UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING AND THE ENVIRONMENT

Institute of Sound and Vibration Research

Factors Affecting Speech Recognition in Noise and Hearing Loss in Adults With a Wide Variety of Auditory Capabilities

by

Sheetal Purushottam Athalye

Thesis for the degree of Doctor of Philosophy

October 2010

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND THE ENVIRONMENT
INSTITUTE OF SOUND AND VIBRATION RESEARCH

Doctor of Philosophy

Factors Affecting Speech Recognition in Noise and Hearing Loss in Adults with a Wide Variety of Auditory Capabilities

By Sheetal Purushottam Athalye

Studies concerning speech recognition in noise constitute a very broad spectrum of work including aspects like the cocktail party effect or observing performance of individuals in different types of speech-signal or noise as well as benefit and improvement with hearing aids. Another important area that has received much attention is investigating the inter-relations among various auditory and non-auditory capabilities affecting speech intelligibility. Those studies have focussed on the relationship between auditory threshold (hearing sensitivity) and a number of suprathreshold abilities like speech recognition in quiet and noise, frequency resolution, temporal resolution and the non-auditory ability of cognition.

There is considerable discrepancy regarding the relationship between speech recognition in noise and hearing threshold level. Some studies conclude that speech recognition performance in noise can be predicted solely from an individual's hearing threshold level while others conclude that other supra-threshold factors such as frequency and/or temporal resolution must also play a role. Hearing loss involves more than deficits in recognising speech in noise, raising the question whether hearing impairment is a uni- or multi-dimensional construct. Moreover, different extents of hearing loss may display different relationships among measures of hearing ability, or different dimensionality.

The present thesis attempts to address these three issues, by examining a wide range of hearing abilities in large samples of participants having a range of hearing ability from normal to moderate-severe impairment. The research extends previous work by including larger samples of participants, a wider range of measures of hearing ability and by differentiating among levels of hearing impairment.

Method: Two large multi-centre studies were conducted, involving 103 and 128 participants respectively. A large battery of tests was devised and refined prior to the main studies and implemented on a common PC-based platform. The test domains included measurement of hearing sensitivity, speech recognition in quiet and noise, loudness perception, frequency resolution, temporal resolution, binaural hearing and localization, cognition and subjective measures like listening effort and self-report of hearing disability. Performance tests involved presentation of sounds via circum-aural earphones to one or both ears, as required, at intensities matched to individual hearing impairments to ensure audibility. Most tests involved measurements centred on a low frequency (500 Hz), high frequency (3000 Hz) and broadband. The second study included some refinements based on analysis of the first study. Analyses included multiple regression for prediction of speech recognition in stationary or fluctuating noise and factor analysis to explore the dimensionality of the data. Speech recognition performance was also compared with that predicted using the Speech Intelligibility Index (SII).

Findings: Findings from regression analysis pooled across the two studies showed that speech recognition in noise can be predicted from a combination of hearing threshold at higher frequencies (3000/4000 Hz) and frequency resolution at low frequency (500 Hz). This supports previous studies that conclude that resolution is important in addition to hearing sensitivity. This was also confirmed by the fact that SII (representing sensitivity rather than resolution) under-predicted difficulties observed in hearing-impaired ears for speech recognition in noise. Speech recognition in stationary noise was predicted mainly by auditory threshold while speech recognition in fluctuating noise was predicted by a combination having a larger contribution from frequency resolution. In mild hearing losses (below 40 dB), speech recognition in noise was predicted mainly by hearing threshold, in moderate hearing losses (above 40 dB) it was predicted mainly by frequency resolution when combined for two studies. Thus it can be observed that the importance of auditory resolution (in this case frequency resolution) increases and the importance of the audiogram decreases as the degree of hearing loss increases, provided speech is presented at audible levels. However, for all degrees of hearing impairment included in the study, prediction based solely on hearing thresholds was not much worse than prediction based on a combination of thresholds and frequency resolution. Lastly, hearing impairment was shown to be multi-dimensional; main factors included hearing threshold, speech recognition in stationary and fluctuating noise, frequency and temporal resolution, binaural processing, loudness perception, cognition and self-reported hearing difficulties. A clinical test protocol for defining an individual *auditory profile* is suggested based on these findings.

Conclusions: Speech recognition in noise depends on a combination of audibility of the speech components (hearing threshold) and frequency resolution. Models such as SII that do not include resolution tend to over-predict somewhat speech recognition performance in noise, especially for more severe hearing impairments. However, the over-prediction is not great. It follows that for clinical purposes there is not much to be gained from more complex psychoacoustic characterisation of sensorineural hearing impairment, when the purpose is to predict or explain difficulty understanding speech in noise. A conventional audiogram and possibly measurement of frequency resolution at 500 Hz is sufficient. However, if the purpose is to acquire a detailed individual auditory profile, the multidimensional nature of hearing loss should not be ignored. Findings from the present study show that, along with loss of sensitivity and reduced frequency resolution ability, binaural processing, loudness perception, cognition and self-report measures help to characterize this multi-dimensionality. Detailed studies should hence focus on these multiple dimensions of hearing loss and incorporate measuring a wide variety of different auditory capabilities, rather than inclusion of just a few, in order gain a complete picture of auditory functioning.

Frequency resolution at low frequency (500 Hz) as a predictive factor for speech recognition in noise is a new finding. Few previous studies have included low-frequency measures of hearing, which may explain why it has not emerged previously. Yet this finding appears to be robust, as it was consistent across both of the present studies. It may relate to differentiation of vowel components of speech. The present work was unable to confirm the suggestion from previous studies that measures of temporal resolution help to predict speech recognition in fluctuating noise, possibly because few participants had extremely poor temporal resolution ability.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

DECLARATION OF AUTHORSHIP

I, ……………………………………………………………………………., [please print name]

declare that the thesis entitled [enter title]

…………………………………………………………………………………………………

…………………………………………………………………………………………………….

and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

this work was done wholly or mainly while in candidature for a research degree at this University;

where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;

where I have consulted the published work of others, this is always clearly attributed;

where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;

I have acknowledged all main sources of help;

where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

none of this work has been published before submission, or [delete as appropriate] parts of this work have been published as: [please list references]

Signed: ………………………………………………………………………..

Date:……………………………………………………………………………….

ACKNOWLEDGMENTS

I am indebted to my supervisor, Professor Mark E Lutman for his guidance and excellent supervision throughout this project. I cannot thank him enough for being so patient and understanding.

I would also like to thank all the partner institutions from HEARCOM project, in particular Linköping University (Sweden), Vrije Universiteit Medical Centre (Amsterdam), Academisch Medisch Centrum (Universiteit van Amsterdam) and Hörzentrum-Oldenburg (Germany) for providing me the data and making both the multi-centre studies possible. Guidance on data input received from Ms. Thamar van Esch (AMC) and help with data acquisition software from Dr. Daniel Berg (HörTech, Oldenburg) is also greatly appreciated. The help and advice on thesis structure received from of Drs. Daniel Rowan and Miyuki Morioka is also gratefully acknowledged.

I am most grateful to the ISVR (Rayleigh Scholarship) and HEARCOM for providing financial support throughout the duration of this project. I would like to thank Sue Robinson and all the other staff in the audiology department at Royal South Hants Hospital at Southampton. I would also like to express my appreciation for all my colleagues in HABC, especially, Maureen Mew, Vicky and Stefan.

The place and my experience wouldn't have been the same without my friends Hadeel, Gouping, Mona and Srikant; where else would I have experienced five people talking in completely different languages on the phone at the same time! I cannot thank them enough for all the interesting discussions, group activities and for making this a fun experience as well as my stay in Southampton a memorable one

I would like to thank my family for their continual support and encouragement. Last, but by no means the least, I wish to thank my husband Mandar for his selfless love, support and patience. The magnitude of his contribution cannot be expressed in a few words, so it is to him that this thesis is dedicated.

Chapter One

Introduction

1.1 Relation between communication, hearing loss and speech recognition in noise

The development of human civilization is made possible to a great extent by man's ability to share experiences, to exchange ideas and to transmit knowledge from one generation to another; in other words – his ability to communicate with other men. Communication is thus the process of exchange of information between two individuals and spoken communication is the most commonly used form of communication. Society is becoming increasingly communication oriented especially in today's man to machine world. Speech and hearing are undoubtedly the two most important aspects of communication. Perhaps the best example of the overwhelming importance of hearing in human society is a comparison of the social attitudes of the blind to those of the deaf. Generally, blind people tend to get along with their fellow human beings despite their handicap. But the deaf, who can still read and write, often feel cut off from society. A deaf person deprived of his primary means of communication tends to withdraw from the world and live within himself.

Thus deafness or hearing impairment has significant impact on communication and hence on one's life. Further, a loss in ability to hear has various aspects and is a multifaceted impairment and can be of different types. The discussion in the thesis is mainly concerned with the sensorineural type of hearing loss. The most severe consequence of this type of hearing loss is the reduced ability to understand speech. For people with hearing loss, this ability is more affected when considered in the presence of noise versus in quiet. This was perhaps best described by Plomp (1978) in developing his speech reception threshold model. According to Plomp, "Every hearing loss for speech can be interpreted as the sum of a loss class A (attenuation), characterized by a reduction of the levels of both speech signal and noise, and a loss D (distortion), comparable with a decrease in speech-to-noise ratio." This attenuation factor is more related to SRT in quiet while the distortion factor is related to SRT in noise *(SRT:*

speech recognition threshold: defined as the dB level at which an individual can hear a certain percentage of words, usually 50% correctly). It is this distortion factor that causes most concern in hearing loss individuals since a hearing aid can usually compensate for class-A hearing losses, primarily in quiet, but not for class-D hearing losses, primarily in noise.

Also it was revealed in a study by Davis (1989) that among the various deficits, the greatest difficulty (26%) faced by adults is to hear speech in the presence of noise. Similarly this difficulty is also considered by many people to be the greatest handicap associated with their hearing impairment (Kramer *et al.*, 1998).

1.2 Basic common rationale for speech in noise tests

The rationale for such tests is that most hearing impairment affects the inner ear and as mentioned previously, is associated with a loss of ability to recognise speech against a background of noise: i.e. an individual needs a more favourable speech-to-noise ratio (SNR) to obtain a criterion level of performance, typically 50% correct. This rationale is based on Plomp's model described above. The model contains two generalised parameters to account for hearing loss. These two parameters describe the hearing loss for speech in quiet and the hearing loss for speech in noise. Hawkins and Stevens (1950) showed that at higher noise levels the threshold of speech in a background of white noise increases at the same rate as the noise level. This finding was generalised for normal hearing listeners and hearing impaired listeners by Plomp in his model. It further becomes evident that the SRTn (speech recognition threshold in noise) of a listener does not depend on absolute presentation level, once it is above absolute threshold, but only on the ratio between speech and noise (Wagener, 2003).

1.3 Psychoacoustic basis for speech-in-noise tests

Fletcher (1940) suggested that the peripheral auditory system behaves as if it contains a bank of band pass filters; with continuously overlapping centre frequencies. These filters are called auditory filters. Fletcher thought that the basilar membrane provided the basis for the auditory filter. Each location on the basilar membrane responds to a limited range of frequencies, so that each different point corresponds to a filter with a different centre frequency. Recent data are consistent with this point of view (Moore, 1986), although the additional amplification by the outer hair cells was not known at the time of Fletcher's work. This mechanism is responsible for the steep filter skirts in normal ears . (Moore, 1986)

Thus, when trying to detect a signal in noise background, the listener is assumed to make use of a filter with a centre frequency close to that of the signal. This filter will pass the signal but remove a great deal of the noise. Only the components in the noise which pass through the filter will have any effect in masking the signal. It is usually assumed that the threshold for the signal is determined by the amount of noise passing through the auditory filter; specifically, threshold is assumed to correspond to the signal-to-noise ratio (SNR) at the output of the filter. This set of assumptions is called the power spectrum model of masking (Patterson & Moore, 1986). However, in hearing impaired individuals these filters widen leading to a lower (worse) signal-to-noise ratio at the output of the filter (Patterson & Moore, 1986).

1.4 Studies related to speech recognition in noise

This difficulty mentioned above is often discussed in association with 'cocktail party effect' which is a general ability to hear out conversations by filtering out or concentrating on speech of one person in group of many or when the ambient noise levels are high. People with hearing impairment struggle to carry out this filtering. Moreover, since we are more sensitive to problems related to speech and communication, it is not surprising that there is a long history of studies devoted to understanding this difficulty. However; the theoretical viewpoints of the different studies tend to vary. In other words, the approach by which each study or group of studies has arrived at the possible explanation for understanding this difficulty is different. One of the pioneering studies in this respect is by Festen and Plomp (1983) who tried to observe relations between auditory functions in impaired hearing using measures of hearing threshold, speech recognition, frequency and temporal resolution. The stand taken by them was that basic properties of the hearing process like frequency resolution, non-linearity etc. were based on theoretical considerations and physiological data. They speculated that a more direct way of understanding the auditory functions was by correlating the results for a number of auditory tests for a group of subjects.

Their study thus put forth two important ideas through which understanding of speech intelligibility in noise can be facilitated. Firstly, they pointed out the variation in performance seen in hearing impaired population leads to significant inter-individual differences and this aspect may help improve our understanding since is not seen or seen to a much lower degree in the normal hearing population (Festen and Plomp, 1981). Secondly, they advocated the use of a test battery approach to further investigate and understand this difficulty.

