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.

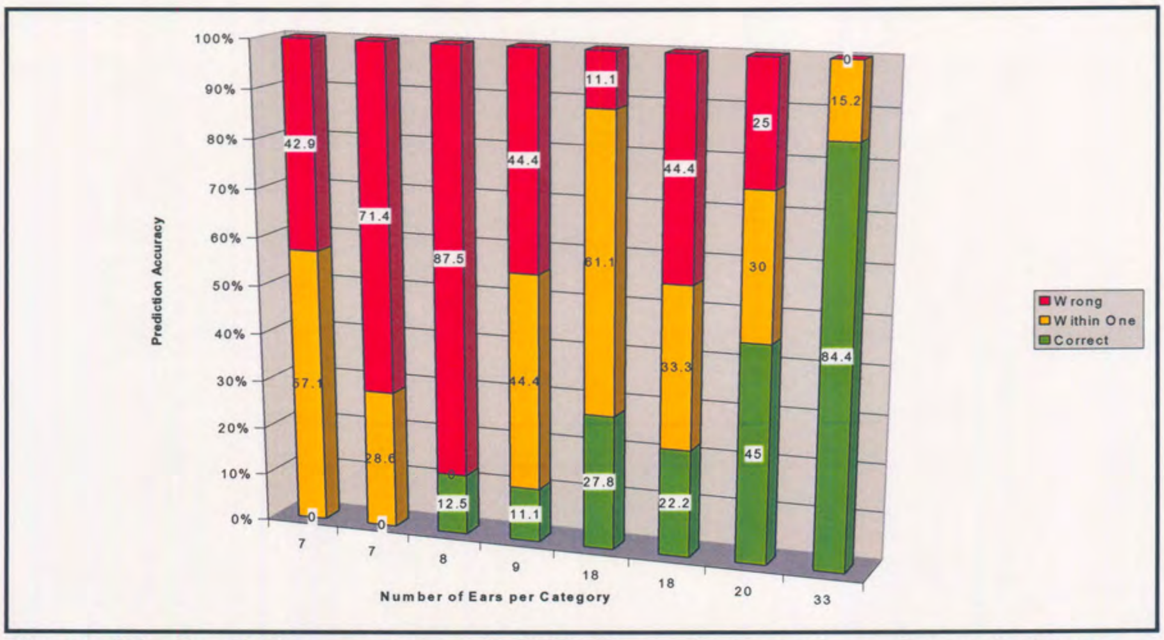


Figure 5.1: Prediction accuracy of 4000 Hz against number of ears in every category.

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.

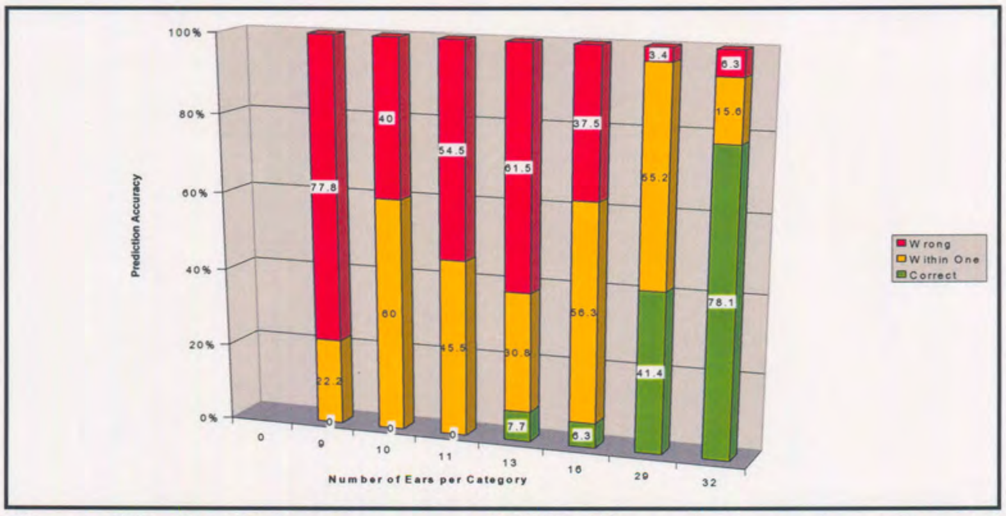


Figure 5.2: Prediction accuracy of 2000 Hz against number of ears in every category.

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.

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	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.

