain frequencies make it easier to make comparisons between DPOAE results and conventional pure tone thresholds. This feature of DPOAE measurement makes it the best-suited emission type to relate to behavioral thresholds (Lee et al. 1993). Although one should be careful to state that DPOAE is a more frequency specific type of emission due to its stimulus frequency specificity, there are studies to support the notion that DPOAEs are frequency specific in the sense that there are good correlations between shapes of audiograms and DPOAE emission spectrums (Gaskill & Brown, 1990; Gorga et al. 1993).

Furthermore, DPOAEs are the only otoacoustic emission type that can be recorded in the presence of a mild to moderate hearing loss. TEOAEs can only classify a person's hearing as normal or impaired (Bonfils, Piron, Uziel & Pujol, 1988). DPOAEs can classify hearing ability as normal, slightly impaired, mildly impaired, moderately impaired, or severely impaired (in cases where no emissions can be measured) (Durrant, 1992; Gaskill & Brown, 1990; Lee et al. 1993). This advantage of DPOAEs allows emission testing of a much larger population with varying hearing sensitivity, making this one of the best reasons to investigate DPOAEs as an additional objective test of hearing.

The use of DPOAEs to predict hearing sensitivity will be reviewed in the next section to better understand the struggle of previous researchers and to set the goals for what such a project might hope to accomplish. It is however a challenging task to compare one study with another because of the numerous DPOAE and demographic features that influence predictive accuracy of hearing sensitivity and that differ greatly in the various studies. Some studies use DPOAE amplitude as predictors (Stover et al. 1996a; Kimberley et al. 1994b), others use DPOAE threshold (Moulin, et al. 1994).

Some use DP Grams (Lonsbury-Martin, 1994), others I/O functions (Lonsbury-Martin & Martin, 1990). Furthermore, the DPOAE frequencies chosen in DPOAE measurement seldom overlap and include different frequency ratios. In the analysis of data some studies correlate PTT frequency to the frequency of f1 (Gaskill & Brown, 1990), f2 (Harris, Lonsbury-Martin, Stagner, Coats & Martin, 1989; Kimberley et al. 1994a; Kimberley et al. 1994b; Kummer et al. 1998), 2f1-f2 or the geometric mean (GM) (Lonsbury-Martin and Martin, 1990; Martin et al. 1990b; Bonfils et al. 1991; Zhao & Stephens, 1998) of the primaries. Most studies use only one frequency in correlation determination but some included adjacent frequencies (Kimberley, et al. 1994a).

It might seem logical to attempt to concentrate the findings of previous researchers into a table to summarize achievements. After several attempts it became clear that the format of such a table would be too complex due to the hundreds of different measurement and analysis techniques and styles of reporting findings. Very few researchers expressed their findings in percentage correct prediction values and percentage false positives and false negatives. Furthermore, researchers used different statistical techniques that add to the rich terminology not expressible in table format. The following discussion will attempt a succinct yet informative overview.

2.3.3 Literature Overview of Studies Attempting to Predict Hearing Sensitivity with DPOAEs.

An early study by Kimberley and Nelson (1989) investigated the correlation between distortion product otoacoustic emissions and hearing threshold. Subjects were selected without regard to age, sex, etiology of hearing loss or pattern of hearing loss. The frequency ratio of the primaries (f2/f1) was 1.2. Distortion product otoacoustic emissions were measured over a stimulus range from 30 dB SPL to 80 dB SPL in 6 dB steps. DPOAE I/O functions were measured covering the frequency range from 700 Hz to 6000 Hz. Kimberley and Nelson (1989) then plotted emission thresholds and auditory thresholds of 21 ears on a scattergram. (Emission thresholds represent the stimulus level required to just raise the emission above the noise floor.) Kimberley and Nelson's (1989) scattergram can be viewed in Figure 2.3. The linear fit shown with the data points has a slope of 1.0 and a correlation coefficient of .86.

The results displayed in this scattergram in Figure 2.3 suggest that DPOAE measurements can predict auditory thresholds within 10 dB over a range from 0 dB SPL to 60 dB SPL. The authors claim that this was the first report of such a precise correlation.

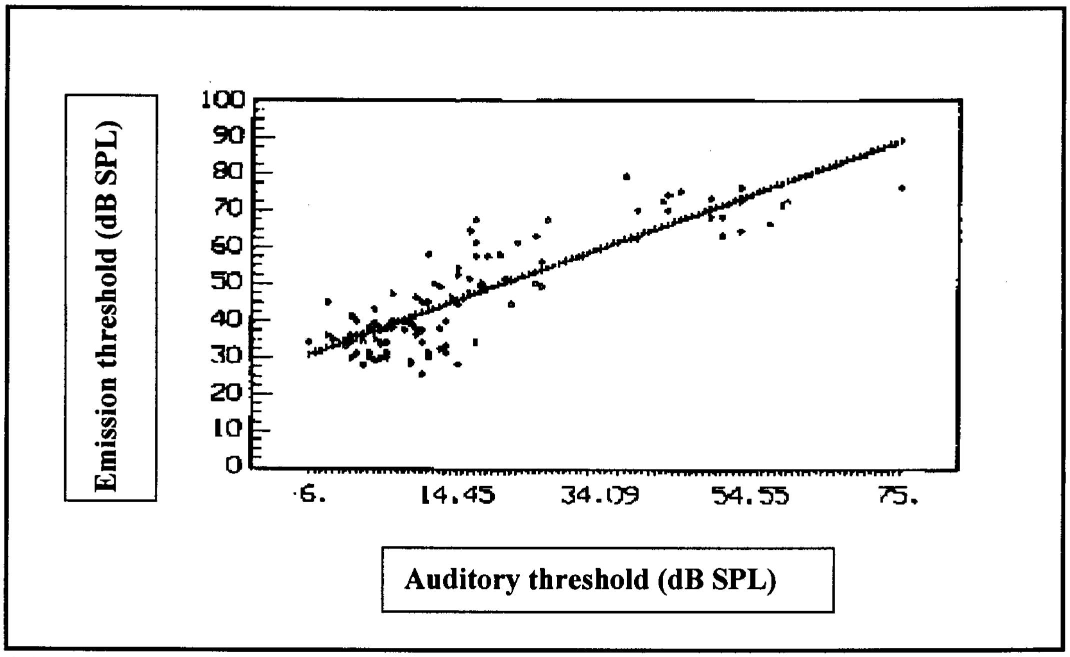


Figure 2.3: Scattergram of emission threshold versus auditory threshold as measured by Kimberley and Nelson, (1989: 368)

Gaskill and Brown (1990) investigated the behavior of the acoustic product in humans and it's relation to auditory sensitivity. They concluded that with certain optimal stimulus parameters (stimulus levels below 60 dB SPL; L1>L2 by 15 dB; f1/f2 = 1.225), half of the subjects showed a statistically significant correlation between DPOAE results and auditory sensitivity at the corresponding f1 stimulus.

Avan and Bonfils (1993) confirmed these findings. The authors conducted a study on DPOAEs in 25 normal hearing and 50 hearing impaired ears. Their results indicated that the DPOAEs evoked by low intensity primary tones (below 62 dB SPL) were strongly correlated with the auditory threshold at the mean frequency of f1 and f2 and that DPOAEs disappear for hearing losses larger than about 30dB. This research also suggests that when low intensity primaries are used, DPOAEs provide frequency specific information on the local cochlear state of the primaries.

A few other studies indicated positive relationships between DPOAEs and pure tone thresholds. Spektor, Leonard, Kim, Jung and Smurzynski (1991) reported a positive qualitative relationship between pure tone thresholds and DPOAE thresholds in 19 children (although these authors did not quantitatively correlate DPOAE thresholds with pure tone thresholds). It seemed that the configuration of the hearing loss correlated well with the frequency pattern of the DPOAEs. Lonsbury-Martin and Martin (1990) assessed DPOAEs in subjects with noise induced hearing loss. They found that DPOAE thresholds provide reasonable good estimates of hearing loss in cases where primary damage to the outer hair cells can be assumed (such as noise induced hearing losses). The authors found a relatively strong correlation between DPOAE thresholds and magnitude of hearing loss. In the subjects they examined, for every 1dB increase in DPOAE threshold, hearing level increased by 1dB. When DPOAE threshold was > 63 dB SPL, the accompanying hearing level was > 20 dB HL. Such a strong correlation between DPOAEs and pure tones in subjects with OHC pathology proves it as an efficient measurement of cochlear functioning. DPOAEs could potentially be successfully applied to other cochlear pathologies such as Meniere's disease and ototoxity.

Gorga, et al. (1993) measured DPOAEs in normal hearing and hearing-impaired human subjects. They investigated the extent to which DPOAEs can be used to correctly distinguish between normal and impaired hearing. DPOAE amplitude was able to distinguish between normal and impaired subjects at 4000 Hz, 8000 Hz and to a lesser extent at 2000 Hz. At 500 Hz, performance was no better than chance, due to high biological noise levels such as breathing and swallowing. They concluded that DPOAE measurement could successfully be implemented to identify high frequency hearing loss, but that it was not an accurate predictor of hearing loss in the lower frequencies.

A study conducted by Stover et al. (1996a) examined the effect of the primary stimulus levels on the ability of DPOAE measurements to separate normal hearing from hearing impaired ears. Clinical decision theory was used to assess both DPOAE threshold and DPOAE amplitude as diagnostic indicators of hearing status. This research suggests that DPOAE threshold and DPOAE amplitude perform equally well in distinguishing normal from impaired hearing but DPOAE amplitude is more suited as a screening method due to shorter testing times. Probst and Hauser (1990) performed similar research in 1990 and concluded that the measurement of DPOAE amplitude alone might fail to detect a mild hearing loss. To determine hearing ability more accurately, more detailed measurements such as I/O functions with DPOAE thresholds should be performed.

Kummer et al. (1998) investigated the growth behavior of DPOAEs and its relationship to auditory sensitivity in 20 normal ears and 15 ears with cochlear hearing loss at the f2 frequency with probe tone levels varying from L2 = 20 - 60 dB SPL. They concluded that this relationship is strongly dependant on stimulus levels. For normal ears, statistically significant correlations could be determined for 14/15 ears when lower stimulus levels were used (L2 = 25 dB SPL) and for 17/20 hearing impaired ears when moderate primary tone levels were used (L2 = 45 dB SPL).

Kimberley et al. (1994b) predicted hearing status in normal and hearing-impaired ears with DPOAEs at six frequencies ranging from 1025 - 5712 Hz. The significance of variables such as DPOAE levels, age and gender were determined in the definition of normal versus abnormal PTTs and then applied to a new set of unfamiliar data to determine their predictive accuracy at each frequency. Classification accuracy of normal hearing varied from 71% at 1025 to 92% at 2050 Hz. Kimberley et al. (1994b) concluded that DPOAE measures can reliably categorize pure tone thresholds as being normal or impaired in a population with varied cochlear hearing status.

Kimberley et al. (1994a) compared an artificial neural network (ANN) approach to multivariate discriminant analysis to classify PTTs with DPOAEs in 229 normal and hearing-impaired ears as normal or impaired. Prediction accuracy varied from 57% correct classification of hearing impairment at 1025Hz to 100% at 2050Hz when normal hearing was defined as PTTs < 20dB HL. Overall classification accuracy was 80% for normal PTTs and 90% for impaired PTTs. They concluded that the neural network approach was more successful in classifying hearing sensitivity due to this technique's ability to model complex relationships or more specifically, the nonlinear relationship between DPOAEs and PTTs. The discriminant analysis technique is restricted to modeling purely linear relationships.

Many previous studies attempted to develop DPOAEs as a possible new objective method of hearing sensitivity prediction. Most researchers attempted to classify hearing status as normal or impaired and did not attempt to predict specific pure tone thresholds for impaired ears. Still, most researchers found it extremely difficult or even impossible to classify impaired PTTs at low frequencies as normal or impaired with DPOAEs (Gorga et al. 1993; Kimberley et al. 1994b; Stover et al. 1996a; Zhao & Stephens, 1998). This unsatisfactory prediction or classification of PTTs with DPOAEs is due to many factors influencing the measurement of DPOAEs (Gorga, et al. 1993; Nieschalk, et al. 1998), the complex nonlinear nature of the measured responses (Lonsbury-Martin, Martin & Whitehead, 1997; Nakajima et al. 1998; Kummer et al. 1998) and shortcomings in data analysis techniques used to date (Kimberley et al. 1994a).

There are also more general issues that influence prediction of PTTs with DPOAEs that should be overviewed to put PTT prediction with DPOAEs in perspective. Chapter three will discuss the issues relating to DPOAE measurement, the complex DPOAE response and analysis of DPOAEs in detail. More general issues contributing to the unsatisfactory prediction of PTTs with DPOAEs will be discussed in the following section.

2.3.4 PTT Prediction with DPOAEs in Perspective.

2.3.4.1 Audiometric Threshold is Determined by Factors Not Included in OAE Generation.

According to David Kemp (1997) the following factors assist in the sensitivity of hearing threshold:

1. An open external auditory meatus.
2. A mobile light and stiff tympanum.
3. A light and well articulated ossicular chain.
4. A mobile and low-loss attachment of the stapes to the oval window.
5. A well formed mobile and low-loss basilar membrane supporting a normal traveling wave.
6. Optimum electrochemical environment of the scala media.
7. Optimum condition of the outer hair cells.
8. Optimum configuration of the outer hair cells (including the medial efferent systems)
9. Optimal coupling of motion within the organ of Corti, especially from basilar membrane to outer hair cell to inner hair cells.
10. Optimum condition and functioning of the inner hair cells.
11. Optimum synaptic function at the inner hair cell- including efferent interaction.
12. Optimal neural transmission out of the inner ear.

13. Optimal mapping and processing of the neural signals reaching the cochlear nucleus.
14. Optimum function of the entire auditory pathway.

Auditory threshold depends on the sum of all 14 factors above. OAE generation depends on only the first eight factors. There is an additional factor in OAE generation that is not present in the hearing threshold factor list. The cochlea delivers a vibratory force to the eardrum and this sound pressure depends on the acoustics of the enclosed ear canal in which the probe microphone is situated and that is different from the open ear canal in PTT testing. Furthermore, in OAE generation the energy travels backwards through the system and such a reverse process might have different acoustic parameters.

At first early studies made no claims that OAEs could predict hearing sensitivity but only that it could distinguish between normal and impaired hearing (Probst, Lonsbury-Martin, Martin & Coats, 1987; Bonfils et al.1990; Collet, Gartner, Moulin, Kauffmann, Disant & Morgon, 1989; Kemp et al. 1990). However, the need for an objective test to aid in the assessment of difficult-to-test populations and the superficial similarity between the audiogram and DPOAE intensity-frequency displays fueled expectations that DPOAEs predicts PTTs spontaneously. OAEs are a sensitive indicator of cochlear dysfunction and are partially correlated with threshold but there are numerous factors influencing PTTs that cannot be measured by OAEs. David Kemp (1997) very effectively summarizes the fact that there is such a seemingly strong correlation between the two measures: "It is an accident of biology that the most common auditory disorders affect this region. Even some retrocochlear disorders exert a negative impact on the cochlea. It is an accident of biophysics that a correlate of sound vibration in the cochlea prior to its arrival at the sensory cells can be so easily recorded." (Kemp, 1997: p19).

The fact that there are so many factors contributing to PTTs that are not measured with OAEs is one of the reasons why researchers struggle to develop OAEs as a diagnostic test of hearing.

2.3.4.2 Nonlinearity is at the Heart of OAE Generation.

When the living human cochlea is stimulated with two pure tones simultaneously, it perceives a combination of tones that can be measured even though these tones are not present in the stimuli. According to Ruggero (1993) only an active process requiring energy can explain the perception of these tones that suggests the presence of significant nonlinearities. In 2f1-f2 DPOAE measurement, there is energy measured at f1, f2, 3f1-f2, 2f1-f2 and sometimes even more. It is tempting to believe that the 2f1-f2 component is sufficient, but it is not (Kemp, 1997). About 80% of all the information in discarded and with the instrumentation up to date, it would be very time consuming to measure all the energy tones. In order to form a model of cochlear nonlinearities to fully understand the dynamics of DPOAEs, enough data will have to be assimilated. Current TEOAE and DPOAE measures only scratch the surface.

2.3.4.3 The Sound Calibration Issue.

Most DPOAE measurement devices on the market today calibrate the DPOAE stimuli by adjusting the speaker voltage levels as a function of frequency to produce a constant SPL as measured by the probe microphone. This in-the-ear calibration method poses a potential problem because of the presence of standing waves in the ear canal that cause variations of stimulus levels at the eardrum as a function of frequency (Lonsbury-Martin et al. 1997). This sound calibration issue causes problems when impaired ears have to be compared to a group of normal ears, or for the prediction of PTTs with DPOAEs. Some agreement has to be reached by investigators, clinicians and DPOAE-testing equipment manufacturers regarding the best method to calibrate stimuli in order to make comparisons between studies and to set a norm for normal DPOAE occurrence in large populations.

Despite these problems, many researchers still hope to develop DPOAEs as an objective PTT prediction procedure. The requirements for such a procedure will be reviewed in the following section.

2.4 Requirements for an Optimal Objective Pure Tone Threshold Prediction Procedure

An extensive literature overview indicated that the "wish list" for an optimal objective procedure includes the following attributes:

2.4.1 Frequency Specificity

A frequency specific test of hearing is needed to predict hearing ability at specific pure tones. The procedure's measurements have to correlate well with conventional pure tone threshold measurements (Harris et al. 1989; Durrant, 1992; Lonsbury-Martin et al. 1992; Lee et al. 1993).

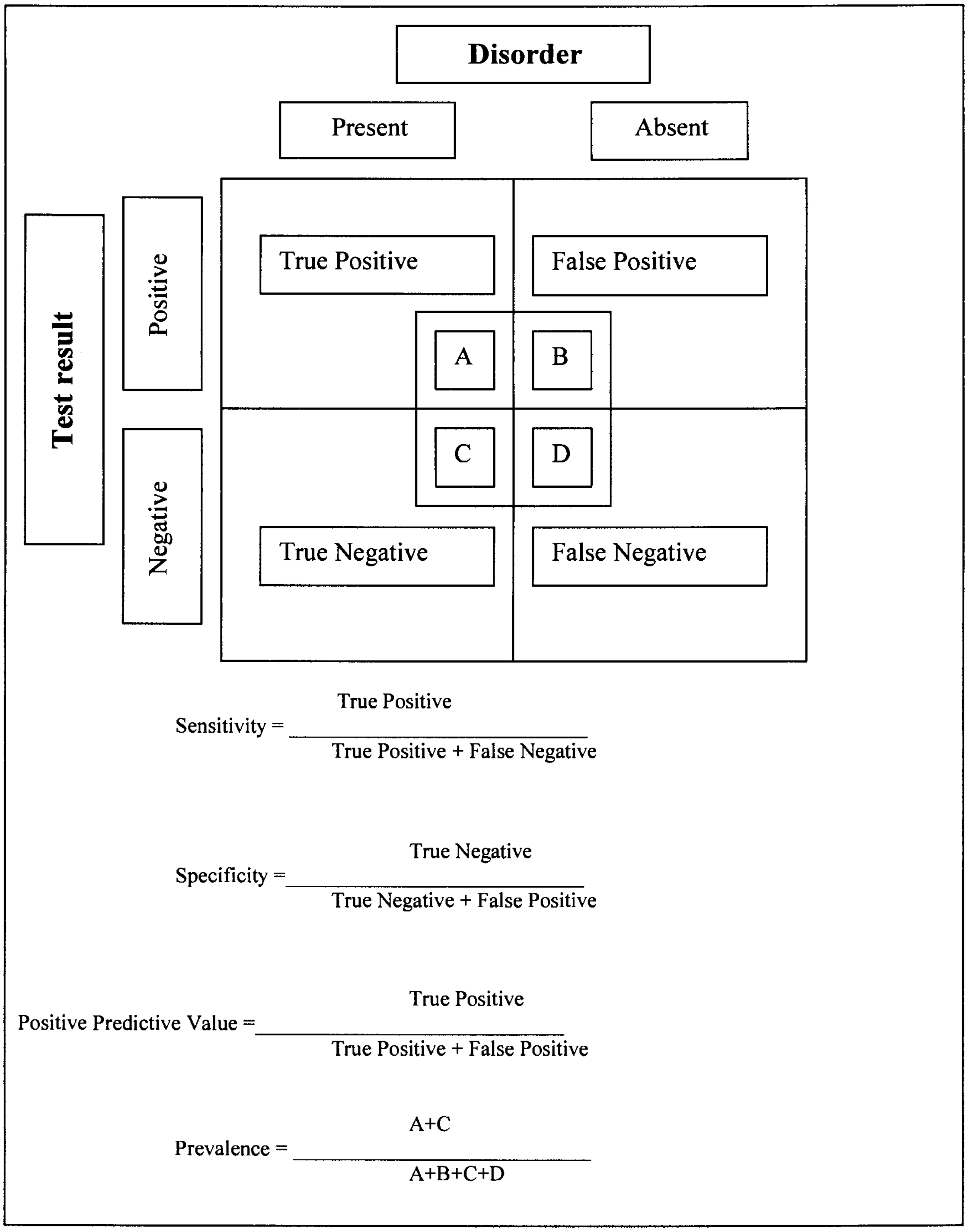
2.4.2 Evaluation of Broad Frequency Span

The measurement procedure must be able to evaluate as many frequencies as possible, (Vinck, Vel, Xu & Cauwenberge, 1995; Whitehead, Lonsbury-Martin, Martin & McCoy, 1996) and not just a one point audiogram as in the case of ABR (Weber, 1994).

2.4.3 Efficiency of the Procedure

The efficiency of a test is an expression of the percentage of individuals who are accurately identified as either normal or abnormal (Glattke et al. 1995). Efficiency can be divided into ***sensitivity*** which is the percentage of impaired ears correctly labeled as hearing impaired (Hussain, Gorga, Neely, Keefe & Peters, 1998) and ***specificity*** which is the number of normal hearing ears correctly labeled as normal (Stevens et al. 1990). Sensitivity is also known as ***the true positive rate*** and specificity as ***the true negative rate***. The ***predictive value*** of a test therefore indicates how many hearing impaired ears tested positive and how many normal ears tested negative.

For practical purposes, these characteristics move in opposite directions as the definition of a positive outcome changes: If sensitivity is raised, specificity is lowered and vice versa. The sensitivity and specificity of a test must be chosen in such a way that most ears with hearing loss is correctly identified without having too many false positive responses. The summary of calculations for test characteristics by Fechtner (1992: p9) can be seen in Figure 2.4.



$$\text{Sensitivity} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}$$

$$\text{Positive Predictive Value} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

$$\text{Prevalence} = \frac{A+C}{A+B+C+D}$$

Figure 2.4: Calculation of test characteristics such as sensitivity, specificity, positive predictive value and prevalence (Fechtner 1992:9).

2.4.4 Good Test-Retest Repeatability

If a person's hearing sensitivity stayed exactly the same from one day to the next, so should the DPOAE responses measured in that ear (Lonsbury-Martin et al. 1992). Another aspect that can be mentioned here is the interaural difference between the two ears of a subject that should be small if the two ears have similar PTTs. According to Tanaka et al. (1987) this aspect is very useful in the identification of unilateral hearing losses and functional unilateral hearing losses.

2.4.5 Differential Diagnosis between Sensory and Neural Hearing Impairment

A test that can accurately discriminate between cochlear and retrocochlear hearing losses would be welcomed (Durrant, 1992; Robinette, 1992; Norton, 1993; Whitehead et al. 1996).

2.4.6 Fast Test Performance

In the assessment of difficult-to-test populations, time is of the essence. DPOAEs have to be measured in a relatively quiet stage, where the subject is still and calm and most difficult-to-test subjects cannot be instructed to cooperate. The test therefore has to be quick to perform (Stevens, et al. 1990; Cane, O'Donoghue & Lutman, 1992; Lonsbury-Martin et al. 1992; Lee et al. 1993; Mauk & Behrens, 1993; Vohr, et al. 1993).

2.4.7 Economic Test Performance

If such a test is extremely expensive to perform, many people for lower socioeconomic groups will refuse the test and turn to less effective but more economic tests. The cost of a test is partly influenced by the ease in which the test can be performed and interpreted. If highly trained and specialized personnel are necessary, the cost will be considerably more (Bonfils et al. 1990; Mauk & Behrens, 1993; Vohr et al. 1993; Whitehead et al. 1996; Kim et al. 1997).

2.4.8 Non-invasive, Comfortable for the Patient

The test is usually performed when the subject is awake and in the case of neonates and children a painless and comfortable method of measurement is needed to ensure a calm and still posture (Bonfils et al. 1990; Lonsbury-Martin et al. 1992; Lee et al. 1993; Mauk & Behrens, 1993; Kim et al. 1997).

2.4.9 Age Differences

A test is required where the responses are not diminished by an increase in age. In other words, an infant should have similar measurable responses as an adult (Quinonez & Crawford, 1997; Abdala, 1998; Lasky, 1998a; Lasky, 1998b; Popelka, Karzon & Clary, 1998).

2.5 DPOAEs as an Optimal Objective Pure Tone Threshold Prediction Procedure

The feasibility of DPOAEs as an optimal objective procedure will be discussed according to the requirements set in the previous section.

2.5.1 Frequency Specificity

Many studies found positive correlations between certain frequencies concerning the measurement of DPOAEs and the frequencies of the PTTs and found DPOAEs to be highly frequency specific in the prediction of PTTs (Kemp et al. 1990; Lee et al. 1993; Nielsen, et al. 1993; Rasmussen, Popelka, Osterhammel & Nielsen, 1993; Kimberley et al. 1994a; Kimberley et al. 1994b; He & Schmiedt, 1997; Kummer et al. 1998; Zhao & Stephens, 1998). Recent research indicated that although the generation of the distortion product due to the interaction of the two primaries is in principle spread out over the whole basilar membrane, it is only the about 1mm around the f2 place that gives maximum contribution to the DPOAE measurement (Mauermann, Uppenkamp, Hengel & Kollmeier, 1999a+b).

2.5.2 Evaluation of Broad Frequency Span

Compared to TEOAEs evoked with click stimuli and ABR, DPOAEs are currently the objective test that evaluates the broadest frequency range: ABR currently evaluates a single area on the basilar membrane in the region of 2000 – 4000 Hz and provides a one point audiogram (Weber, 1994) and this aspect limits this procedure because a passable response can be obtained from normal hearing in these frequencies even if other frequencies are abnormal (Kemp & Ryan, 1993). Balfour et al. (1998) indicated that the attempt to measure a broader frequency range with ABR is limited due to the limited availability of filters and attenuators in ABR equipment in most clinical settings.

The measurement of TEOAS with clicks is currently the most widely used procedure (Harris & Probst, 1991; Prieve, Gorga & Neely, 1996; Avan, et al. 1997) and is limited to frequencies above 1000Hz and below 4000 Hz with most of the energy concentrated in the 2000 – 4000 Hz region (Probst & Harris, 1993). Researchers are investigating the measurement of TEOAEs with tone bursts and find it to be more accurate in the classification of normal and impaired hearing in frequencies higher than 1000 Hz (Vinck et al. 1998). Even though DPOAEs are also limited in the measurement of frequencies lower than 1000 Hz due to the rising noise floor (Durrant, 1992), it can measure frequencies much higher than 8000 Hz and that makes it the emission type that evaluates the broadest frequency range (Whitehead et al. 1996).

2.5.3 Efficiency of the Procedure

The efficiency of DPOAEs as a screening procedure was evaluated by Bonfils, Avan, Landias, Erminy and Biacabe (1997) and although this discussion is about DPOAE as an objective PTT prediction procedure which is more diagnostic, many applicable aspects can be extracted from their study. They investigated what the presentation level of the stimuli should be to have an ideal sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). The ideal discrimination level would allow the separation of all subjects without false negative or false positive responses. Bonfils et al. (1997) identified two test conditions according to what he primary goal of the test is: If the goal is to identify normal hearing (<30dB HL), then DPOAE primary stimulus levels of ≤50 dB SPL will discriminate between subjects with PTTs better or worse than 25-30 dB HL with great sensitivity and specificity. If the objective is to discriminate between profound and severe hearing losses, then primary stimulus levels of ≥ 60 dB SPL will discriminate between subjects with a hearing loss greater or lower than 55-60dB HL with great sensitivity and specificity. The main argument in the measurement of DPOAEs at high intensity levels (such as ≥ 60 dB SPL) is that only passive properties of the cochlea is measured and that it is not a true measurement of OHC functioning and therefore not a true measurement of hearing. Bonfils et al. (1997) argues that these passive emissions are only present in persons with a hearing loss greater than 60dB HL and that there is still an active component measurable in hearing losses <60 dB HL even if high intensity stimuli are used. This aspect makes DPOAEs even more applicable in PTT prediction of hearing impairment over a large decibel range.

2.5.4 Differential Diagnosis Between Sensory and Neural Hearing Impairment

Cane et al. (1992) found that OAEs alone are not sufficient enough to discriminate between cochlear and retrocochlear hearing losses, it should be used in conjunction with other tests such as ABR to determine etiology of sensory-neural hearing loss. Lonsbury-Martin and Martin (1990) discussed two cases where OAEs could discriminate between sensory and neural hearing losses but Robinette (1992) warned that many people have a cochlear condition in conjunction with the retrocochlear hearing loss and that only 18% of retrocochlear losses could be accurately identified with OAEs alone. Telischi, Roth, Stagner, Lonsbury-Martin and Balkany, (1995a) confirmed these findings by reporting that many tumors on the eight nerve causes pressure on the organ of Corti influencing its blood supply. It seems therefore, that DPOAEs can only be useful in differential diagnosis when used in conjunction with acoustic reflex measurements and ABR.

2.5.5 Good Test-Retest Repeatability

Lonsbury-Martin and Martin, (1990) measured high test-retest repeatability with DPOAEs for every ear they tested that demonstrated the same PTTs over a period of time and found it useful to monitor dynamic changes in outer hair cells. Kim, Sun, Jung and Leonard (1997) confirmed these findings. It seems that DPOAEs have excellent repeatability over a long period of time as long as the PTTs stays the same.

2.5.6 Fast Test Performance

The measurement of DPOAEs have been proven to be one of the fastest test available (Bonfils, et al. 1990; Cane, et al. 1992; Lonsbury-Martin, et al. 1992; Mauk & Behrens, 1993; Vohr, et al. 1993; Kim, et al. 1997).

2.5.7 Economic Test Performance

The fact that OAEs require less time to administer than most behavioral tests is one aspect that makes it very economic (Danhauer, 1997). The ease in which it can be measured and interpreted influence the number of specialized personnel necessary to perform the test and therefore also the cost. OAEs are simple to measure and require no advanced technical training (Bonfils et al. 1990; Mauk & Behrens, 1993 Whitehead et al. 1996; Kim et al. 1997). According to Norton and Stover (1994) DPOAEs are technologically the easiest types of emission to measure, being relatively artifact free and requiring no post hoc processing.

2.5.8 Non-invasive, Comfortable for the Patient

DPOAEs are measured by the insertion of a probe tip into the ear canal and is painless and comfortable for the person being tested. It does not require extensive preparation and cleaning of the area where the electrode is placed such as in the case of ABR. The ear canal has to be free of excessive earwax however and the probe should be chosen in such a way that it fits snugly in the ear canal (Bright, 1994). This procedure is not more complex than the probe fitting for tympanometry.

2.5.9 Age Differences

There is some controversy regarding the effect of age on DPOAE levels. According to some researchers, DPOAEs are present at birth (Popelka, et al. 1998) and amplitudes of DPOAEs do not decrease significantly with age, when adjusted for PTTs (Karzon, Garcia, Peterein & Gates, 1994). He and Schmiedt (1996) confirmed these findings and found that differences in DPOAE measurements between neonates and adults are due to sensitivity changes and not due to aging itself. Lasky (1998b) found that I/O functions of newborns and adults were similar, it was only in the fine spectrum where differences could be observed such as a more linear I/O function in adults with saturation at higher primary levels. The amplitudes of DPOAE measurements in adults and neonates were within 1.5 dB of each other for all age groups (Lasky, 1998a). Abdala (1998) found that DPOAEs could even be measured in premature neonates although the fine structure characteristics at 1500 Hz and 6000 Hz were different than measured in adults and suspect that there may be an immaturity in cochlear frequency resolution prior to term birth. No differences were observed at 3000 Hz.

Other researchers attempting to predict PTTs with DPOAEs found that age was definitely a variable that influenced PTT prediction, and that a prediction scheme based on DPOAE level has to be adapted to incorporate subject age (Lonsbury-Martin et al. 1991; Kimberley et al. 1994a; Kimberley et al. 1994b; De Waal, 1998).

DPOAEs can be successfully measured and are present and in all age groups from birth onwards, but in the prediction of PTTs, age has to be included as a significant variable.

It seems that DPOAE measurement meets all the requirements for the development of an optimal PTT prediction procedure. The development of such a procedure also involves a data processing technique that is capable of handling the data sets in the prediction of pure tones from DPOAEs. The following section reviews the requirements for an optimal data processing technique.

2.6 Requirements for an Optimal Data Processing Procedure in the Prediction of PTTs with DPOAEs

The prediction of PTTs with DPOAEs is difficult because of the following attributes of DPOAE measurements:

2.6.1 DPOAEs are Nonlinear in Nature

There are two aspects regarding the nonlinearity of DPOAE measurements. First, the distortion evoked by two-tone stimulation is generated by nonlinear elements that deform the response by creating frequencies that are not in the input signal (Martin et al. 1990b). Second, the increase in amplitude of the distortion does not grow linearly with an increase of amplitude of the stimuli (Nakajima et al. 1998).

The procedure used to analyze DPOAE data, to first identify the correlation between DPOAEs and PTTs and then to apply the correlation to make predictions, should be able to handle nonlinear data sets well.

2.6.2 The Complex Data Set of DPOAEs

There are numerous factors contributing to a complex data set. There are measurement variables such as the choice of f1 and f2, the choice of the primary loudness levels L1 and L2, the ratio of f1 and f2, the definition of a present DPOAE as a certain dB level above the noise floor and the magnitude of the noise floor itself that influence the DPOAE response measured. There are subject variables such as hearing sensitivity, different cochlear impairment types and subject history aspects such as exposure to noise that changes DPOAE responses. Any measured DPOAE is a complex data set influenced by many factors.

A data processing technique is needed that can handle complex data sets by including many variables at a time to determine the correlation between DPOAEs and PTTs and determine the significance of selected variables.

2.6.3 DPOAE Measurements are Often "Noisy"

In the DPOAE spectrum, some of the measurements in the frequency sweep may be absent where others may be preset and some subjects have a higher noise floor measured in the low frequencies. A data processing technique is needed that can make a definite correlation between two data sets even when dealing with absent or noisy data.

2.7 Artificial Neural Networks (ANNs) as a Data Processing Procedure for the Prediction of PTTs with DPOAEs

In the prediction of PTTs with DPOAEs it has been proved that multivariate techniques such as discriminant analysis work better than the more traditional single-variable applications of decision theory (Dorn, Piskorski, Gorga, Neely & Keefe, 1999). It has also been proved that the multivariate technique of a neural network approach is more accurate than discriminant analysis in the prediction of PTTs with DPOAEs (Kimberley et al. 1994a). ANNs have been proven to predict better than discriminant analysis in other field as well, such as bankruptcy prediction (Raghupathi, Schkade & Raju, 1993; Rahimian, Singh, Odom & Shara, 1993) and prediction of the US-$ and DM exchange (Hann & Steurer, 1996).

ANNs excel for a number of reasons: There is less need to determine relevant factors a priori: irrelevant data has such low connection strength that it has no effect on the outcome. Neural networks excel at determining what data is relevant and can cope with numerous factors at the same time. When hundreds of factors are at play, even if some only have a very small effect, neural network models are much more likely to be more accurate for difficult problems than any statistical model (Rahimian, Singh, Thammachote & Virmani, 1993). Neural networks are extremely fault tolerant and can learn from and make decisions based on incomplete data (Nelson & Illingworth, 1991). Even if some of the hardware fails, the neural network system will not be considerably changed. Blum (1992) even suggests training on noisy data to possibly enhance post-training performance. Furthermore, ANNs can deal with nonlinear correlations and has little difficulty outlying data points (Kimberley, et al. 1997).

When it comes to the prediction of PTTs with DPOAEs, a neural network approach is clearly the superior data processing technique to choose (Kimberley et al. 1994a).

2.8 Rationale

Progress in modern technology enabled audiologists to measure the exact degree, configuration, and site of hearing loss and to confirm these findings with a series of objective physiologic procedures, such as tympanometry, the acoustic reflex, ABR, and OAEs in adults and older children. It is in the evaluation of pure tone sensitivity in difficult-to-test populations such as neonates, infants, malingerers and the crucially ill that certain limitations arise and the need for another objective diagnostic evaluation tool was identified. The quest to predict pure tone thresholds accurately with DPOAEs arises therefore not from the need to replace existing conventional behavioral evaluation procedures, but to aid in the assessment of special populations.

An overview of current objective measurements available in audiology indicated that DPOAEs seem to be the most applicable measurement to predict pure tone thresholds objectively, rapidly and over a broad frequency range. Unfortunately, that is only half the case won due to the nonlinear, complex and noisy nature of DPOAE measurements and the inability of conventional statistics to correlate that to pure tone thresholds and make accurate frequency specific predictions. Artificial neural networks were identified as a possible data processing technique to attempt accurate frequency specific predictions.

The study preceding this one (De Waal, 1998) attempted to predict pure tone thresholds with DPOAEs and artificial neural networks. First, PTTs were categorized as normal or impaired (normal defined as < 20 dB HL) with DPOAEs and ANNs and correct classification of normal hearing was 92 % at 500, 87% at 1000, 84% at 2000 and 91% at 4000 Hz. Predictions of impaired hearing was less satisfactory partly due to insufficient data for the ANN to train on and also for similar reasons experienced in all the other studies described in the literature overview of studies attempting to predict hearing sensitivity with DPOAEs (see 2.3.4. PTT prediction with DPOAEs in perspective).

The rationale for this study is to improve prediction accuracy of pure tone thresholds at 500, 1000, 2000 and 4000 Hz in normal and hearing-impaired ears with DPOAEs and artificial neural networks.

It is anticipated that the prediction of PTTs with DPOAEs and ANNs will improve if the amplitude of the distortion product is included as a variable to determine correlations and make predictions. In the previous study, the amplitude of the distortion product was not included as a variable. Furthermore, more extensive experimentation of optimal neural network topologies will be experimented with, such as the optimal number of middle neurons and the structure of input data. Different types of networks will be investigated with as well as different types of network topologies and error tolerance levels. All these aspects that influence prediction accuracy of neural networks will be discussed in chapter three.

2.9 Summary

An overview of pure tone prediction with current objective diagnostic procedures revealed limitations in the evaluation of difficult-to-test populations. It seemed that, despite all the strengths and positive attributes of ABR, tympanometry, MLR, and LLR, a few weaknesses in these procedures made it difficult to measure exact hearing ability and site-of-lesion in populations such as neonates, infants, malingerers, the crucially ill and foreign speakers. ABR is currently the preferred method for diagnostic audiology in special populations (Hall III & Mueller III, 1997) but demonstrates weaknesses such as a limited frequency area in which hearing ability can be determined, lengthy test times, the possibility of sedation and the level of expertise and expense required (Ferraro & Durrant, 1994; Musiek et al. 1994; Robinette, 1994; Weber, 1994). It is therefore with much hope that many researchers turned their investigations to otoacoustic emissions.

Kemp (1978) identified different classes of otoacoustic emissions, depending on the stimuli used to evoke them. Spontaneous otoacoustic emissions (SOAEs) are only prevalent in half of normal hearing persons and can therefore not be implemented as a screening test or diagnostically (Lonsbury-Martin, 1994; Norton & Stover, 1994). Stimulus frequency otoacoustic emissions (SFOAEs) are not currently clinically used due to difficulties in separating in-going stimuli and out-going emitted responses (Lonsbury-Martin & Martin, 1990). Transient evoked otoacoustic emissions (TEOAEs) have been proven as a clinical acceptable hearing screening procedure, but the fact that they are only recordable in normal ears limited their diagnostic hearing testing applications (Kemp & Ryan, 1993; Lonsbury-Martin et al. 1992; Stevens et al. 1990). Distortion product otoacoustic emissions (DPOAEs) on the other hand, revealed many possibilities as a potential test of auditory functioning. First, it has been proven useful in both clinical and research settings, for it is the only emission type that can easily be recorded in many laboratory animals, allowing for experimental control of certain factors (Mills, 1997). Second, it can be measured in ears with a hearing loss of up to 65dB HL, therefore revealing information regarding outer hair cell functioning of hearing-impaired populations as well (Moulin et al. 1994). Third, it is the emission type that can be most easily compared to the conventional audiogram, due to the pure tone frequency nature of the stimuli that can be chosen to stimulate any specific region on the basilar membrane (Durrant, 1992; Lonsbury-Martin & Martin, 1990). Fourth, DPOAEs correlate well with pure tone thresholds and the configuration of the hearing loss (Durrant, 1992; Kimberly & Nelson, 1989; Stover et al. 1996a). Fifth, DPOAEs are not influenced by aspects such as gender and state of consciousness (Cacace et al. 1996; Karzon et al. 1994, Kemp, 1997). Of all the emission types, DPOAEs meet the most requirements for an optimal pure tone prediction procedure.

Many studies described the relationship between DPOAEs and pure tone thresholds (Avan & Bonfils, 1993; Bonfils et al.; 1991; Gaskill & Brown, 1990; Gorga et al. 1993; Kimberley et al. 1994b; Probst & Hauser, 1990; Stover et al. 1996a). Statistical methods used to date, such as multivariate (discriminant) analysis in the case of the study of Kimberley et al. (1994b), but also in all the other studies previously named, indicated a correlation between DPOAE measurements and behavioral pure tones. These studies however, could not predict the actual pure tone thresholds given only the distortion product responses (Lee et al. 1993, Kemp, 1997). The complexity of the data, the numerous variables involved and the possibility of a nonlinear correlation have been some of the reasons why conventional statistical methods could not predict pure tone thresholds given only DPOAEs, but only distinguish between normal hearing and hearing-impaired ears. There are also factors contributing to PTTs that do not contribute to the measurement of OAEs. It is clear that the emission field is still in the early stages of information gathering and that there are many aspects not yet fully understood or agreed upon. Regardless, major expectations concerning clinical applicability of DPOAEs is that DPOAEs will eventually be understood and developed to such an extent that DPOAEs can predict PTTs (Kemp, 1997; Lonsbury-Martin et al. 1997).

The complexities of the measurement, analysis and interpretation of distortion product emissions will be discussed in the following chapter.

Chapter 3: Parameters that Influence Pure Tone Threshold Prediction Accuracy with Distortion Product Otoacoustic Emissions and Artificial Neural Networks

3.1 Introduction

The preceding chapter formulated the need for an objective audiologic procedure to aid in the assessment of difficult-to-test populations. Limitations in current objective procedures inspired the ongoing effort to attempt to predict pure tone thresholds (PTTs) accurately across a wide frequency range. Despite the complex relation between DPOAEs and PTTs, many researchers turned to distortion product otoacoustic emissions as the possible new objective method due to promising predictions of normal hearing, especially in the high frequencies.

Efforts to predict impaired hearing thresholds, and hearing ability at low frequencies have been problematic for several reasons such as difficulties to determine a non-linear correlation between two data sets of which the one is complex and described in neural network terms as "fuzzy" or incomplete. Other relevant issues that contribute to the struggle are the interfering low frequency noise levels caused by subject breathing and electric equipment interference and the fact that pure tone thresholds involve a much broader evaluation of the whole auditory system and not just the evaluation of outer hair cell functioning as in the case of OAEs. Furthermore, the PTT prediction process is made more complex by the large number of critical factors or variables involved in the generation of the stimuli necessary to elicit a DPOAE. These factors are interrelated and influence the amplitude and occurrence of the distortion product. The choice of parameters used to elicit the DPOAE influences the DPOAE data set, therefore also the correlation to be determined between DPOAEs and PTTs and the accuracy of the prediction. An optimal set of parameters has to be identified to attempt to find the best combination of variables to accurately predict PTTs with DPOAEs. Lastly, the efficiency and accuracy of the data processing technique used also influences the PTT prediction process. Conventional statistical methods used in multivariate correlation studies have been found to be limited in their ability to solve complex nonlinear problems where hundreds of factors are at play (Nakajima et al. 1998; Kimberley et al. 1994a). Artificial neural networks (ANNs) have been found to have a superior ability in dealing with correlation determination in noisy nonlinear data sets (Nelson & Illingworth, 1991) and prediction of outcomes where numerous factors influence the data set (Rahimian et al. 1993). There are however many different kinds of networks available with different topologies and training methods and the choice and design of an appropriate network is one aspect that greatly influences the accuracy of prediction of PTTs with DPOAEs.