Most of the studies by other authors that followed have adopted the first (Pavlovic *et al.*, 1984; Nelson *et al.*, 2007) or the second (van Rooij & Plomp, 1990; Jerger *et al.*, 1991; Humes *et al.*, 1994; Divenyi *et al.*, 1997; George *et al*, 2006) approach in order to understand this difficulty of listening to speech in presence of noise, seen in the majority of adult hearing impaired. Though the two approaches cannot be considered exclusive, it is the preference given to one or the other approach that varied.

A study subsequent to Festen and Plomp (1983) was by Pavlovic *et al.,* (1984) who adopted the first idea above. The Pavlovic study or the more recent Nelson *et al.*(2007) study discussed the use of understanding of speech difficulty to improve amplification based on inter-individual differences (Nelson *et al.*, 2007) or articulation index (AI) (Pavlovic *et al.,*. 1984). So they tried to apply their understanding of this deficit to specify suitable amplification.

The others (van Rooij & Plomp, 1990; Jerger *et al.*, 1991; Humes *et al.*, 1994; Divenyi *et al.*, 1997; George *et al.*, 2006, Dreschler and Plomp, 1985; Lutman, 1987) used a test battery approach and tried to understand the different interrelations between the auditory measures, each from a slightly different perspective. Jerger *et al.*(1991) focussed on speech audiometric and neuropsychological measures. They tried to understand the difficulty using a limited set of variables including age, hearing loss and cognition. Humes *et al.*(1994) investigated the same, however using a wide range of speech materials. The study by Divenyi (1997a) used different measures to see the effect of age and lateral asymmetry rather than hearing loss. They attributed the difficulty of understanding speech in presence of noise to auditory segregation or the inability to perform 'auditory scene analyses' whereas George *et al.*(2006) attributed it to suprathreshold deficits. According to George *et al.*(2006) when speech and noise are well above threshold and have similar overall spectra, the distortion component (which gives rise to this difficulty in the first place) is considered to be a reflection of suprathreshold deficits in hearing. Van Rooij and Plomp (1990) investigated relations of speech perception tests with auditive and cognitive components. However the focus was to observe validity and manageability of a test battery comprising the above three components. Similarly, Dreschler and Plomp (1985) studied the various interrelations for a younger group of subjects and Lutman's study did so for subjects with different groups of hearing loss.

From the above studies it becomes evident that understanding of difficulty of speech recognition in noise, in the hearing impaired population broadly constitutes two aspects:

- Understanding speech recognition in noise itself, typically identifying predictors of speech recognition performance
- Understanding relations between different auditory capabilities (thus trying to understand this difficulty by investigating various auditory measures beyond speech recognition, in other words surpassing speech recognition and observing hearing loss more globally)

Having established that, the different studies in the present research are categorised into the two groups, based on above:

Group I	Studies of *explanatory factors* for speech recognition in noise
Group II	Studies examining *interdependence* of the auditory capabilities or *factors*

The definitions of the groups of studies are as follows:

Group I- Studies of explanatory factors

This group of studies is basically related to understanding speech recognition performance in hearing impaired people. The studies included here revolve around explaining speech recognition performance in terms of various other auditory capabilities including hearing threshold level and frequency/temporal resolution. This may include speech recognition scores in different types of noise such as continuous or fluctuating. An underlying question is the extent to which speech recognition in noise can be predicted from hearing threshold levels (the audiogram), or whether other suprathreshold capabilities are required for prediction.

Group II-Studies examining interdependence of auditory factors

Hearing loss is multifaceted and involves capabilities other than speech recognition and hearing threshold. So it needs to be explained fully as a sensory impairment rather than just from the point of view of speech recognition. Thus group II studies focus on exploring this multidimensionality related to hearing loss or hearing difficulties. It is more generalised than group I where the focus is on speech recognition alone. To a great extent, the same studies from group I are reviewed, however from a viewpoint that extends beyond speech recognition in noise.

Thus group I relates to explaining and recovering the different factors and aspects that could possibly underlie deficient speech recognition in noise, while group II relates to discussing and exploring sensorineural hearing loss more globally from the viewpoint of hearing deficits rather than just speech recognition or hearing threshold or one particular auditory ability .

The two groups of studies are now outlined in the following section. They are discussed in greater detail in Chapter 2.

1.4.1 Group I- Studies of explanatory factors

As mentioned above, a number of studies (Festen and Plomp, 1983; Pavlovic, 1983; Dreschler and Plomp, 1985; Lutman, 1987; van Rooij & Plomp, 1990; Jerger *et al.*, 1991; Humes *et al.,* 1994; Divenyi *et al.*, 1997a, b, c; George *et al.*, 2006) have attempted to study speech recognition in noise and its relation to other auditory capabilities including auditory threshold or sensitivity, frequency, temporal, spatial resolution and cognition among others. Other studies have exclusively studied the relation of speech recognition to hearing threshold (e.g. Killion 1997, Killion *et al.*, 2004, Nelson *et al.*, 2007). The various studies have used different methods and have varied implications and findings; however based on their main conclusions they are divided into the following three classes of studies.

Class A: Studies concluding that speech recognition in noise can be predicted well from measures of hearing threshold level (Humes and Roberts, 1990, Jerger *et al.,* 1991; Humes *et al.*, 1994; Divenyi *et al.,* 1997a). In other words, it is unnecessary to consider other (supra-threshold) measures.

Class B: Studies concluding that speech recognition in noise can only be predicted by including measures other than threshold. These studies can be further sub-divided.

Class B: i] Studies concluding that speech recognition in noise can be predicted on average from measures of hearing threshold level, but the wide spread of data suggests that other measures (not measured in the study) may have a role as well (Killion 1997; Killion *et al.*, 2004; Nelson *et al.*, 2007; Pavlovic, 1984; Plomp and Mimpen 1979).

Class B: ii] Studies concluding that speech recognition in noise can only partly be predicted from measures of hearing threshold level; suprathreshold or other measures (obtained in the study) also play an important part **(**Festen and Plomp 1983, van Rooij & Plomp 1990, George etal. 2006, Dreschler and Plomp 1985, Lutman 1987).

The difference between classes B (i) and (ii) is whether the presence of contributing factors additional to hearing threshold (suprathreshold or others) were implicated by inference (i) or actually measured (ii). Thus the class B (i) studies focussed more on implications of the relation between hearing threshold and speech recognition in noise while B (ii) studies focussed more on the relation between speech recognition and suprathreshold and other factors like cognition, age etc. Similarly the main difference between the classes A and B (ii) is the extent of prediction of speech recognition based on hearing thresholds. While class A includes studies that indicate dependence on threshold to a maximum extent with minimal or no contribution from other factors, class B (ii) includes studies that implicate threshold to a lesser extent, with additional contribution from other factors. Thus, they can be looked upon as a spectrum of studies varying in the extent to which suprathreshold abilities are required to explain speech recognition in noise.

This classification is mainly for the purpose of discussion so that the differences and commonalities between them become evident and any obvious discrepancies can be outlined. It should be remembered that there are numerous studies studying various aspects of speech recognition, however the scope of the thesis includes mainly those studies that have attempted to use a test battery approach. Also the focus here is on the interrelations between the different tests rather than the individual tests themselves. Further, the main issue that is addressed collectively by these studies is prediction of speech intelligibility scores based on an individual's hearing threshold level. It can be seen that there is considerable variation in the findings. The class A studies (Jerger *et al.*, 1991; Humes and Roberts, 1990, Humes *et al.*, 1994; Divenyi *et al.*, 1997 a, b) reveal that speech recognition can be predicted to a great extent based on an individual's hearing loss or sensitivity. Humes *et al.*, (1994) quantified this extent as 70-75% of the variance, while Divenyi *et al.*, (1997b) estimated 85% and Jerger *et al.,* (1991) suggested 75%. Further, Divenyi *et al.*; (1997a) found hearing threshold to be major factor that was responsible for differences in auditory performance including speech recognition, especially in the elderly. Thus these studies consider hearing threshold level to be the essential factor for predicting speech recognition in hearing impaired people.

Class B (i) studies (Killion 1997, Killion *et al.*, 2004) have compared hearing loss (pure tone averages) to SNR loss, a term used to denote the increase in SNR needed to achieve a certain percentage of speech recognition (usually 50%). In general, an individual with a hearing loss requires a more favourable or better SNR than a normal hearing individual to perform at a criterion level in a given acoustic environment.

Killion (1997) reported that subjects with a 40 dB pure-tone-average (PTA) hearing loss typically showed a 5 dB SNR loss; those with a 60 dB PTA typically showed a 7 dB SNR loss, increasing to almost 12 dB for an 80 dB PTA. This leads us to believe that there is a monotonic relation between the two. However in other studies (Killion and Niquette, 2000; Killion *et al.*, 2004) it was noted that some subjects with hearing losses of 40 to 60 dB have almost no SNR loss. Thus their data sets revealed a spread of 15 to

20 dB in SNR loss for similar pure tone average losses. In other words, listeners with 40-60 dB pure tone averages show SNR loss ranging between 2-20 dB. This finding was generalised by Killion *et al.* (2004). Other studies (Lyregaard, 1982; Dirks *et al.*, 1982; Killion, 1997; Killion and Niquette, 2000) also support similar findings indicating a wide range of SNR loss in persons with similar pure tone hearing losses. These findings are shown graphically below (re-plotted from the study by Lyregaard *et al.*, 1982)

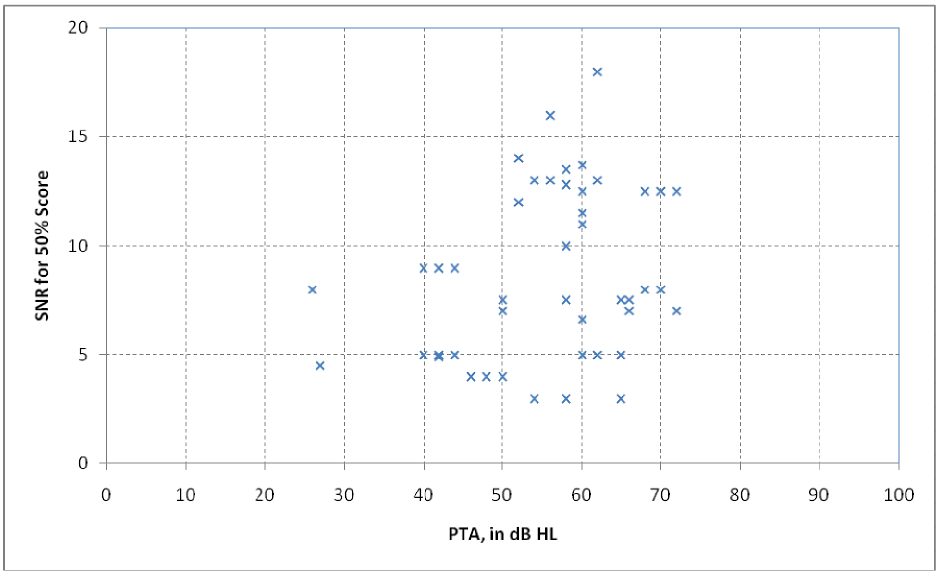


Figure 1.1: Wide variation seen in SNR (dB) as a function of hearing threshold (dB HL) replotted from Lyregaard, 1982.

Pavlovic (1984) used an alternative approach to study speech recognition and hearing threshold. He applied the AI (articulation index) procedure to audiograms of normal and hearing impaired subjects to predict SRT in noise and compared this with measured SRT in noise. He found that the subjects especially with relatively severe loss exhibited a disproportionate loss in speech discrimination compared to that predicted on the basis of AI. They concluded that suprathreshold distortion factors were present in addition to loss of audibility, for the more severe losses. He inferred that, for mild and moderate hearing losses, reduced audibility as represented in the AI model was sufficient to explain SNR loss. In other words, supra-threshold distortions were unimportant for mild/moderate hearing loss. However, he used filtered speech/noise materials which perhaps meant that for some test conditions there was a hidden element of frequency resolution ability and his inferences regarding AI may be an over simplification.

However, in spite of different methods, the class B (i) studies highlight that the relation between speech recognition in noise scores and hearing threshold level includes much unexplained variance and hence it may not be straightforward to predict the score of one based on other for an individual.

In summary, the class B(i) studies (Killion 1997, Killion *et al.*, 2004, Nelson *et al.*, 2007, Pavlovic 1984, Plomp and Mimpen 1979) reported that while normal hearing listeners show more or less similar scores, significant individual differences for speech recognition scores exist among those with sensorineural hearing loss (Pavlovic, 1984). These differences are known to exist with or without any differences in audibility; this variability also becomes greater when they are tested in noise rather than in quiet. In other words speech intelligibility cannot always be predicted well based on only a person's hearing threshold. The variation among individuals appears to be larger than could be explained simply by measurement uncertainty.

Lastly, class B (ii) studies (Festen and Plomp 1983, van Rooij and Plomp 1990, George *et al.,* 2006, Lutman 1987) revealed that speech recognition threshold can be predicted only partially by threshold. There could be numerous other factors responsible. The other factors responsible for the prediction include frequency resolution (Festen and Plomp, 1983), temporal resolution (George *et al.*, 2006) and both frequency and temporal resolution (Lutman 1987, Dreschler and Plomp 1985) and cognition (van Rooij & Plomp 1990). The relative contribution varies in the different studies. It should also be remembered that these factors are not independent and they are also correlated with hearing threshold level, making interpretation complex.

Studies in subclass B(i) especially (Killion 1997, Killion *et al.*, 2004) have exclusively studied hearing threshold and speech recognition. Thus they are limited in their findings in that the two are not proportionately related. The subclass B(ii) studies have gone a step further and were able to observe that speech recognition can be predicted partly by thresholds and determined by statistical inference the other factors that could possibly be playing a role in this prediction.