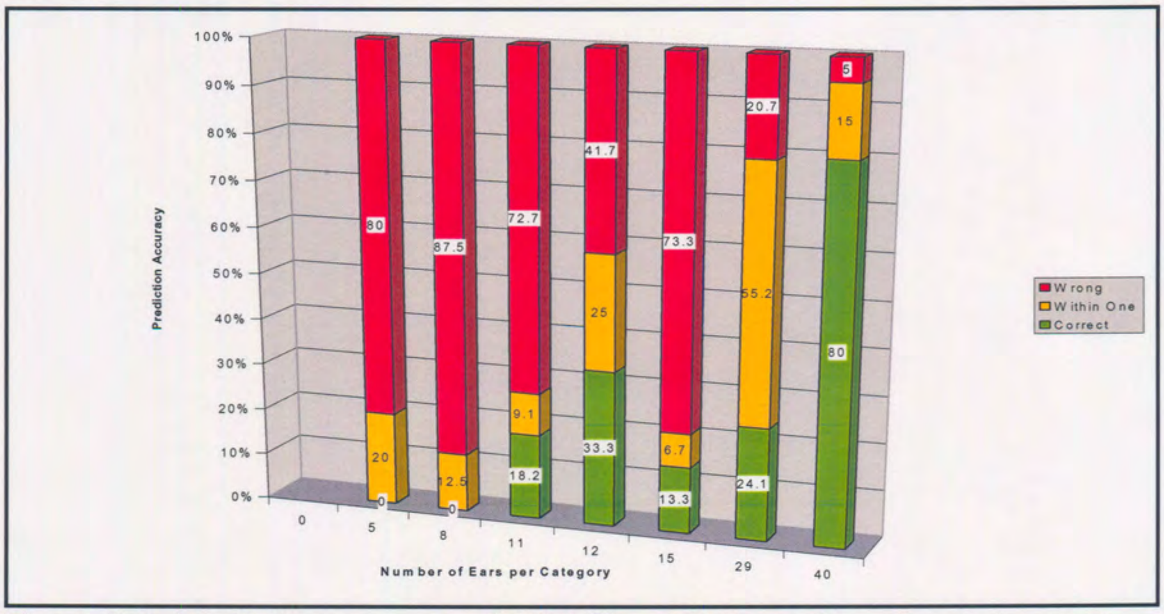


Figure 5.3: Prediction accuracy of 1000 Hz against number of ears in every category.

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.

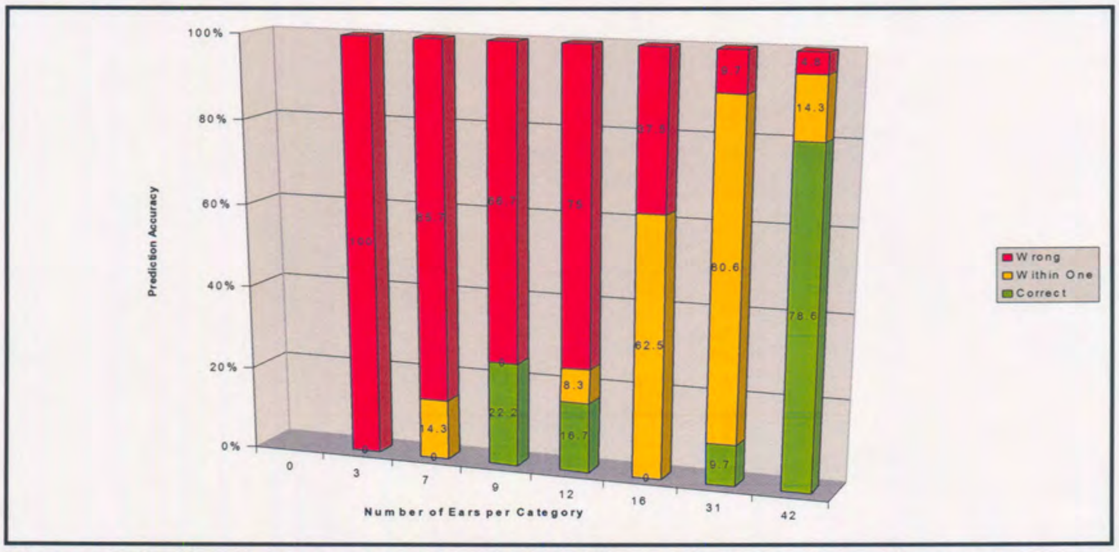


Figure 5.4: Prediction accuracy of 500 Hz against number of ears in every category.