The aim of this chapter is to discuss all the factors that influence prediction accuracy of PTTs with DPOAEs and ANNs. First, all the parameters of the distortion product that play a role in PTT prediction will be discussed and the second half will concern itself with all the factors of neural network choice and design that influence prediction accuracy.

3.2 Stimulus Parameters of DPOAEs

In the generation of DPOAEs, two pure tones are used as stimuli with a frequency ratio that results in a partial overlap of the vibration fields in the cochlea. The ratio of the two stimulus frequencies f1 and f2, as well as their loudness levels, L1 and L2, determine where in the cochlea the maximum stimulation occurs (Kemp, 1997).

3.2.1 The f2/f1 Frequency Ratio

A study by Harris, et al. (1989) investigated which f2/f1 ratio yielded the maximal DPOAE amplitude. They used stimulus frequencies and level ranges that were representative of clinical audiograms and found that on the average, a ratio of 1.22 elicited the largest acoustic distortion products for emissions between 1 and 4kHz.

Nielsen et al. (1993) measured the cubic distortion product at six probe tone frequency ratios varying between 1.15 and 1.40 using equal level primaries of 75 dB SPL. The results showed that a frequency ratio between 1.20 and 1.25 optimizes the amplitude of the distortion product. A frequency ratio between 1.20 and 1.25 is also most applicable to the standard frequencies used in pure tone audiometry.

Other studies that described the optimum frequency ratio included f1/f2 = 1.225 (Gaskill & Brown, 1990), f1/f2 = 1.23 (Avan & Bonfils, 1993) and f1/f2 = 1.3 (Stover, et al. 1996a).

It would therefore seem that a frequency ratio of f1/f2 = 1.2 to 1.3 yields the best DPOAE amplitudes (Avan & Bonfils, 1993; Gaskill & Brown, 1990; Harris et al. 1989, Nielsen, et al. 1993; Stover et al. 1996a).

Another factor that influences DPOAE amplitude, apart from the frequency ratio, is the loudness level ratio of the primaries, namely L1 and L2. It is very important to choose the right frequency and loudness level ratios that yield maximum DPOAE amplitudes. These variables should be chosen in such a manner that the stimulus levels and frequency ranges are representative of clinical audiograms, to enable comparisons between the DPOAEs and pure tone thresholds (Moulin, et al. 1994).

3.2.2 The Loudness Levels of the Primaries

Mills (1997) studied the effect of the loudness levels of the primaries on the distortion product. The author concluded that the cubic distortion emission amplitude is not symmetric, so that given the same L1, higher emission amplitudes can occur for L2 > L1 compared to L1 = L2. Authors such as Stover et al. (1996a) found maximal DPOAE amplitudes when L2 > L1 by 10dB and Gaskill and Brown (1990) L1 > L2 by 15dB. Gorga, et al. (1993) found that 65/55 dB SPL primaries (L1/L2) resulted in maximal separation between normal and impaired ears. Some other studies reported best DPOAE amplitudes for L1=L2, but used very high stimulus levels, such as 75 dB SPL that might have triggered passive emissions from the cochlea (Rasmussen, et al. 1993). To elicit active DPOAE responses with the largest amplitude possible, most researchers recommend L1/L2 ratios in the range of 10-15dB" (Mills, 1997; Stover et al. 1996a; Gaskill & Brown, 1990).

3.2.3 High Versus Low Levels of Stimulation

It seems that there are different mechanisms involved in high and low level stimulated DPOAEs (Harris & Probst, 1997). DPOAEs evoked with low level primaries (< 62 dB SPL) are dominated by active cochlear mechanical processes and are strongly correlated with auditory thresholds. DPOAEs evoked with high level primaries on the other hand, are dominated by passive cochlear mechanics and do not provide frequency specific information on the local cochlear state (Avan & Bonfils, 1993; Kummer et al. 1998; Mills, 1997).

Bonfils et al. (1991) investigated the level effect of the primaries on the distortion product. Equilevel primaries ranging from 84 dB SPL to 30 dB SPL were delivered over a geometric mean frequency range of 485 Hz to 1000 Hz. They found that I/O functions tested with low level primaries (intensities below 60 dB SPL) and frequency ratios around 1.2 showed saturated growth. When primary intensities exceeded 66 dB SPL or when frequency ratios were greater than 1.3 or lower than 1.14, the input output functions became linear without any clear saturating plateau. The authors concluded that DPOAEs generated by primary intensities below 60 dB SPL probably have their origin in the outer hair cells. With high level stimuli however, it is probable that only passive properties of the cochlea contribute to the emission.

Apart from all the parameters that should be specified, there are also two different ways to construct DPOAE testing.

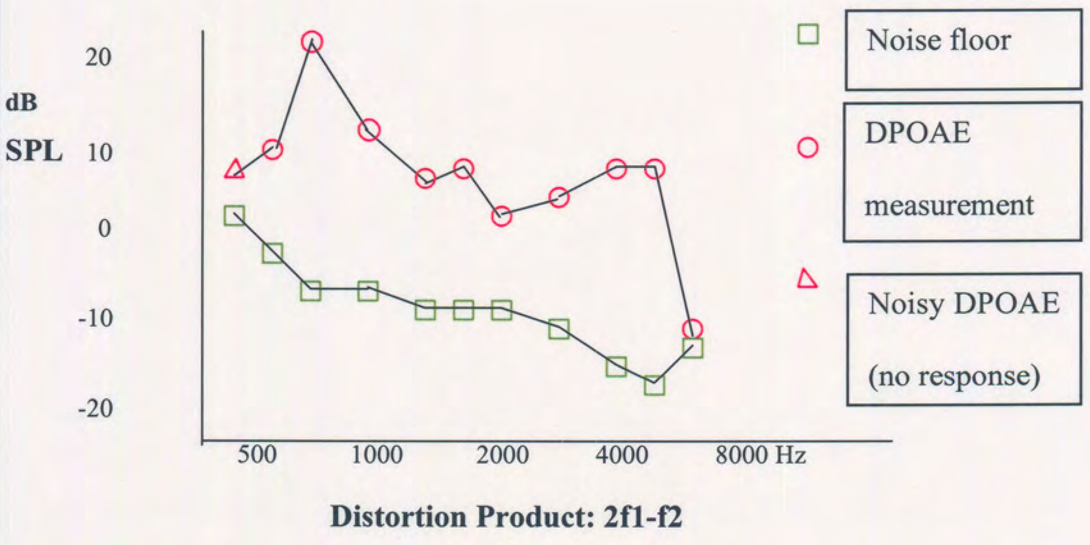
3.2.4 DP Gram Versus I/O Function

In the measurement of DPOAEs, either the frequencies are changed and the loudness level kept constant (this is sometimes referred to as a "distortion product audiogram" or DP Gram) or the frequencies are being kept constant while the loudness level is changed (an input/output function (I/O Function) is obtained). It should be noted that the "distortion product audiogram" does not include the concept of threshold, as does the conventional audiogram in this case.

3.2.4.1 The DP Gram

The first form, namely the DP Gram, depicts DPOAE amplitude as a function of stimulus frequency at a fixed loudness level (Lonsbury-Martin & Martin, 1990). In other words, the loudness level is kept constant and the frequencies are changed.

In this manner, a test cochlea can be evaluated over a large frequency range by comparing the evoked DPOAE amplitudes to average amplitudes determined from a population of normal hearing individuals (Lonsbury-Martin, 1994). The DP Gram is analyzed by comparing DPOAE amplitudes with average amplitudes obtained from normal hearing persons. A DP Gram does not obtain "threshold" information, as does the conventional pure tone audiogram. Figure 3.1 represents a DP Gram of a normal hearing person.



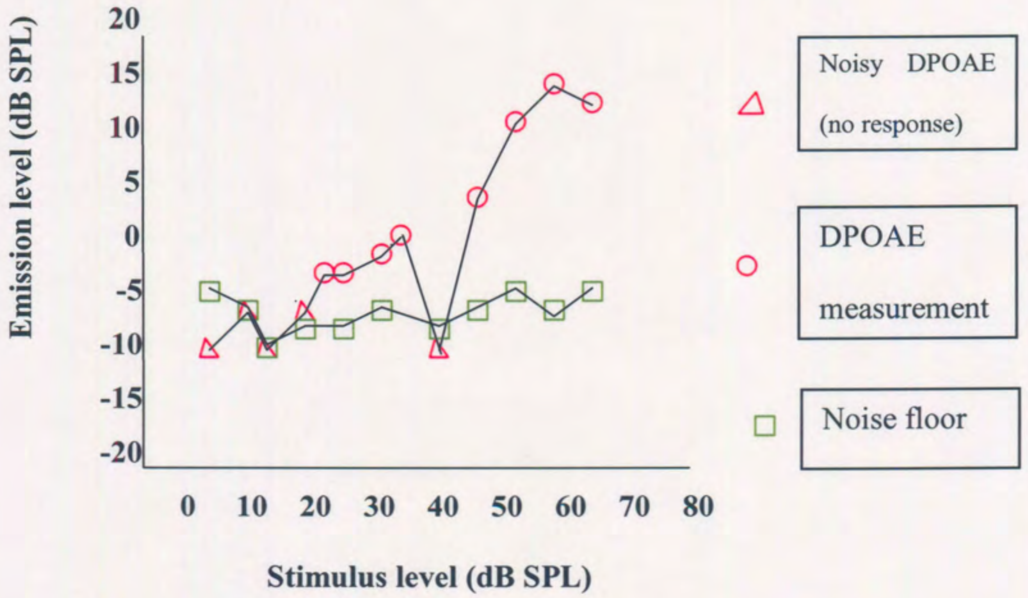
DP Gram of a normal hearing adult's right ear at a loudness level of L1=65 dB SPL, L2=55 dB SPL, in the frequency region of 2f1-f2 from 406 Hz to 4031 Hz.

Figure 3.1: Example of a DP Gram.

3.2.4.2 The I/O Function

The second form of DPOAE measurement, namely the I/O function, can be used to determine the dynamic range of the distortion generation process (Lonsbury-Martin & Martin, 1990). This procedure described the growth of DPOAE amplitude at a constant frequency (Lonsbury-Martin, 1994). An I/O function can be obtained when the frequencies of the primaries are kept constant and the loudness levels are changed (Norton & Stover, 1994).

An example of an I/O Function can bee seen in Figure 3.2.



I/O function of a normal hearing adult. The fixed frequencies are f1= 1660Hz, f2= 2000 Hz and the loudness levels vary from 10 dB to 80dB SPL

Figure 3.2: Example of an I/O function. (From Norton & Stover, 1994:457).

The threshold of a DPOAE depends almost entirely on the noise floor and the sensitivity of the measuring equipment whereas the DPOAE amplitude is greatly influenced by the frequency ratio and decibel ratio of the primaries (Norton & Stover, 1994; Martin et al. 1990b).

To determine the normalcy of an I/O function, the detection threshold (i.e. the stimulus level where the DPOAE reaches a criterion level, for example 3 dB, above the noise floor) is compared to average detection thresholds of normal hearing individuals (Lonsbury-Martin & Martin, 1990). The DPOAE threshold should not be confused with the pure tone audiogram threshold, and cannot be directly compared (Norton & Stover, 1994).

There is not yet clear consensus on the best testing procedure to identify normal and impaired ears. Most researchers use a combination of the two procedures or perform both procedures separately (Martin, et al. 1990a, Spektor et al. 1991 and Smurzynski, Leonard, Kim, Lafreniere, Marjorie and Jung, 1990; Moulin et al. 1994; Kimberley & Nelson, 1989). It seems plausible to gain as much DPOAE threshold and amplitude information as possible by combining the two procedures.

The effect of age and gender on the prevalence of DPOAEs will be discussed next.

3.3 Subject variables: The Effect of Age and Gender on DPOAEs

Subject age and gender influence many aspects of auditory function (Hall III, Baer, Chase & Schwaber, 1993). Within the first decade after the discovery of auditory brainstem response (ABR), many studies were conducted to investigate the influence of age and gender. Significant differences were found between different age and gender groups. Ever since then, these two factors have been routinely taken into consideration in the interpretation of ABR results (Weber, 1994) and are always investigated in new diagnostic audiology fields.

3.3.1 The Effect of Age on the Distortion Product

There is some debate about the effect of age on DPOAEs. Some authors found statistically significant decreases in amplitudes of other emission types such as TEOAEs with increasing age (Norton & Widen, 1990). In the case of DPOAEs, it seems that DPOAEs are present from birth (Popelka, et al. 1998) and is as easily measurable in an infant as in an adult (Lasky 1998b). Some researchers believe that age affects the amplitude of DPOAEs negatively (Lonsbury-Martin et al. 1990) and others argue that age related differences could be attributed to sensitivity changes related with aging, rather than aging itself (He & Schmiedt, 1996). There are also researchers that found that DPOAE amplitudes for adults and neonates are similar, but some differences in the fine structure of the distortion product can be measured (Lasky, 1998a+b). Some of these studies will be discussed briefly.

Lonsbury-Martin et al. (1990) indicated that in the presence of normal hearing (pure tone thresholds lower than 10 dB HL), DPOAE amplitudes and thresholds, especially those associated with high frequency primary tones were significantly correlated with the subject's age. The subjects ranged from 21-30 years of age. It should be noted however, that the authors described the audiograms of the 30 year old subjects as "exhibiting a high frequency hearing loss pattern" (Lonsbury-Martin et al. 1990:10) with hearing thresholds around 10dB HL. The younger subjects had pure tone thresholds of 0-5 dB HL. The lower DPOAE amplitudes and thresholds found in the results of the 30-year-old subjects can therefore be partly explained by higher pure tone thresholds and not solely by the subject's age.

Another study by Karzon, et al. (1994) investigated DPOAEs in the elderly to determine the age effect on DPOAEs. DPOAE results of 71 elderly volunteers ranging from 56-93 years were compared to DPOAE results of normal hearing young adults, age 19-26 years. The authors found that the amplitudes of DPOAEs did not increase significantly with age, when adjusted for pure tone levels. "Although DPOAEs are reduced with age, this effect is largely mediated by age-related loss of hearing sensitivity." (Karzon et al. 1994:604). Avan and Bonfils (1993) confirmed this viewpoint and stated that many of the age related effects were due to high frequency hearing losses even when subjects were "normal" within their age category. He and Schmiedt (1996) also stated that when pure tone thresholds are controlled, there is not a significant aging effect on DPOAE amplitudes and that the negative correlation between DPOAE levels and age is due to changes in hearing threshold associated with aging rather than age itself.

Lasky (1998b) found that I/O functions of newborns and adults were similar; it was only in the fine spectrum where differences could be observed such as a more linear I/O function in adults with saturation at higher primary levels. The amplitudes of DPOAE measurements in adults and neonates were within 1.5 dB of each other for all age groups (Lasky, 1998a). Abdala (1999) found that DPOAEs could even be measured in premature neonates although the fine structure characteristics at 1500 Hz and 6000 Hz were different than measured in adults and suspect that there may be an immaturity in cochlear frequency resolution prior to term birth. No differences were observed at 3000 Hz.

When it comes to the prediction of PTTs with DPOAEs however, some researchers found that age enhanced predictive accuracy considerably (Lonsbury-Martin et al. 1991; Kimberley et al. 1994a; Kimberley et al. 1994b; De Waal, 1998). For all these studies, more accurate PTT predictions were made when subject age was included. It seems that subject age is a very important factor to be included in any prediction scheme based on DPOAE levels. Even though amplitudes of adults and children seem similar, there is much information in the differences measured in the fine structure across different age groups that enhances predictive accuracy of PTTs.

Another potentially relevant factor may be the influence of gender on the prevalence of distortion product otoacoustic emissions.

3.3.2 The Effect of Gender on the Distortion Product

Gender differences have been reported in other emission types. Cacace, et al. (1996) reported spontaneous otoacoustic emissions to be more prevalent in females than males and higher incidence of SOAEs in right ears than left ears. Hall III et al. (1993) indicated that TEOAE amplitudes are significantly larger for females than males.

Lonsbury-Martin et al. (1990) conducted a study to investigate basic properties of the distortion product including the effect of gender on the prevalence of DPOAEs. A comparison of DPOAE amplitudes and thresholds failed to reveal any significant differences except at 4 kHz. Women revealed significantly lower DPOAE thresholds at 4 kHz (about 10 dB lower). The pure tone audiometry thresholds for men and women at 4 kHz were the same. Gaskill and Brown (1990) and Cacace et al. (1996) reported that DPOAEs were significantly larger in female than male subjects tested in the frequency range of 1000- 5000Hz. Both studies however, indicated that the female subjects in their studies had more sensitive auditory thresholds than the males (an average of 2.4 dB better). The differences found between the two groups could therefore not be explained by gender only.

Cacace et al. (1996) attempted to explain some of the reasons why the females had higher amplitudes than the males in the higher frequencies. One reason is the existence of a spontaneous otoacoustic emission (SOAE) in conjunction with DPOAE measurement. Several authors described the effect that a SOAE could have on a DPOAE (Moulin et al. 1994; Probst & Hauser 1990; Kulawiec & Orlando, 1995). If a spontaneous emission exists within 50 Hz of the primary frequencies used to elicit a DPOAE, the spontaneous emission could enhance the DPOAEs amplitude significantly under certain experimental conditions (Kulawiec & Orlando, 1995; Probst & Hauser, 1990). Spontaneous emissions are more prevalent in females than in males and could therefore possibly explain the higher DPOAE amplitudes in females.

This amplitude amplification effect that SOAEs have on DPOAEs cannot always clearly be seen. Cacace et al. (1996) reported that no systematic peaks or notches could be observed in DPOAE responses in the presence of a spontaneous otoacoustic emission in any of the subjects they tested. The mere presence of a SOAE in a frequency region close to the primaries cannot be taken as evidence of amplitude amplification. It is however so, that this gender effect is greatly reduced when only subjects with no SOAEs are considered.

Gender effects on DPOAEs are apparently limited to minor differences in DPOAE amplitudes.

3.4 Which Aspect of the DPOAE can best be Correlated with Pure Tone Thresholds

In the first two decades after DPOAEs were discovered, it was not clear whether it is the f1, f2, the GM frequency or the 2f1-f2 frequency that is actually being stimulated on the basilar membrane. Most authors agreed that DPOAEs appear to be generated in the region stimulated between the primary frequencies, rather than the frequency at the distortion product (Martin et al. 1990b; Kimberley et al. 1994b; Smurzynski et al. 1990; Moulin et al. 1994; Harris et al. 1989). Some studies supported the notion that the generation of the distortion product correlates best with the cochlear place near the geometric mean (GM) of the primaries (Martin et al. 1990b; Lonsbury-Martin & Martin, 1990; Bonfils et al. 1991). These authors concluded that the acoustic distortion product at 2f1-f2 should be correlated with PTTs near the GM of the primaries.

According to research conducted by Kimberley et al. (1994b) and Harris et al. (1989), the features that best correlate with PTTs are those associated with f2 values close to the pure tone threshold frequency. The distortion product, according to these authors, is generated very close to the f2 cochlear place and therefore they correlated PTTs with the f2 frequency of the distortion product.

Recent research on the exact location of the basilar membrane that is simulated with the 2f1-f2 distortion product described a two source model for DPOAE generation (Knight & Kemp, 1999a; Mauermann, et al. 1999a+b; Talmadge, Long, Tubis & Dhar 1998; Shera & Guinan, 1998). According to this theory, there is not just one, but there are two areas on the basilar membrane that contribute to the energy measured in DPOAE testing. The first source of energy comes from the overlap region of the two primary frequencies. Although the waves of the two primaries are spread out over the whole basilar membrane, it is the area about 1mm around the f2 region on the basilar membrane that contributes to most of the energy measured in a DPOAE. This area is known as the "f2 site" (Mauermann, et al. 1999a). DPOAE levels are however not just determined by the health of the cochlea at the f2 place (Talmadge et al. 1998). There is a second source on the basilar membrane that contributes to the energy being measured and it comes from the distortion product wave component that travels apically from the overlap region and is reflected at the 2f1-f2 site, also known as the "re-emission site." The spectral fine structure observed in the ear canal is a reflection of energy coming from both these sources.

The fact that more than one area of the basilar membrane contribute to a DPOAE response influences the method in which a correlation is determined between DPOAEs and PTTs. If the two source model of DPOAE generation is the case, it could be argued that one cannot merely correlate the f2 value or merely the 2f1-f2 value with a PTT frequency, but that the data processing technique has to be able to use both frequencies in the correlation determination process. Artificial neural networks are capable of using any number of frequencies in the correlation determination with one PTT frequency and can determine the significance of each frequency separately. This aspect makes it a very desired data processing technique to use for PTT prediction with DPOAEs.

The following section discusses the artificial neural network as a data processing technique and how it operates in more detail.

3.5 Aspects of the Artificial Neural Network that influences Prediction Accuracy of PTTs

Designing a neural network is somewhat of a mysterious process. The learning process of a neural network is a tedious and painstaking trial-and-error effort. There are no standards for learning algorithms for ANNs, partly because every data set and how the information can be presented to the network is highly unique. Another factor of importance influencing the learning process is the quality of the material that is used to train on, how noisy it is and how significant the correlation is between the data sets.

One has to have a clear understanding of what a neural network is, how it operates, learns and predicts to understand how the design of the network influences the outcome. The following discussion will serve as background to understand the whole process.

3.5.1 Definition of Artificial Neural Networks

Artificial neural networks (ANNs) are a new information processing technique that attempts to simulate or mimic the processing characteristics of the human brain

(Medsker, Turban & Trippi, 1993). An artificial neural network is an algorithm for a cognitive task, such as leaning or optimization, recognition of a pattern or retrieval of large amounts of data (Muller & Reinhardt, 1990). Hiramatsu (1995:58) defined neural networks quite effectively: "A neural network is generally a multiple-input, multiple-output non-linear mapping circuit, which can learn an unknown non-linear input-output relation from a set of examples."

ANNs were inspired by studies of the central nervous system and the brain (Medsker et al. 1993; Klimasauskas, 1993) and therefore share much of the terminology and concepts with its biological counterpart. This biological analogy will be discussed in the next section.

3.5.2 "Anatomy" and "Physiology" of Artificial Neural Networks: A Discussion of Concepts and Terms

Neural networks were initially developed to gain a better understanding of how the brain works. It resulted in computational units, called neural networks, that work in ways similar to how we think the neurons in the human brain work. Several human characteristics such as "learning, forgetting, reacting or generalizing" and also the biological aspects of networks consisting of neurons, dendrites, axons and synapses were ascribed to these artificial neural networks in order to promote understanding of these abstract terms (Nelson & Illingworth, 1991). Some of the terminology of neural networks will be reviewed briefly.

3.5.2.1 Biological Neural Networks

The human brain is composed of cells called neurons and estimates of the number of neurons in the human brain range up to 100 billion (Medsker, et al. 1993). Neurons function in groups called networks. Each network contains several thousand highly interconnected neurons where each neuron can interact directly with up to 20 000 other neurons (Nelson & Illingworth, 1991). This architecture can be described as parallel distributed processing, where the neurons can function simultaneously (Muller & Reinhardt, 1990). In contrast with conventional computers which process information serially, or one thing at a time, the human brain's parallel processing ability enables it to outperform supercomputers in some areas regarding complexity and speed of problem solving such as pattern recognition (Blum, 1992).

A typical biological neuron (Figure 3.3) consists of a cell body containing a **nucleus, dendrites** which provides input to the cell and an **axon**, which carries the output signal from the nucleus (Hawley, Johnson & Raina, 1993). Very often, the axon of one neuron merges with the dendrites of a second neuron. Signals are transmitted through synapses. A synapse is able to increase or decrease the strength of the connection and causes inhibition or excitation of a subsequent neuron (Nelson & Illingworth, 1991). Although there are many different neurons, this typical neuron serves as a functional basis to make further analogies to artificial neural networks.

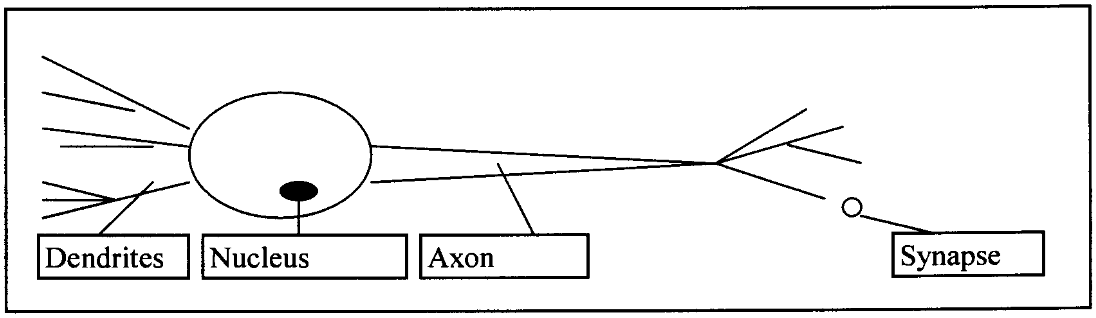


Figure 3.3: A biological neuron (From Medsker, et al. 1993:5)

“Artificial neural networks are based only loosely on biology. We don’t understand how the brain works or what intelligence really is.” (Nelson & Illingworth, 1991: 41) ANNs are an attempt to create a technique to process information in the same fascinating and highly successful way the human mind does. It borrows features from the biological system, to enable scientists to form compact representations of complex problems such as image recognition and forecasting or prediction.

3.5.2.2 Artificial Neural Networks (ANNs)

The building blocks of an artificial neural network are also referred to as neurons, or sometimes as a node, a perceptron or processing element. These artificial neurons (Figure 3.4) bear only a modest resemblance to biological neurons in the sense that they can perform approximately three of the processes we know neurons perform. (Biological neurons perform about 150 processes in the human brain)(Nelson & Illingworth, 1991).

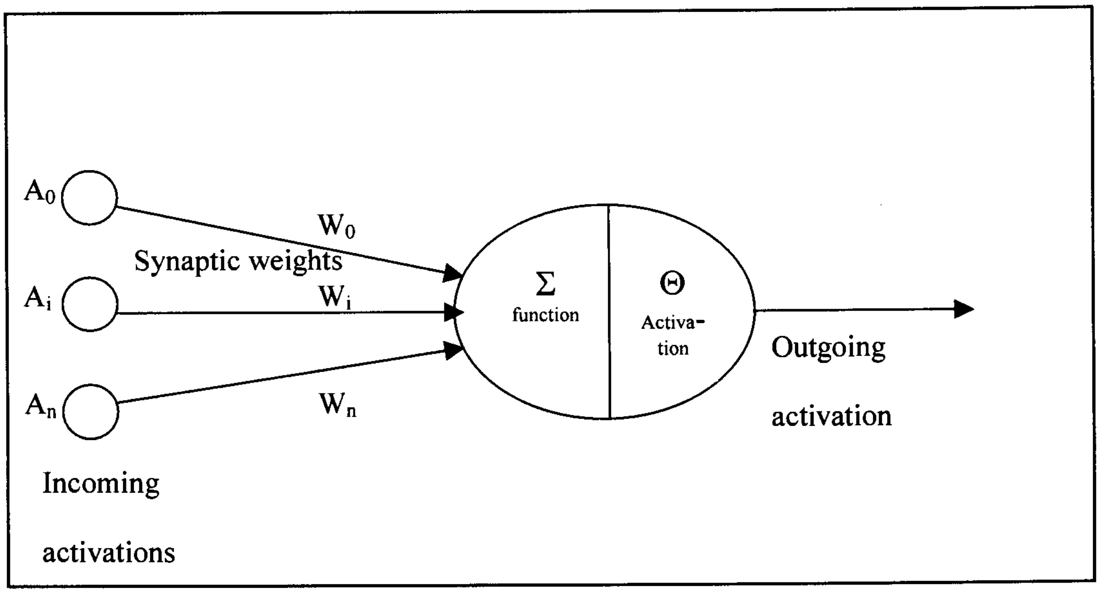


Figure 3.4: An artificial neuron (From Blum, 1992: 37).

An artificial neuron receives an incoming stimulus **(input)** analogous to the impulses that the dendrites of biological neurons receive. The second function is to **calculate** a total for the combined input signals and compare it to some threshold level. Finally, it determines the **output** and sends it out just like a biological neuron sends out an output through its axon (Muller & Reinhardt, 1990).

Several of these artificial neurons or nodes (indicated with circles in Figure 3.5) can be combined to make a **layer** of nodes. Figure 3.5 illustrates the middle or hidden layer been formed, connected to inputs, presented as input one (I_1) to any number of inputs (I_n) and also connected to outputs.

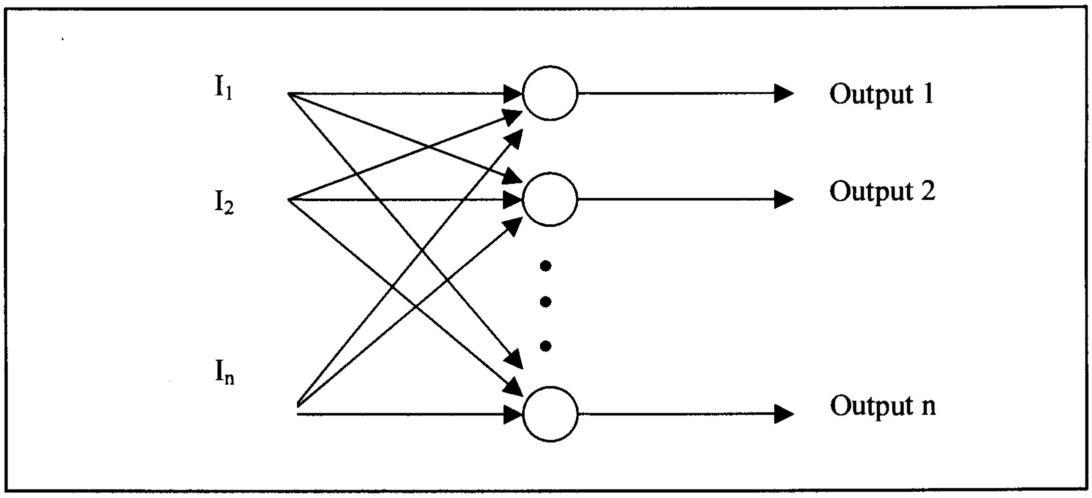


Figure 3.5: Inputs to several nodes to form a layer (From Nelson & Illingworth, 1991: 49).

In this representation, the middle layer is highly interconnected with the inputs (all inputs are connected to all middle level neurons) but only forwardly connected with the outputs. Middle layer neurons can also be highly interconnected to output neurons: the way in which neurons are connected to other layers is specified in the neural network design. The dots in the middle layer suggests that any number of neurons in this layer is possible and is determined by trial-and-error during network training so suit the complexity of the data. To form an artificial neural network, several layers are connected to each other. This is illustrated in Figure 3.6.

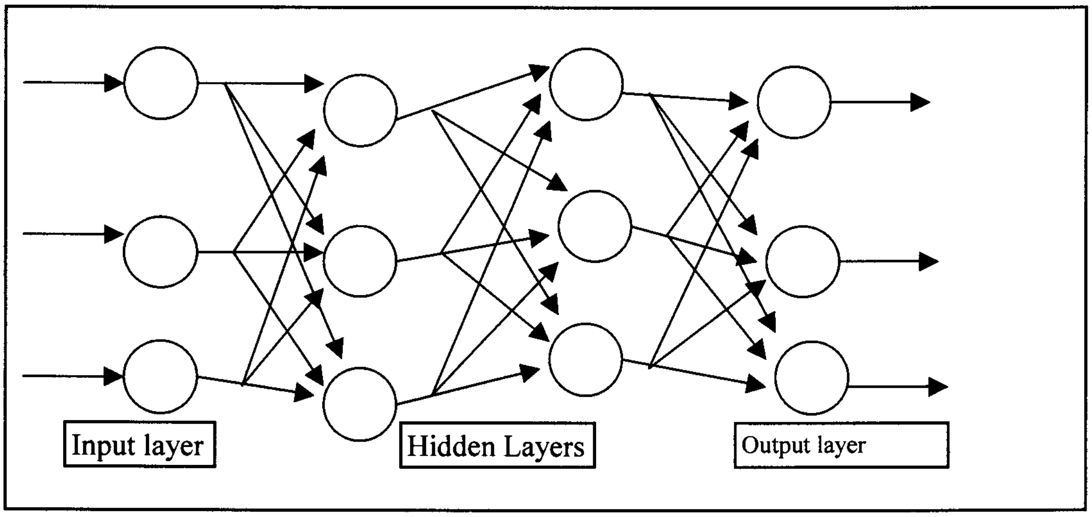


Figure 3.6: Connection of several layers to form a network (From Nelson & Illingworth, 1991:50).

From figure 3.6 it is clear that several different layers can be distinguished. The first layer that receives the incoming stimuli is referred to as the **input layer.** The network's outputs are generated from the **output layer** and all the layers in between are called the **hidden layers or middle layers.** In this four-layered network, all input and middle or hidden layers are highly interconnected with each other.

The "anatomy" of artificial neural networks has just been reviewed. The terminology used in the "physiology" or working of an artificial neural network will be discussed next.

The first layer of neurons, called the input layer, receives the incoming stimulus. The next step is to calculate a total for the combined incoming stimuli. In the calculation of the total of the input signals, there are certain weighting factors: Every input is given a relative **weight** (or mathematical value), which affects the impact or importance of that input. This can be compared to the varying synaptic strengths of the biological neurons. Each input value is multiplied with its weight value and then all the products are added up for a weighted sum. If the sum of all the inputs is greater than the threshold, the neuron generates a signal (output). If the sum of the inputs is less than the threshold, no signal (or some inhibitory signal) is generated. Both types of signals are significant (Blum, 1992; Nelson & Illingworth, 1991). These weights can change in response to various inputs and according to the network's own rules for modification. This is a very important concept because it is through repeated adjustments of weights that the network "learns" (Medsker, et al. 1993).

Medsker, et al. (1993) summarized the crucial steps of the **learning** process of an artificial neural network very effectively:

"An artificial neural network learns from its mistakes. The usual process of learning or training involves three tasks:

1) Compute outputs.
2) Compare outputs with desired answers.
3) Adjust the weight and repeat the process." (Medsker et al. 1993:10)

The learning process usually starts by setting the weights randomly. The difference between the actual output and the desired output is called Δ. The objective is to minimize Δ, or even better, eliminate Δ to zero. The reduction of Δ is done by comparing the actual output with the desired output and by incrementally changing the weights every time the process is repeated until the desired output is obtained.

Hawley, et al. (1993) compared the learning process of an artificial neural system (ANS) with the training of a pet: "An animal can be trained by rewarding desired responses and punishing undesired responses. The ANS training process can also be thought of as involving rewards and punishments. When the system responds correctly to an input, the "reward" consists of a strengthening of the current matrix of nodal weights. This makes it more likely that a similar response will be produced by similar inputs in the future. When the system responds incorrectly, the "punishment" calls for the adjustment of the nodal weights based on the particular learning algorithm employed, so that the system will respond differently when it encounters the same inputs again. Desirable actions are thus progressively reinforced, while undesirable actions are progressively inhibited." (Hawley, et al. (1993:33).

The learning of a neural network takes place in its **training** process. Every neural net has two sets of data, a **training set** and a **test set**. The training phase of a neural network consists of presenting the training data set to the neural network. It is in this training process, that the network adjusts the weights to produce the desired output for every input. The process is repeated until a *consistent* set of weights is established, that work for all the training data. The weights are then "frozen" and no further learning will occur. After the training is complete, the data in the test set is presented to the neural network. The set of weights as calculated by the training set is then applied to the test set. The presentation of the test set is the final stage in the neural network where the answer is given whether it is to predict an outcome, find a correlation, or recognize a pattern (Blum, 1992; Nelson & Illingworth, 1991;Medsker, et al. 1993).

Another term that justifies some explaining is the **programming of a neural network**. "Artificial neural networks are basically software applications that need to be programmed" (Medsker, et al. 1993:22). A great deal of the programming is about training algorithms, transfer functions and summation functions. According to Medsker, et al. (1993) it makes sense to use standard neural network software where computations are preprogrammed. Several of these preprogrammed neural networks are available on the market. Every person using an artificial neural network however, has certain additional programming that needs to be done. It might be necessary to program the layout of the database, to separate the data into two sets, namely, a training set and a test set, and lastly to transfer the data to files suitable for input into the standard artificial neural network.

The basic components of a general neural network have been discussed. The next section will review different types of neural networks.

3.5.3 Different Types of Artificial Neural Networks

There are different types of neural networks, categorized by their topology (the number of layers in the network). To provide just a limited overview of the basic types of neural networks, the single layer network, the two layer network and multi layer networks will be discussed briefly (Rao & Rao, 1995).

3.5.3.1 Single Layer Networks

The single layer network has only one layer of neurons and can be used for pattern recognition. The specific type of pattern recognition in this case is called autoassociation, where a pattern is associated with itself. When there is some slight deformation of the pattern, the network is able to relate it to the correct pattern.

3.5.3.2 Two Layer Networks

Some models have only two layers of neurons, directly mapping the input patterns to the outputs. Two layer models can be used when there is good similarity of input to output patterns. When the two patterns are too different, hidden layers are necessary to create further internal representation of the input signals. Two layer networks are capable of heteroassociation where the network can make associations between two slightly different patterns (Blum, 1992; Nelson & Illingworth, 1991).

3.5.3.3 Multi Layer Networks

Several types of multi layer networks exist. The most common multi layer network is the feedforward network with a backpropagation learning algorithm. According to Rao & Rao (1995), over 80% of all neural network projects in development use backpropagation. "Back propagation is the most popular, effective, and easy-to-learn model for complex, multi layered networks." (Nelson & Illingworth, 1991:121). Most backpropagation networks consist of three layers, an input layer, an output layer and a hidden or middle layer (Figure 3.7). The connections between the layers are forward and are from each neuron in one layer to every neuron in the next layer.

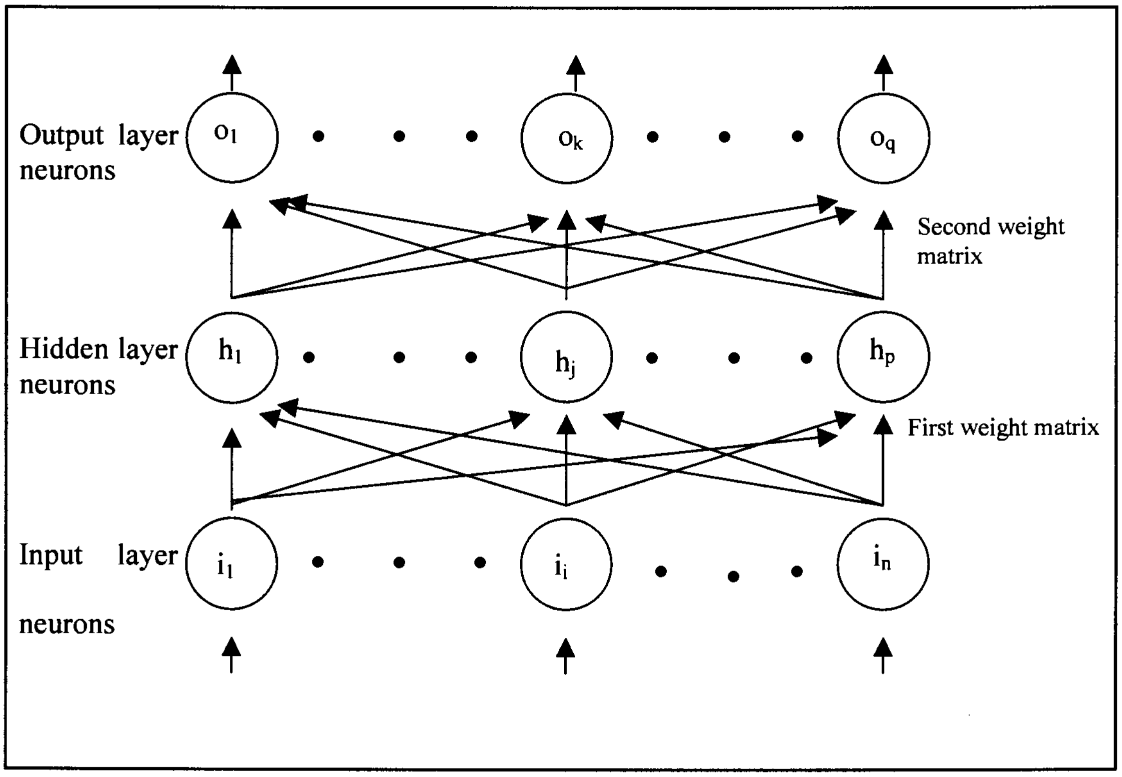


Figure 3.7: Diagram of a feedforward backpropagation neural network (From Blum, 1992: 56).

There are two phases in the learning process of a feedforward backpropagation neural network. The first stage propagates the input signal through the network in a forward manner (from the input layer, to the hidden layer, to the output layer). The second stage is to calculate the difference between the actual output and the desired output and to adapt the output, by changing the weights in the network in a backward manner (from the output layer, to the hidden layer, to the input layer).

The error signals of the output are propagated back into the network for each cycle. At each backpropagation, the hidden layer neurons adjust the weights of connections and reduce the error in each cycle until it is finally minimized (Blum, 1992).

This process is clearly summarized by Nelson and Illingworth as follows: (1991: 122): "The whole sequence involves two passes: a forward pass to estimate the error, then a backward pass to modify weights so that the error is decreased." Backpropagation networks require supervised learning where the network is trained with a set of data (training set) similar to the test set.

Now that the functioning of a neural network is understood, attention can be given to the factors in ANN design that influence prediction accuracy of PTTs with DPOAEs and ANNs.

3.5.4 ANN Factors Influencing Prediction Accuracy of PTTs with DPOAEs

Even when a standard preprogrammed artificial neural network is used, certain parameters has to be specified and can be experimented with to produce a more desired outcome. These parameters include the topology, error tolerance levels and the format of the input data.

3.5.4.1 The Topology

The topology of a network is determined by the number of layers in the network and the number of nodes in each layer.

3.5.4.1.1 Number of Layers in the Network

When there is good similarity between input and output data, only two layers are needed, but when the structure of the input pattern is quite different from the output, hidden layers are needed to create an internal representation from the input signals (Nelson & Illingworth, 1991). The ability of the network to process information increases in proportion to the number of layers in the network. In the design of a neural network, hidden layers can be added one by one until suitable outputs can be achieved. According to Hornik, Stinchcombe and White (1989) however, when a multilayered feedforward network is used, only one hidden layer is enough for any complex problem, provided that there are enough neurons in the hidden layer. According to these authors, failures in feedforward networks with one layer can be attributed to inadequate learning or the presence of a stochastic or random relation rather than a deterministic relation between two data sets. It would therefore seem that a feedforward backpropagation network with three layers, one input layer, one output layer and one hidden layer is sufficient for this application.

3.5.4.1.2 Number of Nodes in each Layer

The number of nodes in the input layer is determined by the amount of data that is fed into the network. For example, if all the present and absent DPOAE responses of 11 frequencies at eight loudness levels serve as input information, then there should be at least 88 input nodes to represent this data numerically. If gender is added as a variable then one more node has to be added to represent gender as either a one or a zero. Every additional input variable needs extra input nodes, and the number of nodes needed is determined by the way in which the data is presented to the network. The input layer therefore only serves as a buffer in which information can be "fanned" through to the next layer (Blum, 1992).

The number of nodes in the output layer is determined by the objective of the neural network and the format in which it is presented. For example, if the objective is to predict a pure tone threshold at a certain frequency and the format is to predict it into one of eight categories of 10dB each, then there will be eight nodes in the output layer. The output layer merely makes the network information available to the outside world (Nelson & Illingworth, 1991).

The determination of the number of nodes in the hidden layer is less straightforward. This number influences network capacity, generalization ability, learning speed and the output response. Fujita (1998) argues that on the one hand it is best to have as many hidden layer neurons as possible for capacity and universality in application to function approximation. On the other hand, from the standpoint of generalization, the number should not be too large for heuristic learning systems in which the best network configuration is unknown beforehand. Too many hidden layer neurons can also reduce the speed of the network considerably. It is difficult to determine the middle level neuron quantity before the learning is done, and it is best to adjust node numbers during learning.

According to Blum (1992), the best size is determined by familiarity with the application. Nelson and Illingworth (1991) describe it as a trial and error effort to determine which size yields optimum results.

3.5.4.2 The Error Tolerance

A feedforward neural network propagates information from the input level to the middle level to the output level, but errors are backpropagated during training. The purpose of the backpropagation of errors is to change the weights between layers to handle the prediction better the next time it encounters the same information. Errors in the output indicates that there are errors in the two sets of weights connected to the hidden layer and are used as a basis for adjustment of the weights between the input and hidden layer and output and hidden layer. The weights connected to the hidden layer have to be adjusted repeatedly until prediction error falls within a specified level. Error tolerance therefore refers to how accurately a network predicts the answer, but also how effectively it trains or learns (Blum 1992; Rao & Rao 1995).