From the above, it becomes evident that although there have been numerous studies investigating the relation between speech recognition in noise and hearing threshold level, the existing empirical evidence is inconsistent and hence inconclusive. Thus a clear discrepancy between the different outcomes of the studies A and B is evident. It is still unclear to what extent speech recognition can be predicted by audibility and whether beyond audibility there could be any other factors responsible for the same. These questions form the basis of the first two aims in the thesis:

I] To try to resolve the discrepancy among studies of classes A and B and thus test whether or not audibility can predict or explain variation satisfactorily in speech recognition scores in noise.

II] If audibility can only partly explain the differences among hearing impaired individuals, what are the other (suprathreshold) factors responsible for the differences. (The issue of suprathreshold auditory capabilities is mainly considered here, while others e.g. cognition are discussed later.)

It should be noted that the discrepancies among the different studies could be due to various reasons including differences in test batteries, methodologies, subjects, statistical techniques etc. These aspects will be discussed in more detail in the next chapter along with significant details of each study. The purpose of the present chapter is to merely outline these discrepancies, which serve to underpin the aims in the thesis. Once these have been outlined here, the next chapter further discusses limitations in various studies and their scope for extension in the present study.

Further it can be noted from the class B studies that the relation between hearing threshold level and speech recognition is not exactly proportionate. Having observed this, there is little evidence to show to what extent this non-proportionality exists across different groups of hearing loss and if there are any possible trends that govern the two aspects. In other words, different auditory capabilities may be important for different degrees of hearing loss. The relationship between speech recognition in noise and threshold/other factors may differ depending on severity of the hearing loss. Clearly, for more severe losses, audibility is more likely to be important at normal conversational voice levels. Very few studies (Lutman, 1987; Pavlovic 1984) in the past have attempted to segregate the different groups of hearing loss and look specifically into performance trends for each one. Further the studies have tended to restrict their inclusion of subjects to those with thresholds up to approximately 50-55 dB. Pavlovic (1984) did include both mild and severe losses. However his study was restricted to predictions from SII and he did not include any actual suprathreshold measures. Lutman's study (1987) mentioned above has a more systematic approach, but the groups reported have overlapping boundaries in terms of their hearing loss. The study pooled the capabilities of (normal + mild) or (mild + moderate) groups as opposed to separate groups in order to increase group numbers. Also the study incorporated use of fixed stimulus presentation levels for measurement as opposed to varying according to hearing loss. Thus, study of speech recognition performance for discrete groups and a wider range of hearing loss (eg: mild, moderate, severe etc) is perhaps necessary to understand these interrelations further. Availability of such systematic information could reveal what could be happening in each group and has been investigated only to a limited extent before.

III] Thus the third aim is to investigate how different threshold/suprathreshold abilities affect speech recognition in noise according to magnitude of hearing loss.

Finally most studies have compared speech recognition performance in hearing impaired using the stationary/continuous noise. There are number of studies (Summers and Molis, 2004; Nelson *et al.*, 2003; Dubno *et al.*, 2002; Bacon *et al.*, 1998; Hagerman, 1995, 2002**)** comparing the performance of individuals in two types of noise (stationary and fluctuating). Some have observed speech recognition performance with other capabilities using stationary noise. However, very few have observed the same using fluctuating noise. In fact, only one recent study by George (2006) has investigated the influence of fluctuating noise systematically showing the importance of temporal resolution for speech recognition in fluctuating noise. Moreover, it is known that while normal hearing listeners perform differently in the two types of noise (i.e. perform better in fluctuating noise), hearing impaired subjects may fail to show this difference in scores (Summers and Molis, 2004; Nelson *et al.*, 2003). They are less able to make use of the additional information present in gaps of the fluctuating noise. (Summers and Molis, 2004; Nelson *et al.*, 2003; Dubno *et al.*, 2002; Bacon *et al.*, 1998; Hagerman, 1995, 2002) This indicates that different factors could be associated with speech recognition in these two different kinds of noise. Fluctuating backgrounds are common in daily life and hence understanding the processes behind this effect is important (Kramer *et al.*, 1996). Thus it would be worthwhile to investigate these effects of fluctuating noise in the present study. In fact, this issue is only a subset of the general issue that hearing disability is multidimensional and cannot be fully characterised by a single dependent variable, such as speech recognition in stationary noise. Ideally multiple variables would be considered. This issue is explored further in the second part of the study.

IV] Thus the fourth aim is to investigate if there are different factors that underlie speech recognition for the two different types of noise: stationary and fluctuating.

The four aims discussed above in Section 1.4.1 are directed towards understanding and explaining the main aspects of speech recognition in noise. The corresponding hypotheses are listed below:

1. Auditory sensitivity alone cannot explain or predict the variation in speech recognition performance in noise across a range of hearing impairment
2. Besides audibility, certain measurable suprathreshold factors can help to explain the variation in speech recognition in noise scores
3. The relative importance of the various factors affecting the speech recognition in noise changes as the degree of hearing loss varies
4. For the two types of noise (stationary and fluctuating), there are different factors responsible for speech recognition scores

An important feature of the approach used to investigate the above aims concerns the presentation levels of the tests. The aim here was to ensure audibility of stimuli in all tests, which was achieved by using higher presentation levels for more impaired ears. It was intended that this would restrict the overwhelming influence of audibility as a factor and allow the influence of secondary auditory capabilities to become evident. Thus testing levels were always suprathreshold as opposed to sometimes close to or below threshold, as in some other studies. This approach is consistent with the idea that, in natural listening conditions, people will adjust signal volume whenever possible. They may increase TV volume or wear hearing aids or adjust the volume or positioning of a telephone headset. Also, speakers may compensate by speaking louder when they address a person known to have impaired hearing. Thus the approach can be characterised as ecological.

1.4.2 Group II-Studies examining interdependence of auditory factors

The discussion until now was restricted to speech recognition and its relation to auditory capabilities. As mentioned in the beginning of the chapter, group II studies are devoted to exploring these relations more broadly. Thus, the aim here is to discuss hearing loss as a complete entity which includes speech recognition and/or numerous auditory capabilities. One of the obvious reasons for doing so, is that none of the studies mentioned above have attempted such a comprehensive approach. But, perhaps it is more meaningful to also see why this is essential. A discussion regarding the auditory or other domains covered by each study is important for the same reasons. The following figure outlines the different domains covered by all the studies included in the discussion plus a set of measures that were not included in any of the studies.

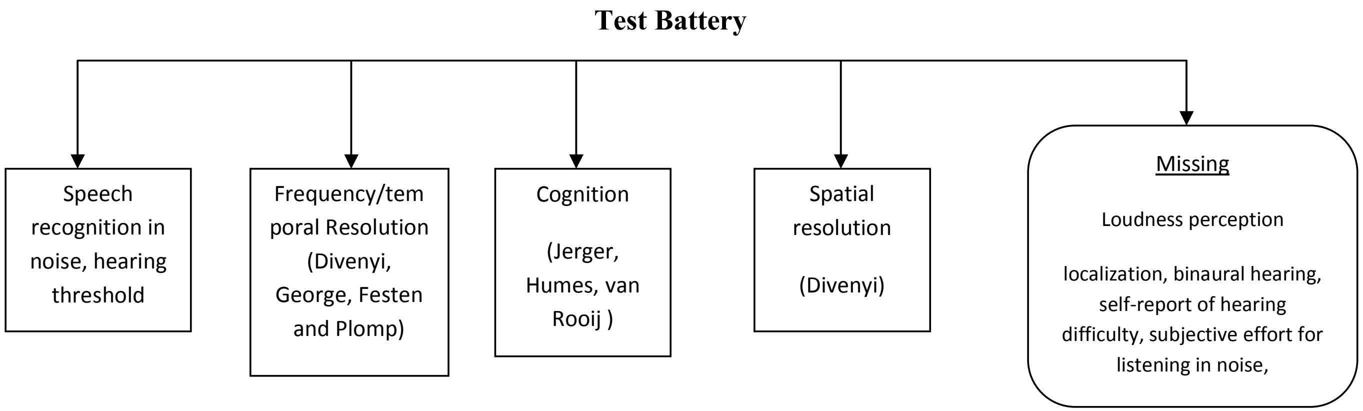


Figure 1.2: Summary of Auditory test domains included in different studies.

As can be seen, the studies have mainly limited themselves to only few domains which to some extent have resulted in incomplete understanding of the auditory functions. While all studies included hearing threshold and speech recognition, one of the other domains is missing in every study. The box on the extreme right shows that there were a number of other auditory and non-auditory (subjective) aspects that have not been paid attention to. This is why an approach that views hearing loss in total is essential since a person with hearing disability is likely to exhibit problems that are related to but outside these aspects that have been investigated earlier. Of course there are individual studies where all these remaining aspects have been studied, but nonetheless usually with exclusion of one or more domains. Thus they have not been observed collectively in one study. Hence, a study including all of these is desirable. Equally important are the relations between speech recognition and the aspects like localization, binaural hearing and subjective measures which are missing. It is also important to realise whether measures of speech recognition can help predict any of these other measures. Evidence of analysis regarding such reverse-predictions where SRT measures are the independent variables are rarely available. This in addition to being a novel concept, will also help understand whether the measures highlighted in the extreme right box are related to hearing threshold and speech recognition or stand independently. In either case, the multiple dimensions of hearing loss would be recovered. In summary, the question of interest is whether the capabilities listed in Fig.1.2 vary independently among hearing impaired people, or do they vary together as a single entity related to severity of hearing loss.

V] Thus the fifth aim is derived from group II studies. This is to explore the multidimensionality of hearing disability. In other words, can experience of hearing loss be related to a single global measure or are multiple measures required?

The consequent hypothesis from group II is:

5. Measures of hearing disability are highly correlated and describe a single underlying dimension. The alternative hypothesis is that they are not highly correlated and can be understood to be multidimensional.

1.5 Motivation of the study

In summary two main questions are investigated in the thesis:

Which auditory factors are responsible for prediction of speech recognition in general across a range of hearing loss?

Is the variation in auditory performance across a range of hearing impairment multidimensional, or can it be approximated by a single unidimensional hearing loss construct?

1.6 Summary of aims and findings

The following table highlights the different aims of the thesis

Table 1.1: Summary of aims put forth in the thesis

Factor	**Aim**
I-Explanatory factor	1.Resolve the discrepancy surrounding the relation between hearing threshold level and speech recognition in hearing impaired and thus investigate if audibility is predominantly or exclusively responsible for variation in speech recognition score of not.
I-Explanatory factor	2. If audibility can explain only in part the differences among speech recognition performance in various hearing impaired, what are the other factors (suprathreshold) responsible?
I-Explanatory factor	3. To investigate if different threshold/suprathreshold factors affect speech recognition in noise with differing magnitude of hearing loss.
I-Explanatory factor	4. To investigate whether the performance of hearing impaired for two types of noise (stationary and fluctuating) helps understand the factors affecting speech intelligibility
II-Multiple dependent factor	5.Can experience of hearing loss be related to a single measure or are multiple measures required ?

1.7 Diagrammatic representation of basic structure and ideas in the thesis

This chapter ends with a figure followed by description of all the chapters in brief.

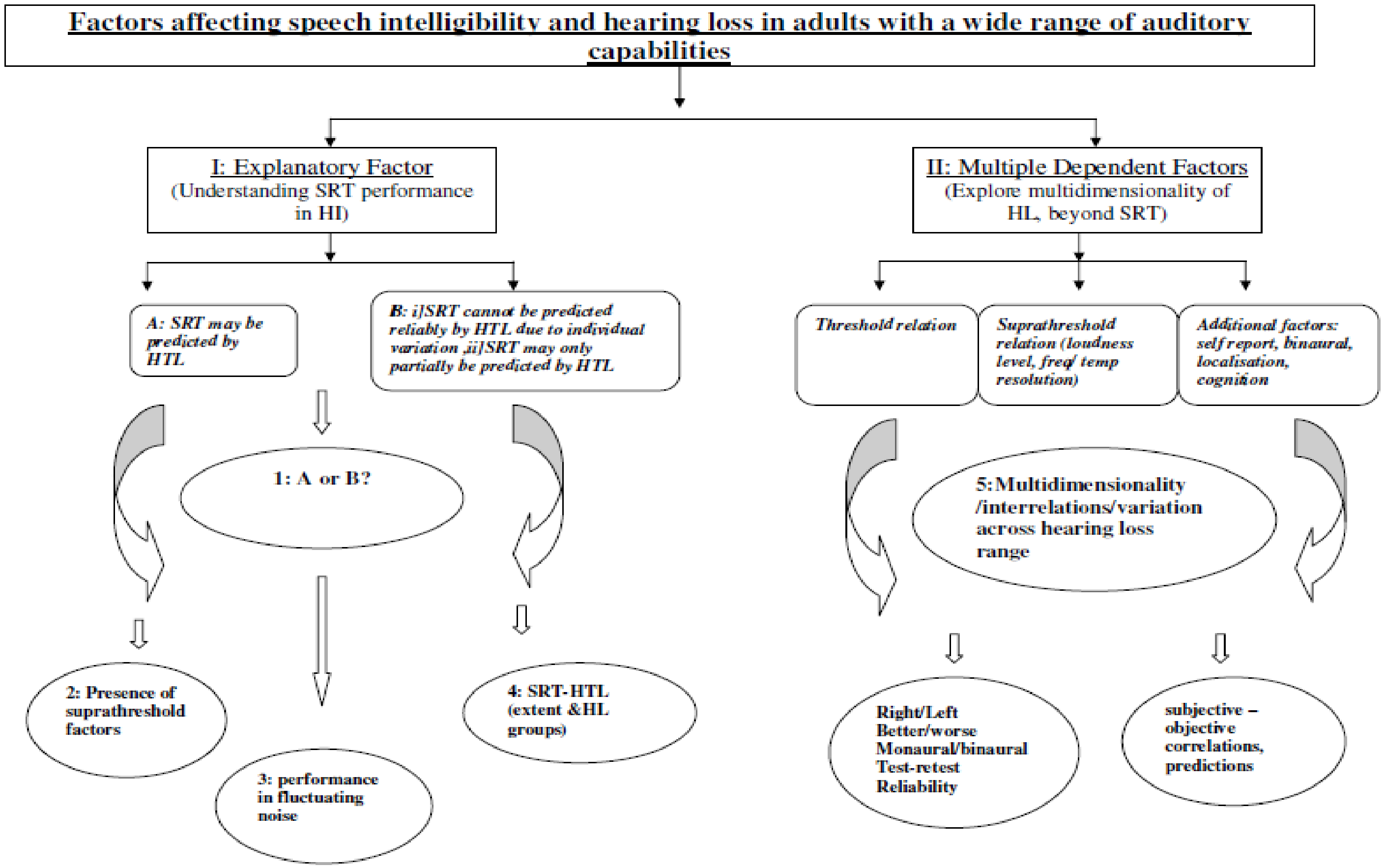


Figure 1.3: Outline of the basic structure and ideas in the thesis.