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.

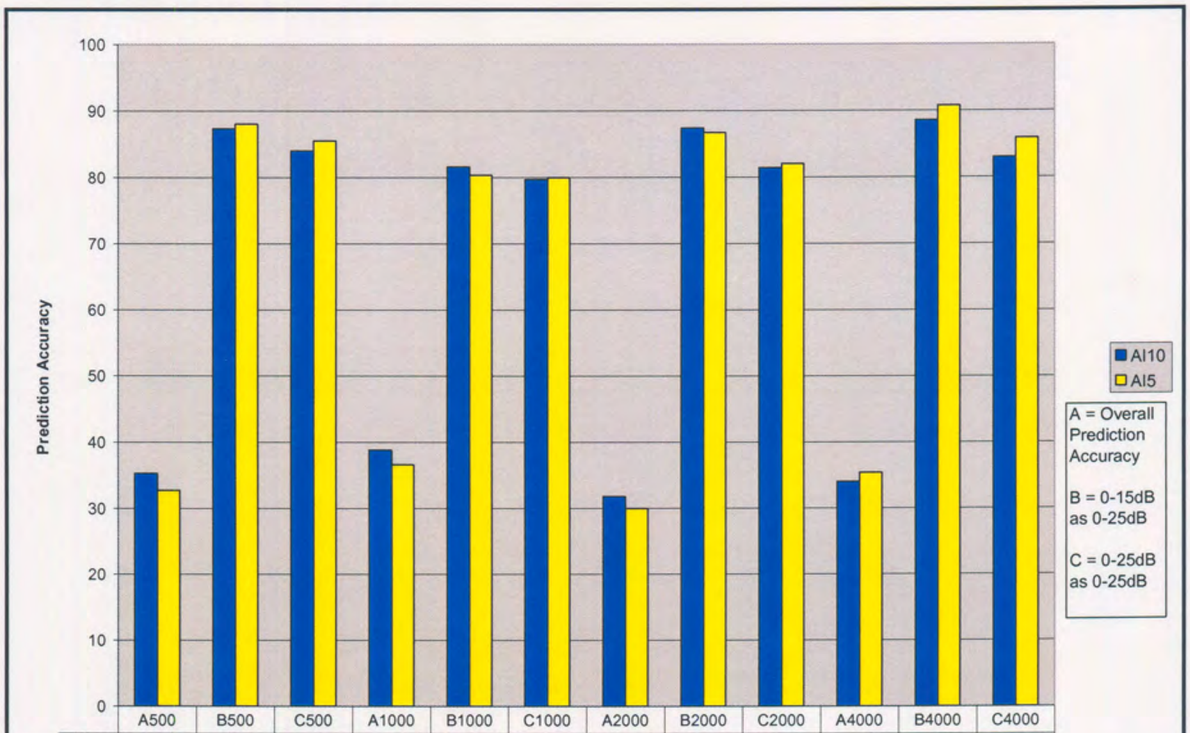
Subject age was always included in ANN training and prediction because it has been found to improve PTT prediction accuracy in a number of previous studies (Lonsbury-Martin et al. 1991; Kimberley et al. 1994a; Kimberley et al. 1994b; De Waal, 1998).

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".