When prediction error is set as close to zero as possible, only answers that are completely correct are accepted. Although it might seem logical to set error tolerance levels as close to zero as possible, it is not always practical, for two reasons. A network with error tolerance of as close to zero as possible trains much longer before accurate enough predictions can be made. Sometimes the training phase becomes so long for each experiment (from hours to days to weeks) that it becomes unpractical to run hundreds of experiments, which is the case when 120 ears have to be predicted at four frequencies. The second disadvantage of very small error tolerance levels is the network's ability to generalize decreases. When a DPOAE data set slightly out of the ordinary has to be predicted, a network with very low error tolerance levels is often incapable of a general prediction and can not reach a training set that falls within the specified error tolerance level.

Error tolerance levels, just as in the case of the number of middle level neurons, have to be experimented with to find the optimal error tolerance level. Due to the fact that each data set, experiment objective and way in which data is presented to the ANN is so unique, there are not yet standards for acceptable error tolerance levels and it has to be determined for each situation by using a trial and error effort (Yuan & Fine, 1998).

3.5.4.3 The Format of the Input Data

"It has been suggested that most of the "black magic" in neural networks comes in defining and preparing the training input set" (Nelson & Illingworth, 1991:154).

Neural networks only deal with numeric input data. All factors that serve as input data has to be numerically transcribed, for example, the gender variable can be predicted with a one or a zero. Sometimes the network requires that the input information be scaled or normalized. For example if DPOAE amplitude serves as input data and could be any number from 0 – 40 dB, it can be scaled by depicting it as a fraction of 40 dB, a DPOAE level of 30 dB would therefore have a value of 0.75. Only one extra input node is needed in this case. Another option for depicting input values is the dummy variable technique where categories are created to depict a certain value and values are depicted with ones and zeros depending on the category in which it falls. In the case of the DPOAE level between 0 – 40 dB, four 10 dB categories can be created, category one depicts DPOAE levels from 0 – 10dB, category two from 11 - 20 dB, category three from 21 - 30 dB and category four from 31 - 40 dB. A DPOAE level of 30 dB would therefore be depicted as 0010, indicating that the DPOAE level falls in the third category. If this method is used, more input nodes are needed depending on the number of categories created to depict the value, in this case four extra input nodes will be needed. With more input nodes, the neural network gets more complex and usually more middle level neurons are needed.

There are many ways in which input data can be presented to the network; the possibilities are as limited as the imagination of the person creating the neural network. Different input strategies often influence the prediction accuracy of the neural network and therefore there has to be experimented with different ways to present the information to the network. All the different ways in which input data was manipulated for this research project will be discussed in detail in Chapter 4, Research Methodology.

3.6 Interpretation and Conclusion

When it comes to the prediction of PTTs with DPOAEs and ANNs, there are many factors influencing the occurrence and levels of DPOAEs and therefore also the correlation that has to be determined between DPOAEs and PTTs and prediction accuracy of the ANN. From an in-depth literature study, the optimal set of stimulus parameters that influence DPOAE occurrence and levels were identified for the measurement of DPOAE. The identified stimulus parameters for DPOAE measurement are as follows:

3.6.1 Factors of the DPOAE Influencing DPOAE Occurrence and Levels:

- A primary f2/f1 frequency ratio of about 1.2 has been proven to elicit largest DPOAE amplitudes between 1 and 4 kHz (Gaskill & Brown, 1990; Avan & Bonfils, 1993; Stover, et al. 1996a).

- The loudness levels of the primaries should preferably be 10 – 15 dB apart (Gorga et al. 1993; Mills, 1997).
- The level of stimulation should not exceed 65 – 75dB to prevent the evaluation of passive properties of the cochlea and to gain more frequency specific information (Avan & Bonfils, 1993; Kummer et al. 1998; Mills, 1997).
- The way in which testing should be constructed is preferably a combination between I/O functions and DP Grams to gain as much information as possible of the DPOAE's threshold and amplitude (Kimberley & Nelson, 1989; Martin, et al. 1990a; Smurzynski et al. 1990).
- The subject variable age seems to have a positive influence for PTT prediction with DPOAEs and should be included in the correlation determination and prediction process (Lonsbury-Martin et al. 1991; Kimberley et al. 1994a; Kimberley et al. 1994b; De Waal, 1998).
- The frequency variable of the DPOAE to correlate with PTTs should preferably include not only the f2 frequency but the 2f1-f2 frequency as well (Mauermann et al. 1999a; Talmadge et al. 1998).

When it comes to the use of artificial neural networks as a data processing technique, several aspects regarding the choice, design and functioning of the network were identified. These aspects influence accuracy of predictions made by the network and are as follows:

3.6.2 Factors of the ANN that Influence Prediction Accuracy of PTTs with DPOAEs:

- From the description of the functioning of a neural network it became clear that a multi-layered ANN is needed for the prediction of PTTs with DPOAEs (Blum, 1992).
- The topology of the network influences prediction accuracy and experimentation is needed to determine the optimal number of neurons or nodes in each layer (Hornik et al. 1989; Nelson & Illingworth, 1991).
- Error tolerance during training and prediction is another factor that influences speed and efficiency of network operation and is also determined by trial-and-error experimentation (Rao & Rao, 1995; Yuan & Fine, 1998).
- There are many ways in which input data can be manipulated and the best way to present input information to the network requires careful consideration and experimentation (Nelson & Illingworth, 1991).

This chapter served as an identification and discussion of all DPOAE and ANN variables that influence PTT prediction accuracy. An optimal set of parameters were identified for the measurement of DPOAEs that will be applied in the testing procedure in the following chapter. However, the process to attempt to predict PTTs with DPOAEs and ANNs involve numerous possibilities in the experimentation to establish optimal neural network configuration and error tolerance levels. There are also different ways to present DPOAE measurements to the network that influence the accuracy of PTT predictions that lead to further necessary experimentation. These experiments to optimize PTT

prediction , as well as the research methodology for the entire research project will be discussed in the following chapter.

Chapter 4: Research Methodology

4.1 Introduction

One very interesting viewpoint on the essence of research methodology was given by Leedy (1993:9). "The process of research, then, is largely circular in configuration: It begins with a problem; it ends with that problem solved. Between crude prehistoric attempts to resolve problems and the refinements of modern research methodology the road has not always been smooth, nor has the researcher's zeal remained unimpeded."

The problem inspiring this research project has already been extensively stated in Chapter 1 and 2. In short, the need for an objective, non-invasive and rapid test of auditory functioning has led to numerous previous studies attempting to develop such a procedure despite the fact that there are many aspects contributing to pure tone thresholds that is not evaluated with otoacoustic emissions. Shortcomings in conventional statistical methods prevented accurate predictions of PTTs with DPOAEs due to the complex non-linear relationship between DPOAEs and PTTs and the noisy nature of DPOAE measurements. A new form of information processing called artificial neural networks (ANNs) was identified as a suitable data processing technique to attempt to solve this problem.

The study preceding this one (De Waal, 1998) attempted to predict pure tone thresholds with DPOAEs and artificial neural networks. First, PTTs were categorized as normal or impaired (normal defined as < 20 dB HL) with DPOAEs and ANNs and correct classification of normal hearing was 92 % at 500, 87% at 1000, 84% at 2000 and 91% at 4000 Hz. Predictions of impaired hearing was less satisfactory partly due to insufficient data for the ANN to train on and also possibly because of lack of experimentation with optimal topologies, error tolerance levels and optimal representation of input data for the neural network.

The aim of this chapter is to describe the research method that developed in the expansion and broadening of the basic work on DPOAEs and ANNs in order to enhance prediction accuracy of PTTs.

4.2 Aims of Research

The aims of this research project are as follows.

4.2.1 Main Aim

To improve prediction of pure tone thresholds (PTTs) at 500, 1000, 2000, and 4000 Hz with distortion product otoacoustic emission (DPOAE) responses in normal and hearing impaired ears with the use of artificial neural networks (ANNs).

4.2.2 Sub Aims

The first sub aim is to determine optimal neural network topology to ensure accurate predictions of hearing ability at 500, 1000, 2000 and 4000 Hz. The number of input nodes and number of output neurons are determined by the number of input- and output data. The number of middle layer neurons however, should be determined by trial and error until the required accuracy of prediction in the training stage is reached.

The second sub aim is to experiment with different ANN error tolerance levels to enhance neural network performance and efficiency during training and prediction.

The third sub aim is to determine if different manipulations of input data into the neural network improves prediction accuracy of PTTs with DPOAEs such as different ways to present the age variable, and DPOAE amplitude to the neural network.

The fourth sub aim is to experiment with the inclusion and omission of noisy low frequency DPOAE data to determine its effect on prediction accuracy.

The last sub aim is to investigate the effect of DPOAE threshold on prediction accuracy with DPOAEs when DPOAE threshold is defined as 1, 2 or 3 dB above the noise floor.

4.3 Research Design

For this research project, the chosen research design was a multivariable correlational study (Leedy, 1993). The correlation between DPOAE measurements and pure tone thresholds (PTTs) was studied by the use of artificial neural networks (ANNs). This correlation was then applied to make predictions of hearing ability in subjects of various ages, demonstrating different levels of sensorineural hearing loss or normal hearing to investigate to what extent DPOAEs can be used as a diagnostic or screening procedure in the objective evaluation of pure tone sensitivity. If DPOAEs can accurately predict pure tone thresholds objectively in a population with varying degrees of sensorineural hearing loss and at different ages, it would be a significant contribution to aid in the evaluation of difficult-to-test populations.

For the purpose of this study, 70 subjects (42 females, 28 males, 8-82 years old) were recruited from a school for the hard of hearing and a private audiology practice. Subjects were evaluated in terms of their pure tone thresholds (PTTs) and DPOAE measurements. The results from these two tests were used to train a neural network to find a correlation between the two data sets, and to use that correlation to make a prediction of PTTs given only the DPOAEs.

The measured variables for this study consisted of:

- PTT measurements at 500, 1000, 2000 and 4000 Hz
- DPOAE responses at eleven 2f1-f2 frequencies ranging from 2f1-f2 = 406 Hz to 2f1-f2 = 4031 Hz

(Both tests were measured in each subject within two hours.)

Controlled variables for this study included:

- The frequencies of the two primaries, f1 and f2, ranging from f1 = 500 Hz to f1 = 5031 Hz, with a primary frequency ratio of 1.2.
- The loudness levels of the primaries ranging from L1 = 70 dB to L1 = 35 dB with a loudness difference of L1 > L2 by 10dB.

Manipulated variables for this study to investigate the effect on PTT prediction accuracy included:

- Subject age presented to the ANN as a 5-year category or a 10-year category (see 4.8.1.3).
- DPOAE threshold defined as 1, 2 or 3 dB above the noise floor (see 4.8.1.5).

- Presentation of the amplitude of the DPOAE to the ANN input as one of four possible methods (AMP 100, AMP 40, ALT AMP or No AMP-see 4.8.1.4).
- The inclusion or omission of noisy low frequency DPOAE results for ANN training (see 4.8.1.1).
- Three different middle level neuron counts for ANN training and prediction (see 4.8.3.4).
- Three different error tolerance levels for ANN prediction and training (see 4.8.3.5).

Neural network results do not consist of predictions of frequencies in a decibel form, but of predictions of PTTs into one of eight possible 10 dB categories. Interpretation of data consists of the analysis of prediction accuracy of the neural network's ability to predict hearing at a specific frequency accurately into a specific 10 dB category.

4.4 Subjects

For this study, data obtained from 70 subjects (120 ears, in some cases only one ear fell within subject selection specification) were used to train a neural network to predict pure tone thresholds given only the distortion product responses. Subjects were recruited from a private audiology practice as well as a school for hard of hearing children. The subjects included 28 males and 42 females, ranging from 8 to 82 years old.

4.4.1 Criteria for the Selection of Subjects

The following criteria were chosen for the selection of subjects.

4.4.1.1 Normal and Impaired Hearing Ability

In order to train a neural network with sufficient data to make an accurate prediction of hearing ability, data across all groups of hearing impairment was needed. For this study, subjects were chosen that had varying hearing ability, ranging from normal to moderate-severely sensorineural hearing impaired. To obtain an equal amount of data in different areas of hearing impairment, data in three different categories of hearing impairment were included, namely normal hearing ability, mild hearing losses and moderately-severe hearing losses.

There are two general classification systems to classify hearing level as being normal or impaired (Yantis, 1994). The first method converts hearing levels into a rating scale based on percentage. A Pure tone threshold average (PTA) for the frequencies 500, 1000, 2000 and 3000 Hz is calculated, 25dB is subtracted (which is assumed to be the normal range) and the answer is multiplied by 1.5% to find percentage of impairment for each ear.

The second approach to describe normal ranges and hearing impairment also uses monaural PTA in the speech frequencies but ads additional descriptors to the different levels. Clark (1981) modified Goodman's (1965) recommendations into the following categories:

-10 to 15dB	Normal hearing
16 to 25dB	Slight hearing loss
26 to 40dB	Mild hearing loss

41 to 55dB	Moderately severe hearing loss
56 to 70dB	Severe hearing loss
91dB plus	Profound hearing loss

For subject selection, the second approach to classification of hearing impairment (as recommended by Clark, 1981) was used. Subjects with normal hearing, slight hearing loss, mild hearing loss and moderately-severe sensorineural hearing loss were included in the study. To divide the subjects into three groups of 40 ears each, the PTT thresholds of the group with normal hearing ranged from 0 dB to 15 dB. The group with slight and mild hearing loss had PTT thresholds that ranged from 16 to 35dB and the moderately-severe hearing-impaired group had PTA's in the range of 36 - 65dB. It should be noted that according to Clark's (1981) specification the moderate hearing loss group only includes hearing losses of up to 55 dB, whereas the severely hearing impaired group extends to 70 dB. DPOAEs have been reported in ears that have a hearing threshold as high as 65dB HL (Moulin, et al. 1994) at the frequencies close to the primaries. It was therefore decided to combine the category of moderate and severe hearing impairment to form the category moderately severe hearing impairment ranging from 36 to 65 dB HL.

The data was divided into three groups merely to ensure that an equal amount of data was obtained in each category. Another modification to Clark's classification system has been made. In addition to the frequencies used by Clark (1981) to determine the PTA, namely 0.5kHz, 1kHz, 2kHz and 3kHz, for this study 4kHz was also taken in consideration in the classification of hearing impairment. The reason for this modification is that DPOAE measurements are required at 4 kHz to predict the pure tone threshold at 4 kHz.

4.4.1.2 Normal Middle ear Functioning

The second selection criterion was normal middle ear functioning. Otoacoustic emissions can only be recorded in subjects with normal middle ear function. Only a very small amount of energy is released by the cochlea and is transmitted back through the oval window and ossicular chain to vibrate the tympanic membrane. Normal middle ear function is crucial to this transmission process (Norton, 1993; Osterhammel, Nielsen & Rasmussen, 1993; Zhang & Abbas, 1997; Koivunen, et al. 2000). The requirement for normal middle ear functioning is also the reason why only sensorineural hearing impaired subjects are included in the impaired hearing group described in 4.4.1.1.

Normal middle ear functioning was determined by otoscopic examination and tympanometry.

4.4.1.3 Good or Normal Attention Span

Only persons that were able to cooperate for approximately an hour were included in the study. Subjects had to be able to follow instructions and sit quietly and still in one position for about forty minutes for DPOAE testing. Subjects demonstrating inadequate ability to follow instructions or cooperate during pure tone audiometry, tympanometry or DPOAE testing were not included in the study. Some of the reasons subjects were excluded from the study in this regard include very young age, ill health and hyperactivity.

4.4.1.4 Criteria Regarding Subject Age and Gender

There is some debate regarding the effect of age on distortion product otoacoustic emissions. In a study by Lonsbury-Martinet al. (1991), a negative correlation between DPOAE measurements and age for subjects 20-60 years was reported. In their report however, it is suggested that this negative correlation is due to changes in hearing threshold associated with aging. A study by He & Schmiedt, (1996) also indicated that the difference in DPOAEs between younger and older subjects can be attributed to the sensitivity changes, rather than the aging itself. According to He and Schmiedt (1996) a 60 year old person with normal hearing (PTA < 15dB) will therefore have the same DPOAEs as a 12 year old with the same pure tone threshold levels.

There was therefore no selection criteria regarding age. The only population that was excluded in this study is the pediatric population, due to differences in middle ear properties such as canal length, canal volume and middle ear reverse transmission efficiency that may cause differences in DPOAE amplitudes (Lasky, 1998a; Lasky, 1998b; Lee, Kimberley & Brown, 1993).

There was also no selection criteria regarding gender. Gaskill and Brown (1990) and Cacace et al. (1996) reported that DPOAEs were significantly larger in female than male subjects tested in the frequency range of 1000- 5000Hz. Both studies however, indicated that the female subjects in their studies had more sensitive auditory thresholds than the males (an average of 2.4 dB better). The differences found between the two groups could therefore not be explained by gender only.

Lonsbury-Martin et al. (1990) conducted a study to investigate basic properties of the distortion product including the effect of gender on the prevalence of DPOAEs. A comparison of DPOAE amplitudes and thresholds failed to reveal any significant differences except a minor difference at 4 kHz.

Gender effects on DPOAEs are apparently limited to minor insignificant differences in DPOAE amplitudes and thresholds and therefore gender was not one of the selection criteria for this study.

Even though subjects were not selected regarding age or gender, a subject's age and gender were used as input information for the neural network. The reason for this is that previous studies that attempted to predict PTTs with DPOAEs found that age enhances prediction accuracy and recommended to use age as a variable in PTT prediction studies (Lonsbury-Martin et al. 1991; Kimberley et al. 1994a; Kimberley et al. 1994b). The previous study by De Waal (1998) also indicated that the combination of age and gender as prediction variables had a greater positive effect on prediction accuracy than the inclusion of age alone.

4.4.2 Subject Selection Procedures

The procedure in which subjects were selected started with a brief interview, following otoscopic examination of the external meatus, tympanometry and pure tone audiometry.

4.4.2.1 Case History and Personal Information

A short interview was performed to obtain a limited case history and some personal information. The research project was also discussed with the subject in a very brief manner and any questions answered. The purpose of the case history was firstly to obtain enough personal information to open a new subject file and obtain the subject's age and gender for later studies of these effects on DPOAEs. Secondly, information regarding hearing status such as any complaints of tinnitus and vertigo, the amount of noise exposure and complaints of middle ear problems was obtained.

In the analysis of data, some subjects may exhibit abnormal DPOAEs in conjunction with normal pure tone thresholds. In a study by Attias, et al. (1995), it was found that in some cases, subjects with normal pure tone thresholds of 0 dB exhibited abnormal otoacoustic emissions, due to noise exposure. The effects of noise exposure can clearly be seen long before the actual hearing loss occurs. This is also true for ototoxic medication (Danhauer, 1997). Cases with exposure to noise, ototoxic medication or subjects with tinnitus and vertigo were included in the research project, this information merely serves as background to formulate reasons for possible abnormalities in DPOAE responses.

Appendix A reviews the aspects that were addressed in the short interview. This interview lasted approximately 10 minutes.

4.4.2.2 Otoscopic Examination

Otoscopic examination of both ears was performed to determine the amount of wax in the ear canal, for excessive wax might block the otoacoustic emission microphone and prevent the reading of a response. The second aspect that was investigated was the light reflection on the tympanic membrane, indicative of a healthy tympanic membrane (Hall III & Chandler, 1994). Otoscopic examination's duration was about 3-5 minutes.

4.4.2.3 Tympanometry

A subject's tympanometry results must have been within the following specifications to be included in the study.

A normal type A tympanogram was one of the criteria for normal middle ear functioning. A type A tympanogram has a peak (or point of maximum admittance) of 0 to -100 daPa. The peak may even be slightly positive, for example +25daPa (Block & Wiley, 1994). A type A tympanogram's static immittance when measured at 226 Hz ranges from about 0.3 to 1.6 cc (Block & Wiley, 1994). Subjects demonstrating type A tympanograms within these specifications were accepted for the study.

Tympanometry was performed in both ears and the duration of the procedure was about 5 minutes.

4.4.2.4 Pure Tone Audiogram

Data obtained from the pure tone audiogram was not only used in the selection of subjects, but also forms part of the measured variables for this study and was used to train the artificial neural network. The determination of the pure tone audiogram will therefore be discussed in detail.

If the subject had normal middle ear functioning, the subject selection procedure continued. A pure tone audiogram was then obtained from the subject. The frequencies that were tested during pure tone air conduction were 125, 500, 1000, 2000, 4000,and 8000 Hz. Even though only 500, 1000, 2000 and 4000 Hz were used to train the neural network, pure tone results at 125 Hz and 8000 Hz could sometimes indicate a slight hearing loss even though hearing at the four middle frequencies was normal. Hearing thresholds at 125Hz and 8000 Hz were never used in the determination of the category in which a subject fell for subject selection (see 4.4.1.1) but were used merely as background information to formulate reasons for possible abnormal DPOAEs.

If a hearing loss was present, or if any of the frequencies except 8000 Hz had a threshold >15 dB, then pure tone bone conduction was also performed to ensure that the hearing loss was of a sensorineural nature. Only subjects with sensorineural hearing losses (no gap between air conduction and bone conduction) were accepted for the study. Threshold determination was in 5dB steps and a threshold was defined as 50% accurate responses at a specific dB level (Yantis, 1994).

Audiograms from subjects were then analyzed. All audiograms indicating normal hearing (500, 1000, 2000, 3000 and 4000 Hz below 15 dB) were included in the first group. Audiograms indicating hearing loss were analyzed in terms of the degree of the hearing loss. Mild hearing losses, indicating a hearing loss between 16 - 35 dB in the frequency region 500 - 4000 Hz were categorized in the second group, namely mild hearing losses. Audiograms indicating hearing losses of 36 - 65 dB in the frequency region of 500 - 4000 Hz were categorized in the third group, namely moderately severe hearing losses. 40 audiograms were included in each category.

The duration of pure tone audiometry was approximately 15 - 20 minutes.

If a subject demonstrated normal middle ear functioning and a pure tone audiogram that could be categorized into one of the three groups, DPOAE measurements were performed within the next hour. This procedure will be discussed in 4.7 "Data collection procedures".

4.5 Description of subjects

Figure 4.1 depicts the gender distribution for subjects included in this study. Figure 4.2 depicts the age distribution of subjects into 10-year categories.

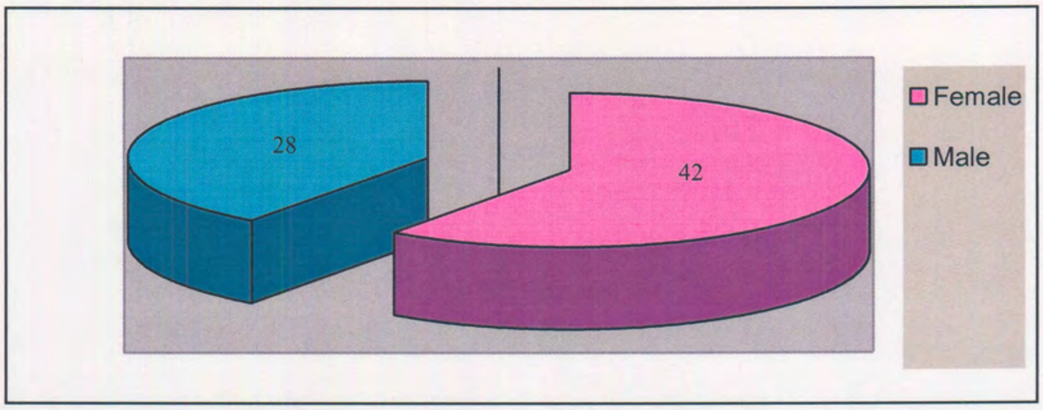


Figure 4.1: Representation of male and female subjects.

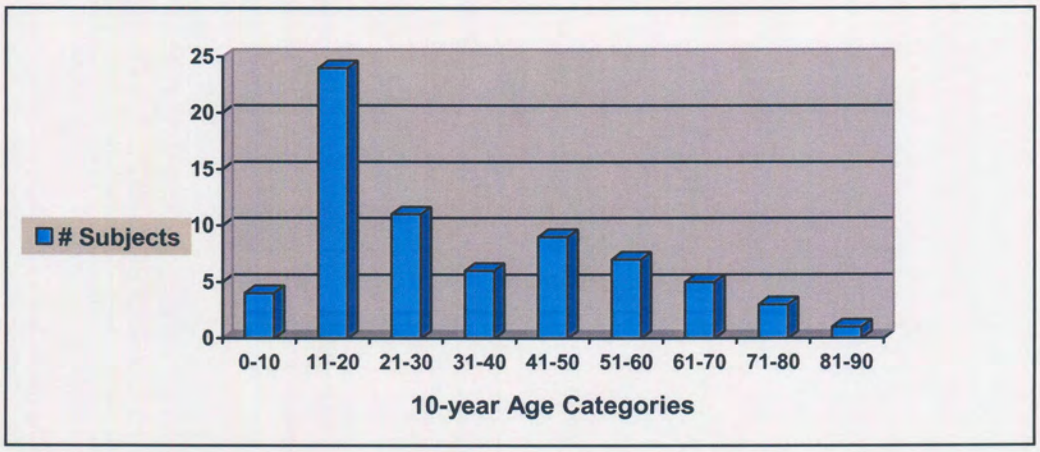


Figure 4.2: Age representation of subjects in 10-year categories.

Table 4.1 indicates the distribution pattern for different types of hearing loss that the 120 ears in the data set exhibited.

Table 4.1: Distribution pattern for different types of hearing loss in the 120 ear data set.

	# Ears Group 1: PTAs 0-15 dB HL	**# Ears Group 2: PTAs 16-35 dB HL**	**# Ears Group 3: PTAs 36- 65dB HL**
Flat audiogram: Not more than 20dB variation between 0.5 – 4 kHz.	40	11	16
Gradual slope: PTTs increases gradually as frequency increases	0	9	24
Ski-slope: Flat configuration up to 2 kHz with >20dB PTT drop in high frequencies	0	10	1
Low frequency loss: 0.5 - 1 kHz more impaired than 2 – 4 kHz	2	0	0
Notch: Notch shaped loss around 1 -3 kHz	4	3	0

4.6 Apparatus

The apparatus for the different sections of research were as follows:

4.6.1 Subject Selection Apparatus

- For otoscopic examination of the external meatus and tympanic membrane an otoscope was used, specifically the **Welch Allyn pocketscope model 211**.

- For tympanometric measurements **the GSI 28 A** middle ear analyzer, calibrated April 1997 was used (Testing was performed in January 1998).

- For determination of auditory pure tone thresholds, the **GSI 60 Audiometer**, calibrated April 1997 was used. The model of the earphones on the audiometer was 296 D 200-2. Pure tone thresholds were measured in a sound proof booth.

4.6.2 Data Collection Apparatus

- The measurement of Distortion Product Otoacoustic Emissions were conducted with a **Welch Allyn GSI 60 DPOAE system** and the probe was calibrated for a quiet room in January, 1998. All measurements were made in a quiet room.

- For determination of auditory pure tone thresholds, the **GSI 60 Audiometer**, calibrated April 1997 was used. The model of the earphones on the audiometer was 296 D 200-2. Pure tone thresholds were measured in a sound proof booth.

4.6.3 Data Preparation Apparatus

- For the preparation of data files, a 600 MHz Pentium computer was used. The software included Excel for Windows 2000.

- For the training of the neural network, the backpropagation neural network from the software by Rao and Rao, 1995 (software supplied in addition to their book) was used. The neural network was trained on three 600 MHz Pentiums.

4.6.4 Data Analysis Apparatus

- Analysis of data was performed in Excel for Windows 2000.

4.6.5 Preliminary Study

The reason for the preliminary study was twofold: Firstly, to confirm subject selection criteria and secondly, which stimulus parameters to use in the measurements of DPOAEs.

4.6.5.1 Confirmation of Subject Selection Criteria

A very large part of the determination of subject selection criteria was based on an extensive overview of related literature. The researcher did however conduct a series of DPOAE measurements on subjects with various categories of hearing ability to confirm current subject selection criteria. Just a few of the interesting cases in the preliminary study will be discussed briefly in Table 4.2.

Table 4.2: Confirmation of subject selection criteria in the preliminary study.

Purpose	Result	Outcome
Confirmation of normal middle ear functioning (Type A tympanogram, and compliance of >0.3cc) as subject selection criteria in cases where PTT results fell within selection criteria but with small variations in tympanometric results.	Case one had perfect PTTs (0dB) but no airtight seal could be obtained as a result of grommets in the tympanic membrane. This subject displayed very high levels of low frequency background noise during DPOAE testing and it was difficult to distinguish the DPOAE responses from the noise floor at most of the low and mid frequencies.	Only cases with air tight seals of the probe in the external meatus to allow measurement of a tympanogram were included in the study.
	Case two had a mild sensorineural hearing loss but displayed compliance measurements of less than 0.3cc during tympanometry. DPOAE responses were virtually indistinguishable from the noise floor due to high levels of low and mid frequency noise.	Only cases demonstrating at least 0.3cc in tympanometry were allowed in the study.

Table 4.2 Continues

Purpose	Result	Outcome
Confirmation of levels for primary tone pairs not to exceed 70 dB SPL.	Some tests revealed that when very high intensity primaries were used (such as 70- 80dB SPL), in some instances one could observe "passive" emissions from the ears of severely hearing impaired subjects. The reason for passive emissions, according to Mills, (1997) is that very high level stimuli can stimulate broad areas of the basilar membrane and phase relations between traveling waves can cause these "passive" emissions that do not correspond well to hearing sensitivity and has poor frequency specificity. In this preliminary study, passive emissions were only observed when stimuli levels were higher than 70dB.	It was therefore decided not to use stimuli levels higher than 70dB.
Confirmation of subject selection criterion that hearing loss should not exceed 65dB HL.	Another aspect that became apparent after a few tests were conducted was the absence of DPOAEs in persons with hearing losses greater than 65dHL. This confirmed studies by Moulin et al. (1994) and Spektor et al. (1991), which found that when stimuli lower than 65dB SPL are used, DPOAEs cannot be measured in ears with a hearing loss exceeding 65dB HL.	Therefore, for this study, only subjects were included with sensorineural hearing losses of up to 65 dB HL.

A second series of tests were conducted to determine optimal stimulus parameters.

4.6.5.2 Determination of Optimal Stimulus Parameters

Most of the stimulus parameters for this study were derived from an in depth literature study. Parameters such as the frequency ratios between the primaries (chosen at 1.2), the loudness levels of L1 and L2 (L2 > L1 by 10dB) and whether to measure DP Grams or I/O functions (eight DP Grams at eight loudness levels were used to gain the same information that both procedures would yield) were selected on recommendation of other previous studies. The choice for these stimulus parameters is described in 4.7.1.2.1 "Specification of stimulus parameters for DPOAE measurement."

There are however a few stimulus parameters that require some experimenting in order to determine applicability and practicality for a certain research project. One such example is the configuration setup, or specifically, the number of frames of data that will be collected in each measurement. The GSI-60 DPOAE system offers two possibilities, a screening option and a diagnostic option. These options will be reviewed in more detail than the section of the preliminary study concerned with confirmation of subject variables in a table format because a thorough understanding of test acceptance conditions is required to clarify later definitions of DPOAE threshold as 1, 2 or 3 dB above the noise floor.

The screening option collects a maximum of 400 frames before stopping each primary tone presentation. Not every test runs up to 400 frames, if a very clear response is measured, the measurement can be made in as little as 10 frames. Test acceptance conditions for the screening configuration are a cumulative noise level of at least –6dB SPL and either a DPOAE response amplitude that is 10 dB above the noise floor or a cumulative noise level of at least –18 dB SPL (GSI-60 manual, p2-44). A maximum of 400 frames is measured, and if no clear response was present, the results are labeled “timed out.”

The diagnostic option runs up to 2000 frames for each primary tone presentation. The minimum number of accepted frames is 128. Test acceptance conditions are that the distortion product minus the average noise floor should be at least 17 dB.

After a few measurements in both configurations it became clear that the diagnostic option requires much more testing time. Testing time of one single DP Gram measured at low level stimuli in the diagnostic configuration could increase testing time up to 12 minutes. Even though the general noise floor was slightly lower during the diagnostic option, it was not practical to conduct 8 DP Grams in each ear with tests lasting 6-12 minutes each. It would take between 60 minutes to 105 minutes to measure one ear alone with DPOAEs. It was therefore not practical to evaluate 120 ears with the diagnostic option. The screening option with a testing time of up to 2 minutes per DP Gram was selected for this study. One ear could be evaluated in about 15 minutes with DPOAEs and the screening procedure yielded very much the same information.

Lastly, the stimulus parameter that required some experimenting was the selection of the frequencies of the primary tone pairs. The GSI-60 DPOAE system has a "Custom DP" function where the examiner can choose any primary frequencies for DPOAE measurement. After a few tests it became clear that care should be taken when selecting primary tones. Not only should the frequency ratio of the primaries preferably be 1.2, but also should frequency values from one tone pair to the next be at least one octave apart to avoid interaction between stimuli (GSI-60 manual, p2-39). The GSI-60 measures the noise floor from the first primary tone pair per group, and if frequency pairs are selected too close to each other, very high levels of noise are being measured. So after a lot of changes in primary tone pairs were made to avoid interaction between stimuli, the researcher ended up with stimuli very similar to the default stimuli of the GSI-60. It was therefore decided to use the default primary frequencies of the GSI-60 for this study by activating all four octaves. (It seems that those stimuli are set as default for a very obvious reason.)

Just for practicality, a few test runs that incorporated the whole data collection procedure were conducted to determine the amount of time required testing each subject. This was determined in order to schedule appointments. As seen in Table 4.3, the whole data collection procedure lasted about an hour. In some cases, especially in the case of subjects with a hearing loss, more time was required for bone conduction but on the average, one hour was sufficient to test one subject.

Table 4.3: Time required testing one subject

Subject history	5 minutes
Audiometry	15 minutes
Otoscopic examination	5 minutes
Tympanometry	5 minutes
DPOAE measurements left ear	15 minutes
DPOAE measurements right ear	15 minutes
Total testing time	**60 minutes**

4.7 Data Collection Procedures

The following procedure was conducted during research.

4.7.1 Data Collection Procedures Conducted during Research

Two sets of data are needed to train a neural network to predict PTTs with DPOAEs: each subject's pure tone thresholds and each subject's DPOAEs.

4.7.1.1 Data Obtained from Pure Tone Audiometry

The necessary pure tone audiometry data has already been obtained during subject selection and the collection procedure for this set of data has been described in the section 4.4.2.4. Pure Tone Audiogram.

4.7.1.2 Data Obtained from DPOAE Measurements

The second set of data that was collected was each subject's DPOAE responses. There are many stimulus parameters that should be specified to be able to repeat this research project and need to be fully described.

4.7.1.2.1 Specification of Stimulus Parameters for DPOAE Measurements

There is a four dimensional space in which the stimulus parameters for DPOAE measurement should be specified (Mills, 1997). The frequencies of the two primary stimulus tones f1 and f2 (f1>f2), the frequency ratio of f2/f1 (how many octaves apart the two frequencies are), the loudness level of f1 (which is L1) and the loudness level of f2 (which is L2). Furthermore, the difference in loudness level between L1 and L2 should also be specified.

4.7.1.2.1.1 The Selection of the Frequencies

In the case of the GSI-60 Distortion Product otoacoustic emissions system, the number of octaves that should be tested can be specified as well as the amount of data points to plot between octaves. The octaves available are 0.5 - 1 kHz; 1-2 kHz; 2-4 kHz and 4 -8 kHz. All of these octaves were selected for DPOAE testing because information regarding all these frequencies was required to make comparisons with the audiogram in the frequency range 500 - 4000 Hz. The amount of data points between frequencies could be any number between 1 and 20. The more data points per octave, the longer the required test time since more frequency pairs are tested between frequencies. The GSI-60 manual suggests 3 data points per octave to be adequate, not increasing the test time too much but yielding enough information regarding DPOAE prevalence between frequencies. In the case of the pure tone audiogram, in-between frequencies were only tested when hearing losses between frequencies varied more than 15 dB (to measure the slope of the hearing loss) and only 1 or in extreme cases 2 in-between frequencies were evaluated. The selection of 3 data points between octaves in the case of DPOAE measurement should therefore be adequate.

The frequencies tested by the GSI-60 when all four octaves are activated and 3 data points per octave is specified amount to 11 frequency pairs. The eleven frequency pairs are presented in Table 4.4.

f1 and f2 are always presented simultaneously as a pair, f1>f2 and the f2/f1 ratio is ~1.2.

Table 4.4: The eleven frequency pairs tested by the GSI-60 DPOAE system when all four octaves are activated.

PAIR	1	2	3	4	5	6	7	8	9	10	11
f1 Hz	500	625	781	1000	1250	1593	2000	2531	3187	4000	5031
f2 Hz	593	750	937	1187	1500	1906	2406	3031	3812	4812	6031

4.7.1.2.1.2 The Selection of the Frequency Ratio of the Primary Frequencies (f2/f1)

Several studies investigated the effect of the frequency ratio on the occurrence of DPOAEs (Cacace et al. 1996; Popelka, Karzon & Arjmand, 1995; Avan & Bonfils, 1993; He & Schmiedt, 1997).

It appears that the frequency ratio of 1.2 - 1.22 is most applicable to a wide range of clinical test frequencies (0.5-8kHz) and a wide range of stimulus loudness levels. A stimulus ratio of f2/f1 = 1.2 was therefore selected for this study.

4.7.1.2.1.3 The Selection of the Loudness Levels of the Primaries, L1 and L2.

There are two ways of eliciting a DPOAE response. Either the frequencies are changed and the loudness level kept constant, this is sometimes referred to as a "distortion product audiogram" (DP Gram), or the frequencies are being kept constant while the loudness level is changed (an input/output function (I/O) is obtained). In this case, several DP audiograms were obtained. All the frequencies selected for all four octaves were presented to the subjects at different loudness levels, starting with maximum loudness levels at L1 = 70 dB; L2 = 60 dB. Loudness levels were decreased in 5 dB steps until DP "thresholds" (lowest intensities where DP responses can be distinguished from the noise floor) for all the frequencies were obtained. The lowest loudness level for the primaries that was tested was L1 = 35 dB; L2 = 25dB. Eight loudness levels were therefore evaluated resulting in eight DP "audiograms" for each ear.

4.7.1.2.1.4 The Relative Loudness Levels of the Primaries (L1 and L2)

An overview of several studies indicated the following loudness level ratios to be most suitable for the detection of DPOAEs: L1 > L2 by 10 dB (Stover et al. 1996a), L1 > L2 by 15 dB (Gorga et al. 1993) and L1 > L2 by 10 - 15 dB (Norton & Stover, 1994). A study by Mills (1997) indicated that more DPOAEs were recorded when L1>L2 than L1 = L2.

A loudness level ratio of L1 > L2 by 10 dB was chosen for this study.

4.7.1.2.1.5 The Criteria for DPOAE Threshold

The detection threshold for a distortion product otoacoustic emission depends almost entirely on the noise floor and the sensitivity of the measuring equipment (Martin et al. 1990b). A distortion product with amplitude less than the noise floor cannot be detected (Kimberley & Nelson, 1990; Lonsbury-Martin et al. 1990). Most researchers specify a DP response to be present if the DP response is 3-5 dB above the noise floor. Harris and Probst (1991:402) specified a DP response as "the first response curve where the amplitude of 2f1-f2 is ≥ 5 dB above the level of the noise floor." Lonsbury-Martin et al. (1990) reported detection thresholds for DPOAE measurements 3 dB above the noise floor. Lonsbury-Martin (1994) set the criterion level for a DPOAE threshold at ≥ 3 dB.

For this study, there will be experimented with detection thresholds for DPOAEs as 1 dB, 2 dB or 3dB above the noise floor to investigate if more accurate PTT predictions can be made with lower detection thresholds.

4.7.1.2.2 DPOAE Testing Procedure

DPOAE measurements were performed directly after the subject selection procedure. Subjects were instructed to sit next to the GSI 60 DPOAE system, not to talk and to remain as still as possible. Subjects were allowed to read as long as they kept their heads as still as possible. First, a new file was opened for the subject. Then the DPOAE probe tip was inserted into the external meatus in such a manner that an airtight seal was obtained.

Eight tests or DP Grams were performed in each ear. Every DP Gram consisted of eleven frequency pairs. Every frequency pair consisted of two pure tones, f1 and f2 presented to the ear simultaneously. (See Table 4.4 for the eleven frequency pairs). The eleven frequency pairs were presented to the ear in a sweep, one at a time starting with the low frequencies, ending with the high frequencies.

The first DP Gram was conducted on the loudness levels F1 = 70dB SPL, F2 = 60dB SPL. The second DP Gram was conducted 5 dB lower at F1 = 65 dB SPL, F2 = 55 dB SPL. A total of eight DP Grams were conducted, each one 5 dB lower than the previous one. The lowest intensity DP Gram that was performed was F1 = 35 dB SPL, F2 = 25 dB SPL.

The procedure was repeated for both ears if both ears fell within selection criteria. The duration of DPOAE testing of eight DP Grams for one ear was between 15-20 minutes. If a subject was tested binaurally, the duration of DPOAE testing was approximately 30-40 minutes.

4.8 Data Preparation Procedures

In the data preparation process, there were three interrelated processes that happened in parallel and influenced each other in such a way that it is challenging to describe the process with a logical serial or start-to-finish approach. One of the processes was to determine how input information was to be presented to the neural network and which variables or combinations of variables to experiment with. Another process was to determine optimal neural network error tolerance levels and topology, specifically the number of hidden layer neurons. The last process was the creation of data files that serve as input into the neural network that represent all the chosen variables and combinations thereof.

These three processes were highly interrelated: The combination of variables to use and how to present them determined how the data file looked that served as input information into the ANN. The input, or specifically the number of nodes in the input layer necessary to represent all variables, determined the complexity of the network and therefore the number of hidden layer neurons needed as well as suitable error tolerance levels. Failures in network operation and prediction in its turn influenced how new experiments were constructed to present input data in new ways, to include new variables or new combinations thereof, or to experiment with different numbers of middle level neurons and error tolerance levels, all in an attempt to make more accurate predictions.

The three interrelated processes namely the choice of how input data is presented to the network, the creation of the data file and the determination of network topology and error tolerance and will be discussed one at a time.

4.8.1 Experiments to Determine ANN Prediction Accuracy by Manipulating the Input and Output Data

The data that served as possible input information into the ANN was the presence or absence of DPOAE responses, defined by 1dB, 2dB and 3dB thresholds, the DPOAE amplitude of all present responses, subject gender and subject age.

4.8.1.1 DPOAE Data at High and Low Frequencies

For some experiments, DPOAE occurrence at all eleven frequencies and all eight loudness levels (or DP Grams) were used. For some experiments only DPOAE occurrence at the eight high frequencies for all eight DP Grams were used (f1 = 500, 625 and 781 Hz were omitted). DPOAEs measured at the low frequencies are often noisy or absent and these experiments attempted more accurate predictions by omitting the noisy data to prevent pollution of data.

When DPOAE occurrence at all eleven frequencies were used for all eight DP Grams, at least 88 input nodes were needed in the input layer to present this map of all present and absent responses to the ANN. When only the eight high frequencies (starting at f1 = 1000 Hz to f1 = 5031 Hz) for all eight DP Grams were used, only 64 input nodes were needed to present DPOAE responses to the ANN.

4.8.1.2 Subject Gender

Subject gender was always included and always depicted with a one or a zero. Subject gender therefore always added just one input neuron to the input layer.

4.8.1.3 Subject Age

Subject age was always included in the training and prediction of the ANN but different ways were used to present it to the neural network. Subject age in this study varied from 8 – 82 years old. The dummy variable technique was used to depict a subject's age into either a 10-year category, or a 5-year category.

In the 10-year category method, there were ten possible 10-year categories and the subject's age was depicted with zeros and a single one corresponding to the appropriate category: A 12 year old subject was therefore depicted as 01000 00000. When this method was used to depict subject age, ten extra input nodes were needed for the input layer.

In the 5-year category method, there were 20 possible 5-year categories. A 12 year old subject would therefore be depicted as 01000 00000 00000 00000. This method required 20 extra input nodes for the input layer. This method specified subject age more accurately but also made the neural network more complex due to a larger number of input nodes.

4.8.1.4 DPOAE Amplitude

There were four different amplitude experiment types.