1.8 Outline of remaining chapters included in the thesis

Chapter 2-Review of Literature

This chapter follows the same pattern as the present chapter wherein the different aspects of the studies will be discussed as per each aim. It is divided into three sections, the first section critically evaluates the studies put forth in this chapter, the second section discusses the first four aims and the aspects related to it in more detail and the third section includes aim five where the different test domains outlined in this chapter are discussed in detail.

Chapter 3- Methodology: Test description and subject criteria

This chapter includes the rationale and design of experimental methods, description of actual test measures selected for each of the test domains described in chapter two and details of subjects included in the thesis.

Chapter 4- Results and interpretation-I

Includes general and descriptive statistics of the data for study I along with the differences between normal hearing and hearing impaired population. Results from regression and factor analysis using the different tests are also discussed. This chapter also follows the same pattern as the first two wherein the different results and their interpretation are discussed as per each aim.

Chapter 5- Results and interpretation-II

Includes explanation and rationale for study II along with descriptive statistics and results from regression and factor analysis of the data for study II as compared to study I. Both the studies are compared and contrasted and similarities and differences are discussed.

Chapter 6- Discussion and summary

This chapter includes general discussion of all the findings and their overall comparison with other studies along with the clinical implications, conclusions and contribution.

Chapter 7- Conclusion and contribution

This chapter summarizes the final conclusions and novel findings from the two studies.

Having discussed the various gaps in knowledge and understanding of the performance of speech intelligibility in hearing impaired, and motivation behind the different studies being carried out as part of the thesis, the next chapter constitutes a review of literature to underpin the objectives discussed here. The purpose of this present chapter is to give a the framework of the different studies carried out within the thesis, and outline the various objectives and hypothesis but the next chapter will discuss and critically evaluate, the different test domains and the studies mentioned in this chapter in more detail.

Chapter Two
Review of Literature

2.1 Introduction

This section is devoted to a more detailed and critical discussion of the objectives and the different groups of studies mentioned in the first chapter. It is roughly divided into three sections. In the first section all the studies are discussed individually with their specifics as well as being compared overall with others. The discussion further extends to outlining the various limitations in the methodologies of the studies in the literature and possibilities of their improvement in the present study. It also includes discussion of the important auditory domains mentioned at the end of the chapter I and their subsequent relevance. The second section includes additional discussion of aims 1-4. In aim 1, the relation of speech recognition in noise and threshold is further reviewed while aim 2 does the same for suprathreshold abilities like frequency and temporal resolution followed by brief discussions of aims 3 and 4. The final section includes aim 5 exclusively. This is treated as a separate section because, each of the auditory domains discussed in the first section are reviewed here again in terms of the different test methods, procedures and measures available for each in the literature, in order to aid design of a final test battery that can be used in the present study.

2.2 Section I: Discussion of studies from groups A and B and others

A table of the studies that are frequently discussed (classes A, Bi, ii and others) is outlined, including their significant characteristics, in appendix I. These studies will be referred to repeatedly and the reader may wish to refer to this table.

In this section however, the grouping (A, B) is of less significance; it was introduced in the first chapter to emphasize the obvious discrepancies observed in the different studies. The insights of individual studies and their distinction from others is of more significance here. Studies from group A and B (ii) are included here. The pioneering study which revealed a major insight into understanding speech recognition was done by Plomp (1978). His model identified factors that characterize speech recognition in noise and in quiet. Thus according to this model hearing loss for speech can be interpreted as sum of class A (Attenuation) and D (Distortion). This A factor is characterized by reduction in effective levels of signal and noise while the D factor is characterized by effectively decrease in signal to noise ratio. Further, the factors A and D are associated with speech recognition in quiet while it is mainly the D factor that becomes important for speech recognition in noise. The studies that followed (Festen and Plomp 1983, Dreschler and Plomp 1985; Lutman 1987; van Rooij & Plomp 1990; Jerger etal 1991; Humes *et al.*1994; Divenyi *et al* 1997a; George etal. 2006) have in one way or other investigated what aspects influence this D factor associated with speech recognition in noise. They are summarised in the Table 2.1

Table 2.1: Summary of the main findings from the different studies in literature.

Study/Authors	**Test measures influencing speech recognition in noise**	**Other results**
Festen and Plomp 1983	Frequency resolution (1000 Hz)	Speech recognition in quiet is governed by hearing threshold level
Dreschler and Plomp 1985	Mean Frequency resolution (500,1000,2000 Hz), temporal resolution (2000 Hz)	Phoneme perception parameter was related to audiometric slope and hence play a secondary role in speech perception.
Lutman (1987)	Overall : Hearing threshold level (2000 Hz) ,frequency resolution (4000 Hz) and temporal resolution (4000 Hz) but varies in different groups of hearing loss	Speech test performance was not related to age once psycho-acoustical variables were accounted for
van Rooij and Plomp (1989,1990,1991)	Hearing threshold level (2000 Hz),cognition (reduced mental efficiency)	Both auditive and cognitive components correlated with age ,but their balanced contribution to speech perception did not change with age
Humes and Roberts (1990)	Hearing threshold level (average 1000,2000,4000 Hz)	Speech recognition scores of individual hearing impaired listeners revealed large differences in performances among them
Jerger *et al.*, (1991)	Hearing threshold level (average 1000,2000,4000 Hz) mainly, cognition (very minimal)	For dichotic testing both hearing loss and cognition accounted for significant variance, age accounted for significant variance for SSI (synthetic sentence identification)
Humes *et al.,* (1994)	Hearing threshold level (average 1000,2000,4000 Hz) mainly, cognition (very minimal)	
Divenyi *et al.,* (1997a,b,c)	Hearing threshold level (average 1000,2000,4000 Hz) mainly, followed by speech spatial resolution	Age- related deficits found in measures of auditory resolution: frequency, temporal, spatial.
George *et al.,* (2006)	Hearing threshold level (average 500,1000,2000 Hz),temporal resolution, age	Problems in speech intelligibility due to supra-threshold deficits are more prominent in modulated maskers than in stationary

From the above it becomes evident there is a general agreement among the majority of studies that hearing threshold level (HTL) especially at higher frequencies predicts SRT with few exceptions. Only the extent varies (refer Chapter I, 1.41: Class Bii) but the issue of how the prediction is improved is more debatable, whether by frequency and/or temporal resolution or effects of cognition or age.

Following the main findings, other issues relating to the studies are now discussed. These include differences/shortcomings in the test methods, subjects recruited, statistical methods; additional auditory/non-auditory domains tested other than SRT, HTL and frequency/temporal resolution etc. They are also discussed hierarchically so that any improvements as compared to previous studies can be noted.

Following the SRT model (Plomp, 1978), the study that attempted systematic investigation of different auditory functions in relation with speech recognition in noise was by Festen and Plomp (1983). Their study established the differences for such interrelations between hearing threshold, speech recognition, frequency/temporal resolution by studying both normal hearing (Festen and Plomp, 1981) and hearing impaired individuals. The study also established the important parameters within each test domain that should be considered while observing the interrelations by carrying out Principal Component Analysis (PCA). Measures of mean audiometric loss (250- 4000 Hz) and audiogram slope (difference in dB 4000−500 Hz) were found to be most reliable to represent the hearing sensitivity measurements. Additional measurements to test the non-linearity of the cochlea were also included. The study also revealed that in normal hearing individuals there is a trade-off between frequency and temporal resolution abilities which was absent in the hearing impaired. However, the study includes only correlation and PCA. No prediction analyses such as linear regression were included. Such analysis is especially important for when trying to uncover the influence of one parameter over another among a significant number of correlations. Further the sample of the study was limited to only 22 hearing impaired subjects which perhaps is not sufficient for factors analysis where the number of subjects has to be at least five times the number of test parameters. All measurements except the audiogram included testing only at 1000 Hz and there was some redundancy in the test measurements especially for frequency and temporal resolution. However, the study includes a good description of numerous concerns relating to doing such interrelation studies such as test reliability, influence of presentation level, minimizing bias in results etc.

The study following the above was by Dreschler and Plomp (1985). It was very similar to the one above and included similar measures of speech recognition, frequency and temporal resolution. However, they improved on two aspects; test measurements were conducted at three frequencies (500, 1000, 2000 Hz) and included regression analysis. Additional measures of loudness and vowel perception were present. Vowel/phoneme perception was hypothesized as a measure bridging the gap between tone perception (hearing threshold) and speech perception (sentences), but was found to play a limited role. But one major difference between their study and others was the age group of subjects included. They tested hearing impaired adolescents aged 13-20 years while all the others included an adult/elderly population. Lastly the study found a co-occurrence of poor frequency and temporal resolution in relation to speech recognition (which means both of them were affected due to poor speech recognition), as opposed to a trade-off relation in normal listeners in the study above by Festen and Plomp (1981) (which showed that if temporal resolution was .affected, frequency resolution was not and vice versa)

Lutman (1987) similarly studied various psycho-acoustical measures for a group of normal hearing and mild-moderate hearing impaired subjects. Additional measures included intensity resolution, temporal integration, two tone suppression, distortion and adaptation along with other common ones (refer 2.3.3 for details) . The sample size included 88 hearing impaired subjects and hence was larger than the two studies above. This study focussed on confounding effects of age. It also observed the different abilities possibly affected in the different severities of hearing loss and hence will be discussed further in relation to aim 2.

The next set of studies were done by van Rooij and Plomp (1989, 1990, 1991) which focussed on relations of auditory and cognitive factors to speech recognition. Thus additional measures included cognitive tests. This set of studies hence varied in the perspective that the relative importance of auditive (sensitivity, frequency and temporal resolution) and cognitive factors which was not investigated in any of the studies above. These results were also further analysed for the confounding variable of age. The test methods were very similar to those used by Festen and Plomp (1983) and by Dreschler and Plomp (1985) with only variation being the frequencies at which the test measurements were carried out (800 Hz and 2400 Hz). Statistical techniques included correlation and PCA as seen with studies above along with an additional canonical correlation analysis (CCA) and Multivariate Analysis of Variance (MANOVA) discussed later in 2.2.1.7.

Jerger *et al.*, (1991) studied the relation between hearing threshold, speech recognition, cognition and age. Frequency and temporal resolution were not tested which was one of the main drawbacks of the study. However, the sample size was 200 elderly subjects, the largest among all the studies. Canonical correlation analysis and multiple regression analysis were used for data analysis.

Humes *et al.,* (1994) also studied the relation between hearing threshold, speech recognition, cognition and age. However the main aim was to study these associations using a wide range of speech material as well as using different presentation levels and background conditions. This study also did not include frequency and temporal resolution measures. Additional measures however included auditory discrimination which was not included by others as well as 27 cognitive measures which were more numerous than the other two studies (Jerger *et al.*, 1991; van Rooij and Plomp, 1990). Correlations, PCA and canonical correlation analysis were used for data analysis.

Divenyi *et al.,* (1997 a, b and c) investigated the audiological correlates of speech recognition in the elderly. One of the main features of the study was inclusion of spatial resolution measures not included in other studies. Also it included more tests of sentence recognition in noise at the sentence level, while most others included those for phonemes or word recognition. Test measurements were carried out at a frequency of 1000 Hz and for wideband noise. Further the focus of the study was to uncover the auditory measures affected in elderly due to age as opposed to hearing loss.

George *et al.,* (2006) studied primarily the factors responsible for release of masking in fluctuating noise. Thus this study was perhaps the first to incorporate the use of fluctuating noise for speech recognition in order to study its relation with other auditory capabilities. They further included a group of subjects with simulated hearing loss additional to the usual two groups of normal hearing and hearing impaired. This approach made it possible to investigate whether signal audibility or suprathreshold deficits were responsible for the reduced benefit from masker modulations observed in hearing impaired subjects. Results revealed that reduced audibility was only partially responsible while temporal resolution and age were the other factors that accounted for this reduced benefit. Their study also observed the effects of presentation level on suprathreshold deficits. Thus the study mainly differed in terms of additional group of subjects with simulated loss as well as the focus which consisted of investigating the aspects of masker modulations as opposed to studying general interrelations between threshold, speech recognition and other suprathreshold abilities.