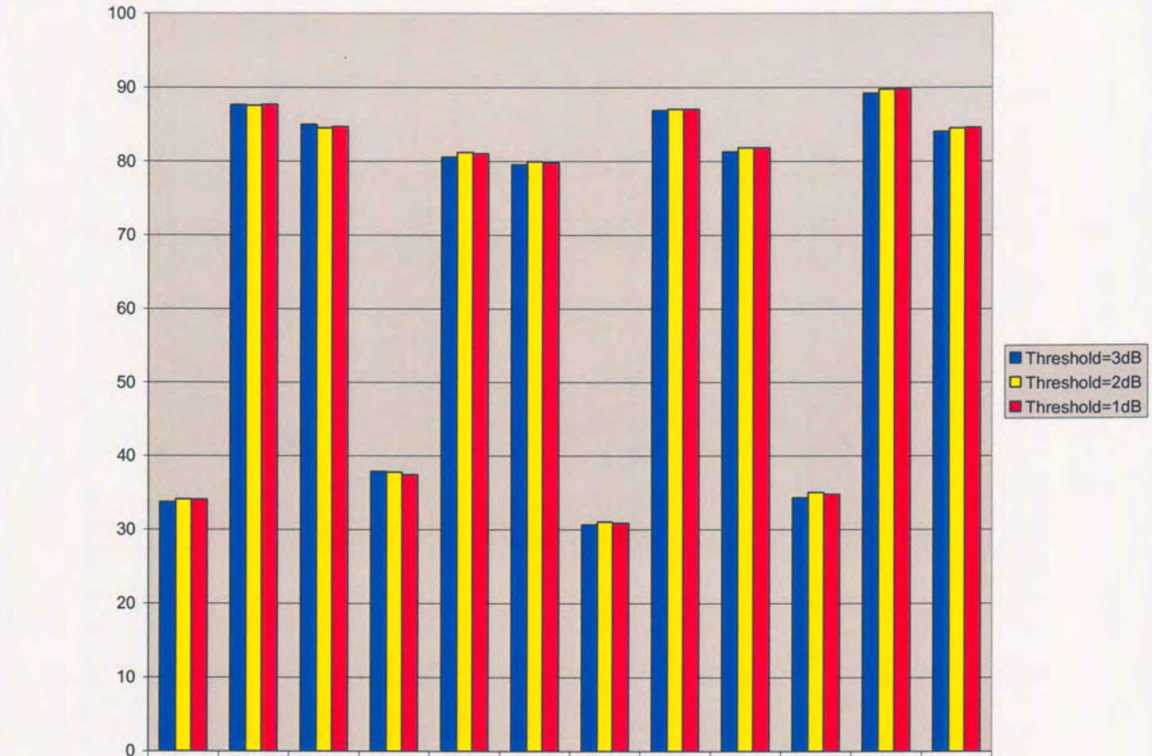
5.7.2 The Effect of DPOAE Threshold Defined as 1, 2 or 3 dB Above the Noise Floor on PTT Prediction Accuracy.

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.



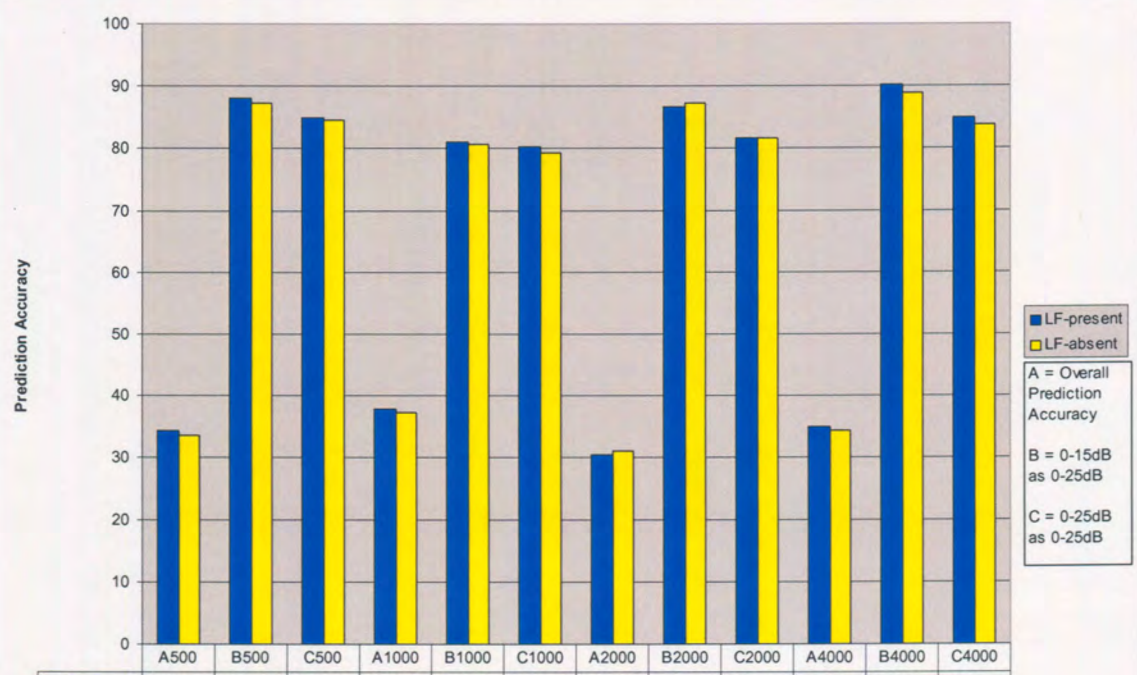
	A0500	B0500	C0500	A1000	B1000	C1000	A2000	B2000	C2000	A4000	B4000	C4000
Threshold=3dB	33.7172	87.6621	85.0138	37.8681	80.5625	79.5069	30.625	86.8819	81.2639	34.2778	89.1944	84.0139
Threshold=2dB	34.1181	87.5625	84.5139	37.7639	81.1875	79.9653	31.0069	87.0833	81.8472	35.0208	89.75	84.4861
Threshold=1dB	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:

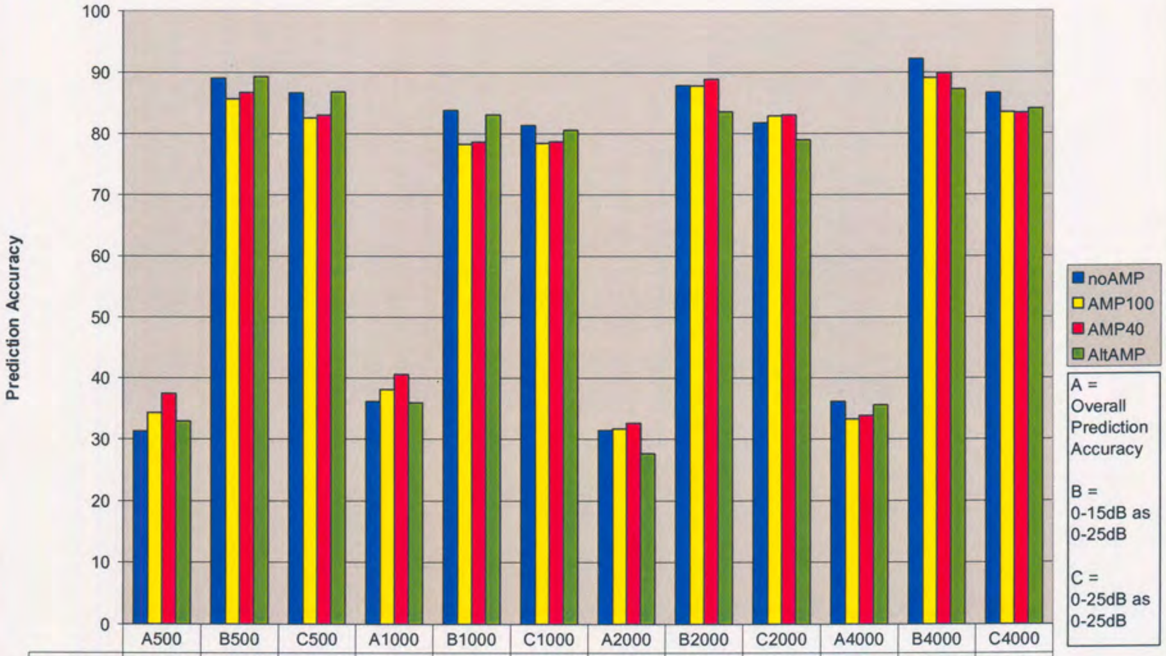
- AMP 100 presented the amplitude of the DPOAE to the ANN as a fraction of 100 (see 4.8.1.4.1).
- AMP 40 presented the amplitude of the DPOAE to the ANN as a fraction of the largest DPOAE response measured in this study (39dB)(see 4.8.1.4.2).
- ALT AMP depicted DPOAE amplitude with the dummy variable 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.