4.8.1.4.1 AMP 100

The first amplitude representation of the DPOAE response was depicted as a fraction of 100 (This experiment is referred to as **AMP 100**). Instead of depicting the presence or absence of a response with a one or a zero, the magnitude of the response was used. A present DPOAE response of 30 dB's input into the neural network would therefore be 0.3. The same 88 input nodes were used that depicted presence or absence of a response, only now with a value indicating the amplitude of the DPOAE. This method of amplitude representation caused the neural network to spend much more time to converge (to reach the required error tolerance level for every ear in the training set). It took about 2 hours per experiment for the network to converge, which is incredibly long if 120 ears have to be predicted at 4 frequencies. 960 hours (40 days) were needed just to reach the optimal error tolerance level before prediction can begin. Some of the experiments were run with this method of amplitude representation but other more effective ways were needed to present amplitude to the ANN.

4.8.1.4.2 AMP 40

The second amplitude representation of the DPOAE response was depicted as a fraction of the largest DPOAE amplitude measured in this population of subjects in other words a percentage (This experiment is referred to as **AMP 40**). (The largest DPOAE response ever measured in this population of subjects was 39dB.) This experiment also used the original 88 input nodes that depict DPOAE occurrence but instead of just a zero indicating absence or a one indicating a present response, the magnitude of the response as a fraction of 40 was used. A 30 dB DPOAE was therefore depicted as 0.75. For AMP 40 convergence was much faster, only about 40 minutes per experiment.

4.8.1.4.3 ALT AMP

The third amplitude representation of the DPOAE response was depicted with the dummy variable technique by indicating into which one of four 10 dB categories the amplitude fell (This experiment is referred to as **ALT AMP**). A 30 dB DPOAE was depicted as 0010. For this experiment, every one of the 88 input nodes had to receive four categories to indicate the category in which the amplitude fell. This increased the number of nodes in the input layer needed to represent this information with four times. An experiment involving all 11 frequencies for all eight DP Grams therefore needed 352 input nodes, instead of the usual 88. This drastic increase in input neurons contributed to a much more complex neural network that required more middle level neurons. For this experiment, the middle layer neurons were always doubled to compensate for the large quantity of input data.

4.8.1.4.4 No AMP

The last amplitude experiment was when the amplitude of the DPOAE was omitted (This experiment is referred to as **No AMP**). The usual 88 input nodes were used and a DPOAE response was indicated as present with a one and absent with a zero. The presence of a DPOAE response is defined as a certain dB level above the noise floor. This brings us to the next experiment type, regarding the threshold of a DPOAE.

4.8.1.5 DPOAE Threshold

Harris and Probst (1991) and Krishnamurti (2000) defined DPOAE threshold as DPOAE response ≥ 5 dB above the noise floor, According to Lonsbury-Martin & Martin (1990) the DPOAE should be 3 dB above the noise floor to be regarded as present.

For this research project, it was decided to use different thresholds for DPOAE responses namely 1dB, 2dB and 3dB above the noise floor. This threshold reduction had more present DPOAE responses as a result. If the 1dB and 2dB thresholds yield more valid DPOAE responses, the network will be able to make more accurate predictions. If the extra responses gained are not valid but just part of the noise floor, prediction accuracy will not be increased but may be decreased.

4.8.1.6 Number of Ears or Data in Every Output Category to be Predicted

From the previous study (De Waal, 1998) it became apparent that the number of ears in every category to be depicted had a great influence on prediction accuracy of the neural network. The reason for this is that the network needs adequate representation in every category to learn the correlation between DPOAEs and PTTs to make an accurate prediction. In some instances in the previous study, certain categories had very little hearing-impaired data such as in the case of 500 Hz for example. Many of the subjects with hearing losses had normal hearing at 500 Hz (such as subjects demonstrating ski slopes). Category 7 in the case of the 500 Hz prediction had only data for one ear. Category 6 had only data for six ears and category 5 only data for five ears. It could be possible that the neural network did not have sufficient data in every category to train on and this aspect influenced the accuracy of the prediction. To test the significance of the number of ears in every category, it was decided in the previous study to enlarge the categories depicting hearing impairment to 15 dB, in order to attempt to include more hearing-impaired data in every category. It was referred to as **scenario five**, and hearing ability was divided in five categories. Categories that depicted normal hearing spanned 10 dB whereas categories that depicted hearing impairment spanned 15 dB. The five categories are presented in Table 4.5.

Table 4.5 The five categories of hearing ability for scenario five.

Category 1	0 – 10 dB HL
Category 2	11 – 20 dB HL
Category 3	21 – 35 dB HL
Category 4	36 – 50 dB HL
Category 5	51 – 65 dB HL

The significance of the number of ears in every category was also tested for this research project. The best experiments for each frequency were selected after the completion of ANN training and prediction and were run in this scenario five method, by enlarging the categories depicting hearing loss to 15 dB for the output of the neural network. The input data, number of middle level neurons, error tolerance, dB threshold above the noise floor and presentation of the age and amplitude variables to the network were kept exactly the same, only in this scenario, the output of the network was changed to predicted hearing loss into three possible 15 dB categories in stead of the usual seven. It will also be referred to as scenario five method in the present study.

Lastly, one aspect that was experimented with was the amount of data or number of pure tone thresholds in the input data of every category.

4.8.1.7 Number of PTTs in every Input Category for ANN Training

Pure tone thresholds are routinely evaluated in 5 dB increments (Hall III & Mueller III, 1997), as was also the case in this study. The possibilities for pure tone threshold values are therefore always rounded up to an increment of 5 dB. In the previous study (De Waal, 1998), all the first categories of all experiments spanned 0 – 10 dB. This implied that pure tone threshold values of 0 dB, 5 dB and 10 dB were included in this category, a total of three possible measurements from the audiogram. All second categories in the previous study always spanned 11-20dB, but since thresholds are only evaluated in 5 dB increments, the possible values to be included in the second category only consisted of measurements obtained at 15dB and 20dB, therefore only two possible measurements from the audiogram. This lead to an uneven distribution of the number of measurements in every category that possibly lead to poorer predictions of categories with less input information for the ANN to train on.

For this study, it was decided to have an equal number of possible thresholds in every category to ensure optimal distribution of input data across all categories for the network to train on. Two possible threshold values from the audiogram were allowed into every category. Category one therefore consisted of data from ears that exhibited threshold values at 0 dB and 5 dB, category two consisted of PTT values of 10 dB and 15 dB and so forth. The PTT data distribution for each category can be seen in Table 4.6.

Table 4.6: PTT data distribution for every input category

	PTT data permitted into each category (dB HL)	
Category 1	0 dB	5 dB
Category 2	10 dB	15 dB
Category 3	20 dB	25 dB
Category 4	30 dB	35 dB
Category 5	40 dB	45 dB
Category 6	50 dB	55 dB
Category 7	60 dB	65 dB
Category 8	70 dB	75 dB

For the previous study, the first two categories were evaluated to investigate prediction accuracy when it comes to the separation of normal hearing and hearing impaired ears. Normal hearing was defined as 0 – 20 dB, according to the definition of normal hearing by Jerger, (1980). For the present study, the first three categories were investigated to determine how accurately the network could separate normal from hearing impaired ears (normal = 0 – 25 dB HL) according to the definition of Goodman (1965), which is also the recommendation of the American Academy of

Otolaryngology and the American Council of Otolaryngology (AAO-ACO) in 1979 for normal hearing.

A few experiments were run with the same PTT distribution as the previous study (De Waal, 1998) of three values in the first category and two in every category thereafter for two reasons: The first reason was to be able to make valid comparisons between the previous and present study. To make accurate comparisons between category one of the previous study and category one of the present study, the PTT distribution for ANN training have to be the same. The second reason was to accommodate Jerger (1980)'s definition of normal hearing, which is 0 – 20 dB HL and spanning the first two categories of this procedure.

Another process in the data preparation involved transcribing raw data into data files suitable for ANN input.

4.8.2 Creation of a Data File for each Ear at each Frequency

The way in which the raw data was transcribed into files was constructed in such a way that each ear had its own file for every frequency. Each ear therefore had four files depicting information at 500, 1000, 2000 and 4000 Hz. A file is merely a row of numbers, depicting the test results in a certain order. Table 4.7 represents a raw data set for one DP Gram. 8 DP Grams for each ear were conducted. The complete raw data set for one ear would therefore have 88 rows of data under each column number. The column numbers in the top row is explained to indicate which measurement that column represents in the section following Table 4.7.

Table 4.7: Example of a raw data set for one DP Gram

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	1	R	500	70	593	60	406	8	0	N	0	0	5	0	0	0	24	F
1	1	R	625	70	750	60	500	9	-1	T/O	0	0	5	0	0	0	24	F
1	1	R	781	70	937	60	625	14	-6	A	0	0	5	0	0	0	24	F
1	1	R	1000	70	1187	60	812	3	-2	N	0	0	5	0	0	0	24	F
1	1	R	1250	70	1500	60	1000	12	-6	A	0	0	5	0	0	0	24	F
1	1	R	1593	70	1906	60	1281	-1	-9	A	0	0	5	0	0	0	24	F
1	1	R	2000	70	2406	60	1593	13	-7	A	0	0	5	0	0	0	24	F
1	1	R	2531	70	3031	60	2031	5	-8	A	0	0	5	0	0	0	24	F
1	1	R	3187	70	3812	60	2562	7	-9	A	0	0	5	0	0	0	24	F
1	1	R	4000	70	4812	60	3187	8	-6	A	0	0	5	0	0	0	24	F
1	1	R	5031	70	6031	60	4031	5	-6	A	0	0	5	0	0	0	24	F

Explanation of column numbers for Table 4.7:

1	Subject number.
2	Number of DP Gram.
3	Ear that is being tested (right or left).
4	Frequency of f1 in Hz.
5	Loudness level of L1 in dB SPL.
6	Frequency of f2 in Hz.
7	Loudness level of L2 in dB SPL.
8	Distortion product frequency in Hz.
9	Distortion product amplitude in dB SPL.
10	Loudness level of noise floor in dB SPL.
11	Test status (A= accepted, N= noisy, T/O= timed out response).
12	Pure tone threshold of 250 Hz in dB HL.
13	Pure tone threshold of 500 Hz in dB HL.
14	Pure tone threshold of 1000 Hz in dB HL.
15	Pure tone threshold of 2000 Hz in dB HL.
16	Pure tone threshold of 4000 Hz in dB HL.
17	Pure tone threshold of 8000 Hz in dB HL.
18	Subject age.
19	Subject gender.

The program that wrote the raw data into files was named CSV 2 EXP (Comma separated values to experiments) and the C++ code for this program can be seen on the accompanied CD. The newly created data files looked different for every experiment, depending on which variables were chosen for that specific experiment and the way the input data was presented to the neural network. If, for example, all frequencies were used for an experiment and age was presented to the network in 5 year categories, that data file would look different from a data file where only the high frequencies were used or if age was presented to the network in 10 year categories.

Table 4.8 is an example of a fraction of a newly created data file for a "No AMP" experiment where all 11 frequencies were used as input data, threshold was defined as 3 dB above the noise floor, gender was included and age was depicted in 10 year categories to attempt to predict the PTT frequency 500 Hz.

Table 4.8 Example of a fraction of a data file of one experiment to predict 500 Hz. (There are 23 out of 120 ears in the example.)

```
"Subject 1 "Right,0,1,0,0,0,0,0,0,0,0,  0,  0,0,0,0,1,1,1,0,0,0,0, 0,0,0,0,0,1,1,0,0,0,0, 0,0,0,0,1,0,0,1,0,0,0, 0,0,0,0,1,0,0,1,0,0,0, 0,0,0,0,0,1,1,0,0,1,0, 0,1,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,1,0,1,1,0, 0,0,0,1,0,0,0,0,0,0,0,  0,0,0,0,0,1,0,0,
"Subject 1 "Left,0,1,0,0,0,0,0,0,0,0,  0,  0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,0,0,0,0,1, 0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,0,0,1,0,0, 0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,1,1,1,1,0, 0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,0,0,0,0,0,  0,0,0,0,1,0,0,0,
"Subject 2 "Right,0,0,0,0,1,0,0,0,0,0,  1,  0,0,0,0,0,0,0,0,0,0,0, 0,0,1,0,0,0,0,0,1,0,0, 0,0,0,0,0,0,0,0,0,1,0, 0,0,1,0,0,0,0,0,1,0,0, 0,0,1,1,0,0,0,1,0,1,0, 0,0,0,0,1,0,0,0,1,1,1, 0,1,0,0,0,0,1,0,0,0,0, 0,1,0,0,0,1,0,0,0,0,0,  0,0,0,0,0,1,0,0,
"Subject 3"Right,0,0,0,0,1,0,0,0,0,0,  0,  0,0,0,0,0,0,0,0,1,0,0, 0,0,0,0,0,0,0,0,1,0,0, 0,0,0,0,0,0,0,1,1,1,0, 0,0,0,0,0,0,0,1,1,0,0, 0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,0,1,1,1,0, 0,0,0,0,0,0,1,1,1,1,1, 0,0,0,0,0,0,1,1,1,1,1,  1,0,0,0,0,0,0,0,
"Subject 3"Left,0,0,0,0,1,0,0,0,0,0,  0,  0,0,0,0,0,0,0,0,0,1,0, 0,0,0,0,0,0,0,0,0,1,0, 0,0,0,0,0,0,0,0,0,1,0, 0,0,0,0,0,0,0,0,1,1,0, 0,0,0,0,0,1,0,0,1,1,0, 0,0,0,0,0,0,0,1,1,1,1, 0,0,0,0,0,0,1,1,1,1,1, 0,0,0,0,0,0,1,1,1,1,1,  1,0,0,0,0,0,0,0,
"Subject 4"Right,0,1,0,0,0,0,0,0,0,0,  1,  0,0,0,0,0,0,1,0,0,0, 0,0,0,0,0,0,0,0,0,1,0, 0,0,0,0,0,0,0,0,0,1,0, 0,0,0,0,0,0,1,0,0,1,0, 0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,1,0,1,0,0, 0,0,0,0,1,0,0,0,0,1,0,  0,0,0,0,0,1,0,0,
" Subject 5 "Right,0,0,1,0,0,0,0,0,0,0,  0,  0,1,0,0,1,1,1,1,0,1,0, 0,0,0,0,1,0,1,1,1,1,1, 0,1,0,1,1,0,1,1,1,1,1, 0,1,1,1,1,0,1,1,1,1,1, 0,0,0,1,1,1,1,1,1,1,0, 0,1,1,1,1,1,1,1,1,1,0, 0,0,1,1,1,1,1,1,1,1,0, 0,0,1,0,1,1,1,1,1,1,1,  1,0,0,0,0,0,0,0,
" Subject 5 "Left,0,0,1,0,0,0,0,0,0,0,  0,  0,0,0,0,0,0,0,0,1,1,0, 0,0,0,0,0,0,0,1,1,1,0, 0,0,0,0,0,1,0,1,1,1,0, 0,0,0,0,0,0,0,1,1,1,0, 0,0,0,0,0,0,0,1,1,1,1, 0,0,0,0,0,0,1,1,1,1,0, 0,1,1,0,0,1,0,1,1,1,1, 0,0,0,0,0,0,0,1,1,1,1,  1,0,0,0,0,0,0,0,
" Subject 6 "Right,0,0,1,0,0,0,0,0,0,0,  0,  0,0,1,0,0,0,0,1,1,1,0, 0,0,1,1,1,1,0,0,1,1,1, 0,0,1,1,1,1,1,1,1,1,0, 1,1,1,1,1,1,1,1,1,1,1, 0,0,1,1,1,1,1,1,1,1,1, 0,0,1,1,1,1,1,1,1,1,1, 0,1,1,1,1,1,1,1,1,1,1, 0,1,0,1,1,1,1,1,1,1,1,  1,0,0,0,0,0,0,0,
" Subject 6 "Left,0,0,1,0,0,0,0,0,0,0,  0,  0,0,0,1,1,0,0,0,1,0,0, 0,1,1,1,1,1,0,0,1,1,0, 0,1,1,1,1,1,1,1,1,1,0, 0,1,1,1,1,1,0,1,1,1,1, 0,1,1,1,1,1,1,1,1,1,1, 1,1,1,1,1,1,1,1,1,1,1, 0,1,1,1,1,1,1,1,1,1,1, 1,1,1,1,1,1,1,1,1,1,1,  1,0,0,0,0,0,0,0,
" Subject 7 “Right,0,0,0,0,0,1,0,0,0,0,  0,  0,0,0,1,0,0,1,0,0,1,0, 0,1,1,0,0,1,0,1,1,0,0, 0,0,0,0,0,0,0,0,0,0,1, 0,0,0,0,0,1,1,1,1,0,0, 0,0,1,0,0,0,1,0,0,0,0, 0,0,0,1,0,0,1,0,0,0,1, 0,0,0,1,1,0,0,1,0,1,0, 0,0,0,0,1,0,0,1,1,0,0,  0,1,0,0,0,0,0,0,
" Subject 7 "Left,0,0,0,0,0,1,0,0,0,0,  0,  0,0,0,1,0,0,0,0,0,0,0, 0,0,0,0,0,1,1,0,0,1,0, 0,1,0,0,0,0,0,0,0,1,0, 0,0,1,0,0,0,0,1,0,0,0, 0,0,0,1,0,0,0,0,1,0,0, 0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,0,1,1,0,0, 0,1,0,0,0,1,0,1,0,0,0,  0,0,0,1,0,0,0,0,
" Subject 8 "Right,0,1,0,0,0,0,0,0,0,0,  1,  0,0,0,0,0,1,0,0,0,0,0, 0,0,0,0,0,0,0,0,1,0,0, 0,0,0,0,1,0,0,0,1,0,0, 0,0,0,0,1,1,1,0,0,0,1, 0,0,0,0,0,0,0,1,1,0,0, 0,0,0,1,0,1,0,0,1,1,0, 0,0,0,0,0,0,0,1,0,0,0, 0,0,0,0,0,1,1,1,0,1,0,  0,0,0,0,0,1,0,0,
" Subject  8 "Left,0,1,0,0,0,0,0,0,0,0,  1,  0,0,0,0,0,0,1,0,0,0,0, 0,0,0,0,0,0,0,0,1,1,0, 0,0,0,0,0,0,0,0,0,0,1, 0,0,0,0,0,0,0,1,1,1,0, 0,0,0,0,0,1,0,0,0,0,0, 0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,1,1,1,1,1,0, 0,0,0,1,1,0,1,1,1,1,0,  0,0,1,0,0,0,0,0,
" Subject  9 "Right,0,0,1,0,0,0,0,0,0,0,  0,  0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,1,1,0,0,0, 0,0,0,0,0,0,0,1,1,1,0, 0,0,0,0,0,0,0,1,1,1,0, 0,0,0,0,0,1,0,1,1,1,0, 0,0,0,0,0,0,1,1,1,1,0, 0,0,0,0,0,1,1,1,1,1,1, 0,0,0,0,0,1,1,1,1,1,1,  0,1,0,0,0,0,0,0,
" Subject  9 "Left,0,0,1,0,0,0,0,0,0,0,  0,  0,0,0,0,0,0,0,1,0,1,1, 0,0,0,0,0,0,1,1,1,1,0, 0,0,0,0,0,0,1,1,1,1,0, 0,0,1,1,1,1,1,1,1,1,1, 0,1,1,1,1,0,1,1,1,1,1, 0,0,1,1,1,1,1,1,1,1,1, 1,1,1,1,1,1,1,1,1,1,1, 0,1,1,1,1,1,1,1,1,1,1,  1,0,0,0,0,0,0,0,
" Subject  10 "Right,0,1,0,0,0,0,0,0,0,0,  0,  0,0,0,0,0,0,0,1,0,0, 0,0,0,0,0,0,0,0,1,0,0, 0,0,0,0,0,1,0,1,1,1,0, 0,0,0,0,0,1,0,1,1,1,1, 0,0,0,0,0,0,1,1,1,1,1, 0,0,0,1,1,1,1,1,1,1,1, 0,0,0,0,1,1,1,1,1,1,1, 0,0,0,1,1,1,1,1,1,1,1,  0,1,0,0,0,0,0,0,
" Subject  11 "Right,0,1,0,0,0,0,0,0,0,0,  1,  0,0,0,0,0,0,1,0,1,0,0, 0,0,0,0,0,0,0,0,1,1,0, 0,0,1,1,1,0,1,0,1,1,0, 0,0,1,0,0,0,1,1,1,1,0, 0,0,1,0,0,1,1,1,1,1,0, 0,0,0,1,0,1,1,1,1,1,0, 0,0,0,0,1,1,1,1,1,1,1, 0,0,1,1,1,1,1,1,1,1,1,  0,0,1,0,0,0,0,0,
" Subject  11 "Left,0,1,0,0,0,0,0,0,0,0,  1,  0,0,0,0,0,0,0,0,1,0,0, 0,0,0,0,0,0,0,1,1,0,0, 0,0,0,0,1,1,0,1,1,1,0, 1,0,0,0,1,0,0,1,1,1,1, 0,0,0,1,1,1,1,1,1,1,0, 0,1,1,1,1,1,1,1,1,1,1, 0,0,1,0,1,1,1,1,1,1,0, 0,0,1,1,1,1,1,1,1,1,1,  0,0,1,0,0,0,0,0,
" Subject  12 "Right,0,0,0,1,0,0,0,0,0,0,  0,  0,0,0,0,0,0,1,0,1,0,0, 1,0,0,1,0,1,0,0,0,0,1, 0,0,1,0,1,0,0,1,0,0,1, 0,0,1,1,0,0,0,1,1,0,0, 0,0,1,1,1,1,1,1,0,0,1, 1,1,1,1,1,1,1,1,0,0,0, 0,1,1,1,1,1,1,1,1,1,0, 0,1,1,1,1,1,1,1,0,1,0,  1,0,0,0,0,0,0,0,
" Subject  12 "Left,0,0,0,1,0,0,0,0,0,0,  0,  0,0,0,0,1,1,0,1,0,0,1, 0,1,0,1,1,0,0,1,1,0,0, 0,0,1,1,0,0,0,1,0,0,1, 0,1,1,1,0,0,1,1,0,0,1, 0,1,1,1,0,1,0,1,1,1,0, 1,1,1,1,1,0,0,1,0,1,0, 1,1,1,1,1,1,1,1,0,0,0, 1,1,1,1,1,1,1,1,1,1,0,  1,0,0,0,0,0,0,0,
" Subject  13 "Right,0,0,1,0,0,0,0,0,0,0,  1,  0,0,0,0,0,0,0,1,0,0, 0,0,0,0,1,0,1,1,0,0,0, 0,0,0,1,1,1,1,1,1,1,0, 1,1,1,0,1,0,0,1,1,1,0, 1,0,0,1,1,1,1,1,1,1,0, 0,0,0,1,1,1,1,1,1,1,0, 0,0,1,1,1,1,1,1,1,1,0, 0,1,1,1,1,1,1,1,1,1,0,  1,0,0,0,0,0,0,0,
" Subject  13 "Left,0,1,0,0,0,0,0,0,0,0,  1,  0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,0,0,0,1,0, 0,0,0,1,1,1,1,0,0,0,0, 0,1,1,1,1,0,1,1,1,1,0, 1,1,1,1,1,0,1,1,1,1,0, 0,1,1,1,1,0,1,1,1,1,0, 1,1,1,1,1,1,1,1,1,1,1, 1,1,1,1,1,1,1,1,1,1,1,  0,1,0,0,0,0,0,0,
```

After the data manipulation and creation of data files it became clear what the requirement for neural network topology is.

4.8.3 Experiments to Determine Neural Network Topology and Error Tolerance Levels

For this project a three-layer backpropagation neural network was chosen. One input layer presented all data to the network, one output layer gave the prediction of pure tone threshold at a given frequency and one hidden layer with a set of weights on each side of it connected the input and output layers. According to Hornik et al. (1989), one hidden layer is enough provided that there are enough middle level neurons for the complexity of the problem.

Many of the ideas on requirements for network topology and data manipulation techniques came from trial runs that were done in the previous study (De Waal, 1998). A short overview of the previous study's trial runs will be given to promote understanding of current topology and the history of methods tried and how it influenced the current way of thinking.

4.8.3.1 History of Trial Runs Done in the Previous Study (De Waal, 1998)

At first, a very simple approach was tried: The neural network had 11 input nodes, representing the L1 dB SPL value where the DPOAE threshold was measured. The neural network had to predict hearing ability at 500, 1000, 2000 and 4000 Hz in dB SPL and had 4 output neurons. The number of middle level neurons was set at 20 and the acceptable prediction error during the training period at 5 dB for this test run.

After a few hours it became clear that the neural network was unable to converge during the training period and that no accurate predictions could be made. For the next few trial runs, middle level neurons were increased up to 100 or the acceptable prediction error during the training period were decreased to 1 dB. All these changes did not improve convergence or prediction ability. The reason was lots of missing data due to absent responses: All the lowest L1 values where a DPOAE response was measured were used as input data for the neural network. There were however some of the hearing impaired subjects that did not have any DPOAE responses at certain frequencies, and no DPOAE threshold values were available to use as input data. All these absent DPOAE thresholds were depicted with a "zero". It became clear that the absence of DPOAE thresholds in the hearing impaired population (about 66% of the subjects) called for a different data preparation method.

As a second approach, the input data was manipulated to present absent and present responses in a different way. Up to now, input data consisted of decibel sound pressure level (SPL) quantities, depicting either a DPOAE threshold at a certain L1 value or DPOAE amplitude. Output data also predicted hearing thresholds in decibel sound pressure level (dB SPL) values. For this approach all data was rewritten in a binary format. The presence of a DPOAE response was depicted with a "1" whereas the absence of a DPOAE response was depicted with a "0".

The criteria for the presence of a DPOAE response were that the DPOAE response had to be 3 dB above the noise floor and that the test status had to be "accepted". All responses less than 3 dB above the noise floor or with a test status that was "noisy" or "timed out" were regarded as absent responses. (It should be noted that Kemp (1990) warned that in order to determine if a response is 3 dB above the noise floor, one could not merely subtract the noise floor from the DPOAE amplitude in its decibel form. The two values should be converted back to its pressure value (watt/m^2), and then subtracted.)

It was during this approach that the 88 input nodes, all zeros and ones, depicting DPOAE presence or absence for all eight DP Grams and 11 frequencies were formulated. The only information available to the neural network in this trail run was therefore the pattern of absent and present responses at all eight loudness levels and no information regarding the amplitudes of DPOAEs were available. For the present study, DPOAE amplitude was reintroduced as described in 4.8.1.4 "DPOAE amplitude."

This binary approach offered the first solution to the problem of absent DPOAE results. For the first time all the data could be used and the neural network could be trained with data across all categories of hearing impairment.

The way in which the output data was presented was also changed from dB SPL output at a given frequency to the binary dummy variable technique where PTTs were predicted into one of seven 10dB categories.

The effects of in the inclusion of gender and age variables were determined. Age was presented in the dummy variable technique into one of nine 10-year categories, gender with a one or a zero. The network had 98 input nodes, 140 middle level neurons and seven output neurons for prediction into one of seven 10dB categories.

Prediction error during training was set at 5%. Age had a very positive effect on prediction accuracy. Gender had very little effect. The neural network run that included both variables at the same time had the best prediction accuracy. It was therefore decided to include both these variables in the present study for every neural network run.

A very important aspect to keep in mind is that for the previous study, the network was trained with the data of 119 ears to predict the one remaining ear. This process was repeated 120 times to predict every ear once. This means that a subject's one ear was included in the training set while the other ear was predicted. It is quite possible that a subject's PTTs for both ears might be related, for example in the case of noise exposure, the two ears might look very similar. For this research project, both ears of a subject were removed out of the training set. The network was trained with 118 ears and predicted the remaining two ears one at a time. The following section discusses network topology for the present study.

4.8.3.2 The Number of Nodes in the Input Layer

As described in the input data manipulation section, the number of input data sets determines the number of nodes that are needed in a neural network's input layer. The number of input neurons needed for each experiment was determined by the variables that served as input data as well as the way in which they were represented. Table 4.9 is a summary of how to determine how many input nodes were needed for each type of experiment. The base input of nodes is when low frequency DPOAEs were omitted. The other columns serve as an indication of how many input nodes have to be added for that situation or experiment.

Table 4.9: Determination of the number of input nodes

	Base input # of nodes	**Low Hz Included**	**Age 5 year categories**	**Age 10year categories**	**Gender**
No AMP	64	+24	+20	+10	+1
AMP 100	64	+24	+20	+10	+1
AMP 40	64	+24	+20	+10	+1
ALT AMP	256	+96	+20	+10	+1

Example: When a AMP 40 experiment is run where the low frequencies are included and age was depicted in 5 year categories, the total number of input nodes needed for that experiment is 64 (base number) + 24 (low frequencies included) + 20 (age) + 1 (gender) = 109 input nodes.

4.8.3.3 The Number Neurons in the Output Layer

In the case of the output and hidden layers, the components are being referred to as neurons because of the two layers of connectivity (an input and an output), which gives it the similar structure as a neuron with a synapse on each side.

A network predicted the pure tone threshold (PTT) for just one frequency at a time. The PTT was predicted into one of eight possible 10 dB categories. There were therefore eight output neurons, each with a value between zero and one, indicating the likelihood of the PTT to be in that specific category. The category with the highest value was chosen to predict a PTT for a certain frequency.

4.8.3.4 The Number of Neurons in the Hidden or Middle Layer

The number of neurons in the hidden or middle layer cannot be determined merely by the amount of input or output data but is a function of the diversity of the data (Blum, 1992). The number of middle layer neurons determines the accuracy of prediction during the training period. With an insufficient number of middle neurons, the network is unable to form adequate midway representations or to subtract significant features of the input data (Nelson & Illingworth, 1991). With too many middle neurons the network has difficulty to make generalizations (Rao & Rao, 1995; Nelson & Illingworth, 1991). The number of middle layer neurons was determined by trial and error, based on the accuracy of the prediction during the training period.

Experiments for AMP 100, AMP 40 and No AMP were run with 80, 100 and 120 middle neurons. Prediction accuracy was studied to determine if an increase in middle level neurons had more accurate PTT predictions as a result. In the case of the ALT AMP experiment where the amplitude of the DPOAE was presented with the dummy variable technique, the number of input nodes increased so much that 80, 100 or 120 middle level neurons were not enough. For the ALT AMP experiments, the number of middle level neurons was doubled to compensate for network complexity due to the large number of input nodes. ALT AMP was therefore run with 160, 200, and 240 middle level neurons.

4.8.3.5 Error Tolerance Levels

As was described in 3.5.4.2 “Error tolerance”, the error tolerance refers to how accurate the network learns and predicts. This value determines if the weights will be changed to lengthen the training process or if the weights will be frozen to start with the prediction phase. The lower (closer to zero) the error tolerance level, the more accurate the learning and prediction but also the longer the training phase. Another aspect that is influenced by error tolerance levels is the networks' ability to generalize. For error tolerance set close to zero, the network might have difficulty predicting a PTT for a DPOAE set that is slightly out of the ordinary. Higher error tolerance levels might have slightly less accurate predictions but training is faster and generalization is better.

For this study, all experiments were run with error tolerance levels of 0.001 (within 0.1% accurate), 0.002 (0.2%) and 0.003 (within 0.3% accurate). The effect of the difference in prediction accuracy for the various error tolerance levels will be discussed in the chapter interpreting results.

Now that network topology, error tolerance and representation of input data in files are finalized, the network is ready to start the training and prediction processes.

4.8.3.6 The Training and Prediction Phase

The possibilities for experiment configuration were:

- Threshold of DPOAEs specified as 1, 2 or 3 dB above the noise floor.
- Age depicted as 10-year or 5-year increments.
- Amplitude depicted as ALT AMP, AMP 100, AMP 40 or No AMP.
- Middle level neurons as 80, 100 or 120 for AMP 100, AMP 40 and No AMP.
- Middle level neurons as 160, 200 or 240 for ALT AMP.

- Error tolerance levels as 0.1%, 0.2% or 0.3%.
- Low frequency DPOAE responses present or absent during training.

If all combinations of variables were run, the number of possible experiments would be 1728 possible combinations. All 1728 were run to determine the optimal set of DPOAE and ANN parameters for the prediction of PTTs. An additional 24 experiments were run: 12 in the "scenario five method" described in 4.8.1.6 "Number of ears or data in every output category to be predicted" to investigate the effect that the number of ears in each category has on prediction accuracy of PTTs. The other 12 were run with the same PTT input distribution as the previous study (De Waal, 1998) described in 4.8.1.7 "Number of PTTs in every input category for ANN training" to make comparisons between the two studies possible. That brings the total number of experiments to 1752. Each experiment took 80 minutes to run. 94 days were needed for neural network training and prediction. This process was done in parallel on three 600 MHz Pentiums. A third of the experiments were run on each computer to save time. It therefore took four and a half weeks for training and prediction of the four pure tone threshold frequencies.

The C++ program that fetched every data file and presented it to the neural network for training was called EXP 2 RES: (experiments to results) and the C++ code for this program can be viewed on the accompanied CD.

For the training of the neural network, both ears of a subject were left out to prevent contamination of data due to the inclusion of a related ear. The three-layered backpropagation neural network by Rao and Rao (1995) was used (software was supplied in addition to their book).

4.9 Data Analysis Procedures

At the end of the four and a half weeks, the output data consisted of 1752 predictions of a pure tone threshold at a certain frequency depicted as different values in all of the eight possible 10-dB categories. The 10 dB categories were presented in Table 4.6. An example of the raw output file of the network's predictions is presented in Table 4.10.

Table 4.10: A Fraction of the raw output file of the neural network's predictions. (The example contains 36 out of the 120 ears.)

0	0.176	0.00	0.000	0.00	0.002	0.00	0.989	0.00	0.004	0.00	0.000	1.00	0.000	0.00	0.000	0.00	0.000	ok_73
1	0.175	0.00	0.000	0.00	0.999	0.00	0.000	0.00	0.000	1.00	0.017	0.00	0.009	0.00	0.003	0.00	0.000	ok_51
2	0.134	0.00	0.000	0.00	0.003	0.00	0.000	0.00	0.381	0.00	0.010	1.00	0.000	0.00	0.000	0.00	0.000	ok_55
3	0.120	1.00	0.041	0.00	0.001	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.002	0.00	0.000	ok_108
4	0.076	1.00	0.392	0.00	0.004	0.00	0.007	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_88
5	0.127	0.00	0.000	0.00	0.158	0.00	0.116	0.00	0.001	0.00	0.000	1.00	0.001	0.00	0.001	0.00	0.000	ok_54
6	0.000	1.00	0.997	0.00	0.001	0.00	0.001	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_77
7	0.001	1.00	0.992	0.00	0.002	0.00	0.000	0.00	0.000	0.00	0.001	0.00	0.000	0.00	0.000	0.00	0.000	ok_62
8	0.000	1.00	1.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_70
9	0.000	1.00	1.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_79
10	0.176	0.00	0.993	1.00	0.000	0.00	0.004	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_57
11	0.167	0.00	0.000	0.00	0.889	0.00	0.001	1.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_72
12	0.175	0.00	0.000	0.00	0.039	0.00	0.985	0.00	0.001	0.00	0.009	1.00	0.010	0.00	0.005	0.00	0.000	ok_61
13	0.123	0.00	0.000	0.00	0.292	1.00	0.059	0.00	0.000	0.00	0.003	0.00	0.000	0.00	0.000	0.00	0.000	ok_126
14	0.177	0.00	1.000	1.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_78
15	0.000	1.00	0.999	0.00	0.003	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_57
16	0.131	0.00	0.079	1.00	0.012	0.00	0.343	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_68
17	0.142	0.00	0.545	0.00	0.001	1.00	0.000	0.00	0.000	0.00	0.002	0.00	0.000	0.00	0.002	0.00	0.000	ok_60
18	0.167	0.00	0.875	0.00	0.148	1.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_63
19	0.000	1.00	0.997	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_64
20	0.120	1.00	0.308	0.00	0.664	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_80
21	0.003	1.00	0.973	0.00	0.001	0.00	0.004	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_79
22	0.167	0.00	0.907	1.00	0.014	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_63
23	0.000	1.00	0.998	0.00	0.002	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_90
24	0.000	1.00	1.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_81
25	0.121	1.00	0.033	0.00	0.049	0.00	0.006	0.00	0.000	0.00	0.003	0.00	0.000	0.00	0.000	0.00	0.000	ok_90
26	0.176	1.00	0.012	0.00	0.998	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.001	0.00	0.000	0.00	0.000	ok_75
27	0.136	1.00	0.003	0.00	0.445	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_72
28	0.002	1.00	0.988	0.00	0.011	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_64
29	0.054	0.00	0.002	0.00	0.000	0.00	0.000	1.00	0.570	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_75
30	0.199	0.00	0.000	0.00	0.826	0.00	0.916	1.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_77
31	0.168	0.00	0.891	0.00	0.097	0.00	0.026	0.00	0.001	1.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_58
32	0.000	1.00	1.000	0.00	0.000	0.00	0.003	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_69
33	0.126	1.00	0.364	0.00	0.784	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_76
34	0.131	1.00	0.366	0.00	0.000	0.00	0.830	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_64
35	0.171	1.00	0.013	0.00	0.907	0.00	0.274	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	ok_103

In order to determine which category the PTT was predicted, the category with the highest value were chosen. The program that performed this task was called RES 2 ANA (results to analysis) and the C++ code for this program can be viewed on the accompanied CD.

The second function of RES 2 ANA was to determine how many predictions were accurate (within the same 10 dB category), how many were one 10 dB category out and how many predictions were wrong (more than one 10 dB category out). These calculations were made for each of the 10 dB categories as well as for the overall prediction ability of the network across all categories for that specific frequency. False positive and false negative predictions were calculated for each category. Another calculation made by RES 2 ANA was to determine how accurately normal hearing (0 – 25 dB) was predicted as normal, and also how accurately very good hearing (spanning 0 – 15 dB) was predicted as normal (within 0 – 25 dB). An example of how the data looked after this step can be seen in Table 4.11. The reason why category eight has no information is because maximum hearing loss at 500 Hz was 65 dB HL and falls in category seven. Category eight was created for 4000 Hz: Nine ears exhibited a PTT of larger than 65dB HL at 4000 Hz. There were therefore no data in category eight at 500, 1000 and 2000 Hz.

The last function of RES 2 ANA was to create a file that was compatible with Microsoft Excel 2000's spreadsheet to be able to use Excel to manipulate data and make visual representations of results.

Table 4.11: RES 2 ANA experiment result representation.

Experiment 62308

AI = 10, LF, Mid = 200, Err = 0.002, Th = 1 dB, Hz = 500, ALT AMP**

category	Correct prediction		One category out		Wrong prediction		False positive	False negative
C1	35/42	83.3%	4/42	9.5%	3/42	7.1%	2%	0%
C2	8/31	25.8%	17/31	54.8%	5/31	16.1%	8%	0%
C3	1/16	6.3%	10/16	62.5%	5/16	31.3%	0%	9%
C4	2/12	16.7%	2/12	16.7%	8/12	66.7%	0%	5%
C5	0/9	0%	1/9	11.1%	8/9	88.9%	0%	4%
C6	0/7	0%	0/7	0%	7/7	100%	0%	1%
C7	0/3	0%	0/3	0%	3/3	100%	0%	1%
C8	0/0	0%	0/0	0%	0/0	0%	0%	0%

Overall correct prediction for all categories	46/120	38%
Overall one category out for all categories	34/120	28%
Overall wrong predictions	40/120	33%
0 – 10 dB predicted as normal (0 – 20 dB)	39/42	92%
0 – 20 dB predicted as normal (0 - 20 dB)	60/73	82%

****Key for Table 4.11:**

AI = Age increment represented as 10 or 5 year categories
LF = Low frequencies present, No LF = Low frequencies absent
Mid = number of middle level neurons
Err = Error tolerance level
Th = Threshold specified as 1, 2 or 3 dB above noise floor
Hz = Frequency to be predicted
ALT AMP = method of amplitude presentation

categorical value. Four different networks were trained for the four prediction frequencies 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz. Data analysis consisted of analyzing the actual and predicted values of all 120 ears and to determine how many were predicted accurately, how many within one 10 dB class and how many were predicted incorrectly. Data was further manipulated in Excel for Windows 2000 to create visual representations.

There are numerous variables that influenced the outcome of this research project. It is quite possible that different DPOAE settings such as other frequency ratios or different loudness levels could reveal different results (Cacace et al. 1996). It is also possible that a different type of neural network or a network with a different topology could affect the results significantly (Nelson & Illingworth, 1991). It was attempted to specify all the stimulus variables that could have an effect on the outcome of this research project in great detail in the preceding chapters.

Chapter 5: Results

5.1 Introduction

Many studies investigated the correlation between distortion product otoacoustic emissions and pure tone thresholds (Durrant, 1992, Avan & Bonfils, 1993; Gaskill & Brown, 1990; Gorga et al. 1993; Probst & Hauser, 1990; Stover et al. 1996a; Gorga et al. 1996). All these studies reported very strong correlations between hearing ability and DPOAE measurements for high frequencies and a decline in correlation for lower frequencies. At 500 Hz, many researchers reported that the correlation was so poor that normal hearing could not be distinguished from impaired hearing due to noisy, missing and incomplete data (Gorga et al. 1993; Stover et al. 1996a; Probst & Hauser, 1990; Gorga et al. 1996). Other researchers took the process a step further and attempted to categorize hearing status in normal hearing and hearing-impaired populations and predict it as normal or impaired with DPOAE's (Kimberley et al. 1994a; Kimberley et al. 1994b; Moulin et al. 1994). In order to create perspective for the results of the present study, prediction accuracy of normal hearing from a few other studies that attempted to predict normal hearing across a range of frequencies will be summarized. Even though prediction frequencies do not overlap for all studies, it is possible to see tendencies for prediction accuracy in frequency regions and get an idea of expected success rates. The summary is given in Table 5.1.

Table 5.1: A comparison of studies: Prediction accuracy of normal hearing with DPOAEs.

	Kimberley et al. 1994a	**Kimberley et al. 1994b DP alone**	**Kimberley et al. 1994b DP + Age**	**Moulin et al. 1994**	**De Waal 1998**
500 Hz	*	*	*	*	92%
706 Hz	*	*	*	52.9%	*
1000 Hz	*	*	*	73.2%	87%
1025 Hz	92%	90%	90%	*	*
1413 Hz	*	*	*	75.6%	*
1464 Hz	88%	86%	87%	*	*
2000 Hz	*	*	*	*	84%
2050 Hz	83%	84%	83%	81.5%	*
2826 Hz	*	*	*	*	*
2880 Hz	70%	80%	83%	*	*
4000 Hz	*	*	*	79.4%	91%
4052 Hz	69%	88%	88%		*
5712 Hz	76%	80%	86%		*
*** Frequency not predicted in the research project**					

The purpose of this chapter is to present all the results obtained from all 1752 experiments in this research project in a logical way. The main goal of this project was to improve PTT prediction at 500, 1000, 2000 and 4000 Hz with DPOAEs and ANNs and each frequency's results will be given separately and in comparison to the previous study (De Waal, 1998). Sub goals for this study was to determine how certain variables of the subject, depiction of input data into the network and ANN configuration influenced prediction accuracy of PTTs and the results of these influences will follow the predictions at each frequency.

For the comparison of prediction accuracy of the present and previous study, certain aspects regarding the differences in methodology for each project should be clarified. The present study will be referred to as the 2000 study, the previous study as the 1998 study.

5.2 Aspects regarding Differences in Methodology for the Present (2000) and Previous (1998) study.

In the previous (1998) study, two types of experiments were performed; the one type predicted PTTs into one of seven 10 dB categories and the other type predicted PTTs into one of five categories. The last type of experiment was referred to as the "scenario five method".

In both methods of the 1998 study, the first category spanned 0-10 dB HL and the second category 11-20 dB HL - thus all PTT inputs depicting threshold information at 0 dB, 5 dB and 10 dB were placed in the first category and all PTTs depicting threshold information at 15 dB and 20 dB HL were placed in the second category.

Even though both categories seemingly only spanned 10dB, there was an uneven distribution of input data as was described in 4.8.1.7 in the previous chapter: Category one received three input thresholds and every subsequent category only two. The present (2000) study corrected this uneven distribution of input data and ensured that the PTT information of only two thresholds was allowed in every category. As seen in Table 5.2 depicting the results for the present study for 4000 Hz, categories described in the top row indicates the two thresholds for every category.

This correction however, makes a straightforward comparison between the two studies difficult for two reasons: First, the categories do not overlap anymore, and do not represent the same input or output decibel ranges. Second, in the previous (1998) study, normal hearing was defined according to Jerger (1980)'s definition which is normal hearing = 0 – 20 dB HL and was determined by the first two categories. For the present (2000) study, normal hearing is defined according to Goodman (1965)'s definition which is normal hearing = 0 – 25 dB HL and depicted by the first three categories. This definition of normal hearing was also recommended by the American Academy of Otolaryngology and the American Council of Otolaryngology (AAO-ACO) in 1979. To determine prediction accuracy of normal hearing for the present study, the first three categories will therefore be investigated.