2.2.1 Critical evaluation of different aspects of research methods used in studies

The following text summarises some of the aspects discussed as well as highlights others in view of the various features and limitations of the studies discussed above.

2.2.1.1 Subjects

The subject sample for hearing impaired included in the various studies ranged from 20-50 individuals. However, in order to study the correlations between different auditory measures, a larger sample is desirable in order to increase the reliability of the test measures as well as to ensure validity if data is to be subjected to more specific analysis, like regression or factor analysis (ideally subject number should be 5-10 times, the number of measures included). This is only achieved by Jerger *et al.*, (1991) (200 subjects) and Lutman (1987) who included 88 subjects. Similarly, the type of hearing loss cases included for all the studies was cochlear loss, but the degree of hearing loss did not exceed 55-60 dB for most studies. Thus findings for a wider range of hearing loss are limited.

2.2.1.2 Unilateral/bilateral measurements

Most of the studies (Pavlovic 1987/83, Humes *et al.,* 1994, George *et al.,* 2006, Festen and Plomp 1983) included only one ear for measurements, while only three studies (Divenyi *et al.*, 1997a, Jerger *et al.*, 1991, Lutman, 1987) measured both left and right ears. Only Lutman (1987) actually compared the similarities/ differences between the two ears, although this is not detailed in the brief publication. There are no published comparisons of better/ worse ears in any of the studies. Such comparisons are simple yet essential to realise the different aspects of auditory functioning. Similarly binaural listening was studied only by Divenyi *et al.* (1997a). The study observed age to be a factor influencing the perceptual separation of speech and noise and also that this elderly group exhibited a right-ear advantage for central auditory tests and a slight left ear advantage for peripheral resolution. Such implications are important and have been researched in only limited studies including the above. Further what has not been investigated is the binaural advantage using speech across a wide range of hearing losses.

2.2.1.3 Frequencies tested

Most of the studies (Festen and Plomp 1983, George *et al.*, 2006, Divenyi *et al.*, 1997a, b, c) have included test measures (e.g. frequency/temporal resolution) centred at only one frequency (1000 Hz), limiting the findings to this frequency. There are one or two exceptions (Dreschler and Plomp, 1985; Lutman,1987). It is desirable to investigate other frequencies, especially representing high and low frequencies as well as broadband measures.

2.2.1.4 Presentation levels

The choice of signal levels constitutes a general problem for the measurements on hearing impaired participants. The level should be sufficiently above the threshold so that all the suprathreshold capabilities (speech recognition, frequency/temporal resolution etc.) can be assessed. However, it should not reach intolerable levels which is possible with a number of subjects with narrow dynamic range. In general, two approaches are incorporated in the above studies to decide on the presentation level: fixed (Jerger *et al.*, 1991; Humes *et al.*, 1994; Divenyi *et al.*, 1997 a, b, c; Lutman, 1987) or individually adjusted (Festen and Plomp, 1983; Dreschler and Plomp, 1985; van Rooij and Plomp, 1990). While fixed level has the advantage of comparing all the subjects at equal levels, it limits inclusion of subjects with higher degrees of hearing loss which was evident in the study by Divenyi *et al.*, 1997 (a, b, c) where the hearing loss did not exceed 50 dB or so. Adjusted levels on the other hand permit this.

While using a fixed level makes data clearly interpretable, statistically, it has the disadvantage of measuring normal hearing participants at a rather loud and possibly unfamiliar level, while measuring many hearing impaired participants at levels that may not be high enough and overcome audibility problems. This therefore limits the testing to moderately hearing impaired only, which was also seen in most studies above. The advantage of using a fixed level is seen in measuring abilities such as frequency resolution which is level dependent. On the other hand, using adjusted levels of presentation ensures that the stimulus is always audible and may be considered to emulate listening via a hearing aid. According to Wagener *et al.,* (2003), the SRT depends only on the SNR as long as one takes care that noise presentation level is high enough to be audible. Thus it becomes possible to obtain optimal comparable accuracy of measurement for all listeners at SRT. But with fixed SNR this accuracy depends greatly on the hearing ability. This may yield a problem in rather severe hearing impaired listeners since the required noise level may exceed the safety or technical limitations. But in these listeners it is anyway questionable whether or not it is reasonable to perform speech intelligibility tests in noise.

2.2.1.5 Test-retest reliability/repeatability

The test-retest reliability for the speech tests assures that for a given individual, the scores will be essentially the same at any point when measured for the same criteria, but at different point in time. Mainly, within subject SD (standard deviation) plays a role in determining it. Further it helps ensure that correlations are not affected substantially by measurement error and hence are valid (Festen and Plomp, 1983). Only two studies (Festen and Plomp, 1983; Dreschler and Plomp, 1980) have included this.

2.2.1.6 Redundant measures

A careful review of the studies (Divenyi *et al.*, 1997 a, b, c; Jerger *et al.*, 1991, Humes *et al.,* 1994) reveals that not only limited test measures were investigated as above, but also redundant measures for the same test domain were included. Humes *et al.,* (1994) included two tests of cognitive and auditory processing each. Jerger *et al.*, (1991) obtained four speech recognition scores and six neuropsychological measures while Divenyi *et al.*, (1997 a, b, c) measured speech recognition scores on ten tests. Such repeated test measures may increase reliability and validity but may also reduce the time efficiency of the study by giving redundant scores. A balance needs to be found between these opposing benefits.

2.2.1.7 Statistical techniques

Overall the methods used by the studies include simple correlations (Spearman/Pearson), canonical correlation analysis, MANOVA (Multivariate Analysis of Variance), multiple regression, principal component analysis and factor analysis. All these methods basically examine variation among a number of variables to give a relation between them. They thus strive to explain the association between the two variables as well as dependence of one on others. According to Rummel (1975), when the question is presence or absence of association between two variables, ANOVA is usually preferred, but the degree and direction (positive–negative) of the same is perhaps best quantified by correlations. Similarly both regression and canonical correlation analysis yield a mathematical function connecting the variables to establish their dependency. But, if the concern is the dependence of one variable on a set of two or more variables, a delineation of this question is in terms of multiple regression. Alternatively, if the concerns are the dependence of a set of two or more variables on a set of two or more variables, then a delineation of this question is in terms of canonical analysis (Rummel, 1975). This could be seen in the studies discussed above such as the CCA (Canonical Correlation Analysis) is used by van Rooij and Plomp (1989, 1990) to study the influence of a set of auditive and cognitive components on another set of phoneme and speech perception components (hence one set of variables on another set). Alternatively, studies like Lutman (1987) and Dreschler and Plomp (1985) used multiple regression to study the influence on a single dependent variable (speech recognition in noise) of a set of variables including frequency resolution, temporal resolution etc.

Principal component analysis and factor analysis are typically used to analyze groups of correlated variables representing one or more common domains. Principal components analysis is a form of factor analysis used to find optimal ways of combining variables into a small number of subsets to identify the structure underlying such variables and to obtain scores to estimate latent factors themselves.

Also to some extent, the use of statistical techniques depends upon the ease of use and availability of tools. The output obtained from canonical correlation analysis is relatively difficult to interpret (Thompson, 1984) as well as the fact that it cannot be easily computed on popular tools like SPSS. On the other hand, the results from correlation, regression analysis and factor analysis can be readily obtained from most packages.

Now, having outlined the features and limitations of many previous studies, the present study attempts to overcome/substitute/modify/improve in the following ways:

- Inclusion of a robust test battery with measures that cover the relevant auditory and other dimensions with minimal redundancy
- Inclusion of a larger sample than in most studies
- Inclusion of a wider range of hearing loss, than included in most studies
- Testing for frequencies other than 1000 Hz, especially representing low and high frequencies as well as broadband measures
- Use of a multicentre approach that achieves uniformity of measurements and greater representativeness and also assisting with larger sample size
- Measuring speech recognition in both stationary and fluctuating noise
- Investigating speech recognition in noise scores in different groups of hearing loss
- Ensure adequate reliability of the test measures to minimise measurement error
- Use of hearing loss dependent presentation levels to ensure that stimulus is always audible to allow test subjects with a wide range of hearing abilities and minimise the direct effects of audibility on other test measures
- Inclusion of various ear comparisons like right-left, better-worse, monaural-binaural
- Inclusion of self-report measures as well as performance measures to relate results to experience reported by hearing impaired people

Along with above measurements, another important parameter to be assessed is comparing the hearing threshold and SRT scores with that of Speech Intelligibility Index (SII). Many studies before (Pavlovic 1987/83, George et al, 2006) have included this measurement in order to ensure the relation between the two measures more objectively. A brief discussion of SII is included in the following section.

2.2.1.8 Speech Intelligibility and Articulation Index

The intelligibility of speech refers to the accuracy with which a normal listener can recognise a spoken word or phrase. There are several available methods of predicting speech intelligibility within an enclosure, most common being the articulation index (AI). Most of the methods to predict intelligibility are based on the same fundamental principle, determining a ratio between the intensities of received speech signal and the interfering noise. It is this basic signal-to-noise relationship upon which speech intelligibility is deemed to depend - the higher the ratio, the greater the intelligibility. The intelligibility also depends on SNR in bands weighted and summed across frequencies and speech shaped noise tends to have the same SNR in all bands, so effectively masks speech.

In order to be able to predict the speech intelligibility under such masking conditions, French and Steinberg (1947) and Kryter (1962) initiated a calculation scheme, known as the Articulation Index (AI)

<u>The Articulation Index (AI)</u>

This value is basically a continuous measure ranging from 0.0 to 1.0 based on calculations of the signal to noise ratios in five octave bands (with centre frequencies of 0.25, 0.5, 1, 2 and 4 kHz). It is possible to obtain a more accurate calculation based upon 1/3rd octave band sound pressure levels (based on work by Kryter), however, this requires more detailed knowledge of both the speech and noise spectra. Since the speech level usually refers to the long term value for normal speakers, octave spectra are normally sufficient for simple calculations.

Calculation of the AI consists of three basic steps.

The measurement of the effective signal-to-noise ratio for each octave band.

Applying a weighting factor to each ratio and clipping to ensure that maximum contributions occur at +18 dB and minimum at −12dB.

Summing the weighted value.

Thus the articulation index can be calculated from the following equation

$$AI = \frac{G_{[i]}}{30dB} \sum_{i=1}^{5} (Lsa - Lna + 12)dB \qquad (1)$$

where *AI* is articulation index $G_{[i]}$ represents the weighting factor for each octave band given in the table below and *Lsa* and *Lna* are effective speech spectrum and noise levels.

Frequency (Hz)	Weighing Factor (G[i])
250	0.072
500	0.144
1000	0.222
2000	0.327
4000	0.234

Since its revision in 1997, the method has been adjusted to include adjustment for an individual's hearing threshold level at each frequency and is named the Speech Intelligibility Index (SII). (New method accepted as the ANSI S3.5-1997 standard). SII can be calculated by:

$$SII = \sum_{i=1}^{n} I_i A_i \qquad (2)$$

where *n* is the number of individual frequency bands used for computation, *I* is the frequency importance function, *A* is band audibility calculated from spectrum levels of noise and speech.

Thus for a given speech-in-noise condition, the SII is calculated from the speech spectrum, the noise spectrum, and the listener's hearing threshold. Both speech and noise signal are filtered into frequency bands. Within each frequency band the factor audibility is derived from the signal-to-noise ratio in that band indicating the degree to which the speech is audible. For this purpose, hearing threshold level is represented by addition of an internal noise, sufficient to raise the masked threshold of a normal ear to the required threshold level. Since not all frequency bands contain an equal amount of speech information i.e., are not equally important for intelligibility, bands are weighted by the so-called band importance function. The band-importance function indicates to which degree each frequency band contributes to intelligibility. It depends on the type of speech material involved e.g., single words or sentences, and other factors. Finally, the SII is determined by accumulation of the audibility across the different frequency bands, weighted by the band importance function. The resulting SII is a number between zero and unity. The SII can be seen as the proportion of the total speech information available to the listener. An SII of zero indicates that no speech information is available to the listener; an SII of unity indicates that all speech information is available. Model parameters have been chosen such that the SII is highly correlated to intelligibility. However, when SII is based on the long-term average spectra of speech, it does not take into account the short-term transients or changes over time. Thus the SII is able to explain the speech intelligibility in stationary noise, but not for fluctuating noise. However, Rhebergen and Versfeld (2005) introduced the Extended Speech Intelligibility Index (ESII) which makes it possible to apply the SII in fluctuating backgrounds by calculating and averaging the SII calculated in short time frames.

2.2.2 Comparison of auditory domains and test measures used in different studies

The use of different test domains and measures has been discussed at some length in section 1.4.2 in Chapter One and the various domains were was outlined in Figure 1.2. Thus the range of domains studied was another significant aspect where studies differed. This is understandable since the selection of tests usually depends on the availability of tests in different languages, different aims etc. Of course, some tests used by Festen and Plomp (1983), Dreschler and Plomp (1985), van Rooij and Plomp (1989, 1990, 1991) were similar but then they differed on other aspects like the age of subjects recruited, frequencies tested, hypothesis etc. This was also true of studies that investigated the effects of cognition (van Rooij and Plomp 1989, 1990, 1991; Jerger *et al.*, 1991; Humes *et al.*, 1994) or age. The auditory domains common to all as well as the additional ones have been discussed above. Also, recapping from figure 1.2, it was observed that certain additional aspects like localization, binaural measurements are included in only a few studies. Also none of these studies have included any subjective/self-report measures while cognition was included in only two studies. Each of these domains conveys information about a different and potentially important aspect of the auditory system. Their relevance is described below:

Binaural Hearing: This allows us to hear sounds accurately and more naturally. It gives us a sense of direction. It also allows better differentiation of speech and noise when they are spatially separated.