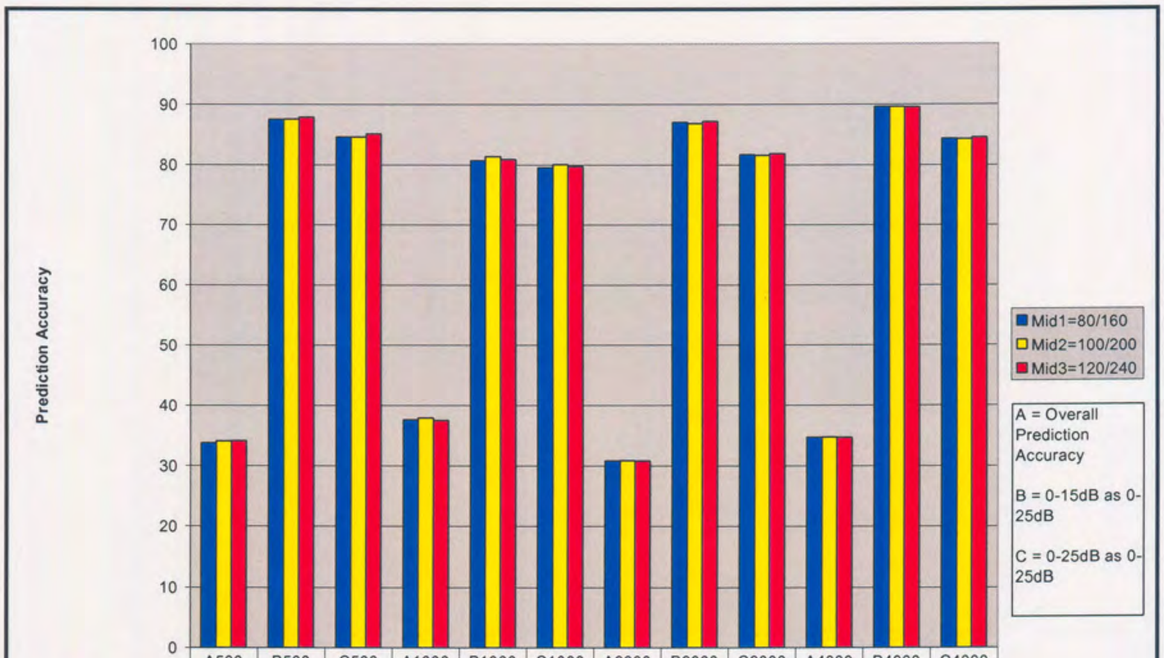
	A500	B500	C500	A1000	B1000	C1000	A2000	B2000	C2000	A4000	B4000	C4000
noAMP	31.25926	89.00926	86.62963	36.13889	83.78704	81.34259	31.39815	87.83333	81.75	36.11111	92.22222	86.57407
AMP100	34.27778	85.65741	82.53704	38.10185	78.26852	78.43519	31.66667	87.80556	82.89815	33.23148	89.05556	83.44444
AMP40	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.

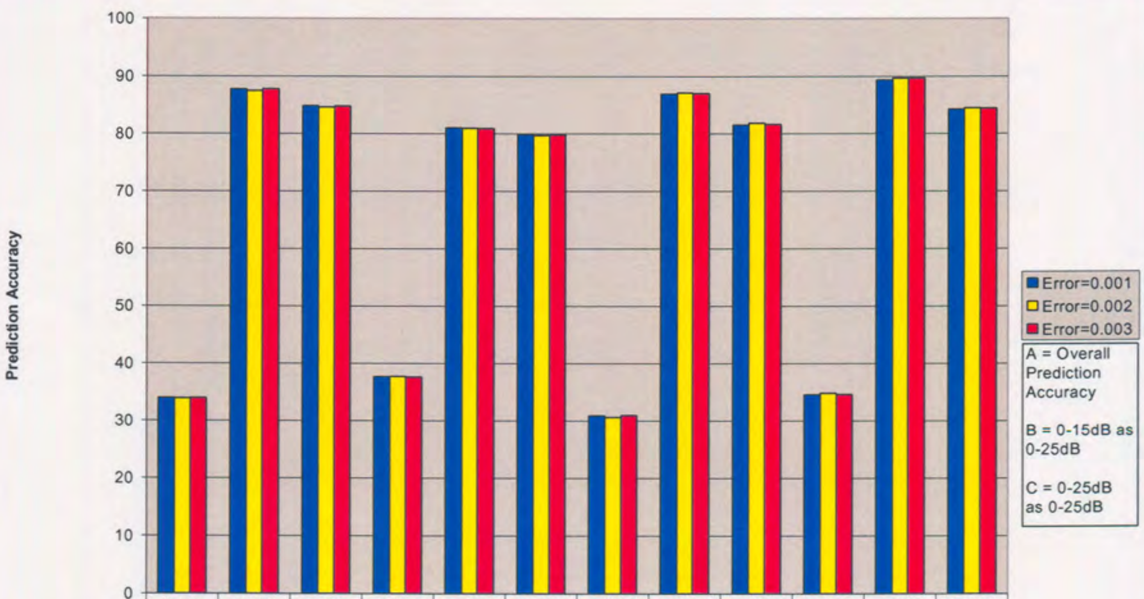
There was virtually no effect of middle neuron count on prediction accuracy for normal or impaired hearing. Prediction accuracy for the three categories was always within 1%.

"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.