Just for the sake of completeness, the three best experiments for each frequency were identified and were run in the PTT distribution method of the previous (1998) study where the first category (0 – 10 dB) received three inputs and all subsequent 10dB categories only two inputs. The reason for this was to investigate if normal hearing according to the definition of Jerger (1980) (0 - 20 dB HL) could be predicted more accurately based on all the other subject-, DPOAE- and ANN-variables that were experimented with. Results for these experiments to enable direct comparisons will be given for each frequency. It should be noted however, that this distribution correction of input thresholds possibly had a positive effect on prediction accuracy and that this comparison does not incorporate that possibility.

The results for the prediction of specific frequencies will be given in descending order, 4000 Hz first and 500 Hz last.

5.3 The Prediction of 4000 Hz.

Frequency specific results will be divided into results from the present study, and a comparison of results to the previous study.

5.3.1 Results Obtained from the Present Study for the Prediction of 4000 Hz.

The best prediction of 4000 Hz was obtained from experiment 19301. In this experiment, age was presented to the network in 5 dB categories, low frequency DPOAEs inputs were present, the number of middle neurons was 80, error tolerance 0.002, DPOAE threshold was defined as 2dB above the noise floor and the No AMP experiment type was used, in other words the amplitude of a DPOAE response were omitted. The results for the prediction accuracy for each category, false positive and negative responses and number of ears in every category are presented in Table 5.2.

Table 5.2: Present study: 4000 Hz predicted into one of eight 10dB categories (Experiment 19301).

Categories	1 (0 + 5dB)	2 (10 +15dB)	3 (20+25dB)	4 (30+35dB)	5 (40+45dB)	6 (50+55dB)	7 (60+65dB)	8 (70+75dB)
Correct	84.8%	27.8%	0%	12.5%	0%	45%	22.2%	11.1%
10dB out	15.2%	61.1%	57.1%	0%	28.6%	30%	33.3%	44.4%
Wrong	0%	11.17%	42.9%	87.5%	71.4%	25%	44.4%	44.4%
0-15dB predicted as 0-15dB	92%							
0-15dB predicted as 0-25dB	96%							
0-25dB predicted as 0-25dB	93%							
False positive responses				False negative responses				
	0%	1%	1%	3%	2%	2%	5%	1%
# ears in category	33	18	7	8	7	20	18	9

Normal hearing (0 - 25dBHL)(Goodman, 1965) was correctly predicted and separated from impaired hearing at 4000 Hz 93% of the time. Very good hearing (0 - 15 dB HL) was accurately predicted as very good (0 – 15 dB HL) 92% of the time. Very good hearing (0 – 15 dB HL) was accurately predicted as normal (0 – 25 dB HL) 96% of the time.

One aspect that should be kept in mind is the sensitivity and specificity of any procedure that might potentially be used as a hearing screening or diagnostic procedure. The sensitivity of a test refers to the test's ability to correctly identify subjects with a hearing loss whereas the specificity refers to the test's ability to correctly identify normal hearing (Konkle & Jacobson, 1991). Sensitivity and specificity is tied directly with the predictive value of a test. The more sensitive a test, the better it's negative predictive value, and the more specific a test, the better it's positive predictive value (Schwartz & Schwartz, 1991). The sensitivity is therefore affected by the number of false negative responses. (A false negative response is when a subject with a hearing loss is predicted as having normal hearing.) Specificity on the other hand, is affected by the number of false positive responses. (False positive responses refer to the number of subjects with normal hearing that has been identified as having a hearing loss.)

The false positive and false negative responses for each category can be seen in Table 5.2. The significance of the low occurrence of false negative and false positive responses will be discussed in Chapter 6.

Prediction accuracy for categories depicting hearing impairment was less satisfactory. It seems that prediction accuracy is greatly influenced by the number of ears in a specific category. The reason for this is that the neural network needs as much information as possible in every category (enough examples in every category) to make accurate predictions learned on previous examples. Category three (20dB and 25dB) and category five (40dB and 45 dB) for example had only seven ears in both categories and were never predicted accurately. Category six (50 dB and 55 dB) had 20 ears and was predicted accurately 45% of the time. Figure 5.1 summarizes the effect that the number of ears in every category had on prediction accuracy.

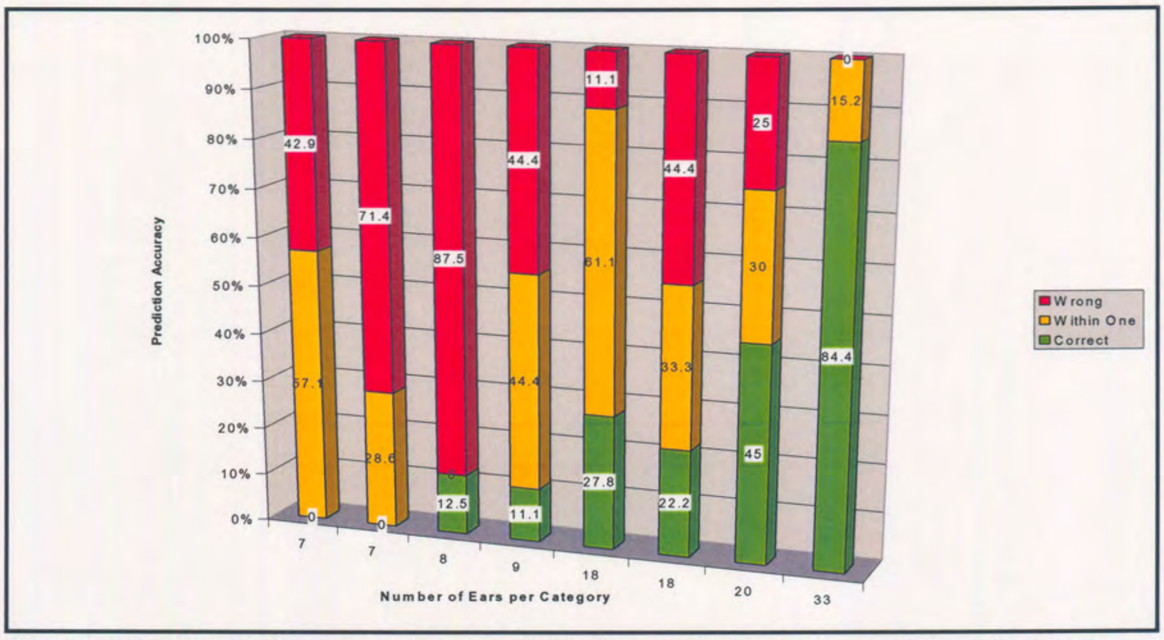


Figure 5.1: Prediction accuracy of 4000 Hz against number of ears in every category.

This aspect will be discussed in more detail in Chapter 6, but it seems that the unsatisfactory predictions of hearing impaired categories were greatly influenced by a lack of data in every category, and not just by errors in topology or lack of correlation between DPOAEs and PTTs.

The next section will discuss prediction accuracy of 4000 Hz for the few experiments that were run (described in 5.2) to enable the direct comparison between the present study (2000) and the previous study (1998).

5.3.2 Results of Present Study (2000) in Comparison to the Previous Study (1998) for 4000 Hz.

Prediction accuracy for the PTT distribution of three PTTs in the first category and two PTTs in every subsequent category was very similar for the present study (2000) and previous study (1998). Category two, three and six were predicted more accurately in the 2000 study. Normal hearing defined as 0-20 dB HL was predicted slightly better in the 2000 study (90% instead of 89%). One noteworthy improvement was the lower incidence of false positive responses in the 2000 study. The false positive rate for the present 2000 study (normal ears predicted as hearing impaired) was 3% for the two categories combined instead of 9% for the 1998 study. Results for prediction accuracy for all categories and false positive and negative rates are summarized in Table 5.3.

Table 5.3: Comparison of previous study (1998) in black and present study (2000) in red: 4000Hz predicted into 10dB categories.

Categories	1 (0-10dB)		2 (11-20dB)		3 (21-30dB)		4 (31-40dB)		5 (41-50dB)		6 (51-60dB)		7 (61-70dB)		8 (71-80dB)	
Correct	94%	92%	0%	14%	13%	25%	0%	0%	25%	0%	41%	55%	26%	20%	-	0%
10dB out	0%	2%	71%	57%	0%	13%	11%	11%	50%	75%	41%	18%	37%	40%	-	0%
Wrong	6%	6%	29%	29%	87%	62%	89%	89%	25%	25%	18%	27%	37%	40%	-	100%
0-10dB predicted as <20dB					94%						93%					
0-20dB predicted as 0-20dB					89%						90%					
False positive responses					**False negative responses**											
	6%	2%	3%	1%	2%	3%	1%	2%	0%	0%	0%	0%	3%	4%	-	1%
# ears in category	47	47	7	7	8	8	9	9	8	8	22	22	19	15	0	4

Prediction accuracy for the 2000 study scenario five method was generally slightly worse than the 1998 study. The false negative rate for the present study however, is lower (5% instead of 10%).

The results for prediction accuracy into the five categories of the scenario five method for both studies are summarized in Table 5.4.

Table 5.4: Comparison of previous study (1998) in black and present study (2000) in red: 4000Hz predicted in the scenario five method into 5 categories.

Categories	1 (0-10dB)		2 (11-20dB)		3 (21-35dB)		4 (36-50dB)		5 (51-65dB)	
Correct	92%	94%	14%	0%	17%	9%	15%	8%	68%	63%
10 dB out	2%	0%	57%	43%	25%	8%	85%	84%	15%	10%
Wrong	6%	6%	29%	57%	58%	83%	0%	8%	17%	27%
0-10dB predicted as <20dB	94%					93%				
0-20dB predicted as 0-20dB	91%					87%				
False positive responses					**False negative responses**					
	7%	2%	3%	3%	1%	4%	1%	0%	3%	7%
# ears in category	47	47	7	7	12	12	13	13	41	41

5.4 The Prediction of 2000 Hz.

Results will be divided into results from the present study, and a comparison of results to the previous study.

5.4.1 Results Obtained from the Present Study for the Prediction of 2000 Hz.

The best prediction of 2000 Hz was obtained from experiment 10301. In this experiment, age was presented to the network in 5 dB categories, low frequency DPOAEs inputs were present, the number of middle neurons was 80, error tolerance

0.002, DPOAE threshold was defined as 3dB above the noise floor and the No AMP experiment type was used, in other words the amplitude of a DPOAE response were omitted.

Normal hearing (0 - 25dBHL)(Goodman, 19965) was correctly predicted and separated from impaired hearing at 2000 Hz 88% of the time. Very good hearing (0 - 15 dB HL) was accurately predicted as very good (0 – 15 dB HL) 90% of the time. Very good hearing (0 – 15 dB HL) was accurately predicted as normal (0 – 25 dB HL) 95% of the time.

The results for the prediction accuracy for each category, false positive and negative responses and number of ears in every category are presented in Table 5.5.

Table 5.5: Present study: 2000 Hz predicted into one of eight 10dB categories (Experiment 10301).

Cate-gories	1 (0 + 5dB)	2 (10 +15dB)	3 (20+25dB)	4 (30+35dB)	5 (40+45dB)	6 (50+55dB)	7 (60+65dB)	8 (70+75dB)
Correct	78.1%	41.1%	0%	0%	0%	6.3%	7.7%	-*
10dB out	15.6%	55.2%	60%	22.2%	45.5%	56.3%	30.8%	-
Wrong	6.3%	3.4%	40%	77.8%	54.5%	37.4%	61.5%	-
0-15dB predicted as 0-15dB	90%							
0-15dB predicted as 0-25dB	95%							
0-25dB predicted as 0-25dB	88%							
False positive responses				**False negative responses**				
	1%	0%	4%	4%	4%	4%	3%	-
# ears in category	32	29	10	9	11	16	13	-
*** There were no ears in category eight, largest hearing loss measured for 2000 Hz was 65dB HL.**								

The false positive and false negative responses for each category can be seen in Table 5.5. The significance of the low occurrence of false negative and false positive responses will be discussed in Chapter 6.

Prediction accuracy for categories depicting hearing impairment was once again less satisfactory and greatly influenced by the number of ears in a specific category. Even the accuracy of prediction into an adjacent category seems to be dependant on the number of ears in a category. Category four (30dB and 35dB) had 9 ears and even though it was never predicted accurately, it was predicted into an adjacent category 22% of the time. Category five (40 dB and 45dB) had two ears more and was predicted into an adjacent class 45.5% of the time. Figure 5.2 summarizes the effect that ear count had on prediction accuracy for 2000 Hz.

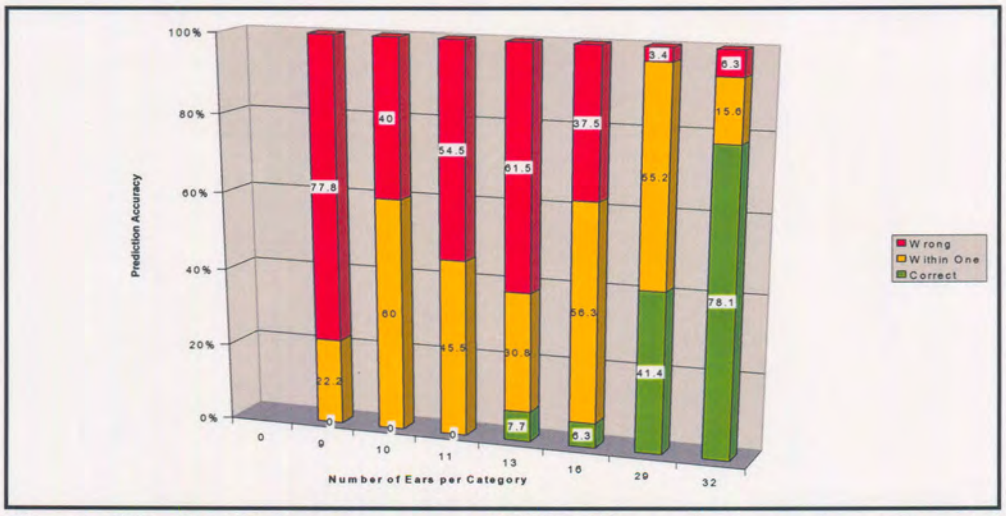


Figure 5.2: Prediction accuracy of 2000 Hz against number of ears in every category.

The next section will discuss prediction accuracy of 2000 Hz for the direct comparison between the present study (2000) and the previous study (1998).

5.4.2 Results of Present Study (2000) in Comparison to the Previous Study (1998) for 2000 Hz.

Prediction accuracy for the PTT distribution of three PTTs in the first category and two PTTs in every subsequent category was very similar for the present study (2000) and previous study (1998). Category two were predicted slightly more accurately in the 2000 study (from 15% – 20%) and category six considerable more accurate (from 19% to 44%). Normal hearing defined as 0-20 dB HL was predicted better in the 2000 study (88% instead of 82%). Another noteworthy improvement was the lower incidence of false positive responses in the 2000 study. The false positive rate for the present 2000 study (normal ears predicted as hearing impaired) was 6% for the two categories combined instead of 21% for the 1998 study. Results for prediction accuracy for all categories and false positive and negative rates are summarized in Table 5.6.

Table 5.6: Comparison of previous study (1998) in black and present study (2000) in red: 2000Hz predicted into 10dB categories.

Categories	1 (0-10dB)		2 (11-20dB)		3 (21-30dB)		4 (31-40dB)		5 (41-50dB)		6 (51-60dB)		7 (61-70dB)		8 (71-80 dB)	
Correct	88%	79%	15%	20%	0%	0%	0%	0%	24%	24%	19%	44%	0%	0%	*	*
10dB out	6%	17%	55%	50%	29%	29%	11%	22%	29%	47%	37%	44%	33%	33%	*	*
Wrong	6%	4%	30%	30%	71%	71%	89%	78%	47%	29%	44%	12%	67%	67%	*	*
0-10dB predicted as <20dB					94%						95%					
0-20dB predicted as 0-20dB					82%						88%					
False positive responses					**False negative responses**											
	6%	1%	15%	5%	3%	1%	1%	4%	3%	1%	3%	0%	1%	0%	*	*
# ears in category	48		20		7		9		17		16		3		*	*
**** There were no ears in category eight, largest hearing loss measured for 2000 Hz was 65dB HL.**																

Prediction accuracy for the 2000 study in the scenario five method was generally slightly worse than the 1998 study. Prediction of normal hearing as 0 – 20 dB HL (Jerger, 1980) was considerably worse, 83% instead of 95%. The false negative and false positive rates are similar.

The results for prediction accuracy into the five categories of the scenario five method for both studies are summarized in Table 5.7.

Table 5.7: Comparison of previous study (1998) in black and present study (2000) in red: 2000Hz predicted in the scenario five method into 5 categories.

Categories	1 (0-10dB)		2 (11-20dB)		3 (21-35dB)		4 (36-50dB)		5 (51-65dB)	
Correct	88%	90%	15%	10%	8%	8%	24%	19%	37%	37%
10 dB out	8%	6%	45%	55%	67%	50%	48%	62%	47%	47%
Wrong	4%	4%	40%	35%	25%	42%	28%	19%	16%	16%
0-10dB predicted as <20dB	96%					95%				
0-20dB predicted as 0-20dB	95%					83%				
False positive responses					**False negative responses**					
	4%	1%	8%	8%	1%	5%	5%	3%	3%	2%
# ears in category	48	48	20	20	12	12	21	21	19	19

5.5 The Prediction of 1000 Hz.

Results for 1000 Hz will also be divided into results from the present study, and a comparison of results to the previous study.

5.5.1 Results Obtained from the Present Study for the Prediction of 1000 Hz.

The best prediction of 1000 Hz was obtained from experiment 68509. In this experiment, age was presented to the network in 5 dB categories, low frequency DPOAEs inputs were present, the number of middle neurons was 200, error tolerance 0.003, DPOAE threshold was defined as 3dB above the noise floor and the ALT AMP experiment type was used, as described in 4.8.1.4.3 in chapter four. The results for the prediction accuracy for each category, false positive and negative responses and number of ears in every category are presented in Table 5.8.

Table 5.8: Present study: 1000 Hz predicted into one of eight 10dB categories (Experiment 68509).

Categories	1 (0 + 5dB)	2 (10 +15dB)	3 (20+25dB)	4 (30+35dB)	5 (40+45dB)	6 (50+55dB)	7 (60+65dB)	8 (70+75dB)
Correct	80%	24.1%	33.3%	0%	13.3%	0%	18.2%	*
10dB out	15%	55.2%	25%	20%	6.7%	12.5%	9.1%	*
Wrong	5%	20.7%	41.7%	80%	73.3%	87.5%	72.7%	*
0-15dB predicted as 0-15dB	86%							
0-15dB predicted as 0-25dB	89%							
0-25dB predicted as 0-25dB	88%							
False positive responses				**False negative responses**				
	0%	5%	1%	2%	10%	5%	5%	*
# ears in category	40	29	12	5	15	8	11	*
*** There were no ears in category eight, largest hearing loss measured for 1000 Hz was 65dB HL.**								

Normal hearing (0 - 25dBHL)(Goodman, 19965) was correctly predicted and separated from impaired hearing at 1000 Hz 88% of the time. Very good hearing (0 - 15 dB HL) was accurately predicted as very good (0 – 15 dB HL) 86% of the time. Very good hearing (0 – 15 dB HL) was accurately predicted as normal (0 – 25 dB HL) 89% of the time.

The false positive and false negative responses for each category can be seen in Table 5.8. The significance of the occurrence of false negative and false positive responses will be discussed in Chapter 6.

Prediction accuracy for categories depicting hearing impairment was once again less satisfactory. All categories with an ear count of less than 10 were never predicted accurately. Figure 5.3 summarizes the effect that the number of ears in every category had on prediction accuracy.

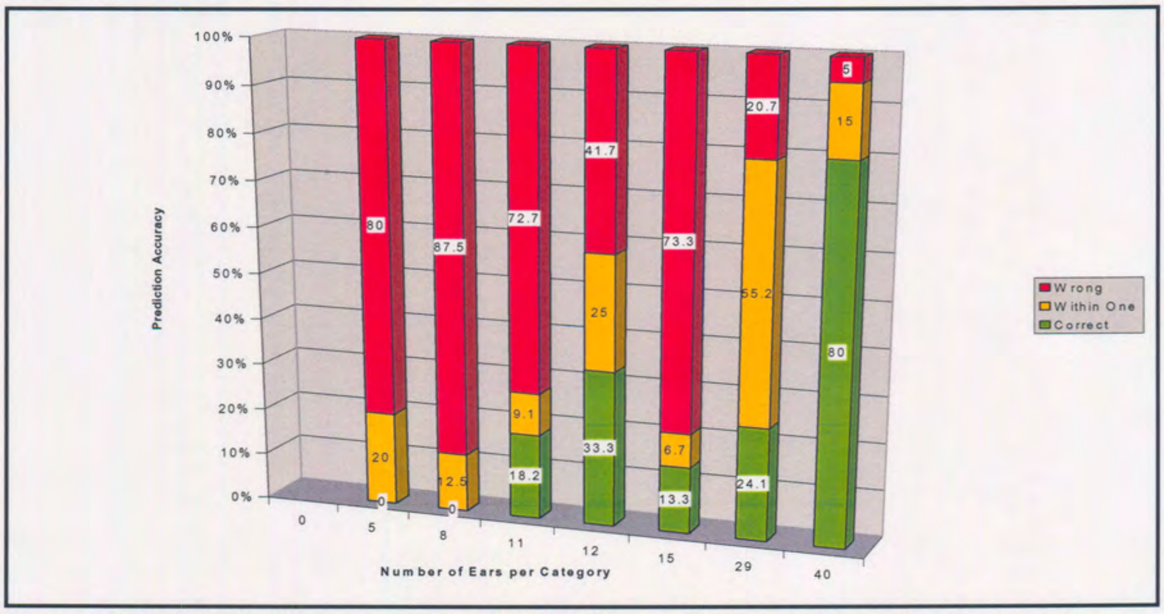


Figure 5.3: Prediction accuracy of 1000 Hz against number of ears in every category.

This aspect will be discussed in more detail in Chapter 6, but it once again seems that the unsatisfactory predictions of hearing impaired categories were greatly influenced by a lack of data in every category, and not just by errors in topology or lack of correlation between DPOAEs and PTTs.

The next section will discuss prediction accuracy of 1000 Hz for the direct comparison between the present study (2000) and the previous study (1998).

5.5.2 Results of Present Study (2000) in Comparison to the Previous Study (1998) for 1000 Hz.

Results for prediction accuracy for all categories and false positive and negative rates are summarized in Table 5.9.

Table 5.9: Comparison of previous study (1998) in black and present study (2000) in red: 1000Hz predicted into 10dB categories.

Cate-gories	1 (0-10dB)		2 (11-20dB)		3 (21-30dB)		4 (31-40dB)		5 (41-50dB)		6 (51-60dB)		7 (61-70dB)		8 (71-80 dB)	
Correct	92%	95%	23%	18%	0%	0%	0%	0%	31%	19%	13%	0%	14%	0%	*	*
10dB out	3%	2%	44%	33%	33%	33%	67%	33%	13%	13%	25%	63%	0%	0%	*	*
Wrong	5%	3%	33%	50%	67%	67%	33%	67%	56%	68%	62%	37%	86%	100	*	*
0-10dB predicted as <20dB					95%						96%					
0-20dB predicted as 0-20dB					84%						83%					
False positive responses					**False negative responses**											
	8%	1%	12%	9%	3%	5%	1%	0%	4%	8%	0%	2%	2%	3%	*	*
# ears in category	59	59	18	18	9	9	3	3	16	16	8	8	7	7	*	*
**** There were no ears in category eight, largest hearing loss measured for 1000 Hz was 65dB HL.**																

Prediction accuracy for the PTT distribution of three PTTs in the first category and two PTTs in every subsequent category was very similar for the present study (2000) and previous study (1998). Category one was predicted slightly more accurately in the 2000 study (from 92% – 95%) but subsequent categories were predicted less accurately. Normal hearing defined as 0-20 dB HL was predicted slightly less accurately in the 2000 study. The incidence of false positive responses in the 2000 study decreased and the false negative responses increased.

Prediction accuracy for the 2000 study in the five category scenario was generally worse than the 1998 study. Prediction of normal hearing as 0 – 20 dB HL (Jerger, 1980) was considerably worse, 81% instead of 87%. The results for prediction accuracy into the five categories of the scenario five method for both studies are summarized in Table 5.10.

Table 5.10: Comparison of previous study (1998) in black and present study (2000) in red: 1000Hz predicted in the scenario five method into 5 categories.

Categories	**1 (0-10dB)**		**2 (11-20dB)**		**3 (21-35dB)**		**4 (36-50dB)**		**5 (51-65dB)**	
Correct	93%	88%	22%	11%	0%	0%	37%	16%	27%	14%
10 dB out	5%	3%	39%	50%	67%	56%	5%	37%	20%	33%
Wrong	2%	9%	39%	39%	33%	44%	58%	47%	53%	53%
0-10dB predicted as <20dB	98%					91%				
0-20dB predicted as 0-20dB	87%					81%				
False positive responses					**False negative responses**					
	9%	4%	9%	7%	2%	5%	3%	7%	3%	5%
# ears in category	59	59	18	18	9	9	19	19	15	15

The results for 500 Hz will follow next.

5.6 The Prediction of 500 Hz

Results for the prediction of 500 Hz will be divided into results from the present study, and a comparison of results to the previous study.

5.6.1 Results Obtained from the Present Study for the Prediction of 500 Hz.

For the prediction of 500 Hz, experiment 62313 revealed the greatest separation of normal hearing (0 – 25 dB HL) (Goodman, 1985) and accurately predicted normal hearing 94% of the time. This is an exceptionally good prediction of normal hearing, especially for 500 Hz since so many other research projects have struggled with the prediction of normal hearing at 500 Hz in the past (Kimberley et al. 1994a; Kimberley et al. 1994b; Stover et al. 1996a; Gorga et al. 1993). The significance of this finding will be discussed in Chapter 6. For experiment 62313, age was presented to the network in 5 dB categories, low frequency DPOAEs inputs were absent, the number of middle neurons was 240, error tolerance 0.002, DPOAE threshold was defined as 1dB above the noise floor and the ALT AMP experiment type was used. The results for the prediction accuracy for each category, false positive and negative responses and number of ears in every category are presented in Table 5.11. False positive and false negative responses are low and the significance thereof will be discussed in Chapter 6.

Table 5.11: Present study: 500 Hz predicted into one of eight 10dB categories (Experiment 62313).

Catego-ries	1 (0 + 5dB)	2 (10 +15dB)	3 (20+25dB)	4 (30+35dB)	5 (40+45dB)	6 (50+55dB)	7 (60+65dB)	8 (70+75dB)
Correct	78.6%	9.7%	0%	16.7%	22.2%	0%	0%	*
10dB out	14.3%	80.6%	62.5%	8.3%	0%	14.3%	0%	*
Wrong	4.8%	9.7%	37.5%	75%	66.7%	85.7%	100%	*
0-15dB predicted as 0-15dB	75%							
0-15dB predicted as 0-25dB	95%							
0-25dB predicted as 0-25dB	94%							
False positive responses				**False negative responses**				
	0%	2%	1%	6%	5%	4%	0%	-
# ears in category	42	31	16	12	9	7	3	-
* There were no ears in category eight, largest hearing loss measured for 500 Hz was 65dB HL.								

Prediction accuracy for categories depicting hearing impairment was also poor. Figure 5.4 summarizes the effect that the number of ears in every category had on prediction accuracy. This aspect will be discussed in more detail in Chapter 6.

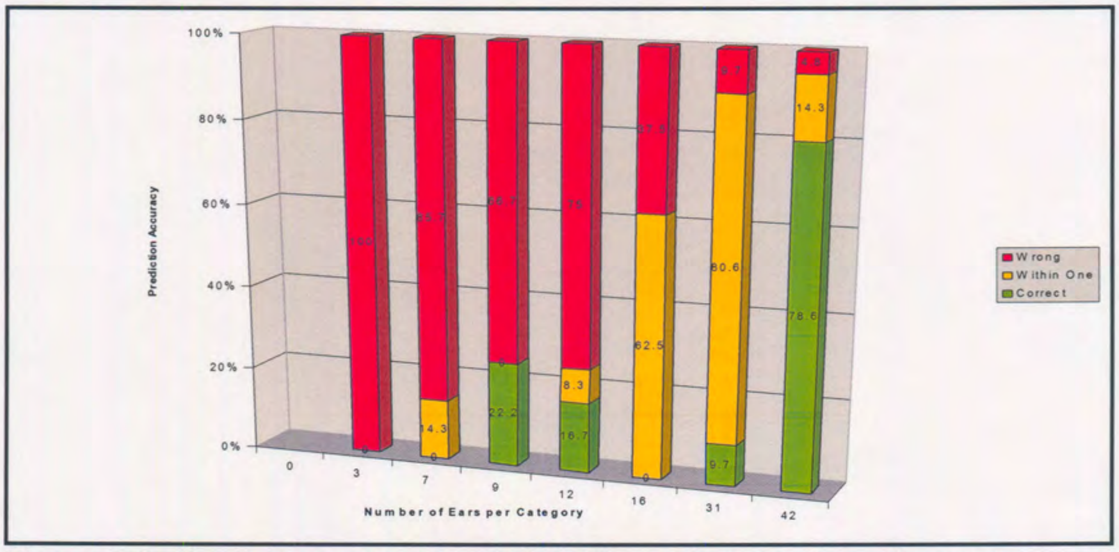


Figure 5.4: Prediction accuracy of 500 Hz against number of ears in every category.

The next section will discuss prediction accuracy of 500 Hz for the direct comparison between the present study (2000) and the previous study (1998).

5.6.2 Results of Present Study (2000) in Comparison to the Previous Study (1998) for 500 Hz.

Prediction accuracy was usually better for most categories for the present study (2000). Normal hearing defined as 0-20 dB HL was predicted better in the 2000 study (90% instead of 87%). Another improvement was the much lower incidence of false positive responses in the 2000 study. The false positive rate for the present 2000 study (normal ears predicted as hearing impaired) was 6% for the two categories combined instead of 20% for the 1998 study. False negative rates were slightly poorer for the 2000 study. Results for prediction accuracy for all categories and false positive and negative rates are summarized in Table 5.12.

Table 5.12: Comparison of previous study (1998) in black and present study (2000) in red: 500Hz predicted into 10dB categories.

Cate-gories	1 (0-10dB)		2 (11-20dB)		3 (21-30dB)		4 (31-40dB)		5 (41-50dB)		6 (51-60dB)		7 (61-70dB)		8 (71-80 dB)	
Correct	82%	84%	19%	31%	0%	0%	22%	28%	0%	0%	0%	0%	0%	0%	*	*
10dB out	15%	13%	50%	58%	75%	75%	11%	11%	20%	0%	0%	0%	0%	0%	*	*
Wrong	3%	3%	31%	11%	25%	25%	67%	61%	80%	100	100	100	100	100	*	*
0-10dB predicted as <20dB					97%						96%					
0-20dB predicted as 0-20dB					87%						90%					
False positive responses					**False negative responses**											
	12%	1%	8%	5%	1%	0%	3%	9%	2%	4%	3%	5%	0%	0%	*	*
# ears in category	60	60	26	26	4	4	18	18	5	5	6	6	1	1	*	*
**** There were no ears in category eight, largest hearing loss measured for 500 Hz was 65dB HL.**																

Prediction accuracy for the scenario five method of the 2000 study was generally slightly better for all categories expect category 2. Prediction of normal hearing as 0 – 20 dB HL (Jerger, 1980) was worse however, 87% instead of 92%. The false positive rates are better, for the 2000 study, 8% for the two categories combined instead of 18% for the 1998 study. False negative responses are slightly worse.

The results for prediction accuracy into the five categories of the scenario five method for both studies are summarized in Table 5.13.

Table 5.13: Comparison of previous study (1998) in black and present study (2000) in red: 500Hz predicted in the scenario five method into 5 categories.

Categories	**1 (0-10dB)**		**2 (11-20dB)**		**3 (21-35dB)**		**4 (36-50dB)**		**5 (51-65dB)**	
Correct	80%	83%	31%	27%	13%	33%	25%	25%	0%	15%
10 dB out	13%	10%	65%	54%	47%	27%	33%	8%	14%	14%
Wrong	7%	7%	4%	19%	40%	40%	42%	59%	86%	71%
0-10dB predicted as <20dB	93%					93%				
0-20dB predicted as 0-20dB	92%					87%				
False positive responses					**False negative responses**					
	11%	3%	7%	5%	3%	7%	3%	6%	0%	3%
# ears in category	60	60	26	26	15	15	12	12	7	7

The next section will present all the results obtained for experimentation with subject,- DPOAE- and ANN-variables that were investigated to determine their effect on prediction accuracy.

5.7 Subject-, DPOAE- and ANN-Variables Experimented with to Determine Optimal PTT Prediction Accuracy.

Results for each variable will be given separately.

5.7.1 The Effect of the Subject Variable AGE Presented to the Network in 5 year or 10 year Categories on PTT Prediction Accuracy.

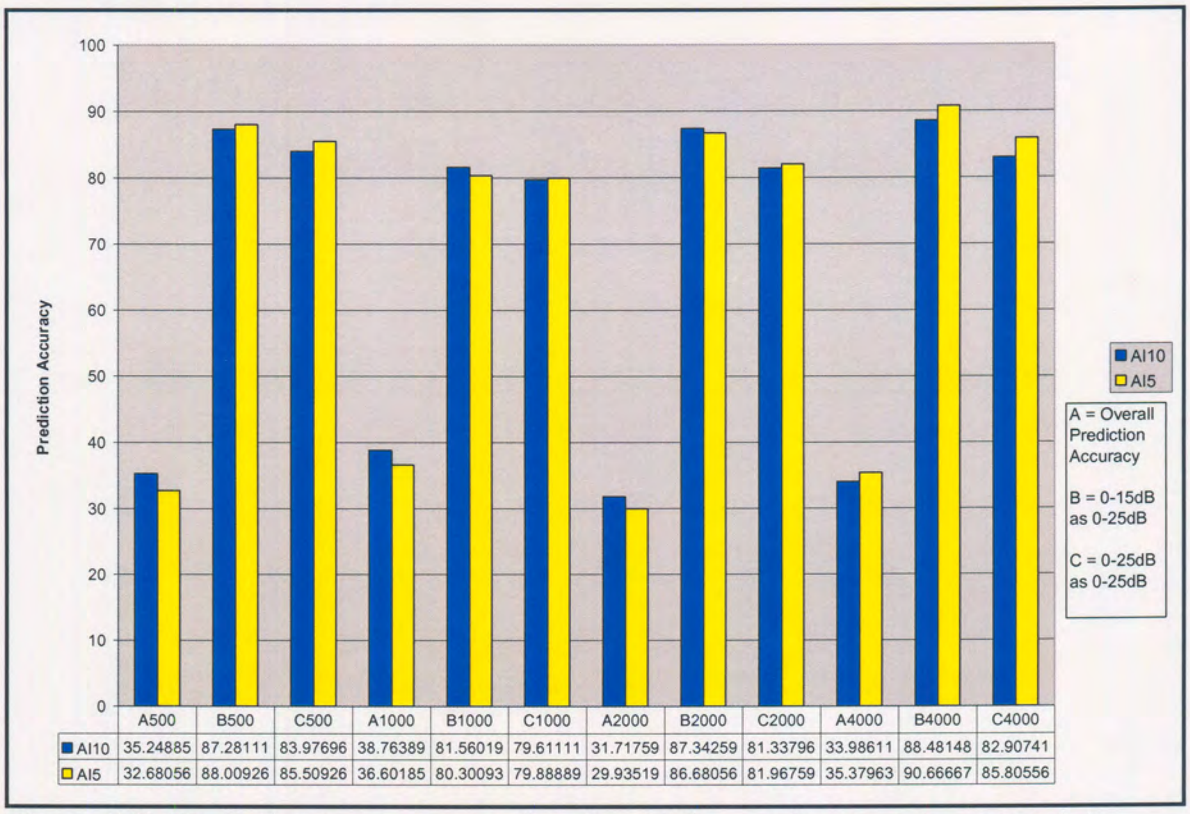
Subject age was always included in ANN training and prediction because it has been found to improve PTT prediction accuracy in a number of previous studies (Lonsbury-Martin et al. 1991; Kimberley et al. 1994a; Kimberley et al. 1994b; De Waal, 1998).

Different ways were used to present subject age to the network with the dummy variable technique, either with 10-year increments or with 5-year increments. This concept was described in 4.8.1.3 “Subject age” in the previous chapter.

To present the network with a subject’s age within 5 years might seem like a more accurate age presentation than the 10-year category method’s less specific presentation.

The 5 year increment method however, had a great increase on the number of input neurons, quantity of input data to deal with and therefore also the complexity of the topology of the network.

PTT prediction accuracy of the two methods is summarized in Figure 5.5



	A500	B500	C500	A1000	B1000	C1000	A2000	B2000	C2000	A4000	B4000	C4000
■ AI10	35.24885	87.28111	83.97696	38.76389	81.56019	79.61111	31.71759	87.34259	81.33796	33.98611	88.48148	82.90741
□ AI5	32.68056	88.00926	85.50926	36.60185	80.30093	79.88889	29.93519	86.68056	81.96759	35.37963	90.66667	85.80556

Figure 5.5: Prediction accuracy as a function of the age increment presented to the ANN input.

Differences in prediction accuracy of PTTs at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz of the two methods were always within 5%. It seems like the advantage gained by specifying age increments more specifically were lost in the creation of a more complex data set to train with. This finding will be discussed in detail in Chapter 6.

Overall prediction accuracy (the mean value for all eight categories) is depicted by "A", prediction of very good hearing (0 – 15 dB HL) as normal (0 – 25 dB HL) is depicted by "B" and separation of normal hearing as 0 – 25 dB HL as defined by Goodman (1965) is depicted by "C".

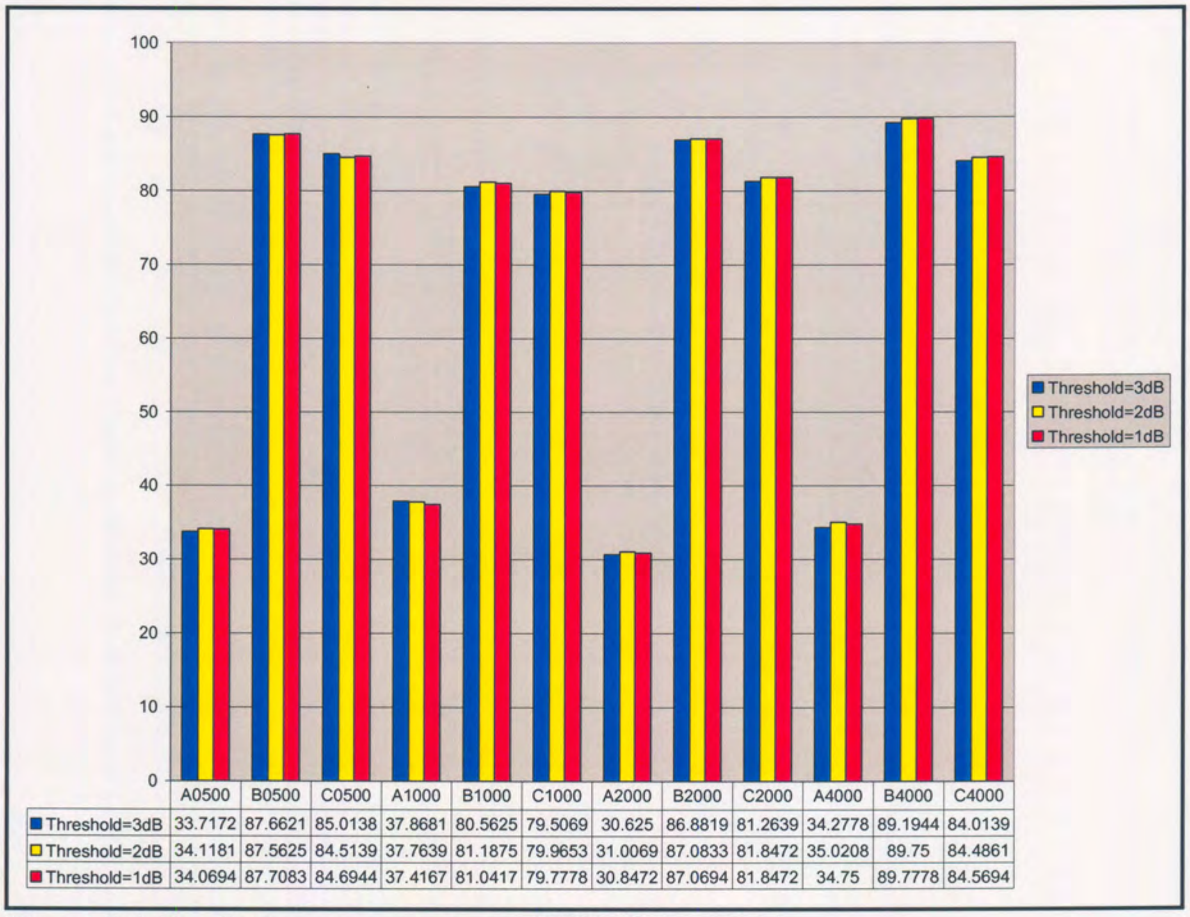
5.7.2 The Effect of DPOAE Threshold Defined as 1, 2 or 3 dB Above the Noise Floor on PTT Prediction Accuracy.

A distortion product with amplitude less than the noise floor cannot be detected (Kimberley & Nelson, 1990; Lonsbury-Martin et al. 1990). Most researchers specify a DP response to be present if the DP response is 3-5 dB above the noise floor. Harris and Probst (1991) and Krishnamurti (2000) specified a DP response as ≥ 5 dB above the level of the noise floor. Lonsbury-Martin (1994) set the criterion level for a DPOAE threshold at ≥ 3 dB.

The criteria for the presence of a DPOAE response are that the test status has to be "accepted" and a specified dB level above the noise floor. For this research project, one of two criteria had to be met before test status was "accepted": either the cumulative noise level is at least –18 dB SPL, or the DPOAE amplitude is 10dB above the noise floor. About half of the tests run (47%) had noise levels low enough to pass the first criterion of test acceptance based on cumulative noise levels of at least –18dB SPL. For all these tests, DPOAE threshold was experimented with as 1, 2 or 3 dB above the noise floor to determine differences in prediction accuracy.

All responses with a test status that was "noisy" or "timed out" were regarded as absent responses. (It should be noted that Kemp (1990) warned that in order to determine the threshold of a DPOAE response, one could not merely subtract the noise floor from the DPOAE amplitude in its decibel form. The two values should be converted back to its pressure value (watt/m^2), then subtracted.)

The PTT prediction accuracy for DPOAE threshold defined as 1, 2 and 3dB above the noise floor is summarized in Figure 5.6.



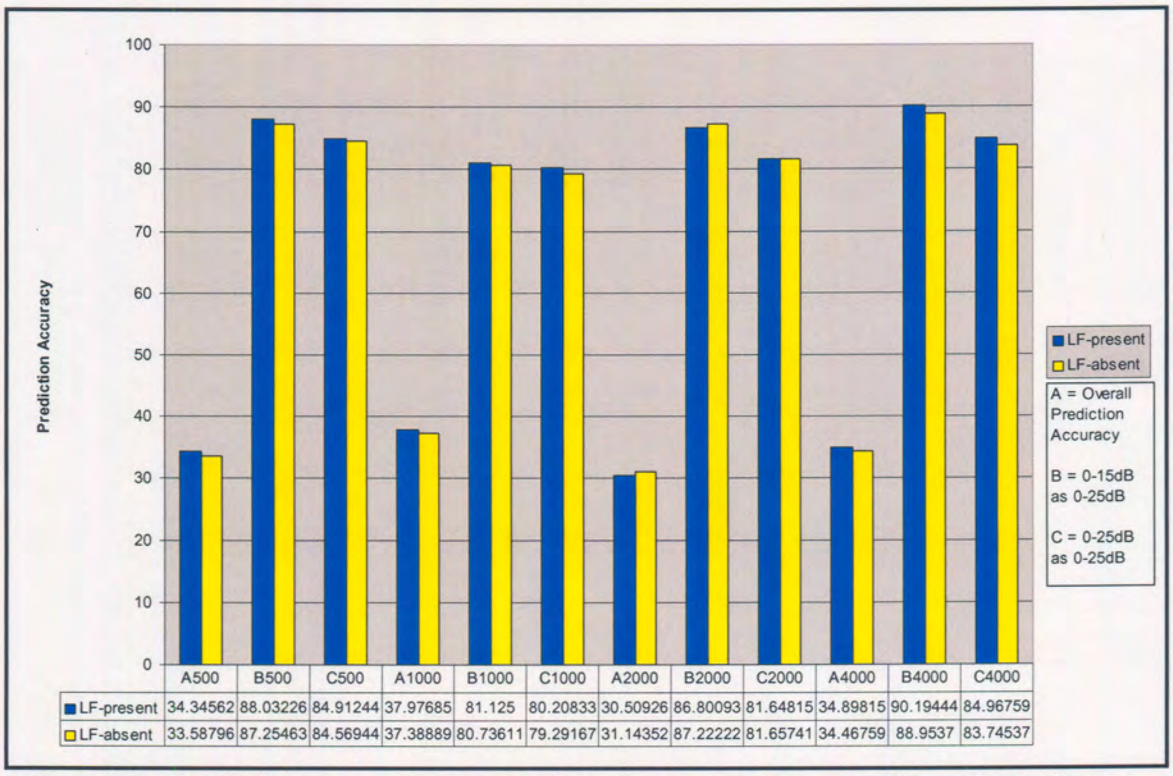
	A0500	B0500	C0500	A1000	B1000	C1000	A2000	B2000	C2000	A4000	B4000	C4000
■ Threshold=3dB	33.7172	87.6621	85.0138	37.8681	80.5625	79.5069	30.625	86.8819	81.2639	34.2778	89.1944	84.0139
□ Threshold=2dB	34.1181	87.5625	84.5139	37.7639	81.1875	79.9653	31.0069	87.0833	81.8472	35.0208	89.75	84.4861
■ Threshold=1dB	34.0694	87.7083	84.6944	37.4167	81.0417	79.7778	30.8472	87.0694	81.8472	34.75	89.7778	84.5694

Figure 5.6: Prediction accuracy versus DPOAE threshold.