Localization: Listeners in everyday life need to identify who is starting to speak when they are in a group of people. This enables them to turn towards the speaker to maximize their use of binaural hearing; they will also benefit from being able to lip read and see facial expressions and gestures made by the speaker. Localising sounds in the environment is also important for safety and feelings of security.

Loudness perception: This measure can test for loudness recruitment, often associated with cochlear hearing loss. It refers to a condition in which growth of loudness with increasing level is more steep than normal. This reduces the effective range of sound levels normally heard. It also allows for measuring Most Comfortable Level (MCL), Uncomfortable Level (UCL) and other related aspects.

Self-report of hearing difficulty: These measures allow a person to describe difficulties that go beyond speech in quiet and speech in noise recognition and hence they are of interest in case performance measures miss them. They allow us to see how subjects actually perceive hearing loss as opposed to its objective evaluation.

Listening effort: An individual with hearing impairment may achieve 100% recognition in scores yet have to exert greater listening effort to maintain listening performance than normal hearing individuals (Downs, 1982). It is this effort that will be investigated here. This aspect of hearing disability may help to explain differences between performance and self-report measures.

As can be seen, the focus of most studies has been on hearing threshold, speech recognition in noise, frequency resolution and temporal resolution, with cognition and spatial resolution in some others. Further the studies have restricted their measurements in other ways. For example, Festen and Plomp (1983) focussed on measurement of auditory capabilities at 1000 Hz and did not include any other frequencies. Also, what perhaps lacked was a broad balance of different domains so that there is minimal redundancy, as well as a good mix of subjective and objective measures.

Thus the present research attempts to include all the test domains shown in Figure 1.2 (Chapter One) in one test battery with inclusion of measurements at both high and low frequencies. Of course it is never possible to cover all aspects in detail, but the aim is that the test battery is broad enough to cover at least the main parameters in the auditory and other relevant areas. Thus, in view of the above discussion, final test domains included in the present research are:

- Pure tone audiometry (hearing threshold)
- Speech intelligibility (speech recognition in quiet and noise) for stationary & fluctuating noise
- Frequency resolution
- Temporal resolution
- Loudness perception
- Cognition (lexical access) (for details refer 3.2.9, Appendix VIII)
- Spatial perception: localization, binaural hearing
- Self-report of hearing difficulties and impact on everyday life
- Effort required understanding speech in the presence of noise

(measures are included at a low and high frequency: 500 and 3000 Hz wherever appropriate)

2.3 Section II: Additional discussion of aims 1-4

2.3.1 Aim one: Speech recognition and hearing threshold level

The relation between the two has been discussed at length in the first chapter. This part includes discussion of studies in group B (i) unlike section I. A more detailed review of these studies leads to some common observations, summarised in the following table. The studies are compared to the attenuation (audibility) and distortion (clarity) factors in Plomp's model (1978) mentioned in the beginning of the chapter. The table identifies four stereotypical hearing states (columns) and characterises them in terms of typical hearing difficulties.

Table 2.2: Summary of common observations of group B studies.

Type	**I**	**II**	**III**	**IV**	
Type of loss	Normal hearing	Hearing loss	SNR loss	Both SNR+Hearing Loss	
Clinical manifestation	Normal audibility and clarity	Loss of audibility (sensitivity)	Loss of clarity	Loss of both, audibility + clarity	
Difficulties	No difficulty in both quiet and noise	More Difficulty hearing in quiet (& hence sometimes in noise if at low dB levels)	More Difficulty hearing in noise even if high dB levels	Difficulty hearing in both noise and quiet	
				IV a	**IV b**
				Audibility + clarity affected to similar extent (i.e. increased hearing loss, increased SNR)	Audibility and clarity not affected to same extent (i.e. increased hearing loss, decreased SNR or vice versa)
		Benefit from conventional hearing aids	Limited benefit from conventional hearing aids	Benefit from hearing aids to some extent only if signal-to-noise ratio is improved	Benefit from hearing aids to some extent only if signal-to-noise ratio is improved

Though the above classification is quite broad and there could be more than one type present in the clinical population, it is usually the type IV both 'a' and 'b' that causes the maximum concern and in particular 'b' wherein the listeners most commonly complain that 'they can hear but not understand'. Such difficulties are most commonly seen in individuals with sensory hearing loss especially due to ageing (presbycusis).

Other studies (Jerger *et al.*, 1989) revealed high correlations between hearing sensitivity and speech understanding measures. Similarly various studies (Plomp and Mimpen, 1979; Duquesnoy, 1983; Duquesnoy and Plomp 1980; Gelfand *et al.*, 1986) supported the finding that hearing loss status is the most significant factor determining speech intelligibility in quiet and noise. On the other hand, various other studies have proposed explanations for variations in speech recognition scores such as effects of aging, overall cognitive and personality factors (Nelson *et al.*, 2009).

So again, the relation between the two needs to be established further and the present research attempts to do the same as part of aim 1.

2.3.2 Aim 2-Relations between speech recognition and suprathreshold measures

As mentioned in the first chapter as well as in aim 1, most of the studies in group Bii revealed presence of other factors while exploring the relation between speech recognition, hearing threshold and other measures. These include various suprathreshold measures such as frequency resolution (Festen and Plomp, 1983), temporal resolution (George *et al.*, 2006), spatial resolution (Divenyi *et al.*, 1997 b, c) as well as cognition (Jerger *et al.*, 1991) and sometimes loudness recruitment. Frequency and temporal resolution in relation to speech recognition are more of interest in this section while the others will be discussed more in aim five.

2.3.2.1 Reduced frequency resolution (selectivity)

Many people with a sensorineural hearing loss have difficulty separating sounds of different frequencies when they are presented simultaneously (Ludvigsen and Kuk, 2001). This loss of frequency selectivity also manifests itself as excessive upward spread of masking (Ludvigsen and Kuk, 2001). The end result includes increased difficulty with speech understanding in noise, because frequency components become more difficult to resolve in environments with competing signals (Ludvigsen and Kuk, 2001).

a) Frequency resolution and critical band concept:

Fletcher (1940) measured the threshold of a sinusoidal signal as a function of the bandwidth of a band pass noise filter. He suggested that the peripheral auditory system behaves as if it contained a bank of band pass filters (with continuously overlapping centre frequencies) called 'auditory filters'. According to him the basilar membrane provided the basis for the auditory filters. Thus, each location on the basilar membrane responds to a limited range of frequencies, so that each different point corresponds to a filter with a different centre frequency. Data reported by Moore (1986) were found to be consistent with this. Thus, when trying to detect a signal in noise background, the listener is assumed to make use of a filter with a centre frequency close to that of the signal. This filter will pass the signal but remove a great deal of the noise. Only the components in the noise which pass through the filter will have any effect in masking the signal. It is usually assumed that the masked threshold for the signal is determined by the amount of noise passing through the auditory filter; specifically, threshold is assumed to correspond to the signal-to-noise ratio at the output of the filter. This set of assumptions is called the power spectrum model of masking (Patterson & Moore, 1986).

In the band-widening experiment described above by Fletcher (1940), it was observed that further increase in noise bandwidth results in more noise passing through the auditory filter, provided the noise bandwidth is less than the filter bandwidth. However, once the noise bandwidth exceeds the filter bandwidth; further increases in noise bandwidth will not increase the noise passing through the filter. This bandwidth at which the signal threshold ceases to increase is called the 'critical bandwidth'.

b) Methods to estimate the shape of the auditory filter

Most methods including the psychophysical tuning curve (PTC) (Moore, 1986) and notched noise method (Patterson, 1976) for estimating the shape of the auditory filter at a given centre frequency are based on the assumptions of the power spectrum model of masking. The threshold of a signal whose frequency is fixed is measured in the presence of a masker whose spectral content is varied. It is assumed, that the signal is detected using the single auditory filter which is centred on the frequency of the signal, and that threshold corresponds to a constant signal-to-masker ratio at the output of that filter. However, the PTC method can give rise to off-frequency listening (Moore, 1986) since it assumes that only one auditory filter is involved. The listener might make use of more than one filter especially when the masker frequency is above the signal frequency. He does this by attending to a filter centred just below the signal frequency. The notched noise method is more useful since it prevents this off-frequency listening by making use of a noise masker band stop or notch centred at the signal frequency. The signal is fixed in frequency and the deviation of each edge of the noise is from the centre frequency is denoted by Δf. The width of the notch is varied and the threshold is determined as the function of the notch width. As the width of the spectral notch is increased, less and less noise passes through the auditory filter and the masked threshold decreases.

c) Frequency resolution in hearing impaired ears

According to Moore (1986), the perceptual consequences of a loss in frequency selectivity are many and variable. The first major consequence is a greater susceptibility to masking by interfering sounds when we are trying to detect a signal in a noisy background. A second, but related, difficulty is that of perceptually, separating two or more simultaneously presented sounds. When the auditory filter is broader than normal, it is much more difficult to hear out one voice from a mixture of voices. Thus holding a conversation when two people are talking at once can be very difficult for the hearing impaired person Moore (1986). A third difficulty arises in the perceptual analysis of complex sounds, such as speech or music. Moore also reviewed several studies (Zwicker and Schorn, 1978; Florentine *et al.*, 1980; Pick *et al.*, 1977) which revealed flatter psychophysical tuning curves (PTC) and broader auditory filters in cases of cochlear impairment as compared to normals.

d) Level dependence of frequency resolution

From above it can be seen that many hearing impaired listeners appear to have poor frequency resolution. And this occurs due to broadening of the auditory filters. At the same time it also known that as signal level is increased, the spectral representation in the basilar membrane becomes broader (Allen, 2000). Thus, because hearing impaired subjects require signals to be presented at high sound-pressure levels, the deterioration of frequency resolution found for them may be to some extent a result of level effects. There has been some evidence that frequency resolution becomes poorer with increasing signal level (Evans 1977). However study by Wightman and Raz (1980) revealed otherwise. They found a difference of only 2 dB between the thresholds of frequency resolution measures at high and low probe levels with one subject even producing a narrow curve at high levels. They suggested that there is only a slight reduction in frequency selectivity at high levels and hence it is unlikely that poor frequency resolution in these subjects is due to high signal levels.

e) Frequency resolution and speech intelligibility

Several investigators have found significant links between speech intelligibility and frequency resolution (Dreschler and Plomp, 1985; Patterson *et al.*, 1982; van Rooij *et al.*, 1989; Glasberg and Moore, 1989; Ter Kuers *et al.*, 1992) while others have not (Dubno and Schaefer, 1992; Lutman and Clark, 1986) since they found other factors like hearing threshold level or age to be more important.

2.3.2.2 Reduced temporal resolution

Temporal resolution refers to the ability to distinguish consecutive pulses as separate events. Time is a very important dimension in hearing since all sounds change over time. Furthermore, for sounds which convey information, such as speech and music, much of the information appears to be carried in the changes themselves, rather than in the parts of the sounds which are relatively stable (Moore, 1986).

a) Methods to measure temporal resolution (Moore, 1986).

Gap detection: This is the most common method of measuring temporal resolution. The threshold of detecting a gap in broadband noise provides a simple and convenient measure of temporal resolution since the long-term magnitude spectrum of broadband white noise remains the same if the noise is briefly interrupted. Usually a two-alternative forced-choice (2AFC) procedure is used: the subject is presented with two successive bursts of noise and either the first or the second burst at random is interrupted to produce the gap. The task is to indicate which burst contained the gap.

Temporal modulation transfer function (TMTF): In this method, white noise is sinusoidally amplitude modulated, and the threshold for detecting the modulation is determined as a function of modulation rate. This function relating threshold to modulation rate is called TMTF.

Forward masking: Forward masking is when the masker is presented first and the signal follows it. It is often used to test temporal resolution ability. Basically, a loss of temporal resolution gives rise to more forward masking i.e., the masker will decay more slowly after the termination of the masking sound, thus decreasing the perceived gap size.

b) Temporal resolution in hearing impaired people

Reduced temporal resolution is known to adversely affect masking release (which refers to improved signal to noise ratio in fluctuating noise compared to stationary noise seen in normal hearing individuals and absent in hearing impaired) (George *et al.*, 2006). Studies (Festen and Plomp 1990; Glasberg and Moore, 1992; Festen, 1993; Dubno, 1992) have shown reduced temporal resolution in terms of either reduced gap detection or smaller masking release or reduced ability to take advantage of temporal dips etc. The temporal dips arise because there are moments when the overall level of the competing speech is low, for example during the brief pauses in the speech or during production of low-energy sounds such as m, n, k or p . Peters *et al.* (1998). During these temporal dips the SNR is high and this allows brief 'glimpses' to be obtained of the target speech . Peters *et al.* (1998)

c) Temporal resolution and speech intelligibility

Some studies (Tyler *et al.*, 1982; Dreschler and Plomp, 1985; Irwin and McCauley, 1987, George *et al.*, 2006) investigating the relation between these have found a link between the two another study (Festen and Plomp, 1983) have not since they found frequency resolution to be more important.

Analysis of specific signal attributes such as frequency, intensity and duration is partly done by the central auditory system (Albeck *et al.*, 1992; van Rooij & Plomp, 1990). Extracting a signal from a competing background of noise is also intrinsically done by central mechanisms (Albeck *et al.,* 1992). Thus deterioration of neurons in the central auditory nervous system can limit both frequency and temporal resolution for more complex signals which may contribute to reduced speech recognition performance (Philips *et al.,* 2000).