Results for prediction accuracy for DPOAE thresholds defined as 1, 2 or 3dB above the noise floor revealed a very interesting and significant finding: There is virtually no difference in prediction accuracy for DPOAE threshold as 1, 2 or 3 dB above the noise floor. A DPOAE response only 1 dB above the noise floor, can predict normal and impaired hearing across all frequencies as well as a DPOAE threshold of 3 dB above the noise floor. The significance of this interesting find will be discussed in more detail in Chapter 6.

5.7.3 The Effect of the Emission or Inclusion of Low Frequency DPOAE Information for ANN Training on PTT Prediction Accuracy

For this research project, it was decided to experiment with the inclusion or omission of low frequency DPOAE input data to investigate if certain frequencies could be predicted more accurately by omitting noisy low frequency DPOAEs. For experiments where low frequency DPOAE data was omitted, f1= 500 Hz, 625 Hz and 781 Hz were omitted in the input data to train the neural network on. The summary for PTT prediction accuracy with DPOAEs and ANNs where noisy low frequency emissions were omitted or present can be seen in Figure 5.7.



	A500	B500	C500	A1000	B1000	C1000	A2000	B2000	C2000	A4000	B4000	C4000
■ LF-present	34.34562	88.03226	84.91244	37.97685	81.125	80.20833	30.50926	86.80093	81.64815	34.89815	90.19444	84.96759
□ LF-absent	33.58796	87.25463	84.56944	37.38889	80.73611	79.29167	31.14352	87.22222	81.65741	34.46759	88.9537	83.74537

Figure 5.7: Prediction accuracy versus the presence or absence of low Frequency DPOAEs.

A comparison of results of PTT prediction accuracy for low frequencies present or absent revealed no significant difference. Differences in prediction accuracy were always within 2%. It seems like the "noisy" low frequency DPOAEs had virtually no effect on the training or prediction capabilities of the neural network, which confirms the viewpoint of Blum (1992) that neural networks excel in dealing with noisy incomplete data. This finding will be discussed in more detail in the following chapter.

5.7.4 The Effect of DPOAE Amplitude Presentation to the Neural Network on PTT Prediction Accuracy

For this research project, the amplitude of the DPOAE was presented to the network in four different ways:

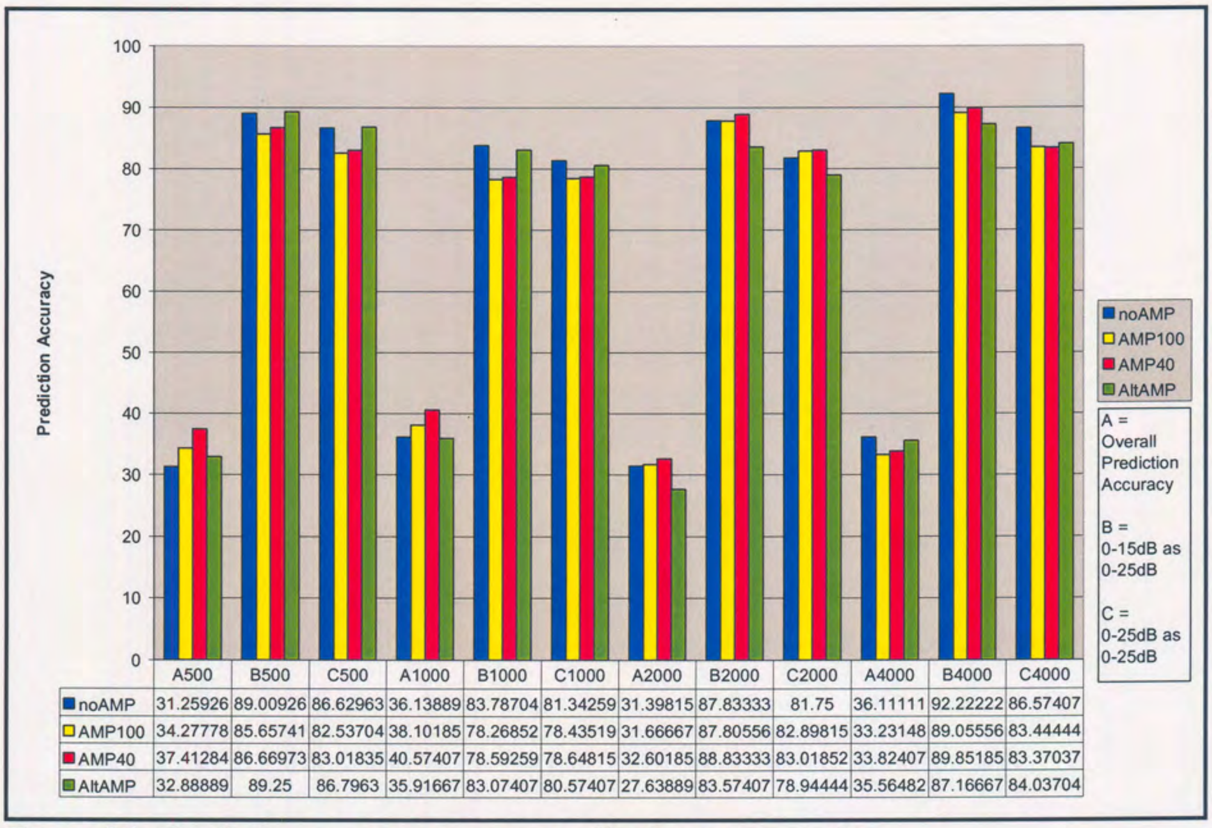
- AMP 100 presented the amplitude of the DPOAE to the ANN as a fraction of 100 (see 4.8.1.4.1).
- AMP 40 presented the amplitude of the DPOAE to the ANN as a fraction of the largest DPOAE response measured in this study (39dB)(see 4.8.1.4.2).
- ALT AMP depicted DPOAE amplitude with the dummy variable technique into one of four possible 10 dB categories (see 4.8.1.4.3).
- No AMP: This method left out amplitude information (see 4.8.1.4.4).

As was described in 4.8.1.4, each method of amplitude representation influenced the ANN in an unique way. AMP 100's neural network had trouble converging (reaching optimal error tolerance levels during training to begin prediction). AMP 40's ANN had the exact same topology than the AMP 100 method but convergence was much faster because the input values were larger (a fraction of 40 instead of a fraction of 100) and the network therefore found it easier to make midway representations to reach error tolerance levels faster. The ALT AMP technique had the one advantage that information was presented to the network in the same fashion than the output predictions, which was by depicting information in categories with the dummy variable technique. Input mode and output mode for that neural network was therefore the same. A Disadvantage of the ALT AMP method however, was that the complexity of the topology of the network increased drastically due to the fact that 352 input neurons were needed to present amplitude in this fashion instead of the usual 88.

The way amplitude was presented to the neural network definitely had an effect on prediction accuracy. Certain patterns are recognizable and seem to depend on the frequency to be predicted, and the decibel range to be predicted (prediction of normal hearing versus overall prediction accuracy across all categories).

The two low frequencies (500 Hz and 1000 Hz) demonstrated the same pattern for prediction accuracy of normal and impaired hearing based on amplitude representation. For overall prediction accuracy across all categories, the AMP 40 method revealed most accurate predictions. For the prediction of normal hearing at low frequencies, the No AMP method where low frequency data was omitted as well as the AMP 40 method provided some of the best results.

For the prediction of the two high frequencies, each frequency demonstrated its own pattern. At 4000 Hz, the No AMP method revealed most accurate predictions for both the separation of normal hearing and prediction accuracy across all categories. For the prediction of 2000 Hz, the AMP 40 method revealed best prediction accuracy for both the identification of normal hearing and prediction accuracy across all categories and ALT AMP always had the poorest prediction accuracy. All these tendencies are summarized in Figure 5.8.



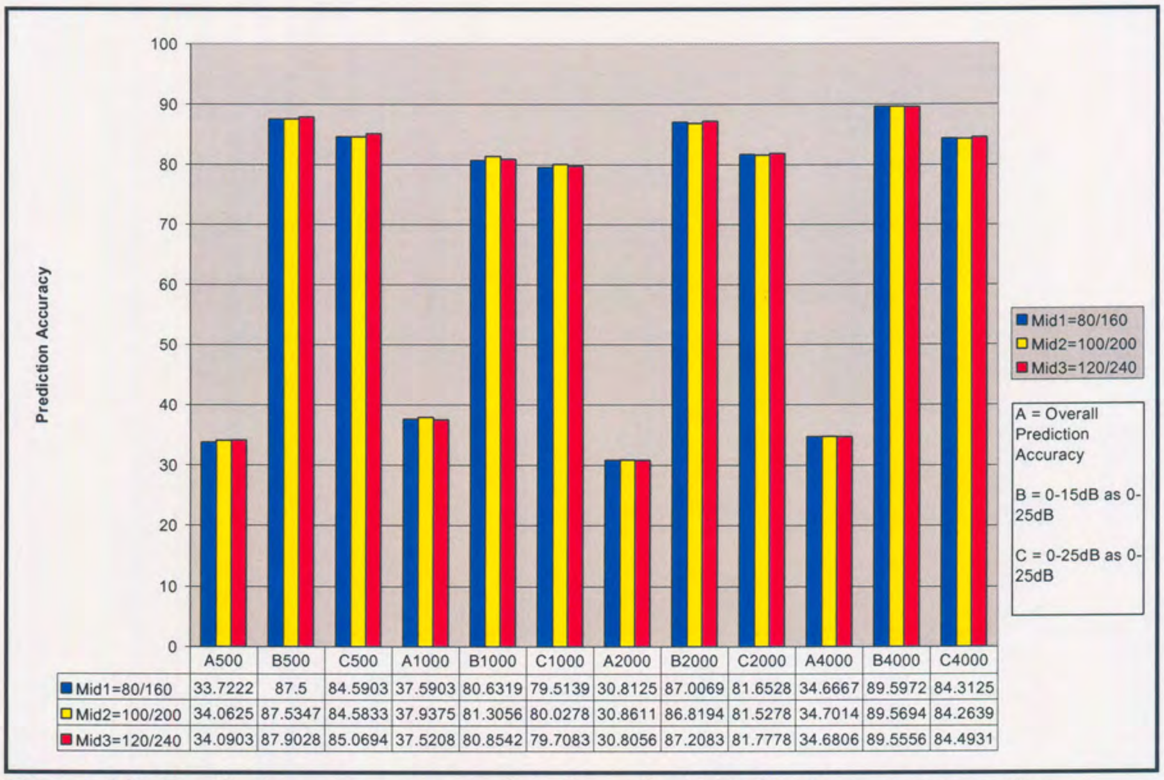
	A500	B500	C500	A1000	B1000	C1000	A2000	B2000	C2000	A4000	B4000	C4000
■ noAMP	31.25926	89.00926	86.62963	36.13889	83.78704	81.34259	31.39815	87.83333	81.75	36.11111	92.22222	86.57407
□ AMP100	34.27778	85.65741	82.53704	38.10185	78.26852	78.43519	31.66667	87.80556	82.89815	33.23148	89.05556	83.44444
■ AMP40	37.41284	86.66973	83.01835	40.57407	78.59259	78.64815	32.60185	88.83333	83.01852	33.82407	89.85185	83.37037
■ AltAMP	32.88889	89.25	86.7963	35.91667	83.07407	80.57407	27.63889	83.57407	78.94444	35.56482	87.16667	84.03704

Figure 5.8: Prediction accuracy versus amplitude representation.

5.7.5 The Effect of the ANN Variable MIDDLE NEURON COUNT on PTT Prediction Accuracy

For this research project, there was experimented with three different middle neuron count possibilities. All experiments were run with 80, 100 or 120 middle level neurons, except ALT AMP experiments that were run with double middle level neuron quantities to compensate for the complexity of topology due to large numbers of input nodes (this was described in 4.8.3.4).

Fig 5.9 summarizes the results obtained for prediction accuracy for different middle level neuron quantities.



	A500	B500	C500	A1000	B1000	C1000	A2000	B2000	C2000	A4000	B4000	C4000
■ Mid1=80/160	33.7222	87.5	84.5903	37.5903	80.6319	79.5139	30.8125	87.0069	81.6528	34.6667	89.5972	84.3125
□ Mid2=100/200	34.0625	87.5347	84.5833	37.9375	81.3056	80.0278	30.8611	86.8194	81.5278	34.7014	89.5694	84.2639
■ Mid3=120/240	34.0903	87.9028	85.0694	37.5208	80.8542	79.7083	30.8056	87.2083	81.7778	34.6806	89.5556	84.4931

Figure 5.9: Prediction accuracy as a function of middle level neuron quantity.

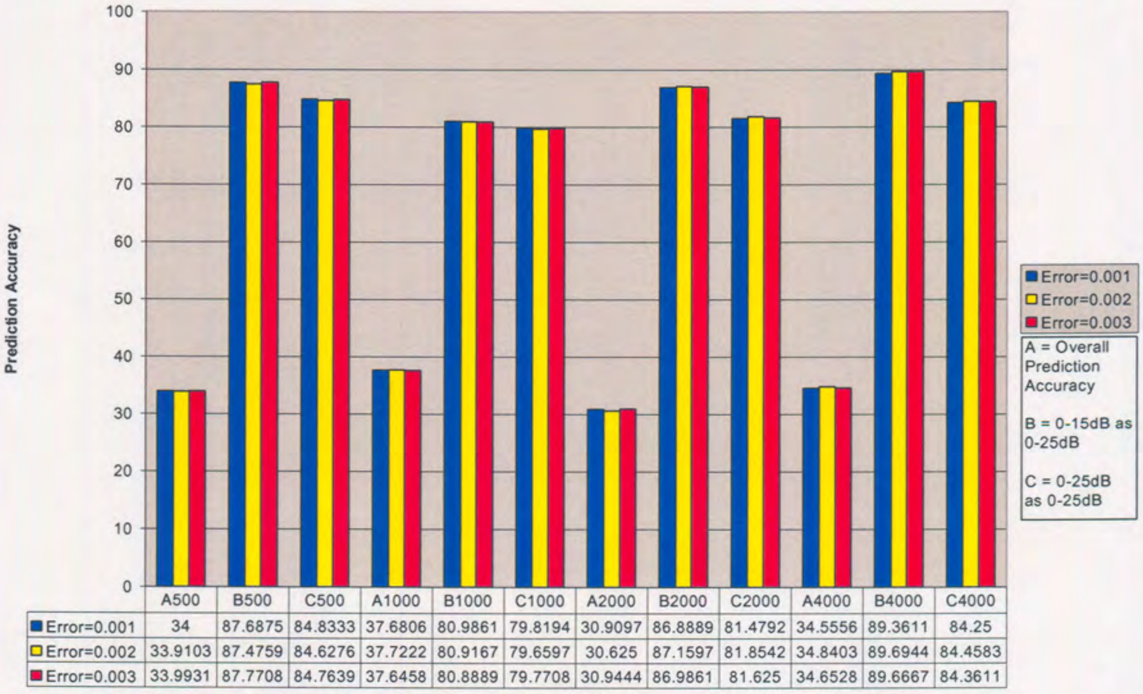
There was virtually no effect of middle neuron count on prediction accuracy for normal or impaired hearing. Prediction accuracy for the three categories was always within 1%.

"Mid 1" represent the lowest category of middle neurons (80 for all experiments except ALT AMP that had 160), "Mid 2" represent the middle category of middle neurons (100 for all experiments except ALT AMP that had 200) and "Mid 3" represents the category with the highest number of middle level neurons (120 for all experiments except ALT AMP that had 240).

5.7.6 The Effect of the ANN Variable TRAINING ERROR SENSITIVITY on PTT Prediction Accuracy

Another aspect that was experimented with, was the training error sensitivity. This refers to the permitted accuracy with which the network learns and predicts and this concept was described in detail in 3.5.4.2 "The error tolerance".

For this study, error tolerance levels of 0.001 (within 0.1% accurate), 0.002 (within 0.2% accurate) and 0.003 (within 0.3% accurate) were experimented with. Results are summarized in Figure 5.10.



	A500	B500	C500	A1000	B1000	C1000	A2000	B2000	C2000	A4000	B4000	C4000
■ Error=0.001	34	87.6875	84.8333	37.6806	80.9861	79.8194	30.9097	86.8889	81.4792	34.5556	89.3611	84.25
□ Error=0.002	33.9103	87.4759	84.6276	37.7222	80.9167	79.6597	30.625	87.1597	81.8542	34.8403	89.6944	84.4583
■ Error=0.003	33.9931	87.7708	84.7639	37.6458	80.8889	79.7708	30.9444	86.9861	81.625	34.6528	89.6667	84.3611

Figure 5.10: Prediction accuracy versus error tolerance levels.

Results in prediction accuracy revealed that there was no significant difference for these three error tolerance levels.

5.8 Summary

Results of this study revealed that the prediction accuracy of normal hearing (defined as 25 dB HL according to Goodman, 1965) was 94% at 500 Hz, 88% at 1000 Hz, 88% at 2000 Hz and 93% at 4000 Hz. Results of this study in comparison to other studies can be seen in Table 5.14 and are indicated in red.

Table 5.14: A comparison of studies to present study: Prediction accuracy of normal hearing with DPOAEs.

	Kimberley et al. 1994a	Kimberley et al. 1994b DP alone	Kimberley et al. 1994b DP + Age	Moulin et al. 1994	De Waal 1998	DeWaal 2000 (present study)
500 Hz	*	*	*	*	92%	94%
706 Hz	*	*	*	52.9%	*	*
1000 Hz	*	*	*	73.2%	87%	88%
1025 Hz	92%	90%	90%	*	*	*
1413 Hz	*	*	*	75.6%	*	*
1464 Hz	88%	86%	87%	*	*	*
2000 Hz	*	*	*	*	84%	88%
2050 Hz	83%	84%	83%	81.5%	*	*
2826 Hz	*	*	*	*	*	*
2880 Hz	70%	80%	83%	*	*	*
4000 Hz	*	*	*	79.4%	91%	93%
4052 Hz	69%	88%	88%		*	*
5712 Hz	76%	80%	86%		*	*

*** Frequency not predicted in the research project**

Prediction accuracy of impaired hearing was less satisfactory and it seems that the number of ears in every category had a greater effect on prediction accuracy than the limitations of ANNs or lack of correlation between DPOAEs and PTTs.

The few experiments that were run to make a direct comparison between the 1998 and 2000 study possible revealed that there were only minor differences found in prediction accuracy at 4000 Hz and 1000 Hz. Better prediction of normal hearing (defined as 0 – 20 dB HL according to Jerger, 1980) in the 2000 study was found at 2000 Hz (predicted with 88% prediction accuracy in the 2000 study opposed to 82% in 1998) and also at 500 Hz (in 2000 predicted with 90% accuracy opposed to 87% in 1998). Better false positive values were obtained in the 2000 study at all four frequencies.

Results for the investigation of the effect of subject-, DPOAE- and ANN-variables revealed very little change in prediction accuracy as a result of the presentation of the age increment (always within 5%), the inclusion of omission of low frequency DPOAE data (always within 2%), middle neuron quantities (always within 1%) and training error sensitivity (always within 1%). The fact that DPOAE threshold did not have a significant effect on prediction accuracy, (always within 1%) is a very significant find that could serve as basis for the argument that DPOAE thresholds may be lowered and defined closer to the noise floor. Lastly, amplitude representation to the network had a more clear effect on prediction accuracy and is dependant on the frequency to be predicted and whether it is a prediction of normal hearing, or overall prediction of all categories. The next chapter will discuss all these findings in more detail and interpret the significance thereof.

Chapter 6: Discussion and Interpretation of Results

6.1 Introduction

The development of objective procedures in audiology came a long way since the 1920s. With the aid of modern technology, audiologists can measure the exact degree, configuration, and site of hearing loss in adults and confirm these findings with a series of objective electrophysiologic and physiologic procedures, such as tympanometry, the acoustic reflex, ABR, and otoacoustic emissions (Northern, 1991). From the overview of the development of objective procedures in audiology in Chapter 2, however, it is evident that there are some weaknesses in current objective diagnostic procedures when it comes to the evaluation of special populations. In the evaluation of neonates from birth to 6 months, the crucially ill, and malingerers, audiologists often have to rely heavily on objective electrophysiologic procedures to determine hearing ability. To determine hearing thresholds with electrophysiologic procedures is often costly, requires a large amount of time as well as highly trained and specialized personnel, and may require sedation. Above all, current objective physiologic procedures, such as ABR, have a limited frequency area in which hearing ability can be determined accurately. There is therefore a definite need for an objective, reliable, rapid, and economic test of hearing that evaluates hearing ability across a range of frequencies to aid in the assessment of difficult-to-test populations.

The distortion product otoacoustic emission has been intensely investigated as a possible new test of hearing (Probst & Hauser, 1990; Gorga et al. 1993; Moulin et al. 1994; Kimberley et al. 1994a; Kimberley et al. 1994b; Stover et al. 1996a). Many studies found a strong correlation between DPOAEs and PTTs despite the fact that results obtained from conventional pure tone audiometry involves an evaluation of the entire auditory system opposed to DPOAE results that involve only an evaluation of cochlear functioning (Gaskill & Brown, 1990; Lee et al. 1993; Vinck et al. 1996; Kummer et al. 1998). Some studies attempted to predict normal hearing with DPOAEs and conventional statistical methods (Kimberley et al. 1994b; Moulin et al. 1994) and artificial neural networks (Kimberley et al. 1994a; De Waal, 1998). Researchers however, experienced difficulty to predict hearing status at low frequencies due to problems with the rising noise floor and the inability of conventional statistics to deal with nonlinear and noisy data sets (Durrant, 1992; Gorga et al. 1996; Gorga et al. 1993; Kimberley et al. 1994b; Stover et al. 1996). This could possibly be attributed to the fact that most studies attempted to correlate normal functioning of a single area on the basilar membrane, such as the f2 frequency place (Durrant, 1992; Harris et al. 1989; Kimberley et al. 1994b), the GM frequency place (Martin et al. 1990b; Lonsbury-Martin & Martin, 1990; Bonfils et al. 1991) or the 2f1-f2 place (Smurzynski et al. 1990) with the corresponding pure tone frequency and used the threshold or amplitude of a single DPOAE to correlate with a pure tone threshold.

The current study investigated the correlation between DPOAEs and PTTs with artificial neural networks, a data processing technique proven superior to conventional statistics when it comes to dealing with noisy non-linear data sets (Kimberley et al. 1994a; Raghupathi et al. 1993; Hann & Steurer, 1996). This study also used the presence or absence and amplitudes of all 11 DPOAE responses to predict hearing ability at a single PTT frequency. The fact that pure tone thresholds are interrelated made it possible to predict hearing status at low frequencies objectively and accurately with this method, for the first time (De Waal, 1998). The present study once again showed that the distortion product otoacoustic emission has enough data at surrounding frequencies to make objective predictions at problematic frequencies such as 500 Hz with a high level of accuracy.

The aim of this chapter is to discuss all findings and interpret the significance thereof. Firstly, the correlation found between DPOAEs and PTTs will be discussed, secondly, the results obtained for the prediction accuracy at the four frequencies 500, 1000, 2000 and 4000 Hz. Thirdly, the significance of results in terms of readiness for broad clinical use will be discussed, with emphasis on the requirements for sensitivity and specificity of screening and diagnostic procedures. Then, to integrate findings of the four frequencies, a few case studies will be discussed where the network made accurate predictions for all frequencies, and also case studies predicted inaccurately with possible reasons for inaccurate predictions. Lastly, all the DPOAE and ANN variables that were experimented with to determine their effect on prediction accuracy will be discussed and interpreted.

6.2 Indication of a Correlation between DPOAE Measurements and Pure Tone Thresholds

Many other studies used statistical techniques to determine the correlation between DPOAEs and pure tone thresholds or described case studies that demonstrated a close relationship between DPOAEs and pure tone thresholds (Gaskill & Brown, 1990; Gorga et al. 1993; Kummer et al. 1998; Lee et al. 1993; Vinck et al. 1996). In the case of statistical methods, a correlation coefficient can be determined and that serves as an indication of the correlation found and its significance. In this study, however, artificial neural networks were used to predict pure tone thresholds. The network extracts necessary information from input stimuli and then forms an internal representation of relations between different data sets by adjustment of the weights of the middle neurons. The neural network then uses the learned representations to make predictions. One of the limitations of a neural network is that it cannot justify the learned relationships and specify exact correlations in terms of strength or significance. By analyzing the accuracy of the predictions, one can make assumptions about the correlation between DPOAEs and pure tone thresholds, but one cannot dissect a neural network to find precise reasons and relationships for accurate predictions. With this aspect in mind, the implied correlation between DPOAEs and pure tone thresholds will be discussed briefly.

If there were no correlation between DPOAEs and pure tone thresholds (PTTs), then the neural network would not have been able to make accurate predictions of hearing ability with more than 50% accuracy. Correct predictions of hearing ability would have been mere chance or at random. If a histogram was drawn to illustrate the prediction accuracy of a data set that is not correlated with another data set at all, one can expect to see an equal number of predictions in every one of the domain values. This would result in a “flat” histogram or a histogram representing random predictions at the various domain values. It would definitely not result in a histogram depicting a normal curve distribution.

The prediction accuracy of this neural network is illustrated in the histograms in Figure 6.1.

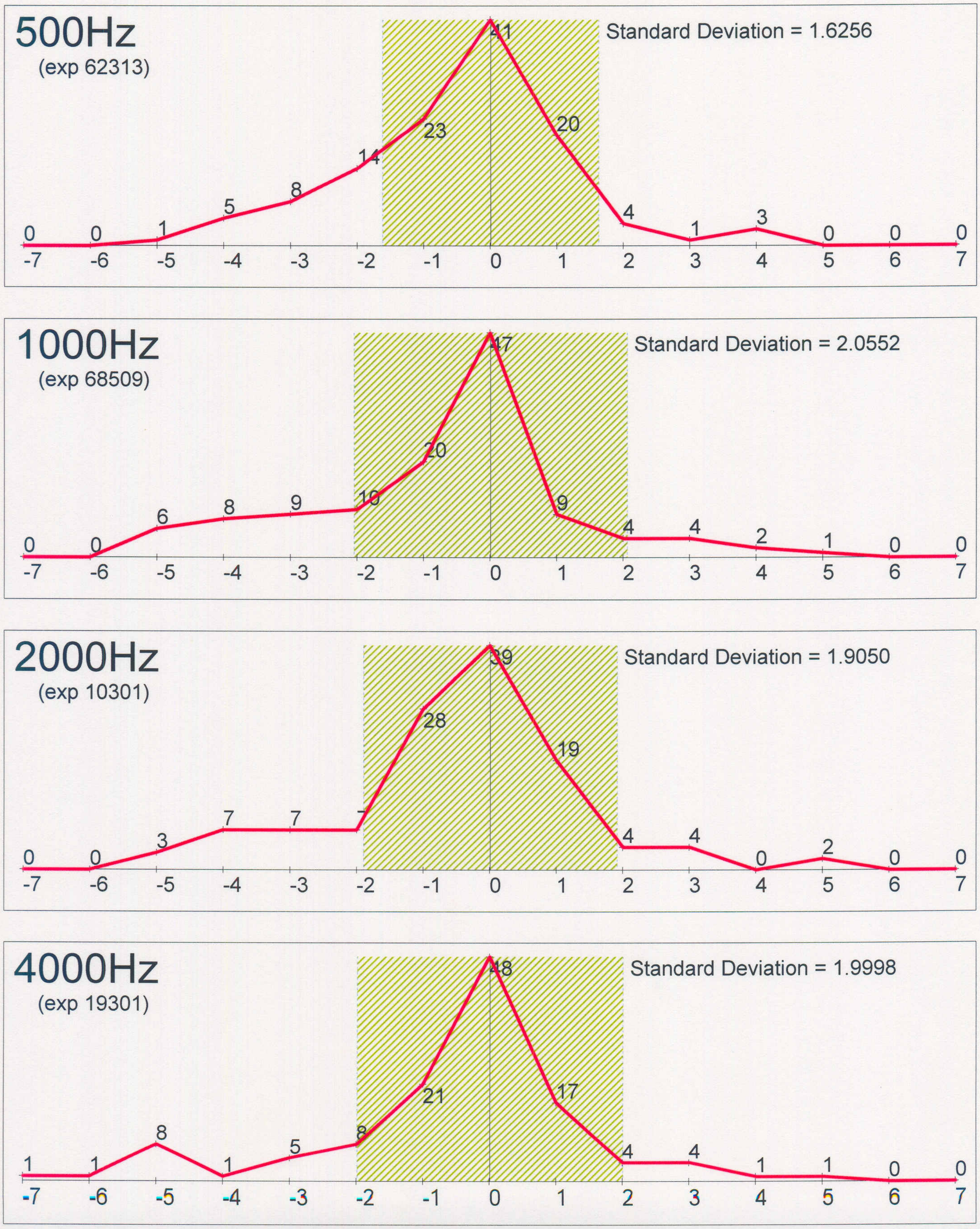


Figure 6.1: Prediction accuracy of PTTs with DPOAEs and ANNs

At first glance the presence of a normal curve distribution can be seen in all these histograms in Figure 6.1. Most ears were predicted accurately within the same class (these ears are indicated at the zero (0) place on the histogram) or within one category of hearing impairment. This is clearly an indication that the neural network found a correlation between DPOAEs and PTTs and used that correlation to make the predictions.

This study therefore confirms the results of many other researchers that the distortion product is strongly correlated with pure tone thresholds in normal hearing and hearing-impaired ears (Gaskill & Brown, 1990; Gorga et al. 1993; Kimberley et al. 1994; Kummer et al. 1998; Lee et al. 1993; Moulin et al. 1994; Vinck et al. 1996).

6.3 Prediction of 500 Hz

The prediction of 500 Hz with distortion product otoacoustic emissions has been problematic for many authors (Gorga et al. 1993; Moulin et al. 1994; Probst & Hauser, 1990; Stover et al. 1996a). Regardless of the loudness level of the primaries, the chosen frequency ratios, loudness level ratios or any other variables that could influence the study, the rising noise floor below 1000 Hz limited the measurement of clear responses at f2 = 500 Hz (Durrant, 1992). Probst and Hauser (1990) attempted to predict hearing ability as normal or impaired in the geometric mean frequency range of 500 Hz to 8000 Hz. Their findings indicated that the majority of normal and near-normal ears had no or small DPOAE amplitudes at 500 Hz and 8000 Hz. No correlations with hearing threshold could be established at these two frequencies. In a study by Stover et al. (1996a), the noise floor for lower frequencies (500 Hz and 707 Hz) was so high, that data at these two frequencies were interpreted as unknown or absent and coded as missing for data analysis. At 500 Hz, the prediction of normal hearing was no better than chance. Gorga et al. (1993) and Moulin et al. (1994) experienced similar problems with very high noise measurements at low frequencies and could also not predict normal hearing at 500 Hz.

The prediction accuracy for hearing ability at 500 Hz for this study yielded promising and interesting results. Normal hearing (0-25dB HL) at 500 Hz was predicted as normal 94% of the time. The improvement in prediction ability of low frequencies in this study can be attributed to two reasons. The first reason includes the different data processing procedure, an artificial neural network with excellent correlation finding and prediction capabilities. The second reason is that pure tone threshold at 500 Hz was not predicted with DPOAE amplitude or threshold at low frequencies, but with a pattern of all present and absent DPOAE responses with their amplitudes across all 11 DPOAE frequencies and all eight DP Grams. Previous studies attempted to correlate the pure tone threshold with either the threshold (or amplitude) of the f2 frequency (Harris et al. 1989; Kimberley et al. 1994), the geometric mean frequency (Bonfils et al. 1991; Lonsbury-Martin & Martin, 1990) or the distortion product frequency (Smurzynski et al. 1990). This study did not use a single point DPOAE measurement to predict a single point pure tone threshold, but used the whole spectrum of emissions to predict a single pure tone threshold. The artificial neural network was able to gain enough information from the whole spectrum of absent and present responses to predict normal hearing ability at 500 Hz correctly 94% of the time.

Another aspect of the prediction of hearing ability at 500 Hz that should be investigated, is the number of false positive and false negative predictions the neural network made. Category one (0 + 5dB) had 0% false positive responses, in other words, subjects with normal hearing were never predicted as having a hearing loss. Category two (10 + 15dB) had 2% false positive responses. Category three (20 + 25dB) had 1% false positive responses. If the three categories are combined to represent normal hearing (0-25dB), the false positive rate is only 3%. This procedure was therefore very specific and almost never predicted a subject with normal hearing as hearing impaired at 500 Hz.

The sensitivity of the procedure (the percentage of hearing impaired ears predicted as normal) is influenced by the number of false negative responses. The false negative rates at 500 Hz for category four (30 + 35dB) was 6%, category five (40 + 45dB) was 5%, category six (50 + 55dB) was 4% and category seven (60 + 65dB) was 0%. The percentage of ears predicted as normal decreased as the category's hearing loss increased. This finding supports results from the study of Probst and Hauser (1990) in which they found that mild hearing losses are often predicted as normal.

The total false negative rate for the combined categories spanning hearing loss was 15%. This value is much lower than reported elsewhere: no study predicted normal hearing at a frequency as low as 500 Hz and reported that false negatives were so high at low frequencies that performance was no better than chance (Gorga et al. 1993; Moulin et al. 1994; Probst & Hauser, 1990; Stover et al. 1996a). Unfortunately, this value might still be considered too high and raises questions regarding the sensitivity of this procedure for screening purposes. It could also be argued that category four (30 + 35dB) represents a mild hearing loss, and that the sensitivity of this procedure to identify moderate and severe hearing losses could be determined by the false negative rates for categories five to seven, which would bring the number down to 9%. Even though this procedure does not perform perfect classification of normal and impaired hearing, the results obtained in this study at 500 Hz clearly shows that DPOAEs and ANNs can predict hearing status in populations with varying degrees of hearing loss objectively with high degrees of accuracy, sensitivity and specificity. (The false positive and false negative rates for all four frequencies in terms of suitability for use in screening and diagnostic audiology tests are further discussed in 6.8 "Results of the present study in perspective: Readiness for broad clinical use".)

Bonfils et al. (1991) investigated objective low-frequency audiometry by distortion product otoacoustic emissions and found that active emissions could be measured as low as 2f1-f2 = 512 Hz. The authors concluded that DPOAEs could be used as an objective low-frequency test of auditory functioning. This study confirms the results of Bonfils et al. (1991). **DPOAEs and ANNs can accurately categorize hearing ability at 500 Hz as normal, 94% of the time.**

It was also attempted to predict impaired hearing at 500 Hz into 10 dB categories. The prediction of impaired hearing at 500 Hz, but also at all the other frequencies (1000 Hz, 2000 Hz and 4000 Hz) was very unsatisfactory and prediction percentages for categories spanning hearing loss indicated poor prediction capabilities of the neural network. Many possible reasons come to mind for this incapability of hearing impaired prediction. The reasons apply to all the frequencies and these reasons will be discussed in the next section.

6.4 Possible Reasons for Poor Prediction of Categories Representing Hearing Loss.

The first reason that can be suggested for the poor prediction of hearing impaired categories is that hearing impaired subjects might demonstrate less clear DPOAE responses that might influence the correlation between DPOAEs and PTTs. This possibility was seen in the DPOAE evaluation where it took the DPOAE equipment longer to test a hearing impaired subject in order to get enough frames of data to regard a test as "accepted" (the criterion for a test to be accepted for the GSI-60 DPOAE system for a screening test is that the DPOAE amplitude had to be 10dB above the noise floor or the cumulative noise level had to be –18dB SPL. The maximum number of frames tested in a screening procedure is 400, and if no clear response is measured in that time, the test is scored "timed out" which means that no response was obtained. This was described in "4.6.5.2 Determination of optimal stimulus parameters"). It is possible that more responses could be obtained from hearing-impaired subjects if the criterion for test acceptance is lowered to 5 dB for example. The lowered criterion for the acceptance of a test as "accepted" could possibly enhance the number of useable responses from hearing-impaired subjects and might therefore enhance prediction accuracy of categories spanning impaired hearing. This aspect should be further investigated.

Another reason for poor prediction accuracy of hearing impaired categories might be that the optimal procedure for data analysis has not yet been identified. I can be hypothesized that another type of neural network with different topologies, learning rules and error tolerances would be able to make more accurate predictions. Or could a complete new form of data processing, such as "genetic programming", inspired by

Darwinian invention and problem solving that "progressively breeds a population of computer programs over a series of generations" (Koza, Bennett, Andre & Keane, 1999:3) to find an optimal solution to a problem, be able to make more accurate predictions? These possibilities have yet to be investigated.

One definite aspect however, that seems to have influenced the prediction accuracy of hearing ability in categories spanning hearing impairment more than neural network capabilities and the underlying correlation between the two data sets, was the number of ears in every category that the neural network had to train on.

Neural networks need enough examples in every category to form representations of how a specific ear's DPOAE type relate to its PTT type in order to make accurate predictions. Even though it was attempted to categorize all audiograms in this study into three groups to ensure that hearing impairment was as well represented as normal hearing, the nature of sensorineural hearing impairment tends to affect certain frequencies more than others, and low frequencies are often normal. In the case of 500 Hz, this leads to the uneven distribution of many ears in categories representing normal hearing and few ears in hearing loss categories. At 4000 Hz however, category six (50 + 55dB) had more ears (a total of 20) and was predicted accurately 45% of the time and within 10dB 30% of the time. Hearing loss at 4000Hz in this category was wrongly predicted only 25% of the time with a false negative rate of only 2%. The same category at 500 Hz had only 7 ears to train on and prediction accuracy was never correct and within 10dB only 14% of the time. If the network had more ears in every category to train on, prediction accuracy might have been considerably better.

To illustrate this concept, figure 6.2 plots the number of ears in every category against prediction accuracy. It clearly demonstrates that the number of ears in every category had an enormous effect on prediction accuracy.

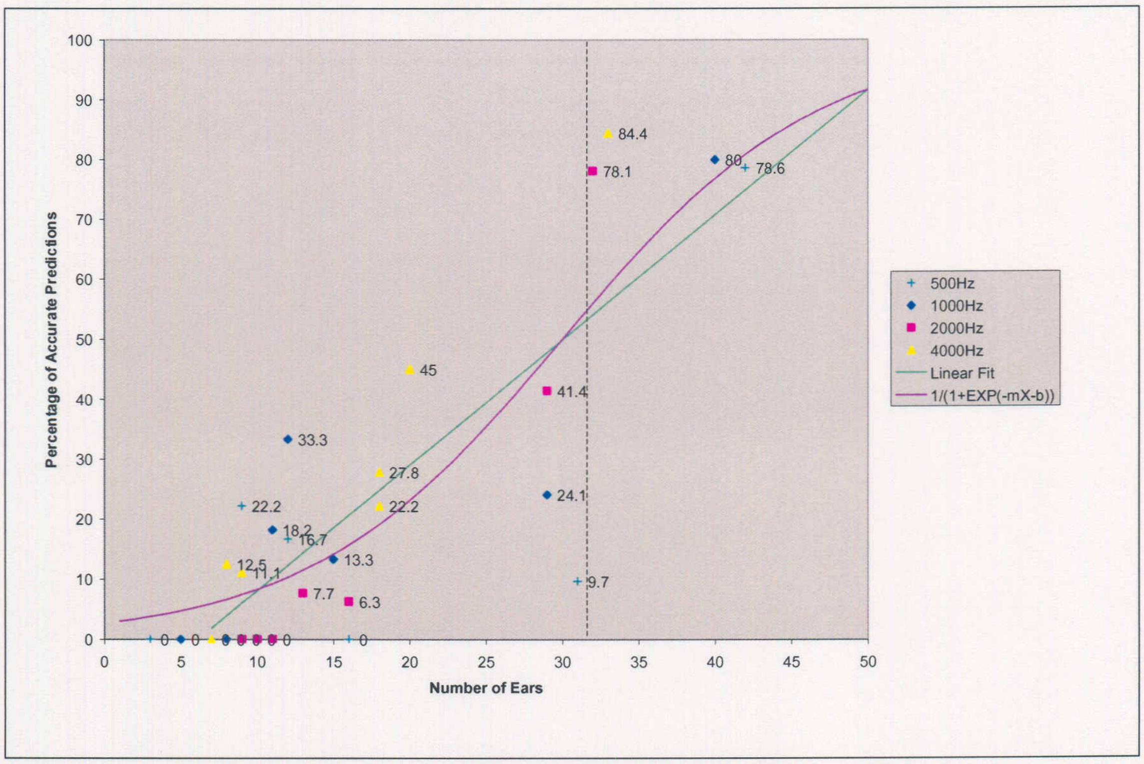


Figure 6.2: Prediction accuracy and ear count correlation.

The aim of this study is not to describe the relationship between the number of ears per category and prediction accuracy thereof. However, Figure 6.2 shows a number of possible relationships merely to illustrate the notion of “more-is-better”. The linear fit (in green) shows that higher number of ears lies significantly above expectation. It is worthy to note that the relationship cannot be linear, since any line with a slope larger than zero will have to cross the 100% accuracy limit at some point. This is of course not possible.

The figure indicates an alternative fit (in purple) of the form 1 / (1 + e^{-mx-b}), that seems better and more intuitive since it starts out low, just like the experimental data, but also asymptotically approach 100% for large numbers of ears. It is also not established if this function has any correlation with the sigmoid function (Blum, 1992:39) (1 / (1 + e^{-x}) that was used to normalize output of the separate ANN layers but it seems to be a more likely option than a linear fit.

For this data set there is a fairly clear threshold at 32 ears per category where prediction accuracy suddenly surges into the 75% and higher region.

Should any of the example relationships hold, it is expected that a 95% or higher accurate predictions could potentially be made if an ANN receives 80 or so ears per category.

The shortage of data in certain categories is probably the main reason for this study's poor prediction capabilities for ears demonstrating hearing impairment and not poor correlation between DPOAEs and PTTs of hearing-impaired ears or incapability of the neural network to deal with this data set.

6.5 Prediction of 1000 Hz

Researchers attempting to predict normal hearing at 1000 Hz with DPOAEs performed better than at 500 Hz, but still reported very high false negative rates and influence of low frequency noise interfering with test measurements (Gorga et al. 1993; Kimberley et al. 1994b; Moulin et al. 1994; Probst & Hauser, 1990). In a study by Kimberley et al. (1994b), hearing ability at 1025 Hz could be accurately predicted as normal 90% of the time. Kimberley et al. (1994b) did not state the false negative rate for 1025 Hz specifically, but an average false negative rate (where hearing-impaired ears were predicted as normal) across the frequency range f2 = 1025 Hz to 5712 Hz of 22%. Moulin et al. (1994) accurately predicted normal hearing at 1000 Hz 73% of the time. False negative responses in this study varied between 12% and 17% with an average of 15%. In the study by Gorga et al. (1993), false negative rates ranged from about 25% to over 60% depending on the hit rate that was selected.

For this study, the accurate prediction of normal hearing (0-25dB) at 1000 Hz was 88%. The false positive rate for the three categories combined is 6%. False negative responses for 1000 Hz were very high, the highest for all four frequencies tested: 2% for C4 (category four)(30 + 35dB), 10% for C5 (category 5)(40 + 45dB), 5% for C6 (50 + 55dB) and 5% for C7 (60 + 65dB). That comes to a total of 22%, which is unacceptably high to regard this procedure as sensitive enough to correctly identify hearing impairment.

Even though the predictions of normal hearing in this study for 1000 Hz were more accurate than stated elsewhere (Gorga et al. 1993; Moulin et al. 1994; Probst & Hauser, 1990), the high incidence of false negative responses influences the sensitivity of this procedure. This high incidence of false negative responses lessens the clinical applicability of this neural network run as a possible screening procedure. (The false positive and false negative rates for all four frequencies in terms of suitability for use in screening and diagnostic audiology tests are further discussed in 6.8 “Results of the present study in perspective: Readiness for broad clinical use”.)

The prediction of specific categories spanning hearing impairment at 1000 Hz was again disappointing. The possible reasons were discussed in 6.4. It seems that the main reason for the network's incapability to predict hearing loss at 1000 Hz also stems from insufficient data in every category. The shortage of ears to train on might even have affected the prediction accuracy of normal hearing: C3 (category three spanning 20 + 25dB) only had 12 ears and were predicted accurately only 29% of the time versus C1(0 + 5dB) that had 40 ears and was predicted accurately 80% of the time. From Figure 6.2 it could be suggested that a study with an expansive data set of more than 80 ears in every category might be able to make predictions of hearing ability at 1000 Hz, normal or impaired, with very high levels of accuracy.

6.6 Prediction of 2000 Hz

The prediction of normal hearing at 2000 Hz in other studies yielded far more promising results than their predictions at 500 Hz and 1000 Hz. Kimberley et al. (1994b) predicted normal hearing at 2050 Hz with 84% accuracy. Mean false negative responses for f2 = 1025 Hz to 5712 Hz was 22%. Moulin et al. (1994) predicted the DPOAE frequency of 1413 Hz (closest to the GM frequency of 2000 Hz) correctly 75.6% of the time with an average false negative response of 15%.

In this study, normal hearing (0-25dB HL) at 2000 Hz could be predicted accurately 88% of the time. The false positive rate for the first three categories combined was 5%, in other words, 5% of normal ears were predicted as hearing impaired at 2000 Hz which makes this procedure quite specific in this frequency region.

False negative rates were 4% for C4 (30 + 35dB), 4% for C5, 4% for C6 and 3% for C7. This comes to a total of 15% for all categories spanning hearing loss. This number is the same as the number found by Moulin et al. (1994) and lower than in the study by Kimberley et al. (1994b) but 15% is still considered rather high and raises concern for the sensitivity of this procedure as a screening procedure to identify hearing loss at 2000 Hz. (See 6.8 "Results of the present study in perspective: Readiness for broad clinical use".)

The influence of ear count in every category on prediction accuracy of hearing impaired categories is depicted in Figure 6.2. Categories that had less than 12 ears were never predicted accurately. C3 (20 + 25dB), which forms part of the representation of normal hearing had only ten ears and was never predicted accurately. It could be speculated that normal and impaired hearing at 2000 Hz would have been predicted more accurately if there were more data in every category.

6.7 Prediction of 4000 Hz

The prediction of normal hearing at 4000 Hz has been a strong point for some of the previous studies. Moulin et al. (1994) successfully predicted the DPOAE frequency of 4000 Hz (primary frequencies are between 5000 Hz and 6000 Hz) as normal 79.4% of the time. These authors predicted the DPOAE frequency of 2826 Hz (primary frequencies between 3500 Hz and 4500 Hz) as normal 82% of the time, with an average false negative rate of 15%. Kimberley et al. (1994b) predicted normal hearing at 4052 Hz accurately 88% of the time with mean false negative rates of 22%.

The prediction of normal hearing at 4000 Hz with the use of DPOAEs and ANNs revealed that normal hearing could be predicted accurately 93% of the time. The false positive rate was only 2%, which makes this a very specific procedure to use for identification of normal hearing at 4000 Hz. Very good hearing (0-15dB) was correctly predicted an normal (0-25dB) 96% of the time. This procedure therefore offers a very rapid and accurate objective measurement of normal hearing.

The false negative rates at 4000 Hz were 3% for C4, 2% for C5, 2% for C6, 5% for C7 and 1% for C8 (70 + 75dB). The total false negative rate for all categories spanning hearing impairment was 13%. This is the lowest incidence of false negative responses for the four frequencies in this study, and also stated elsewhere for 4000 Hz (Moulin, et al. 1994; Kimberley et al. 1994b). This study proved it possible to predict normal and impaired hearing at various frequencies with DPOAEs and ANNs but more research is needed to lower false negative rates to acceptable levels for screening and diagnostic purposes. (See 6.8 "Results of the present study in perspective: Readiness for broad clinical use".)

Prediction of specific classes of hearing impairment at 4000 Hz was once more not satisfactory due to insufficient data in certain categories for the neural network to train on. The ear count versus prediction accuracy correlation for 4000 Hz can be seen in Figure 6.2. C3 (20 + 25 dB) had only seven ears and was never predicted accurately opposed to C1 (0 + 5dB) that had 33 ears and was predicted accurately 85% of the time. It could be speculated that normal and impaired hearing could be predicted even more accurately if there were more ears in every category.

6.8 Results of the Current Study in Perspective: Readiness for Broad Clinical Use.

The need for an objective, rapid, non-invasive, inexpensive, and accurate measurement of hearing inspired the investigation of the distortion product otoacoustic emission (DPOAE) as a possible new test of hearing.

Otoacoustic emissions have been used for a number of clinical uses in audiology. Applications include (a) screening of neonates and infants; (b) differential diagnosis of cochlear versus retrocochlear hearing losses; (c) monitoring of the effects of noise exposure or ototoxic drugs on the outer hair cells; and (d) to monitor fluctuating hearing loss in persons with Meniere's disease (Lonsbury-Martin et al. 1992; Norton & Stover, 1994; Norton & Widen, 1990; Robinette, 1992; Koivunen, et al. 2000).

The distortion product has been proven as an acceptable screening procedure. It is present in all normal hearing ears and even though it is measurable in ears with a hearing loss of up to 65dB HL, the amplitude and threshold of the DPOAE indicate different qualities, revealing hearing impairment (Moulin et al. 1994; Smurzynski et al. 1990). It can be measured non-invasively, objectively and rapidly (Norton & Stover, 1994). It is not significantly affected by gender (Cacace et al. 1996; Lonsbury-Martin et al. 1990) and has good test-retest stability (Cacace et al. 1996). Furthermore, it can be measured over a wide range of frequencies (Bonfils et al. 1991). DPOAEs are not affected by state of consciousness and do not require sedation (Norton & Stover, 1994). Lastly, it is an economic test that yields ear specific information. Many researchers used these attributes to correctly identify normal hearing in populations with varying degrees of hearing ability (Gorga et al. 1993;

Kimberley et al. 1994; Moulin et al. 1994). The only limitation of DPOAEs as a screening procedure is the lack of sensitivity identified in some of the studies. Sensitivity of a screening procedure is affected negatively by a high incidence of false negative responses. False negative responses refer to hearing-impaired ears that are predicted as normal (Schwartz & Schwartz, 1991). Some studies, including the present one, revealed an incidence of false negative responses too high for clinical acceptability (Kimberley et al. 1994; Moulin et al. 1994). According to Brass and Kemp (1994), it is very important to have a very high sensitivity for a first pass screening procedure (low incidence of false negative responses), as close as 100% sensitive as possible. The specificity of a screening procedure (affected by the number of false positive responses), on the other hand, is less important and is quite acceptable even if the test is only moderately specific (such as a specificity of 75%). Brass and Kemp (1994) compared test effectiveness and efficiency quite effectively: "In terms of screening effectiveness, the final number of false negatives is very important as this is the group of those for whom we were screening but missed. In terms of screening efficiency, the final number of false positives is important as this increase the number passed on to and hence the cost of the next stage of screening." (Brass & Kemp, 1994: 386).

It would therefore seem that the number of false negatives in this study is too high for an acceptable screening procedure, even though it is better than reported elsewhere (Gorga et al. 1993; Moulin et al. 1994; Probst & Hauser, 1990; Stover et al. 1996a).

DPOAEs however, have never been used as a diagnostic test of hearing where specific thresholds for frequencies were determined (Kimberley et al. 1994b; Lee et al. 1993). Many researchers mentioned the increasing role that otoacoustic emissions have in diagnostic audiology (Kemp et al. 1990; Martin et al. 1990; Moulin, 2000a, Moulin 2000b) and described the possibility of DPOAEs as a diagnostic audiologic test (Durrant, 1992; Kimberley et al. 1994b; Lee et al. 1993). These authors also stated that additional research is necessary before otoacoustic emissions can be implemented as a diagnostic test of hearing.

This research project attempted to predict hearing ability at 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz with DPOAEs in normal and hearing impaired ears with the use of artificial neural networks. It was attempted to predict specific hearing levels within 10dB for normal and hearing-impaired ears. Even though this study could correctly identify normal hearing quite accurately, the specific predictions of hearing levels at various frequencies were rather disappointing. One possible reason for the poor prediction of categories depicting hearing loss is the number of ears in every category that the neural network had to train on. Chapter 3 explained the learning and training of a neural network in 3.5.2.2, and that every category should have enough data for the neural network to train on, to form adequate representations and to make accurate predictions. To investigate this possibility, the accuracy of the prediction was correlated with the amount of data that the neural network had to train on. (See Figure 6.2.) Results indicated that there seems to be a threshold value for the number of ears in every category: when there were more than 32 ears in a category, prediction accuracy was always above 75%. These results suggests that categories depicting hearing impairment can be predicted accurately, if there is enough data in every category for the neural network to train on. If all the parameters of this research project were kept exactly the same with just one alteration of research design namely, the increase of subjects, hearing ability could be accurately predicted within 10dB from 500 Hz to 4000 Hz, for hearing levels up to 65dB HL.

This find is definitely a great contribution to the development of DPOAEs as a diagnostic test of hearing. Results proved that it is possible to predict hearing ability at frequencies as low as 500 Hz objectively, rapidly, non-invasively, and economically with DPOAEs and ANNs. The research in this study could be used as a foundation stone for further research to develop the distortion product as a new objective diagnostic test of hearing. Such a test would change the field of objective diagnostic audiology as we know it considerably and would be a tremendous aid in the evaluation of difficult-to-test populations.

To integrate results obtained from the prediction of 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz, a few case studies will be investigated where the neural network accurately predicted hearing ability in all four frequencies. A few case studies will also be investigated where the neural network made false predictions with possible reasons for inaccurate predictions. These results are discussed below.

6.9 Case Studies where the Audiogram was Predicted Accurately

The network predicted hearing ability into one of seven 10dB categories. Examples for correct predictions can be seen in Figure 6.3. Subjects that demonstrated very good hearing ability (0-15dB) at all four frequencies were usually predicted accurately. Figure 6.3. depicts the audiogram and predicted categories for ear 35, an ear with normal hearing that was predicted accurately, ear 85, an ear with a hearing loss predicted accurately except for 4000 Hz that was 10dB out and ear 107, an ear of an 80-year-old person, predicted accurately except for 500 Hz that was 10dB out. Subject information, such as age, gender, and complaints of tinnitus or vertigo is presented in Table 6.1.

Table 6.1: Subject information of cases predicted accurately

	Ear 35	**Ear 85**	**Ear 107**
Subject age	61	72	80
Subject gender	Female	Male	Female
Tinnitus?	No	No	High frequency
Vertigo?	No	No	No
Medication	None	None	None
Noise exposure?	None	None	None

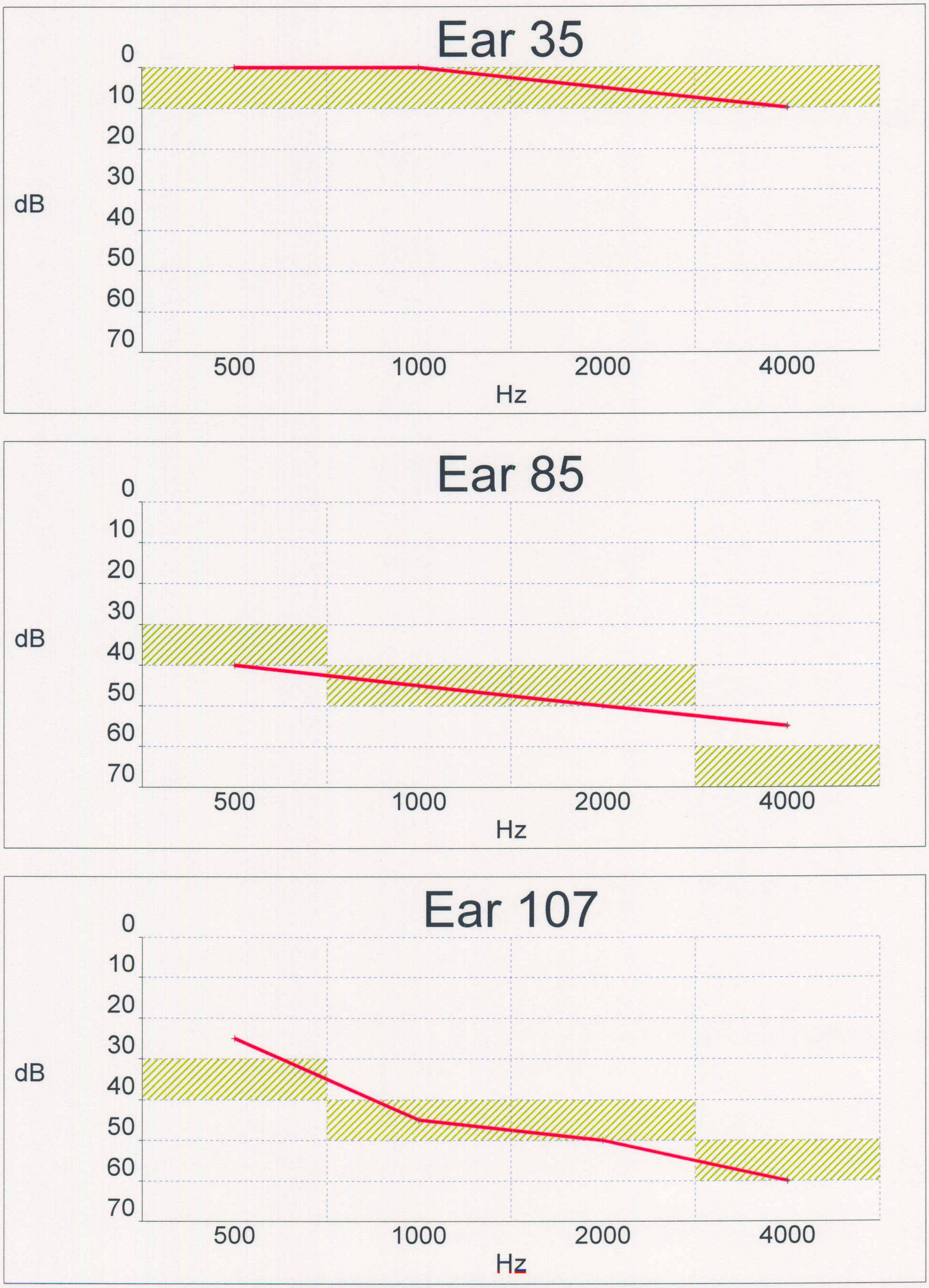


Figure 6.3: Case studies where the audiogram was predicted accurately (prediction in green, audiogram in red).

6.10 Case Studies where the Audiogram was Predicted Inaccurately

Subjects demonstrating hearing loss were sometimes predicted inaccurately. Six examples of inaccurate predictions will be given. Audiograms and predictions can be seen in Figure 6.4. The subject information of these six ears is presented in Table 6.2.

Table 6.2: Subject information of cases predicted inaccurately

	Ear 19	**Ear 33**	**Ear 71**	**Ear 73**	**Ear 74**	**Ear 91**
Subject age	31	35	16	39	39	43
Gender	Female	Female	Male	Male	Male	Male
Tinnitus	No	High Frequency	No	No	No	No
Vertigo	No	No	No	No	No	No
Medication	None	None	None	None	None	None
Noise exposure	None	None	None	20 years	20 years	15 years

6.10.1 Interesting Phenomena in Cases Predicted Inaccurately

With closer analysis of the cases that were predicted inaccurately, it became evident that there were a few circumstances in which the neural network could almost never predict hearing ability correctly. Some of these instances included noise exposure, very mild hearing losses, and possible retrocochlear hearing losses. These cases will be discussed below. It is however, also the case that in some instances, the neural network predicted hearing ability inaccurately for no apparent reason.

Figure 6.4: Case studies where the audiogram was predicted inaccurately (prediction in green, audiogram in red).

6.10.1.1 Subjects Demonstrating Hearing Loss Due to Noise Exposure

Subjects with a large amount of noise exposure revealed poor correspondence between pure tone audiograms and DPOAE measurements. The DPOAE measurements indicated a much larger hearing loss than the pure tone audiograms when DPOAE measurements were compared to the normal range. This confirms the research by Durrant (1992) that indicated that damage in outer hair cells could be measured before the actual hearing loss occurs. Even though these subjects already had noise-induced hearing losses, damage to these subject's outer hair cells indicated a far greater hearing loss than their pure tone audiograms. The neural network therefore predicted hearing ability inaccurately as much more hearing-impaired in most of the cases demonstrating noise-induced hearing losses, such as in the cases of ear 73, 74 and 91.

6.10.1.2 Subjects Demonstrating Very Mild Hearing Losses

Subjects demonstrating very mild hearing losses, especially if only some of the frequencies were hearing impaired, were often predicted as normal. Smurzynski et al. (1990) also referred to a subject with a very mild hearing loss that had normal DPOAEs. Probst and Hauser (1990) and Gorga et al. (1996) found that a mild hearing loss is sometimes not detected with DPOAE measurement. If the DPOAE measurements of a specific ear are within the normal range, the neural network will predict that ear as normal. For this study, all eleven DPOAE frequencies were used for input information. If most of the DPOAE responses were normal, it often happened that the neural network predicted all the frequencies as normal, as in the case of ear 19 and ear 33.

6.10.1.3 A Subjects Demonstrating A Possible Retrocochlear Hearing Loss

Another example of poor correspondence between DPOAE results and the pure tone audiogram was that in a few instances, DPOAE results appeared much better than the hearing ability depicted in the audiogram. This could possibly be one of the small percentages of subjects that could be classified as having a retrocochlear pathology based on otoacoustic emission results (Robinette, 1992). In the case of ear 71, the neural network predicted hearing ability as normal at 500 Hz, 1000 Hz and 2000 Hz, but ear 71 demonstrated a mild hearing loss. (The audiogram and predictions of ear 71 can be seen in Figure 6.4.) The DPOAE measurements, when compared with the normal range of DPOAE measurements, also indicated that DPOAEs were much better than would be expected from an ear with a mild hearing loss. Even though other site-of-lesion tests were not performed on this ear, the fact that the DPOAE measurements and neural network predictions were so much better than expected could be as a result of retrocochlear pathology. The outer hair cells on the basilar membrane in the cochlea could therefore be normal and capable of normal DPOAEs.

The next section discusses all the subject, DPOAE and ANN variables that were investigated to determine their effect on prediction accuracy.

6.11 Subject-, DPOAE- and ANN-Variables Experimented with to Determine Optimal PTT Prediction Accuracy.

The subject variable that was experimented with was AGE, presented to the network in 5-year categories or 10-year categories. DPOAE variables include the threshold of the distortion product defined as 1, 2 or 3dB above the noise floor, the frequency information of the DPOAE to use for ANN input, and the amplitude of the DPOAE presented to the neural network in a number of different ways. ANN variables include experimentation with the number of middle neurons and different error tolerances during training and prediction.

6.11.1 The Effect of the Subject Variable AGE Presented to the Network in 5 year or 10 year Categories on PTT Prediction Accuracy.

It seems that the investigation of the influence of age on the distortion product has been problematic for many authors (Avan & Bonfils, 1993; He & Schmiedt, 1996; Karzon et al. 1994; Nieschalk et al. 1989). The reason for this is, that it is very difficult to determine how much of the differences in the distortion product observed in elderly subjects are due to age, and how much is due to sensitivity changes associated with aging. It seems that these authors agree that the negative correlation between DPOAE levels and age is due to changes in hearing threshold associated with aging rather than age itself. However, some researchers found that when the age variable was included in pure tone prediction studies with DPOAEs, age had a positive effect on prediction accuracy (Lonsbury-Martin et al. 1991; Kimberley et al. 1994a; Kimberley et al. 1994b, De Waal, 1998). The age variable was therefore always included in this study, but in two different ways, either depicted with the dummy variable technique into one of 20 possible 5-year categories, or into one of 10 possible 10-year categories.

The results for differences in prediction accuracy for these two methods were given in Figure 5.5. Prediction accuracy for the two methods was always within 5%. The presentation of age as within 5 years, or within 10 years therefore did not have a great effect on prediction accuracy. There may be more than one reason for this. The first possible reason is the fact that even though age was presented more specifically to the network in the 5-year category method, the extra inputs that were needed to represent this information contributed to a more complex neural network topology. Many more inputs were needed which made the data set more complex and it took the network longer to converge. The advantage of a more specific age presentation might have been lost in the creation of a more complex data set. The second possible reason is that the differences in DPOAEs due to age related changes might not a phenomenon that is greatly influenced by only 5 years: the age related difference between a subject only five years older than another subject with identical PTTs might be too small to detect or to make a big difference in PTT prediction accuracy. It would make an interesting study to investigate the effect of age on prediction accuracy with DPOAEs and PTTs by including age categories larger than 10-year categories, such as 15-year and 20-year categories to find the cutoff point where age is not presented specifically enough anymore to have a positive effect on prediction accuracy.

6.11.2 The Effect of DPOAE Threshold Defined as 1, 2 or 3 dB Above the Noise Floor on PTT Prediction Accuracy.

As was described in 4.6.5.2 “Determination of optimal stimulus parameters”, the acceptance of a DPOAE test was either that the cumulative noise floor had to be –18dB SPL or if noise levels were higher, that the DPOAE amplitude had to be 10 dB above the noise floor. For about half of the tests run (47%), noise levels were low enough that test acceptance was achieved by the –18dB SPL cumulative noise level criteria. For all these cases, a DPOAE threshold had to be determined. Other researchers specified DPOAE threshold as 5 dB above the noise floor (Krishnamurti, 2000) or 3dB above the noise floor (Lonsbury-Martin, 1994). For this research project, it was investigated to see if a lower threshold for cases that passed the first criterion of cumulative noise levels of –18dB SPL would have more accurate predictions as a result. There were experimented with 1, 2 and 3 dB above the noise floor. Results were summarized in Figure 5.6.

It was very interesting to note that there was virtually no effect on prediction accuracy for threshold defined as 1, 2 or 3 dB above the noise floor. A threshold of 1dB could predict hearing ability in normal and impaired ears as well as a threshold of 3dB. All predictions were within 1%. This finding leads the researcher to believe that the definition of DPOAE threshold may be lowered to 1dB. A lower DPOAE threshold might improve prediction accuracy for other studies if valuable DPOAE data close to the noise floor is not discarded but used as valid present responses.

6.11.3 The Effect of the Emission or Inclusion of Low Frequency DPOAE Information for ANN Training on PTT Prediction Accuracy

For experiments where low frequency DPOAE data was omitted, f1= 500 Hz, 625 Hz and 781 Hz were omitted in the input data to train the neural network on to investigate if certain frequencies could be predicted more accurately. The summary for PTT prediction accuracy with DPOAEs and ANNs where noisy low frequency emissions were omitted or present was given in Figure 5.7.

Results with low frequencies present or absent revealed only a slight difference. Differences in prediction accuracy were always within 2%. All frequencies, except

2000 Hz, were even predicted slightly more accurately when low frequency DPOAEs were present. It seems like the "noisy" low frequency DPOAEs had very little effect on the training or prediction capabilities of the neural network, which confirms the viewpoint of Blum (1992) that neural networks excel in dealing with noisy incomplete data. Nelson and Illingworth (1991) described the nature and strengths of artificial neural networks very effectively: "The characteristics of intuition, prediction and statistical pattern recognition allow neural networks to deal with situations in which input data may be fuzzy, incomplete, or ambiguous, or may even have some corrupted data. Because most of the data in this world is inexact, this characteristic becomes highly significant." (Nelson & Illingworth, 1991:69). Blum (1992) even suggested training a network with "noisy" data to enhance generalization capability of the network's predictions. It seems like the presence of noisy low frequency data in ANN training had a slightly positive effect on prediction accuracy.

6.11.4 The Effect of DPOAE Amplitude Presentation to the Neural Network on PTT Prediction Accuracy

For this research project, the amplitude of the DPOAE was presented to the network in four different ways: AMP 100, AMP 40, ALT AMP and No AMP (see 4.8.1.4 for descriptions of characteristics for each representation.) These results were summarized in Figure 5.8.

As was described in 4.8.1.4, each method of amplitude representation influenced the ANN in a unique way. AMP 100's neural network had trouble converging (reaching optimal error tolerance levels during training to begin prediction), probably because the amplitude was depicted as a fraction of 100, which made the number rather small (a DPOAE response of 30 dB which can be regarded as quite a strong DPOAE was depicted as 0.3). The largest DPOAE response measured in this study was 39dB, so no amplitude input in the AMP 100 method were ever depicted as more than 0.4. The significance of the size of DPOAE amplitude might have been lost in this method and that might be one reason why it took the network so long to converge. As a result of this, AMP 100 never had the most accurate prediction at any frequency.

AMP 40's ANN had the exact same topology than the AMP 100 method but convergence was much faster because the input values were larger (a fraction of 40 instead of a fraction of 100) and the network therefore found it easier to make midway representations to reach error tolerance levels faster. The same 30dB response would now be depicted as 0.75. This method created more "room" to distinguish between DPOAE amplitudes of various sizes and indicate the significance thereof. AMP 40 had the most accurate predictions of overall prediction accuracy at 500 Hz, 1000Hz, and 2000 Hz. At 4000 Hz, AMP 40 was second in accuracy, only 2% less accurate than the No AMP method. At 2000 Hz, AMP 40 was most accurate for identification of very good hearing (0-15d) and normal hearing (0-25dB).

The ALT AMP technique had the one advantage that information was presented to the network in the same fashion than the output predictions, which was by depicting information in categories with the dummy variable technique. Input mode and output mode for that neural network was therefore the same. A Disadvantage of the ALT AMP method however, was that the complexity of the topology of the network increased drastically due to the fact that 352 input neurons were needed to present amplitude in this fashion instead of the usual 88. This method was able to predict very good hearing (0-15dB) and normal hearing (0-25dB) most accurately at 500 Hz. ALT AMP came a close second to No AMP in the prediction of very good and normal hearing at 1000 Hz (within 0.8%).

With the No AMP method where the amplitude of the distortion product was left out entirely, most accurate predictions were made at prediction of normal and very good hearing at 1000 Hz and 4000Hz. At 4000 Hz, overall prediction accuracy was also best for the No AMP method.

When it comes to the overall prediction accuracy across all categories spanning normal hearing and hearing loss, the AMP 40 method seems to be the best way to present amplitude to the neural network. When it comes to the prediction of normal hearing (0-25dB) or very good hearing (0-15dB), then the No Amp and ALT AMP methods, work the best.

The finding that the No AMP method showed good prediction for normal hearing is consistent with the findings of Moulin et al. (1994) who concluded that DPOAE threshold is a more sensitive predictor of hearing than DPOAE amplitude. It could be argued that in the case of normal hearing, DPOAE thresholds are low enough to predict hearing ability without any amplitude information, such as in the No AMP method.

When it comes to the prediction of hearing loss, however, it seems that the inclusion of amplitude data had a positive effect. The fact that hearing loss categories were predicted best with the AMP 40 method in this study is consistent with the findings of

Probst and Hauser (1990) who found that DPOAE amplitude is strongly correlated with PTTs in hearing impaired ears.

Future studies that attempt to predict PTTs in hearing-impaired ears should therefore include both the threshold and amplitude information of the distortion product to ensure optimal prediction accuracy.

6.11.5 The Effect of the ANN Variable MIDDLE NEURON COUNT on PTT Prediction Accuracy

As was stated in chapter 4, the size of the hidden layer is function of the diversity of the data (Blum, 1992). The number of middle layer neurons determines the accuracy of prediction during the training period. With an insufficient number of middle neurons, the network is unable to form adequate midway representations or to subtract significant features of the input data (Nelson & Illingworth, 1991). With too many middle neurons the network has difficulty to make generalizations (Rao & Rao, 1995; Nelson & Illingworth, 1991). The number of middle layer neurons was determined by trial and error, based on the accuracy of the prediction during the training period. If the neural network was unable to converge during training, the number of middle level neurons was increased and the prediction attempted again.

Various network runs were conducted with varying numbers of hidden layer neurons, ranging from 20 to 180. With a number of hidden layer neurons below 60, the network was unable to extract significant features from the input data and sometimes would not converge during training. A number of hidden layer neurons more than 160 resulted in poor generalization ability. For this study, three possible middle neuron counts were identified, namely 80, 100 and 120 and experiments were run with all three possibilities to determine its effect on prediction accuracy.

Results for prediction accuracy for the three middle neuron counts were summarized in Figure 5.9. Prediction accuracy for all three middle neuron counts was always within 1%. There was therefore no significant enhancement for prediction accuracy when middle neuron counts were slightly increased.

Many authors agreed that there is not a clear-cut formula for the determination of the optimal number of middle level neurons for an artificial neural network but that it is determined by the complexity of the data and should be determined with trial-and-error experimentation (Rao & Rao, 1995; Blum, 1992; Nelson & Illingworth, 1991; Fujita, 1998).

It seems that all three middle level neuron counts for this study worked equally well and any one of the three middle level neuron counts can be applied to this research project for the prediction of PTTs with DPOAEs.

6.11.6 The Effect of the ANN Variable TRAINING ERROR SENSITIVITY on PTT Prediction Accuracy

Another neural network variable is the acceptable error during training. A neural network learns from its mistakes. The first step in the learning process is to compute the outputs, the second step to compare the outputs with the desired answers and the last step to adjust the set of weights to enable a better prediction the next time. The second step, namely the comparing of outputs with desired answers, can be made in various levels of accuracy. The neural network can be required to make predictions within any percentage of accuracy during the training stage. The normal assumption would be that the more accurate the prediction during training, the more accurate the predictions would be. It is however often found that if training error is set too close to zero, that generalization abilities of the network decreases and that it experience difficulty predicting cases slightly out of the ordinary (Rao & Rao, 1995).

Error tolerance levels for this study were set at 0.001, 0.002 and 0.003. Results were summarized in Figure 5.10. Prediction accuracy for the three tolerance levels was always within 1%. For this network configuration, the three error tolerances performed equally well.

6.12 Summary

Other studies	**This study**
In the investigation of DPOAEs as a possible new hearing screening or diagnostic procedure, many authors found a correlation between DPOAEs and PTTs (Gaskill & Brown, 1990; Gorga et al. 1993; Kummer et al. 1998; Lee et al. 1993; Vinck et al. 1996).	Results of this study confirmed the correlation between DPOAEs and PTTs and found that artificial neural networks found a significant correlation between DPOAEs and PTTs to enable ANNs to predict PTTs accurately given only DPOAEs.
Some researchers predicted hearing ability as normal or hearing-impaired at various frequencies, with more success in the high frequencies (Kimberley et al. 1994; Moulin et al. 1994). There were no reports of prediction at 500 Hz due to difficulties experienced with the rising noise floor below 1000Hz (Gorga et al. 1993; Moulin et al. 1994; Stover et al. 1996a).	Normal hearing ability was accurately predicted at frequencies as low as 500 Hz in the present study. Prediction accuracy for normal hearing was 94% for 500 Hz, 88% for 1000 Hz, 88% for 2000 Hz and 93% for 4000 Hz. These predictions are better than reported elsewhere (Kimberley et al. 1994 a+b; Moulin et al. 1994).

Even though this study could correctly identify normal hearing quite accurately, even at 500 Hz, the specific predictions of impaired hearing levels at various frequencies were rather disappointing.

One possible reason for the poor prediction of categories depicting hearing loss is the limited number of ears in every category that the neural network had to train on. With closer analysis of the possible correlation between prediction accuracy and data quantities, it became clear that the neural network would perform much better with more data to train on.

This finding serves as a strong indication that DPOAEs can be used to predict hearing status in populations with varying degrees of hearing loss objectively and accurately. **The results in this study forms a solid foundation stone for further research to continue the quest for optimal pure tone prediction with distortion product otoacoustic emissions.**

Chapter 7: Summary, Evaluation of the Study and Conclusion

7.1 Summary

An overview of current objective diagnostic procedures revealed that many technological advanced tests exist for the successful evaluation of hearing ability and site of lesion testing in adults. It is in the evaluation of difficult-to-test populations however, that limitations in current objective diagnostic procedures were identified. It seemed that, despite all the strengths and positive attributes of ABR, tympanometry, MLR, and LLR, weaknesses in these techniques made it difficult to measure exact hearing ability and site-of-lesion in populations such as neonates, infants, malingerers, the crucially ill and foreign speakers. Some of these weaknesses include a limited frequency area in which hearing ability can be determined, lengthy test times, the unfortunate possibility of sedation and the level of expertise and expense required for successful implementation of the diagnostic test battery and interpretation of results (Ferraro & Durrant, 1994; Musiek et al., 1994; Robinette, 1994; Weber, 1994). It is therefore with much hope that many researchers turned their investigations to otoacoustic emissions.

Kemp (1978) identified different classes of otoacoustic emissions, depending on the stimuli used to evoke them. Spontaneous otoacoustic emissions (SOAEs) are only prevalent in half of normal hearing persons and can therefore not be implemented as a screening test or diagnostically (Lonsbury-Martin, 1994; Norton & Stover, 1994). Stimulus frequency otoacoustic emissions (SFEs) are not currently clinically used due to difficulties in separating in-going stimuli and out-going emitted responses

(Lonsbury-Martin & Martin, 1990). Transient evoked otoacoustic emissions (TEOAEs) have been proven as a clinical acceptable hearing screening procedure, but the fact that they are only recordable in normal ears limited their diagnostic hearing applications (Kemp & Ryan, 1993; Lonsbury-Martin et al., 1992; Stevens et al., 1990). Distortion product otoacoustic emissions (DPOAEs) on the other hand, revealed many possibilities as a potential test of auditory functioning. First, it has been proven useful in both clinical and research settings, for it is the only emission type that can easily be recorded in many laboratory animals, allowing for experimental control of certain factors (Mills, 1997). Second, it can be measured in ears with a hearing loss of up to 65dB HL, therefore revealing information regarding outer hair cell functioning of hearing-impaired populations as well (Moulin et al., 1994). Third, it is the emission type that can most easily be compared to the pure tone audiogram, due to the pure tone nature of the stimuli that can be chosen to represent pure tone audiogram frequencies (Durrant, 1992; Lonsbury-Martin & Martin, 1990). Fourth, DPOAEs correlate well with pure tone thresholds and the configuration of the hearing loss (Durrant, 1992; Kimberly & Nelson, 1989; Stover et al., 1996a). Fifth, DPOAEs are not significantly influenced by gender (Cacace et al., 1996; Karzon et al., 1994).

Many studies described the relationship between DPOAEs and pure tone thresholds (Avan & Bonfils, 1993; Bonfils et a; 1991; Gaskill & Brown, 1990; Gorga et al., 1993; Kimberley et al., 1994b; Probst & Hauser, 1990; Stover et al., 1996a). Statistical methods used to date, such as multivariate (discriminant) analysis in the case of the study of Kimberley et al., (1994b), but also in all the other studies previously named, indicated a correlation between DPOAE measurements and behavioral pure tones. These studies however, could not predict the actual pure tone thresholds given only the distortion product responses (Lee et al., 1993). The complexity of the data, the numerous variables involved and the possibility of a nonlinear correlation have been some of the reasons why conventional statistical methods could not predict pure tone thresholds given only DPOAEs, but only distinguish between normal hearing and hearing-impaired ears (Kimberley et al. 1994a; Kimberley et al. 1994b).

For this study, a mathematical model, called **artificial neural networks** (ANNs), was used to investigate the relationship between DPOAE measurements and pure tone thresholds. This technique has excellent correlation finding capabilities, even in the case of a non-linear correlation. The neural network was used to predict pure tone thresholds given only the distortion product responses.

ANNs were initially developed to gain a better understanding of how the human brain functions (Nelson & Illingworth, 1991). It is an algorithm for a cognitive task, such as learning or pattern recognition (Muller & Reinhardt, 1990) and can be used to make predictions based on learned correlations from previous similar data (Blum, 1992). Various disciplines became interested in the use of ANNs to address complex problems in the last two decades, ranging from cognitive psychology, physiology, medicine, computer science, electrical engineering, economy and even philosophy. ANNs have been successfully applied to the field of Speech Pathology for speech recognition purposes (Metz, et al., 1992) and Audiology to predict normal hearing ability (Kimberley et al. 1994a). It was therefore with great expectations that neural networks were applied to the field of diagnostic Audiology (Kimberley et al. 1994a; De Waal, 1998).

The rationale for the current study was to further investigate DPOAEs as a possible new objective test of hearing by improving prediction accuracy of PTTs with DPOAEs and ANNs. Factors that influenced DPOAEs levels were identified (such as the frequency ratio and levels of the primaries) and were controlled to ensure optimal DPOAE responses. Factors that influenced ANN efficiency were identified (such as middle level neuron count, input format and error tolerance levels) and were manipulated to reach optimal prediction accuracy levels. The correlation between the measured variables of DPOAEs and PTTs were studied with artificial neural networks and then used to predict hearing ability at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz with DPOAEs.

Data was obtained from 70 **subjects** (120 ears, in some cases only one ear fell in the subject selection criteria), 28 males and 42 females, ranging from 8 to 82 years old. Selection criteria included normal hearing, sensorineural hearing losses of varying degrees and normal middle ear functioning. The subject selection procedures included a short case history, otoscopic examination, tympanometry and a pure tone audiogram.

The distortion product has numerous stimulus variables that should be specified to ensure optimal testing conditions. The choice of **stimulus parameters** in this study was based on an extensive literature study and confirmed in the preliminary study. For this research project, eight DP Grams at 5dB intervals ranging from L1=70dB SPL to

L1=35dB SPL were measured. A frequency ratio of 1.2 was selected for the two primaries and the loudness level ratio of the two primaries was L1>L2 by 10dB. The frequency range of F1 = 500 to F1 = 5031 was tested. The **criterion for DPOAE presence** was that test conditions had to be "accepted" which implies either a DPOAE response of 10dB above the noise floor or a cumulative noise level of – 18 dB SPL. For the cases where the noise level was low enough to pass test acceptance conditions for a – 18 dB SPL noise floor, **DPOAE threshold** was defined as 1, 2 or 3 dB above the noise floor to investigate the effect of a lower threshold criterion as stated in the literature on prediction accuracy (the lowest criterion in literature was set by Lonsbury-Martin, 1994, as 3dB above the noise floor).

Eight tests or DP Grams were performed in each ear. Every DP Gram consisted of eleven frequency pairs. Every frequency pair consisted of two pure tones, f1 and f2 presented to the ear simultaneously.

A **backpropagation network** was chosen for this study for two reasons: **1)** A nonlinear correlation is suspected between DPOAE thresholds and traditional pure tone thresholds. Metz, et al. (1992) reported the backpropagation neural network to be very successful in dealing with nonlinearities that potentially occur in complex data sets. According to Blum (1992), the backpropagation neural network is capable of nonlinear mappings and able to generalize well. **2)** The purpose of this study is to predict pure tone thresholds with distortion product thresholds with the use of neural networks. According to Blum, (1992) and Tam and Kiang, (1993), the backpropagation neural network is highly applicable in the areas of forecasting and prediction.

Several different **trial runs** were conducted to determine neural network topology and the way the input data should be presented to the neural network. The input data was presented to the neural network in a binary mode, and the data pattern of all absent and present DPOAE responses as well as DPOAE amplitude, subject age and gender served as input stimuli. Hearing ability was divided into categories and the neural network had to predict hearing ability into one of the 10dB categories.

Data analysis consisted of analyzing the actual and predicted values of all 120 ears and to determine how many were predicted accurately, how many within one class and how many were predicted incorrectly.

Results indicated that normal hearing ability could be distinguished from hearing-impaired hearing quite accurately, on all selected frequencies and as low as 500 Hz. Many researchers failed to predict normal hearing ability at 500 Hz due to the rising of the noise floor at the lower frequencies (Gorga et al., 1993; Moulin et al., 1994; Probst & Hauser, 1990; Stover et al. 1996a). **In this study, normal hearing ability at 500 Hz was predicted accurately 94% of the time. Normal hearing at 1000 Hz was correctly identified 88% of the time, at 2000 Hz 88% of the time and at 4000 Hz 93% of the time.**

In the **previous study** (De Waal, 1998), normal hearing at 500 Hz was predicted with 92% accuracy, 1000 Hz with 87%, 2000 Hz with 84% and 4000Hz with 91% accuracy. False positive and negative rates were also higher than in the present study. It seems that differences in neural network topology and input data manipulation from the previous study to the present, had a positive effect on prediction outcome. It should also be noted that the accuracy of the previous study's results could have been influenced by the fact that a subject's related ear was included in the training set, the neural network had an "unfair" advantage to be trained with a subject's one ear that might look very similar to the ear to be predicted, such as in the case of noise exposure. For the present study, both ears of a subject were removed from the prediction set. Despite this disadvantage, more accurate predictions were made at all four frequencies with less false positive and false negative responses.

The **false negative values** in this study are lower than reported elsewhere (Gorga et al., 1993; Moulin et al., 1994; Probst & Hauser, 1990; Stover et al. 1996a), but are still considered too high for optimal sensitivity and specificity for screening and diagnostic purposes.

The improvement of predictions of normal hearing at the four frequencies from the previous study to the present study can be attributed to a few reasons: First, the previous study struggled to incorporate the amplitude of the DPOAE response into training data and eventually used only the presence or absence of a response as neural network input. The present study experimented extensively with ways to incorporate DPOAE amplitude and successfully presented this information to the network. Secondly, there were experimented more extensively with neural network topologies, and ways to present input data to the network and more optimal procedures have been discovered and used in the prediction of PTTs with DPOAEs. It could be speculated that even more extensive experimentation could lead to better results than obtained in the present study.

Predictions of categories depicting hearing impairment were not satisfactory. Subjects were initially selected in such a manner that there were 40 ears with normal hearing, 40 ears with mild hearing loss and 40 ears with moderately severe hearing loss. The distribution of hearing loss at the four frequencies, however, resulted in an unequal amount of data in the different categories. Some categories were poorly represented and the neural network did not have enough data to train on, resulting in inaccurate predictions. In retrospect it also seems that the **number of subjects** included in this study, was too small to provide enough ears in every prediction category. Figure 6.2 clearly demonstrates that there is a threshold for the optimum number of ears in a category to enable a neural network to make accurate predictions. When less than about 32 ears were present in a category, prediction accuracy was very poor (less than 50%), but all categories with more than 32 ears showed a significant increase in prediction accuracy (more than 75%). If this hypothesis is true, then very accurate prediction of hearing ability from 500 - 4000 Hz for hearing losses up to 65dB HL can be expected with DPOAEs and ANNs where categories are well represented. **Despite the limited data set and shortage of ears in most categories, normal hearing was still identified with high levels of accuracy with DPOAEs and ANNs and this study effectively proved that DPOAEs can be developed as a diagnostic test of hearing that meet all the requirements for such a procedure.**

Interesting cases were identified that had irregular neural network predictions. Cases that were exposed to long periods of noise were always predicted as more hearing impaired, cases with possible retrocochlear hearing losses were predicted as having normal hearing due to their normal emissions and cases with minimal hearing losses were sometimes predicted as normal. These irregularities once again stress the importance of the case history as part of the diagnostic battery.

7.2 Evaluation of Research Methodology

The evaluation of research methodology will discuss the research design, the validity and reliability of this study and some limitations that were identified.

7.2.1 The Research Design

The research design chosen for this study was a multivariable correlational study. The correlation between the measured variables of DPOAEs (DPOAE thresholds at eleven 2f1-f2 frequencies) and PTTs (thresholds at 500, 1000 2000 and 4000 Hz) was studied with artificial neural networks and then used to predict hearing ability in normal and sensorineural hearing impaired ears up to 65 dB HL. Factors that influenced DPOAEs levels were identified (such as the frequency ratio and levels of the primaries) and were controlled to ensure optimal DPOAE responses. Factors that influenced ANN efficiency were identified (such as middle level neuron count, input format and error tolerance levels) and were manipulated to reach optimal prediction accuracy levels. DPOAE responses, DPOAE amplitude, subject age and subject gender formed the input of a feedforward neural network with a backpropagation learning algorithm. The network was trained with data of 118 ears to "learn" the correlation between DPOAE and PTT responses, and then used the learned correlation to predict an unknown subject's PTTs given only the subject's DPOAE responses. It has been proved that it is possible to predict audiometric thresholds with DPOAEs and ANNs, based on the correlation that the network found between the two sets of data and then used to make the prediction. The multivariable correlational study method in this research project was therefore applied successfully.

7.2.2 Validity and Reliability

Ventry and Schiavetti (1980) identified several factors that can influence the validity and reliability of data. The validity of the data can be divided into internal validity and external validity.