As stated when considering aim one, the results from the different groups of studies to some extent varied due to the different test batteries used. However, some of the test measures including SRT, HTL, frequency and temporal resolution were commonly studied in Festen and Plomp, (1983); George *et al.,* (2006); Divenyi *et al.,* (1997 a, b, c), while the studies by Jerger *et al.,* (1991) and Humes *et al.,* (1994) did not test for frequency and temporal resolution. However, it is worthwhile to note that in the former studies where the similar test measures were compared, the results were different. Festen and Plomp (1983) observed that frequency resolution is closely allied to speech recognition in noise. On the other hand, George *et al.* (2006) concluded that temporal resolution was an important factor for determining SRT in fluctuating noise, whereas Peters *et al.,* (1998) concluded both spectral and temporal dips are important in understanding speech in the presence of background noise. Study by Bernstein (2010) like George *et al.* (2006) also investigated why hearing impaired listeners do not receive as much benefit to speech intelligibility from fluctuating maskers, relative to stationary noise, as normal hearing listeners. It was suggested that this difference may arise as a consequence of differences in the signal-to-noise ratio (SNR) at which HI and NH listeners are tested. The Extended Speech Intelligibility Index (ESII) was fit to NH data, and then used to make FMB (fluctuating masker benefit) predictions for a variety of results in the literature. Using this approach, reduced FMB for HI listeners and NH listeners presented with distorted speech was accounted for by SNR differences in many cases. HI listeners may retain more of an ability to listen in the gaps of a fluctuating masker than previously thought.

Thus, whether frequency resolution or temporal or both are responsible for reduced speech recognition performance is still debated and the problem is that they tend to co-vary. So the studies in the present research will contribute towards this debate in aim 2.

2.3.3 Aim 3: Speech recognition in different groups of hearing loss

The studies discussed so far have investigated relations between speech recognition and threshold measurements as reviewed in aim 1 or with suprathreshold measurements as in aim 2. In other words, the focus of all studies was on observing interrelations on a specific group of hearing impaired people. This subject group of hearing impaired people usually consisted of individuals with a range of hearing loss not exceeding moderate. The study by Lutman (1987) observed the outcome of the FAAF (Four Alternative Auditory Feature Test) speech test in noise in relation to other psychoacoustic measures including the sensitivity (audiogram), frequency resolution(PTC), temporal resolution (gap detection), intensity resolution (intensity difference limen), temporal, integration, suppression (two tone), distortion ($2f_1$–f_2) and adaptation. Results from multiple regression suggested that for normal and mild groups combined, upward spread of masking and gap detection both in the 4000 Hz region were the best predictors. In mild and moderate groups combined, gap detection at 2000 Hz and average sensitivity at 2000 Hz and 4000 Hz were best. However, after partialling for sensitivity, gap detection was not a significant predictor. Pavlovic (1984) also attempted to study speech recognition and used SII predictions for different groups of losses. He inferred from the representation in his SII model that suprathreshold distortion factors were present in addition to loss of audibility, for the more severe losses while for mild and moderate hearing losses, reduced audibility was sufficient to explain SNR loss. However, much of his inference was related to the fact that he used filtered speech materials as discussed in chapter one [section 1.41, class B(i) studies].

In Lutman's study, the findings obtained were for combined groups of hearing loss like normal-mild or mild-moderate rather than for systematic investigation of discrete groups. It is possible that auditory measures can be differentially affected as the magnitude of hearing loss differs. Whereas Pavlovic's study used only SII predictions and did not measure any suprathreshold measures, in the present study, the focus will be on specific groups of hearing loss and their relation with other suprathreshold abilities.

2.3.4 Aim 4: Speech recognition in different types of noise

Various studies that have compared the use of stationary and fluctuating noises revealed both greater repeatability for speech recognition for stationary noises than for fluctuating noises, and larger differences between the scores of normally hearing and hearing impaired listeners for fluctuating noises than for stationary noises (Summers and Molis, 2004; Nelson *et al.*, 2003; Dubno *et al.*, 2002; Bacon *et al.*, 1998; Hagerman, 1995, 2002; Eisenberg *et al.,* 1995; Gustafson and Arlinger, 1993; Takahashi and

Bacon, 1992; Bronkhorst and Plomp, 1992; Festen and Plomp, 1990, George *et al.,* 2006). From these studies it appears that hearing impaired listeners are poorer in using the speech information that is physically present in the gaps of the fluctuating noises. Part of this effect appears to be related to reduced audibility (elevated thresholds); the reminder is often interpreted as being caused by a loss of temporal resolution and possibly counteracted by informational masking. In fact, in the study by George *et al.* (2006) temporal resolution and age were considered to be responsible for the reduced benefit from masker release in hearing impaired ears while frequency resolution was ruled out as an explanatory suprathreshold deficit. Bacon *et al.,* (1998) also reported a relationship between pure-tone thresholds and the size of the masking release from fluctuating noise and a high negative correlation of 0.75 between the two for hearing impaired listeners and 0.83 when normal listeners were included.

However, only one recent study by George *et al.,* (2006) has compared both the types of noise to observe if different factors influence speech recognition in the two noises. Thus the findings are limited and more study of this aspect is required which will be covered in chapter four.

2.4 Multidimensionality of hearing loss (Aim 5)

As outlined in the first chapter, this aim concerns exploring relations beyond speech recognition and hearing threshold, so that hearing loss can be viewed in a more complete perspective. Also in the latter section it was suggested that the use of a complete test battery will aid in such a perspective. The selection of the different test domains was further discussed in the same light based on review of the relevant studies in section I. In this section each of the domains will be discussed along with some background on each.

2.4.1 Hearing sensitivity

The absolute threshold of hearing is the minimum sound level of a pure tone that an average ear with normal hearing can hear in a noiseless environment. It is commonly referred to as hearing threshold level (HTL). It measures an individual's auditory sensitivity and also suggests the extent to which less intense components of the everyday sounds such as speech may be inaudible. Pure tone audiometry (PTA) is standard clinical measure included in all auditory test batteries. Various suprathreshold deficits including speech recognition in noise and quiet, frequency resolution, temporal resolution may be correlated with HTL since most of them are found to be more or less dependent on the extent of HTL.

2.4.2 Speech recognition in noise

As discussed in the beginning of the chapter one, the ability to recognise speech in the presence of interfering noise was considered by many people to be the greatest handicap associated with their hearing impairment (Kramer *et al.*, 1996). Its importance as part of a test battery is established. Testing speech recognition in noise has greater face validity than in quiet since it approximates everyday listening conditions where presence of ambient noise is common experience. Further, speech recognition tests need to be a compromise between realistic environment and reproducibility in the clinical setting. Actually, realism of levels of both speech and noise are essential for an accurate estimate of any deficient measures. Further, any speech test has to be language specific and optimized for the general and clinical population. Such speech in noise tests using sentence material have been developed in various languages and are being constantly developed in different places. A table outlining some of the commonly used speech recognition tests in different languages is included in Appendix II.

Again, as mentioned previously, speech recognition in both stationary and fluctuating noise is important as both types of background are commonly experienced in everyday life. The following figure shows the performance of normal hearing and hearing impaired individuals on different types of interfering backgrounds. As can be seen, there is a clear difference in the performance between the two groups as well for different types of noise in the normal hearing. The latter differences are minimal for the hearing impaired group. Such observations in different interfering backgrounds help uncover the underlying auditory impairments. Thus inclusion of speech recognition in different types of noises such as stationary and fluctuating, is as important as studying its interactions with other auditory and non-auditory capabilities, which was frequently highlighted in the previous discussions.

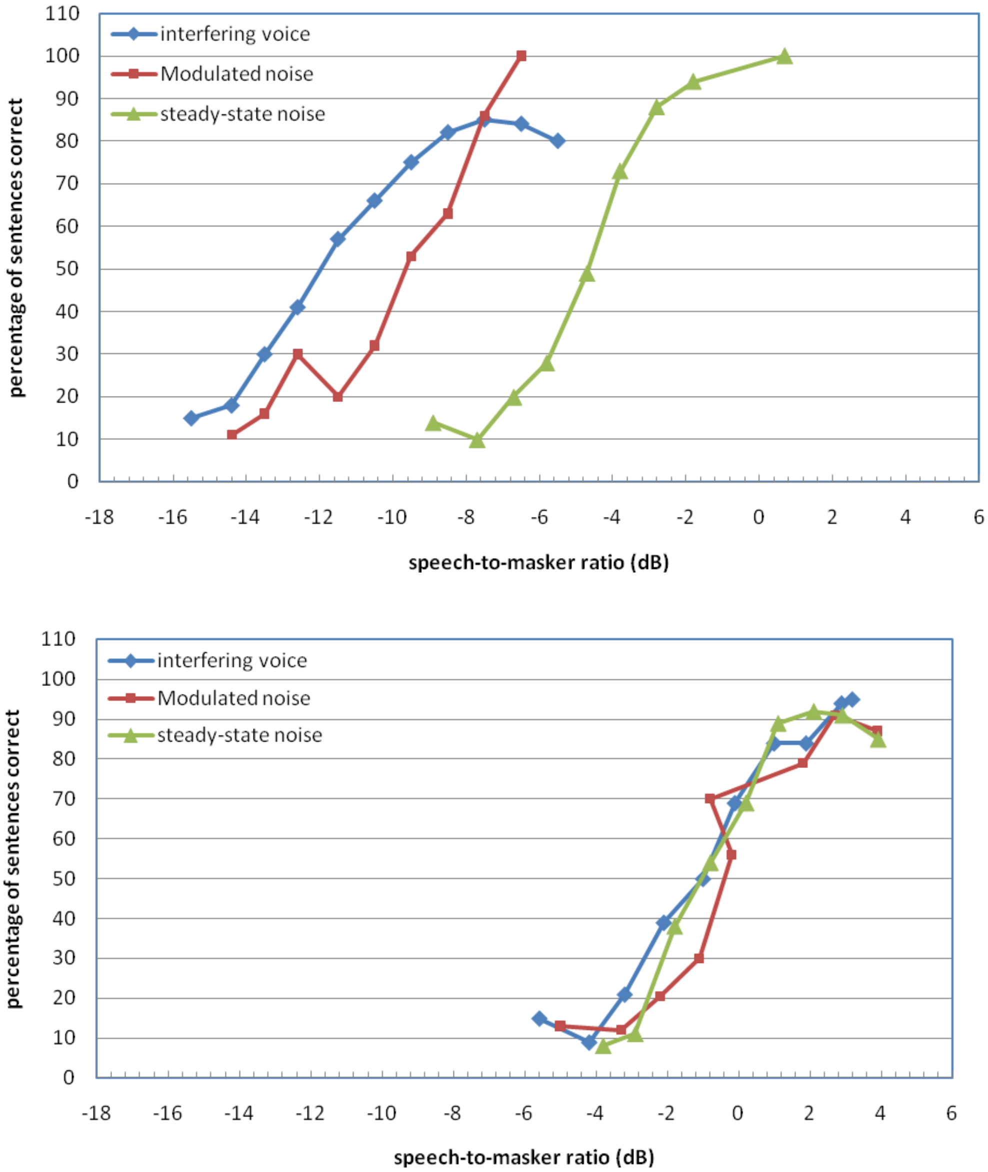


Figure 2.1: Average discrimination curves for sentences presented in steady-state noise, two band modulated noise and interfering noise for normal (upper panel) and hearing impaired listeners (lower panel). Re-plotted from Festen and Plomp (1990).

2.4.3 Loudness perception

Loudness scaling is popularly used as a measure for diagnosis of loudness recruitment in clinical audiology. Loudness scaling basically reveals the relation of loudness percept to the intensity of the sound stimulus. The use of categorical rather than magnitude scaling procedures for clinical use seems more practical because of relatively short completion time and simplicity for the patient (Allen *et al.*, 1990). Though magnitude estimation procedures are possible, they may not be reliable enough for use clinically due to a wide range of educational backgrounds, ages and hearing losses found in the general population (Studebaker and Scherbecoe, 1988). The Categorical Loudness Units basically define how loud a stimulus is perceived in terms of 'soft' and 'loud' rather than in the ratios of loudness of different stimuli.

Most loudness procedures (e.g. Allen *et al.*, 1990; Elberling and Nielson 1993; Ricketts and Bentler, 1996, Cox *et al.*, 1997; Rasmussen *et al.*, 1998; Keidser *et al.*; 1999) use a pre-measurement phase, which determines the individual's auditory dynamic range before the actual data collection phase begins, which is time consuming. However, the procedure of Brand and Hohmann (2002) omits this phase thus increasing time efficiency. Also instead of using ascending level sequences (Cox *et al.*, 1997; Keidser *et al.*, 1999) which cause significant bias of the loudness function estimates, the above method uses randomized levels which avoid an accumulation of biases.

Further, many loudness procedures (Cox *et al.*, 1997; Launer *et al.,* 1996; Elberling, 1999 etc) have used seven categories except Brand and Hohmann (2002) who used 11 categories (seven labelled and four interleaved) and Kießling (1996) who used 13. The 11 categories as total range of responses was considered a compromise between feasibility in clinical set up and precision (Brand and Hohmann, 2002). The most commonly investigated frequency is 500 Hz (Allen *et al.*, 1990; Kießling, 1996) followed by broadband noise (Ricketts and Bentler, 1996) and 3000 Hz (Keidser *et al.*, 1999).

Thus the procedure by Brand and Hohmann proves to have advantages since there is omission of the pre-measurement phase included in most procedures, thus increasing time efficiency as well as use of randomized levels which avoid an accumulation of biases. This was considered a compromise between feasibility in clinical set up and precision.