Internal validity deals with factors such as history, where the amount of time elapsed between the first and last test could include certain factors such as medication or treatment which could affect the readings of the second test differently than the first test. To preclude the influence of this fact on test data, the pure tone audiogram, tympanogram and distortion product measurement were performed in one session, lasting about an hour.

Internal validity also deals with instrumentation. The accuracy of the data obtained for the pure tone audiogram is a result of how well the audiometer was calibrated, how recently the audiometer has been calibrated, and the cooperation of the subject (Leedy, 1993). The audiometers used in this research project (calibrated annually) were calibrated less than a year before this project. Pure tone thresholds were double checked with speech reception thresholds when poor cooperation of the subject was suspected, the instructions for pure tone audiometry was repeated and a threshold was determined as 3 responses out of 6 stimuli presented.

The GSI 60 DPOAE system was calibrated for a particular quiet room. Regarding the fit of the probe, closure was obtained or DPOAE testing even though closure is not considered necessary but helpful by some authors (Bright, 1994). The closure fit of the probe reduced any external noise.

Another factor that influences internal validity according to Ventry & Schiavetti, (1980) is the differential selection of subjects. The subjects selected for this study were divided into three groups, normal hearing, slight to mild hearing losses and moderately severe sensorineural hearing losses. The subjects were selected carefully to ensure that no other factors than a sensorineural hearing loss is present that could influence the test data such as middle ear pathology in this case. Tympanograms were interpreted carefully to ensure normal middle ear pressure. Subjects that had normal hearing but no tympanogram due to a perforation in the tympanic membrane were not included in the study. The selection of subjects was strictly according to the subject selection criteria as set out in Chapter 4.

Reliability deals with the accuracy of the data obtained (Leedy, 1993) or precision of measurement (Ventry & Schiavetti, 1980). Reliability can be assessed by examining the stability and consistency of the test or measure. Gaskill & Brown (1990) conducted a study to investigate stability and reproducibility of DPOAE Grams over time and with different ear probes. These authors found DPOAE measurements to be extremely stable over time and that different probe fits do not significantly influence DPOAE measurements. DPOAE measurements therefore seem to be reliable. The fact that DPOAEs are so reliable makes it an ideal procedure to monitor cochlear function in Meniere's disease, the administering of ototoxic medication or during surgery of structures close to the cochlea (Cane, Donoghue & Lutman, 1992; Subramaniam, Henderson & Spongr, 1994; Teleschi, Roth, Stagner, Lonsbury-Martin, 1995a; Teleschi, Widick, Lonsbury-Martin & McCoy, 1995b).

Neural networks are also very reliable. Two neural network runs with exactly the same inputs yield exactly the same results (Blum, 1992).

Hall III et al. (1993), identified more factors influencing measurement and analysis, and therefore the validity and reliability of the study. First, it is important to determine the status of the middle ear and external ear canal, for DPOAEs depend on both an inward and outward propagation of stimulus energy. These two factors were carefully assessed during subject selection procedures. Second, the measurement parameters for DPOAEs should be carefully chosen to ensure optimal measurement conditions. In this study, measurement parameters were chosen after an extensive literature study and confirmed with finds during the pilot study.

The last aspect that could have an effect on the validity of this research project is human error during data preparation, analysis, and processing. Human error was eliminated or reduced where possible by electronic preparation, processing and analysis of data. DPOAE results were read into excel directly from the GSI-60 DPOAE system database to eliminate human error during the creation of subject files in data preparation. The computer extracted data that was used for the training of the neural network. Data analysis, where the correct answers were compared to the predicted answers, was also conducted on the personal computer to eliminate human error. Even the Figures, depicting prediction accuracy and correlation between number of ears and prediction accuracy were done on the computer with data directly from excel.

According to Leedy (1993) validity investigates the end results of the measurement. "Are we really measuring what we think we are measuring?" (Leedy, 1993:41). This research project attempted to find a correlation between DPOAEs and PTTs with ANNs and to use that correlation to predict PTTs with DPOAEs. It can be stated with reasonable certainty that this research project did in fact do what it was intended to do. Reliability, according to Leedy (1993) deals with the accuracy of the measurement. All measurements in this study were measured as accurately as technology currently allows on calibrated equipment.

7.2.3 Limitations of the Study

A few limitations were identified in this study. These are all aspects that should be kept in mind in the interpretation of results.

First, as stated previously, some of the categories depicting hearing impairment, were not represented adequately by the amount of data that the neural network had to train on. Even though subjects were initially selected to include an equal number of ears in three different categories of hearing ability, the pattern distribution of many sensorineural hearing losses is such that hearing loss is more prevalent in the higher frequencies than in the lower frequencies (Yantis, 1994). This resulted in an unequal number of ears in the categories that the neural network had to predict. Some of the categories were represented so poorly, that the neural network did not have enough data to train on. The network could not form adequate midway representations of the hearing ability of a subject in a category where only a few examples were present. To address this problem, more subjects should be included in neural network studies to ensure more data in every category.

The second limitation is the fact that this study did not investigate every possible neural network type and configuration available to determine their effectiveness as predictors of hearing ability. There are so many combinations of neural network configurations available, and even though numerous combinations were tried and tested for this application, it cannot be stated with certainty that this network type and configuration is the optimal choice. It is quite possible that better results can be obtained with other neural network types, or different topologies.

The third limitation is the high incidence of false negative responses recorded in this study. This high incidence of false negatives influences the sensitivity, and therefore the clinical acceptability of DPOAEs as a potential screening or diagnostic procedure. Further research is necessary to investigate possible neural network runs with different topology, different inputs and better measurement of DPOAEs to attempt to lower this high rate of false negatives. Changes in network topology had a significant improvement on false negative values from the previous study to the present, and further research might possibly reveal that DPOAEs can be used with ANNs to predict PTTs within acceptable levels of false negative responses for diagnostic purposes.

The last limitation identified in this study is the length of time needed for DPOAE measurement to obtain adequate information for one ear. The way in which this research project was constructed, 8 DP Grams were conducted in each ear. The pattern of all present and absent DPOAE responses from all eight DP Grams was used as input information. The duration of one DP Gram was about 2 minutes. It therefore took about 15 minutes per ear to obtain the necessary information. Even though it is still only half the time that is required to obtain a single threshold for one ear in ABR testing (Weber, 1994), it could be argued that 15 minutes is not such a rapid test of auditory functioning as was hoped for.

7.3 Recommendations for Future Research

There are a number of recommendations for studies attempting to improve PTT prediction accuracy of DPOAEs with ANNs using the method described in this study:

- The first recommendation is to increase the number of subjects. Figure 6.2 indicated that the “threshold” number of ears in every category to enable prediction accuracy of more than 75% is around 32 ears. With around 80 or more ears in a category, prediction accuracy could be expected to be very accurate, possibly more than 95%.
- The second recommendation is that the application of neural networks to this particular field of Audiology should be further investigated. Neural networks offer so many possibilities. It is possible that different types of networks or different types of configurations would yield more accurate predictions or lower false negative values. Slightly higher error tolerance levels such as 1, 2 or 3% may be able to generalize better and may improve prediction accuracy. It is also possible that other areas in Audiology could also benefit from artificial neural networks.

- The third recommendation is to experiment with lower "acceptance" criteria for DPOAE measurements, possibly a cumulative noise level slightly higher than –18 dB SPL (such as –10 dB SPL) or a DPOAE level slightly lower than 10dB above the noise floor (such as 5dB). If the larger number of responses obtained in this method is invalid and part of the irregular noise floor, then prediction accuracy of the neural network will be less accurate. In this case, the higher standards for test acceptance conditions can always be reintroduced, and neural network runs can continue with the higher standards, such as a DPOAE level of 10dB or a cumulative noise level of –18dB SPL and invalid responses can be discarded. If however, neural network prediction is more accurate due to the presence of more responses regarded as accepted, it could possibly influence the efficiency of DPOAE testing as it is currently conducted. When testing with higher "acceptance" criteria levels, responses regarded as "timed out" or "noisy" are lost forever. When testing with lower criteria, the validity of responses can always be determined later. It is possible that this technique might identify new levels for acceptable DPOAE measurement and improve prediction accuracy.

7.4 General Implications of the Study and Concluding Remarks

Audiologists are currently relying heavily on objective audiological tests to assess hearing ability in difficult-to-test populations. There are however, still many limitations in current objective procedures despite the enormous progress in the last few decades. Some of these limitations include the limited frequency area of objective hearing assessment, the expenses, time and expertise required, and the possibility of sedation. It is with much hope that many researchers turned to the investigation of

DPOAEs as a possible new rapid, objective, accurate and cost effective test of auditory functioning. The distortion product has been proven as an acceptable screening procedure. Otoacoustic emissions however, have never been used as a diagnostic test of hearing where specific thresholds for frequencies were determined due to shortcomings in conventional statistical methods (Kimberley et al., 1994; Lee et al., 1993).

The investigation of DPOAEs indicated strongly that DPOAEs are suitable as a diagnostic audiologic test of hearing. It is suggested that pure tone thresholds can be accurately predicted within 10dB as low as 500 Hz and for hearing levels of up to 65dB HL with ANNs. The successful application of ANNs in this field of Audiology opened the door to the development of an objective, rapid, accurate and economical test of hearing to aid in the assessment of difficult-to-test populations. It is strongly believed that this breakthrough will play a leading role in the efficiency with which the pediatric population will be assessed in the next decade.

REFERENCES

Abdala, C., 1998. "A developmental study of distortion product otoacoustic emission (2f1-f2) suppression in humans". *Hearing Research*, 121, pp. 125-138.

Abdala, C., 1999. "Distortion product otoacoustic emission (2f1-f2) amplitude growth in human adults and neonates". *Journal of the Acoustical Society of America* 107(1), pp.446-456.

Attias, J., Furst, M., Furman, V., Haran, I.R., Horowitz, G. & I. Bresloff, 1995. "Noise-Induced Otoacoustic Emission Loss With or Without Hearing Loss". *Ear & Hearing*, 16(6), pp. 612-618.

Avan, P. & P. Bonfils, 1993. "Frequency Specificity of Human Distortion Product Otoacoustic Emissions". *Audiology*, 32, pp. 12-26.

Avan, P., Elbez, M. & P. Bonfils, 1997. "Click-evoked otoacoustic emissions and the influence of high-frequency hearing losses in humans". *Journal of the Acoustical Society of America*, 101(5), pp. 2771-2777.

Balfour, P.B., Pillion, J.P. & A.E. Gaskin, 1998. "Distortion Product Emission and Auditory Brain Stem Response Measures of Pediatric Sensorineural Hearing Loss with Islands of Normal Sensitivity". *Ear & Hearing*, 19, pp. 436-472.

Block, M.G. & T.L. Wiley, 1994. "Overview and Basic Principles of Acoustic Immittance Measurements." in Katz, J. (Ed) 1994. *Handbook of Clinical Audiology*. 4th Ed. London: Williams & Willkins.

Blum, A. 1992. *Neural networks in C++. An Object-Oriented Framework for Building Connectionist Systems*, NY: John Wiley & Sons.

Bonfils, P., Avan, P., Francois, M., Marie, P., Trotoux, J. & P. Narcy, 1990. "Clinical Significance of Otoacoustic Emissions: A Perspective". *Ear & Hearing*, 11 (2), pp. 155-158.

Bonfils, P., Avan, P., Landias, P., Erminy, M. & B. Biacabe, 1997. "Statistical Evaluation of Hearing Screening by Distortion Product Otoacoustic Emissions". *Ann Otol Rhinol Laryngol,* 106, pp. 1052-1062.

Bonfils, P., Avan, P., Londero, A., Trotoux, J. & P. Narcy, 1991. "Objective Low-Frequency Audiometry by Distortion-Product Acoustic Emissions." *Archives of Otorhinolaryngology Head and Neck Surgery,* 117, pp. 1167-1171.

Bonfils, P., Piron, J.P., Uziel, A. & R. Pujol, 1988. "A correlative study of evoked otoacoustic emission properties and audiometric thresholds". *Archives of Otorhinolaryngology,* 245, pp. 53-56.

Bonfils, P & A. Uziel, 1989. "Clinical Applications of Evoked Acoustic Emissions: Results in Normally Hearing and Hearing-Impaired Subjects". *Ann Otol Rhinol Laryngol,* 98, pp. 326-331.

Brass, D. & D.T. Kemp, 1994. "Quantitative Assessment of Methods for the Detection of Otoacoustic Emissions." *Ear & Hearing*, 15, pp.378-389.

Bright, K.E., 1994. "Helpful Hints for OAE Measurement." *The Hearing Journal,* 47, pp. 10-61.

Brownell, W. 1990. "Outer Hair Cell Electromotility and Otoacoustic Emissions". *Ear & Hearing,* 11 (2), pp. 82-92.

Brunt, M.A., 1994. "Tests of Cochlear Function." in Katz, J. (Ed) 1994. *Handbook of Clinical Audiology.* 4th Ed. London: Williams & Willkins.

Cacace, A.T., McClelland, W.A., Weiner, J. & D.J. McFarland, 1996. "Individual Differences and the Reliability of 2f1-f2 Distortion-Product Otoacoustic Emissions: Effect of Time-of Day, Stimulus Variables, and Gender". *Journal of Speech and Hearing Research,* 39, pp. 1138-1148.

Cane, M.A., Donoghue. G.M. & M.E. Lutman, 1992. "The Feasibility of Using Oto-acoustic Emissions to Monitor Cochlear Function during Acoustic Neuroma Surgery." *Scandinavian Audiology*, 21, pp. 173-176.

Clark, J.G. 1981. "Uses and abuses of hearing loss classification." *Asha*, 23, pp. 493-500.

Collet, L., Gartner, M., Moulin, A., Kauffmann, I., Disant, F. & A. Morgon, 1989. "Evoked Otoacoustic Emissions and Sensorineural Hearing Loss." *Arch Otolaryngol Head Neck Surg*. 115, pp. 1060-1062.

Cope, Y. 1995. "Objective hearing tests". In McCormick, B. (Ed) 1995. *The Medical Practitioners Guide to Pediatric Audiology*. New York: Cambridge University Press.

Dallos, P. 1997. "Outer Hair Cells: The Inside Story" *Ann Otol Rhinol Laryngol*, 106, pp. 16-22.

Danhauer, J.L., 1997. "How Otoacoustic Emissions Testing Can Change Your Audiology Practice." *The Hearing Journal*, 50, pp. 62-69.

De Boer, E. 1983. "Power amplification in an active model of the cochlea. Short-wave case". *Journal of the Acoustical Society of America*, 73, 577-579.

De Waal, R. 1998. "The Use of Artificial Neural Networks to Predict Pure Tone Thresholds in Normal and Hearing-Impaired Ears with Distortion Product Otoacoustic Emissions." In Partial Fulfillment for the Degree M. Communication Pathology, Faculty of Arts, University of Pretoria, Pretoria, South Africa.

Dorn, P.A., Piskorski, P., Gorga, M.P., Neely, S.T. & D.H. Keefe, 1999. "Predicting Audiometric Status from Distortion Product Otoacoustic Emissions Using Multivariate Analyses". *Ear & Hearing*, 20, pp. 149-163.

Durrant, J.D. 1992. "Distortion-Product OAE Analysis: Is It Ready For Broad Clinical Use?" *The Hearing Journal*, 45 (11), pp. 42-45.

Elberling, C., Parbo, J., Johnsen, N.H. & P. Bagi, 1985. "Evoked Otoacoustic Emission: Clinical Application". *Acta Otolaryngology*, 421, pp. 77-85.

Ewing, I. & A. Ewing 1944. The Ascertainment of Deafness in infancy and early childhood." *Journal of Laryngology*, 59, pp. 309-333.

Fechtner, J.W. 1992. "Screening in the 1990's: Some Principles and Guidelines". In Bess, F.H. & J.W. Hall (Ed) 1992. *Screening Children for Auditory Function.* Nashville, Tennessee.

Ferraro, J.A. & J.D. Durrant, 1994. "Auditory Evoked Potentials: Overview and Basic Principles." in Katz, J. (Ed) 1994. *Handbook of Clinical Audiology.* 4th Ed. London: Williams & Willkins.

Fujita, O. 1998. "Statistical estimation of the number of hidden units for feedforward neural networks". *Neural Networks*, 11, pp. 851-859.

Gaskill, S.A. & A.M. Brown, 1990. "The behavior of the acoustic distortion product, 2f1-f2, from the human ear and its relation to auditory sensitivity". *Journal of the Acoustical Society of America,* 88(2), pp. 821-839.

Glattke, T.J., Pafitis, I.A., Cummiskey, C. & G.R. Herer, 1995. "Identification of Hearing Loss in Children and Young Adults Using Measures of Transient Otoacoustic Emission Reproducibility". *American Journal of Audiology,* 4, pp. 71-86.

Goodman, A. 1965. "Reference zero levels for pure-tone audiometer." *ASHA*, 7, pp. 262-263.

Gorga, M.P., Neely, S.T., Bergman, B. & K.L. Beauchaine, 1993. "Otoacoustic emissions from normal-hearing and hearing-impaired subjects: Distortion product responses". *Journal of the Acoustical Society of America,* 93(4), pp. 2050-2060.

Gorga, M.P., Stover, L., Neely, S.T. & D. Montoya, 1996. "The use of cumulative distributions to determine critical values and levels of confidence for clinical distortion product otoacoustic emission measurements." *Journal of the Acoustical Society of America,* 100(2), pp. 968-977.

Hall III, J.W., Baer, J.E., Chase, P.A. & M.K. Schwaber, 1993. Clinical application of otoacoustic emissions: What do we know about factors influencing measurement and analysis?" *Otolaryngology- Head and Neck Surgery,* 110(1), pp. 22-34.

Hall III, J.W. & D. Chandler, 1994. "Tympanometry in Clinical Audiology." in Katz, J. (Ed) 1994. *Handbook of Clinical Audiology.* 4th Ed. London: Williams & Wilkins.

Hall III, J.W. & H.G. Mueller III, 1997. *"Audiologists' Desk Reference. Vol I. Diagnostic Audiology Principles, Procedures, and Practices".* San Diego: Singular Publishing Group Inc.

Hann, T.H. & E. Steurer, 1996. "Much ado about nothing? Exchange rate forecasting: Neural Networks vs. linear models using monthly and weekly data". *Neurocomputing,* 10, pp. 323-339.

Hardy, W.G., Hardy, J.B. & C.H. Brinker, Frazier, T.M. & A Doughtery, 1962. "Auditory screening of infants." Annals of Otology, Rhinology and Laryngology, 71, pp. 759-766.

Harris, F.P., Lonsbury-Martin, B.L., Stagner, B.B., Coats. A.C. & G.K. Martin, 1989. "Acoustic Distortion Products in Humans: Systematic changes in amplitude as a function of f2/f1 ratio." *Journal of the Acoustical Society of America,* 85(1), pp. 220-229.

Harris, F.P, & R. Probst, 1991. "Reporting Click-Evoked and Distortion-Product Otoacoustic Emission Results with Respect to the Pure-Tone Audiogram". *Ear & Hearing,* 12 (6), pp. 399-405.

Harris, F.P. & R. Probst 1997. "Otoacoustic Emissions and Audiometric Outcomes". In Robinette, M.S. & T.J. Glattke (Ed) 1997. *Otoacoustic Emissions, Clinical Applications*. New York: Thieme.

Harrison, W.A. & S.J. Norton, 1999. "Characteristics of Transient Evoked Otoacoustic Emissions n Normal-Hearing and Hearing-Impaired Children." *Ear & Hearing*, 20, pp. 75-86.

Hawley, D., Johnson, J & D. Raina, 1993. "Artificial Neural Systems: A New Tool for Financial Decision Making". In Trippi, R.R. & E. Turban (Ed) 1993. *Neural Networks in Finance and Investing: Using Artificial Neural Networks to Improve Real-World Performance*. Chicago: Probus Publishing Company.

He, N & R. Schmiedt, 1996. "Effects of aging on the fine structure of the 2f1-f2 acoustic distortion product". *Journal of the Acoustical Society of America,* 99 (2), pp. 1002-1015.

He, N & R. Schmiedt, 1997. "Fine structure of the acoustic distortion product: Effects of primary level and frequency ratios". *Journal of the Acoustical Society of America,* 101(6), pp. 3554-3565.

Hiramatsu, A. 1995. "Training Techniques for Neural Network Applications in ATM". *IEEE Communications Magazine*, October 1995, pp. 58-63.

Hornik, K., Stinchcombe, M., & H. White, 1989. "Multilayer Feedforward Networks are Universal Approximators". *Neural Networks*, 2, pp. 359-366.

Hurley, R.M. & F.E. Musiek, 1994. "Effectiveness of Transient-Evoked Otoacoustic Emissions (TEOAEs) in Predicting Hearing Level". *Journal of the American Academy of Audiology,* 5 (3), pp. 195-203.

Hussain, D.M., Gorga, M.P., Neely, S.T., Keefe, D.H. & J. Peters 1998. "Transient Evoked Otoacoustic Emissions in Patients with Normal Hearing and in Patients with Hearing Loss". *Ear & Hearing*, 19, pp. 434-449.

Jerger, J.F., Burney, P., Mauldin, L. & B. Crump 1974. "Predicting hearing loss from the acoustic reflex". *Journal of Speech and Hearing Disorders*, 39, pp.11-21.

Jerger, J. & S. Jerger, 1980. "Measurement of hearing in adults." In Paparella, M.M. & D.A. Shumrick (Eds). Otolaryngology, 2nd ed. Philadelphia: W.B. Saunders.

Karzon, R.K., Garcia, P., Peterein, J.L. & G.A. Gates. 1994. Distortion Product Otoacoustic Emissions in the Elderly." *The American Journal of Otology*, 15(5), pp. 596-605.

Kemp, D.T. 1978. "Stimulated acoustic emissions from within the human auditory system". *Journal of the Acoustical Society of America,* 64: 1386-1391.

Kemp, D.T. 1979. "Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea. *Archives of Otorhinolaryngology*, 224, pp. 37-45.

Kemp, D.T. 1997. "Otoacoustic Emissions in Perspective". In Robinette, M.S. & T.J. Glattke (Ed) 1997. *Otoacoustic Emissions, Clinical Applications*. New York: Thieme.

Kemp, D.T. & S. Ryan, 1993. "The Use of Transient Evoked Otoacoustic Emissions in Neonatal Hearing Screening Programs". *Seminars in Hearing,* 14(1), pp. 30-44.

Kemp, D.T., Ryan, S. & P. Bray, 1990. "A Guide to the Effective Use of Otoacoustic Emissions". *Ear & Hearing,* 11(2), pp. 93-105.

Kim, D.O., Sun, X.M., Jung, M.D. & G. Leonard. 1997. "A New Method of Measuring Distortion Product Otoacoustic Emissions Using Multiple Tone Pairs: Study of Human Adults". *Ear & Hearing*, 18, pp. 277-285.

Kimberley, B.P., Brown, D.K. & J.B. Allen 1997. "Distortion Product Emissions and Sensorineural Hearing Loss". In Robinette, M.S. & T.J. Glattke (Ed) 1997. *Otoacoustic Emissions, Clinical Applications*. New York: Thieme.

Kimberley, B.P., Hernadi, I., Lee, A.M. & D.K. Brown, 1994(b). "Predicting Pure Tone Thresholds in Normal and Hearing-Impaired Ears with Distortion Product Emission and Age". *Ear & Hearing,* 15(3), pp. 199-209.

Kimberley, B.P., Kimberley, B.M & L. Roth. 1994(a). "A Neural Network Approach to the Prediction of Pure Tone Thresholds with Distortion Product Emissions". *ENT Journal,* 73(11), pp. 812-823.

Kimberley, B.P. & D.A. Nelson, 1989. "Distortion product emissions and sensorineural hearing loss". *The Journal of Otolaryngology,* 18 (7), pp. 365-369.

Klimasauskas, C. 1993. "Applying Neural Networks". In Trippi, R.R. & E. Turban (Ed) 1993. *Neural Networks in Finance and Investing: Using Artificial Neural Networks to Improve Real-World Performance.* Chicago: Probus Publishing Company.

Knight, R.D. & D.T. Kemp, 1999(a). "Indications of different distortion product otoacoustic emission mechanisms from a detailed f1,f2 area study". *Journal of the Acoustical Society of America,*107(1), pp. 457-473.

Knight, R.D. & D.T. Kemp, 1999(b). "Relationships between DPOAE and TEOAE amplitude and phase characteristics". *Acoustical Society of America,* 106(3), pp. 1420-1435.

Koivunen, P., Uhari, M., Laitakari, K., Alho O.P. & J. Luotonen, 2000. "Otoacoustic Emissions and Tympanometry in Children with Otitis Media." *Ear & Hearing*, 21, pp. 212-217.

Konkle, D.F. & J.T. Jacobson, 1991. "Hearing Loss in Newborns, Infants, and Young Children." in Rintelmann, W.F. (Ed) 1991. *Hearing Assessment.* Austin, Texas: Pro-ed.

Kossl, M. & G.S. Boyan, 1998. "Acoustic Distortion Products from the Ear of a Grasshopper." *Journal of the Acoustical Society of America.* 104(1), pp. 326-335.

Koza, J.R., Bennett, F.H., Andre, D. & M.A. Keane, 1999. "Genetic Programming III." San Francisco: Morgan Kaufmann Publishers.

Kraus, N., Kileny, P. & T. McGee, 1994. "Middle Latency Auditory Evoked Potentials." In Katz, J. (Ed) 1994. *Handbook of Clinical Audiology.* 4th Ed. London: Williams & Willkins.

Krishnamurti, S., 2000. "Relationship between Distortion Product Otoacoustic Emissions and Frequency Modulation Difference Limens." *Ear & Hearing* , 21, pp. 305-309.

Kulawiec, J.T. & M.S. Orlando, 1995. "The Contribution of Spontaneous Otoacoustic Emissions to the Click Evoked Otoacoustic Emissions." *Ear & Hearing*, 16, pp. 515-520.

Kummer, P., Janssen, T. & Arnold, W. 1998. "The level and growth behavior of the 2f1-f2 distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss." *Journal of the Acoustical Society of America*, 103(6), pp. 3431-3444.

Lasky, R.E., 1998(a). "Distortion product otoacoustic emissions in human newborns and adults. I. Frequency effects." *Journal of the Acoustical Society of America*, 103(2), pp. 981-991.

Lasky, R.E., 1998(b). "Distortion product otoacoustic emissions in human newborns and adults. II. Level effects." *Journal of the Acoustical Society of America*, 103(2), pp. 992-1000.

Lee, A., Kimberley, B. & D. Brown, 1993. "Developing Distortion Product Emission Measurements for Clinical Applications". *Journal of Speech Language Pathology and Audiology*, 17 (1), pp. 15-25.

Leedy, P.D, 1993. "*Practical Research, Planning and Design.*"5th Ed. NY: Macmillan Publishing Co.

Lonsbury-Martin, B.L. 1994. "Introduction To Otoacoustic Emissions." *The American Journal of Otology,* 15 (1), pp. 1-3.

Lonsbury- Martin, B.L., Cutler, W.M. & G.K. Martin, 1991. "Evidence for the influence of aging on distortion product emissions in humans." *Journal of the Acoustical Society of America*, 89, 1749-1759.

Lonsbury-Martin, B.L., Harris, F.P., Stagner, B.B., Hawkins, M.D. & G.K. Martin, 1990. "Distortion Product Emissions in Humans I: Basic Properties in Normally Hearing Subjects". *Ann Otol Rhinol Laryngol,* 99, pp. 3-14.

Lonsbury-Martin, B.L. & G.K. Martin, 1990. "The Clinical Utility of Distortion-Product Otoacoustic Emissions". *Ear & Hearing,* 11 (2), pp. 144-154.

Lonsbury-Martin, B.L., Martin, G.K. & M.L. Whitehead. 1997. "Distortion Product Otoacoustic Emissions". In Robinette, M.S. & T.J. Glattke (Ed) 1997. *Otoacoustic Emissions, Clinical Applications*. New York: Thieme.

Lonsbury-Martin, B.L., McCoy, M.J., Whitehead, M.L. & G.K. Martin, 1992. "Otoacoustic Emissions: Future Directions For Research and Clinical Applications". *The Hearing Journal,* 45 (11), pp. 47-52.

Martin, F.N. 1984. "*Introduction to Audiology*."3rd Ed. Prentence Hall Inc, Englewood Cliffs, N.J. 07632.

Martin, G.K., Ohlms, L.A., Franklin, P.H., Harris, F.P. & B.L. Lonsbury-Martin, 1990(a). "Distortion Product Emissions in Humans III. Influence of Sensorineural Hearing Loss." *Ann Otol Rhinolaryngol*, 99, pp. 30-42.

Martin, G.K., Probst, R. & B.L. Lonsbury-Martin, 1990(b). "Otoacoustic Emissions in Human Ears: Normative Findings". *Ear & Hearing,* 11 (2), pp. 106-120.

Mauermann, M., Uppenkam & B. Kollmeier 1999(a). "Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (DPOAE) fine structure in humans. I. Fine structure and higher-order DPOAE as a function of the frequency ratio f2/f1". *Journal of the Acoustical Society of America,* 106(6), pp 3473-3483.

Mauermann, M., Uppenkamp, S., P.J.W. van Hengel & B. Kollmeier 1999(b). "Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (DPOAE) fine structure in humans. II. Fine structure for different shapes of cochlear hearing loss. *Journal of the Acoustical Society of America,* 106(6), pp. 3484-3491.

Mauk, G.W. & T.R. Behrens. 1993. "Historical, Political and Technological Context Associated with Early Identification of Hearing Loss". *Seminars in Hearing.* 14(1), pp. 1-13.

Medsker, L., Turban, E. & R. Trippi, 1993. "Neural Network Fundamentals for Financial Analysts". In Trippi, R.R. & E. Turban (Ed) 1993. *Neural Networks in Finance and Investing: Using Artificial Neural Networks to Improve Real-World Performance.* Chicago: Probus Publishing Company.

Metz, D.E., Schiavetti, N. & S.D. Knight, 1992. "The Use of Artificial Neural Networks to Estimate Speech Intelligibility From Acoustic Variables: A Preliminary Analysis". *Journal of Communication Disorders,* 25, pp. 43-53.

Mills, D.M, 1997. "Interpretation of distortion product otoacoustic emission measurements. I. Two stimulus tones". *Journal of the Acoustical Society of America,* 102(1), pp. 413-429.

Moulin, A. 2000a. "Influence of the primary frequency ratio on distortion product otoacoustic emissions amplitude. I. Intersubject variability and consequences on the DPOAE-gram." *Journal of the Acoustical Society of America,* 107(3), pp. 1460.

Moulin, A. 2000b. "Influence of the primary frequency ratio on distortion product otoacoustic emissions amplitude .II. Interrelations between multicomponent DPOAEs, tone burst evoked OAEs and spontaneous OAEs." *Journal of the Acoustical Society of America,* 107(3), pp. 1471.

Moulin, A., Bera, C. & L. Collet, 1994. "Distortion Product Otoacoustic Emissions and Sensorineural Hearing Loss". *Audiology,* 33, pp. 305-326.

Muller, B. & J. Reinhardt, 1990. *Neural Networks: An Introduction.* London: Springer-Verlag.

Musiek, F.E., Berenstein, S.P., Hall III, J.W. & M.K. Schwaber, 1994. "Auditory Brainstem Response: Neurodiagnostic and intraoperative applications." in Katz, J. (Ed) 1994. *Handbook of Clinical Audiology.* 4th Ed. London: Williams & Willkins.

Nakajima, H.H., Mountain, D.C. & A.E. Hubbard. 1998. "Nonlinear characteristics of electrically evoked otoacoustic emissions". *Hearing Research,* 122, pp. 109-118.

Nelson, M.M. & W.T. Illingworth, 1991. *A Practical Guide to Neural Nets.* New York: Addison-Wesley Publishing Company, Inc.

Nielsen, L.H., Popelka, G.R., Rasmussen, A.N. & P.A. Osterhammel, 1993. "Clinical Significance of probe tone frequency ratio on distortion product otoacoustic emissions." *Scandinavian Audiology,* 22, pp. 159-164.

Nieschalk, M., Hustert, B. & W. Stoll, 1998. "Distortion Product Otoacoustic Emissions in Middle-Aged subjects with Normal versus Potentially Presbyacusic High-Frequency Hearing Loss." *Audiology*, 37, pp. 83-99.

Northern, J.L. 1991. "*Hearing in Children.*" Baltimore, MD: Williams & Willkins.

Northern, J.L. & S.A. Gabbard, 1994. "Acoustic Reflex Measurements." in Katz, J. (Ed) 1994. *Handbook of Clinical Audiology.* 4th Ed. London: Williams & Willkins.

Norton, S.J. 1993. "Application of Transient Evoked Otoacoustic Emissions to Pediatric Populations". *Ear & Hearing,* 14(1), pp. 64-73.

Norton, S.J., Schmidt, A.R. & L.J. Stover, 1990. "Tinnitus and Otoacoustic Emissions: Is There a Link?". *Ear & Hearing,* 11 (2), pp. 159-166.

Norton, S.J. & L.J. Stover, 1994. "Otoacoustic Emissions An Emerging Clinical Tool." in Katz, J. (Ed) 1994. *Handbook of Clinical Audiology.* 4th Ed. London: Williams & Willkins.

Norton, S.J. & J.E. Widen, 1990. "Evoked Otoacoustic Emissions in Normal-Hearing Infants and Children: Emerging Data and Issues". *Ear & Hearing,* 11 (2), pp. 121-127.

Odom, M & R. Shara, 1993. "A Neural Network Model for Bankruptcy Prediction". In Trippi, R.R. & E. Turban (Ed) 1993. *Neural Networks in Finance and Investing: Using Artificial Neural Networks to Improve Real-World Performance.* Chicago: Probus Publishing Company.

Ohlms, L.A., Lonsbury-Martin, B.L. & G.K. Martin 1990. "Acoustic-distortion products: Separation of sensory from neural dysfunction in sensorineural hearing loss in human beings and rabbits." *Otolaryngology-Head and Neck Surgery,* 104(2), pp. 159-174.

Osterhammel, P.A., Nielsen, L.H. & A.N. Rasmussen, 1993. "Distortion Product Otoacoustic Emissions, The Influence of the Middle Ear Transmission". *Scandinavian Audiology,* 22, pp. 111-116.

Popelka, G.R., Karzon, R.K. & E.M. Arjmand, 1995. "Growth of the 2f1-f2 Distortion Product Otoacoustic Emission for Low-Level Stimuli in Human Neonates". *Ear & Hearing,* 16(2). pp. 159-165.

Popelka, G.R., Karzon, R.K. & R.A. Clary. 1998. "Identification of Noise Sources That Influence Distortion Product Otoacoustic Emission Measurements in Human Neonates. *Ear & Hearing,* 19, pp. 319-328.

Prieve, B.A., Gorga, M.P. & S.T. Neely. 1996. "Click-and tone-burst otoacoustic emissions in normal-hearing and hearing-impaired ears". *Journal of the Acoustical Society of America,*99(5), pp. 3077-3086.

Probst, R. & R. Hauser, 1990. "Distortion Product Otoacoustic Emissions in Normal and Hearing-Impaired Ears." *American Journal of Otolaryngology*, 11, pp. 236-243.

Probst, R. & F.P. Harris. 1993. "Transiently evoked and Distortion Product Otoacoustic Emissions". *Arch Otolaryngol Head Neck Surg*, 119, pp. 858-860.

Quinonez, R.E. & M.R. Crawford. 1997. "Electrophysiologic changes in preterm neonates: Auditory brain stem responses and distortion product otoacoustic emissions". *Ann Otol Rhinol Laryngol*, 106, pp. 721- 728.

Raghupathi, W., Schkade, L, & B. Raju, 1993. "A Neural Network Approach to Bankruptcy Prediction". In Trippi, R.R. & E. Turban (Ed) 1993. *Neural Networks in Finance and Investing: Using Artificial Neural Networks to Improve Real-World Performance.* Chicago: Probus Publishing Company.

Rahimian, E., Singh, S., Thammachote, T & R. Virmani, 1993. "Bankruptcy Prediction by Neural Network". In Trippi, R.R. & E. Turban (Ed) 1993. *Neural Networks in Finance and Investing: Using Artificial Neural Networks to Improve Real-World Performance.* Chicago: Probus Publishing Company.

Rao, V. & H. Rao, 1995. *"C++ Neural Networks and Fuzzy Logic."* 2nd Ed. NY: MIS:Press.

Rasmussen, A.N., Popelka, G.R., Osterhammel, P.A. & L.H. Nielsen, 1993. "Clinical significance of relative probe tone levels on distortion product otoacoustic emissions." *Scandinavian Audiology*, 22, pp. 223-229.

Robinette, M.S. 1992. "Otoacoustic Emissions In Cochlear Vs. Retrocochlear Auditory Dysfunction". *The Hearing Journal,* 45 (11), pp. 32-34.

Robinette, M.S. 1994, "Integrating Audiometric Results." in Katz, J. (Ed) 1994. *Handbook of Clinical Audiology.* 4th Ed. London: Williams & Willkins.

Ruggero, M.A. 1993. "Distortion in those good vibrations". *Current Biology,* 3, pp. 755-758.

Ruth, R.A., 1994. "Electrocochleography." in Katz, J. (Ed) 1994. *Handbook of Clinical Audiology.* 4th Ed. London: Williams & Willkins.

Salchenberger, L., Cinar, E. & N. Lash, 1993. "Neural Networks: A New Method for Predicting Thrift Failures". In Trippi, R.R. & E. Turban (Ed) 1993. *Neural Networks in Finance and Investing: Using Artificial Neural Networks to Improve Real-World Performance.* Chicago: Probus Publishing Company.

Schwartz, D.M. & J.A. Schwartz, 1991. "Auditory Evoked Potentials in Clinical Pediatrics." in Rintelmann, W.F. (Ed) 1991. *Hearing Assessment.* Austin, Texas: Pro-ed.

Shera, C.A. & J.J. Guinan. 1998. "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs". *Journal of the Acoustical Society of America,* 105(2) pp. 782-798.

Siegel, J.H., 1995. "Cross-Talk in Otoacoustic Emission Probes." *Ear & Hearing,* 16(2), pp. 150-156.

Smurzynski, J., Leonard, G., Kim, D.O., Lafreniere, D.C. Marjorie & M.D. Jung, 1990. "Distortion Product Otoacoustic Emissions in Normal and Impaired Adult Ears." *Arch Otolaryngol Head Neck Surg,* 116, pp. 1309-1316.

Spektor, Z., Leonard, G., Kim, D.O., Marjorie, Jung, M.D. & J. Smurzynski, 1991. "Otoacoustic Emissions in Normal and Hearing-Impaired Children and Normal Adults". *Laryngoscope,* 101. Pp. 965-976.

Stach, B.A., Wolf, S.J. & L. Bland, 1993. "Otoacoustic Emissions as a Cross-Check in Pediatric Hearing Assessment: Case Report". *Journal of the American Academy of Audiology,* 4, pp. 392-398.

Stevens, J.C., Webb, H.D., Hutchinson, J., Connell, J., Smith, M.F. & J.T. Buffin, 1990. "Click Evoked Otoacoustic Emissions in Neonatal Screening". *Ear & Hearing,* 11 (2), pp. 128-133.

Stover, L., Gorga, M.P. & S.T. Neely, & Montoya 1996(a). "Toward optimizing the clinical utility of distortion product otoacoustic emission measurements". *Journal of the Acoustical Society of America,* 100(2), pp. 956-967.

Stover, L. Neely, S.T. & M.P. Gorga, 1996(b). "Latency and multiple sources of distortion product otoacoustic emissions". *Journal of the Acoustical Society of America,* 99(2). pp. 1016-1024.

Subramaniam, M., Henderson, D. & V. Sponger, 1994. "The Relationship Among Distortion Product Otoacoustic Emissions, Evoked Potential Thresholds, and Outer Hair Cells Following Interrupted Noise Exposures." *Ear & Hearing*, 15(4), pp. 299-309.

Suzuki, T & Y. Ogiba, 1961. "Conditioned orientation reflex audiometry." *Archives of Otolaryngology*, 74, pp. 84-90.

Talmadge, C.L., Long, G.R., Tubis, A. & S. Dhar. 1998. "Experimental confirmation of the two-source interference model for the fine structure of distortion product otoacoustic emissions". *Journal of the Acoustical Society of America,* 105(1), pp. 275-292.

Tam, K. & M. Kiang, 1993. "Managerial Applications of Neural Networks: The Case of Bank Failure Prediction". In Trippi, R.R. & E. Turban (Ed) 1993. *Neural Networks in Finance and Investing: Using Artificial Neural Networks to Improve Real-World Performance.* Chicago: Probus Publishing Company.

Tanaka, Y., O-Uchi, T., Arai, Y. & J. Suzuki. 1987. "Otoacoustic Emission as an Indicator in Evaluating Inner Ear Impairments". *Acta Otolaryngol* (Stockh), 103, pp. 644-648.

Teleschi, F.F., Roth, J., Stagner, B.B., Lonsbury-Martin, B.L. & T.J. Balkany, 1995(a). "Patterns of Evoked Otoacoustic Emissions Associated With Acoustic Neuromas." *Laryngoscope,* 105, pp. 675-682.

Teleschi, F.F., Widick, M.P., Lonsbury-Martin, B.L. & M.J. McCoy, 1995(b). "Monitoring Cochlear Function Intraoperatively using Distortion Product Otoacoustic Emissions. " *The American Journal of Otology*, 16(5), pp. 597-608.

Ueda, H. 1998. "Do click-evoked otoacoustic emissions have frequency specificity? *Journal of the Acoustical Society of America,* 105(1), pp. 306-310.

Ventry, I.M. & N. Schiavetti, 1980. *Evaluating Research in Speech Pathology and Audiology.* London: Addison-Wesley Publishing Company.

Vinck, B.M., Cauwenberge, P.B., Corthals, P. & E. De Vel. 1998. "Multi-variant Analysis of Otoacoustic Emissions and Estimation of Hearing Thresholds: Transient Evoked Otoacoustic Emissions". *Audiology*, 37, pp. 315-334.

Vinck, B.M., Vel, E.D., Xu, Z.M. & P.B.Van Cauwenberge 1995. "Distortion Product Otoacoustic Emissions: A Normative Study." *Audiology*, 35, pp. 231-245.

Vohr, B.R., White, K.R., Maxon, A.B. & M.J. Johnson. 1993. "Factors affecting the interpretation of transient evoked otoacoustic emission results in neonatal hearing screening". *Seminars in Hearing*, 14(1), pp. 57-72.

Weber, B.A., 1994. "Auditory Brainstem Response: Threshold Estimation and Auditory Screening." in Katz, J. (Ed) 1994. *Handbook of Clinical Audiology.* 4th Ed. London: Williams & Willkins.

Whitehead, M.L., Lonsbury-Martin, B.L., Martin, G.K. & M.J. McCoy. 1996. "Otoacoustic Emissions: Animal Models and Clinical Observations". In Van De Water, T.R., Popper, A.N. & R.R. Fay (Ed), 1996. *Clinical Aspects of Hearing*. New York: Springer-Verlag.

Yantis, P.A. 1994. "Puretone Air-Conduction Threshold Testing". in Katz, J. (Ed) 1994. *Handbook of Clinical Audiology.* 4th Ed. London: Williams & Willkins.

Yuan, J.L. & T.L. Fine, 1998. "Neural-Network Design for Small Training Sets of High Dimension". *IEEE Transactions On Neural Networks*, 9(2), pp. 266-280.

Zhang, M. & P.J. Abbas, 1997. "Effects of middle ear pressure on otoacoustic emission measures". *Journal of the Acoustical Society of America,* 102(2), pp. 1032-1037.

Zhao, F. & D. Stephens. 1998. "Analysis of Notches in Audioscan and DPOAEs in Subjects with Normal Hearing". *Audiology,* 37, pp. 335-343.

APPENDICES:

Appendix A: The interview

PERSONAL INFORMATION

Name: **Subject file #**

Date of birth: ____/____ /____ **Gender** **M** **F**

YY MM DD

1.1 INFORMATION REGARDING HEARING STATUS

Complaints of a hearing loss?

If Yes, what is the degree of the hearing loss?

When was the onset of the hearing loss?

What was the cause of the hearing loss?

Is there a history of hearing loss in the family?

If yes, what was the cause: genetic /trauma /unknown?

Complaints of current middle ear problems?

If Yes, what is the current status of the middle ear problem, for example, does the subject experience any hearing loss, pain or fluid discharge.

What was the frequency of past middle ear infections.

Any allergies?

Complaints of tinnitus? If yes, what is the perceived pitch and loudness level of the tinnitus?

Complaints of vertigo? If yes, how severe and how frequent?

Has the subject been exposed to high noise levels?

If yes, amount of noise exposure:

Type of noise exposed to for example gun shots, machinery, loud music.

What types of medication does the subject currently use?