2.4.4 Frequency and Temporal Resolution

Frequency and temporal resolution of the auditory system together dictate its ability to discriminate complex acoustic signals, including speech sounds. Natural sounds also feature combined frequency-temporal patterns (Supin, 1997). It is therefore essential to study the resolution of such frequency-temporal patterns. Further, impaired frequency resolution results in impaired discrimination of formants and vowels while masking of syllables occurs with impaired temporal resolution which in turn can jeopardise speech communication (Schorn and Zwicker, 1990).

Although there are many studies that have measured either frequency resolution (e.g. Leeuw and Dreschler, 1994; Rosen *et al.*, 1998; Noordhoek *et al.*, 2001) or temporal resolution (e.g. Eddins 1999, 2001; Noordhoek *et al.*, 2001) extensively, few have measured both in the same participants because of the time constraints. Very few studies have tried to measure them using a single combined procedure.

Supin (1997) measured combined frequency-temporal resolution of hearing in normal hearing listeners using rippled noise stimulation in conjunction with a phase-reversal test at octave frequencies. In the test, the participants have to detect phase reversals (interchanges of peaks and valleys in frequency domain) for different reversal rates and ripple densities. The ripple-density resolution limits were constant at phase-reversal rates below 2-3/s and diminished at higher phase-reversal rates. However, a large learning effect was found when measurements were conducted with feedback; subjects appeared to discriminate based on spectral coloration (distortion) instead of phase reversals. It was tried to eliminate this effect by omitting feedback. However, it could not be shown whether this change in paradigm really eliminated this extra cue, since the long-term spectra of alternating and non-alternating stimuli are always different.

Phillips *et al.* (2000) evaluated frequency and temporal resolution in a group of three types of elderly listeners: normal hearing, hearing loss with good speech-recognition skills and hearing loss with poor speech recognition skills for simple and complex stimuli, as well as syllable recognition in quiet and noise. Results revealed that the group of listeners with hearing loss and poor word recognition did not differ from those with hearing loss and good word recognition on both spectral and temporal resolution for simple stimuli. However, frequency resolution was compromised for listeners with poor word recognition abilities when targets were presented in the context of complex signals which supported their hypothesis that unusual deficits in word-recognition performance among elderly listeners were associated with poor spectral resolution for complex signals.

Larsby and Arlinger (1997) suggested a method for evaluation of temporal, spectral and combined frequency-temporal resolution. Masked thresholds of tone pulses at signal frequencies of 500, 1000, 2000 and 4000 Hz in four different noises were measured: broadband continuous noise, noise with spectral gaps around the signal frequency (bandwidth: 1/10, 1/3 and 1-octave), noise with temporal gaps (coinciding with the signals) and noise with both spectral and temporal gaps as shown in Figure 2.2 shown below.

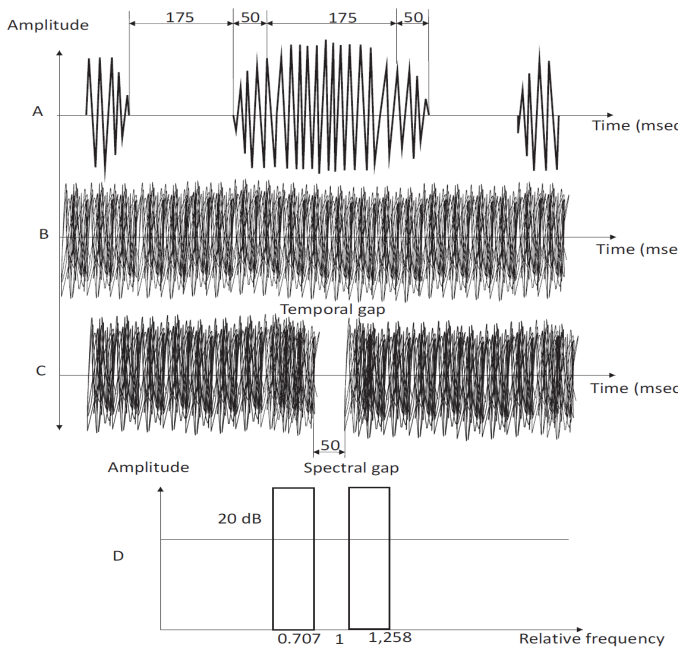


Figure 2.2: Adapted from Larsby and Arlinger (1997): a) pulsed sinusoidal test signal the octave band noise was modified to provide different masking conditions: b) with no gap, c) with spectral gap, d) with temporal gap

A Bekesy tracking procedure was used with varying test tone. Results revealed that release of masking increased with increasing temporal or spectral gaps and more so if both gaps were combined. Maximum release of masking was obtained with a 100 ms temporal gap combined with 1-octave notch filter. The overall difference in release of masking between the young and elderly normal hearing was 2 dB.

The method presented by Larsby and Arlinger (1999) was used by them to measure spectral, temporal and combined resolution in two groups of listeners (normal hearing and hearing impaired). Since many hearing impaired listeners have sloping hearing losses, different masking effects at different frequencies will occur. To avoid such problems, the method was modified by replacing the wide-band pink noise with an octave band noise as masker. It was demonstrated that hearing-impaired subjects show less release of masking than normal-hearing subjects, the release of masking is inversely related to the degree of hearing impairment. The test-retest reliability was reasonably good, comparable to or better than what is found for regular hearing threshold measurements using a tracking method. There was no need for training for the test.

These combined frequency and temporal measurement methods prove to be very useful mainly due to their time efficiency and clinical feasibility.

2.4.5 Tests of cognition

The speech process yields not only the sensation of an incoming stimulus, but also its processing and interpretation in the context of previous experiences. Information about how things are related and categorized, for example contextual, lexical, syntactic and semantic information, is stored in the long-term memory. Controlling top-down processes work in parallel with stimulus-driven bottom-up processes in every information-processing stage (Hallgren 2005). Thus hearing includes both audition and cognition.

Any kind of distortion or limitation of an incoming stimulus (e.g. in difficult listening situations, such as noise and reverberation) or distortion because of a hearing impairment, makes the process more dependent on top-down processing. The situation becomes more cognitively demanding than normal. Some cognitive functions deteriorate as a consequence of hearing impairment in severely hearing impaired and deaf people. More precisely, it has been shown that the phonological ability declines when auditory stimulation is reduced over a longer period (Anderson and Lyxell, 1998; Andersson, 2001). Further, it is well known that many cognitive functions decline with age in the later part of life. Different noise sources put different demands on cognitive skills in the individual. In the complex process of speech understanding the listener depends on peripheral hearing as well as central auditory and cognitive functions. For speech processing in noise these cognitive functions are likely to be especially important since the noise partly masks the speech signal. Several studies have shown the importance of cognitive skills in speech processing tasks (Gatehouse *et al.*, 2003; Lunner, 2003; Lyxell *et al.*, 2003; Pichora-Fuller, 2003).

Cognitive functions also appear to be important in order to make use of amplification in modern hearing aids with advanced signal processing. In recent studies it has been argued that individual cognitive prerequisites interact with different signal processing algorithms in determining the benefit obtained from hearing aids (Gatehouse *et al.*, 2003; Lunner, 2003).

In studies of speech recognition in hearing impaired listeners across a range of ages, results may be confounded by cognitive effects. Therefore, it is desirable to include direct measures of cognitive function so that these effects can be accounted for in the analysis.

2.4.6 Binaural hearing

Binaural hearing in simple terms means listening using two ears. By its nature, the auditory system is equipped to listen and extract cues from signals coming from both the ears (binaural) as opposed to just one (monaural) in order to aid in better hearing. (Ross 2006). Binaural redundancy, as when the brain receives same information from both ears independently, has also been cited as a binaural advantage. Because of the brain's ability to synthesize dissimilar information arriving from the two ears, the overall (two-ear) perception is usually greater than that occurring from each ear separately (Ross 2006). In other words, when speech and noise are presented binaurally to an observer, the intelligibility of speech in a given amount of noise is higher, especially when the speech and noise seem to arrive from different places. This ability becomes especially important while listening to a speech/ signal in the presence of background noise, since central auditory structures are able to suppress interfering noises while focusing on the speech of just one person in a noisy environment, a phenomenon known as the cocktail party effect. (Ross 2006)

The potential advantages of binaural hearing in people with impaired hearing might include: (i) improved hearing in a quiet background when the speech reaching the ears is at or a little above the auditory thresholds—this is due to summation of the sound energy equivalent to an increase of 3 to 6 dB, (ii) improved speech discrimination—by summation of information content from the two ears, when their hearing losses are dissimilar in frequency distribution, (iii) enhanced localization of the speech source, and (iv) improved ability to hear speech in a background of noise MacKeith and Coles (2007). With binaural hearing one ear is nearly always nearer the source of the desired signal than the other and is therefore in a better position with regard to the relative levels of the signal and noise—the 'head shadow effect MacKeith and Coles (2007)'. Further, the additional sounds arriving at the far ear provide the brain with information through time of arrival differences and intensity differences at the two ears that enables it to process the speech and noise signals separately, with an apparent unmasking of the speech—the 'squelch effect' MacKeith and Coles (2007) . Hence inclusion of binaural hearing when considering multidimensional aspects of hearing is essential. As pointed in first chapter, it has only rarely been included in multidimensional studies such as Divenyi (1997).

Binaural advantages are most commonly measured in terms of release of masking or masking level difference (MLD). MLD are most widely studied for 500 Hz tone detection in presence of distracting noise (Wilson *et al.*, 2003; Olsen, 1976; Poth *et al.*, 1992) but have been also explored using speech in the presence of speech shaped noise (Poth *et al.*, 1992) or distracter sentences (Cameron *et al.* 2006). Other than type of stimuli, it is also studied as a measure of different speaker locations or separations (whether free field or through earphones), ranging from the classical 180° (Hirsh, 1950) to various other locations including 60° (Freyman *et al.*, 1999), 90° (Cameron *et al.*, 2006) amongst others. On an average, the more the spatial separation between the signal and the distracter, the higher are MLD values.

The abbreviation ILD is commonly used for two related phenomena. ILD as Interaural Level Difference refers to the level cue that is present in binaural experiments. In the context of this study we use ILD for speech as the Intelligibility Level Difference that is defined as the benefit for speech intelligibility due to binaural effects. The ILD in the present research thus attempts to explore the effects of binaural hearing and spatial separation. Thus ILD used in this way is in effect MLD for speech.

2.4.6.1 Intelligibility level difference (ILD)[1]

The intelligibility level difference (ILD) quantifies the benefit that a listener has from separating speech and noise sources. The ILD, for example is the difference between the

[1] .(source:http://hearcom.eu/prof/DiagnosingHearingLoss/AuditoryProfile/BinauralIntegration.html).

binaural SRT when speech is presented from the front and noise is presented from the side (S_0N_{90}) and the binaural SRT when both speech and noise are presented from the front (S_0N_0). Because of the benefits achieved from the head shadow effect and from binaural processing in the auditory system, the separation of speech and noise sources can lead to an improvement of the SRT. This benefit is estimated by the ILD test, and it is about 6-12 dB in normally hearing subjects. ILD thus represents the release of masking from using the all dichotic cues available to the listener, which are predominantly (i) the better SNR in the ear opposite to the noise (this is a monaural cue) and (ii) the binaural "squelch" from utilising the differences between signals across the ears Lutman (2008).

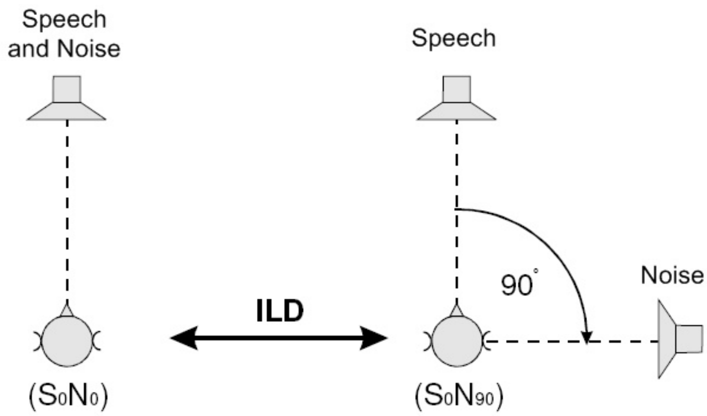


Figure 2.3: Diagrammatic representation of ILD as difference between the S_0N_0 and S_0N_{90}conditions.(source:http://hearcom.eu/prof/DiagnosingHearingLoss/AuditoryProfile/BinauralIntegration.html).

2.4.6.2 Binaural intelligibility level difference (BILD) [2]

For differentiating between the head shadow effect and binaural processing in the brain, the binaural intelligibility level difference test can be used. The SRT in a binaural situation with S_0N_{90} presentation is compared to the SRT in the same situation but with plugging of the ear that is directed towards the noise source. Because of the benefits achieved from binaural processing, the results without the plug (binaural presentation) can be better than those with the plug (monaural presentation) by as much as 3-6 dB in normally hearing subjects. This difference is called the BILD. The 'plugging' here is

[2] (source:http://hearcom.eu/prof/DiagnosingHearingLoss/AuditoryProfile/BinauralIntegration.html).

related to blocking the ear from the noise so that monaural measurements can be obtained which can be achieved by switching off the headphone on the side of the ear that is pointed towards the noise. This restricts the release of masking from monaural cues on the ear nearer to the noise. By obtaining the difference between the SNR outcome in this test condition and the SNR outcome in the ILD test condition (b), the contribution to unmasking from dichotic (binaural) cues, or "squelch", can be derived.

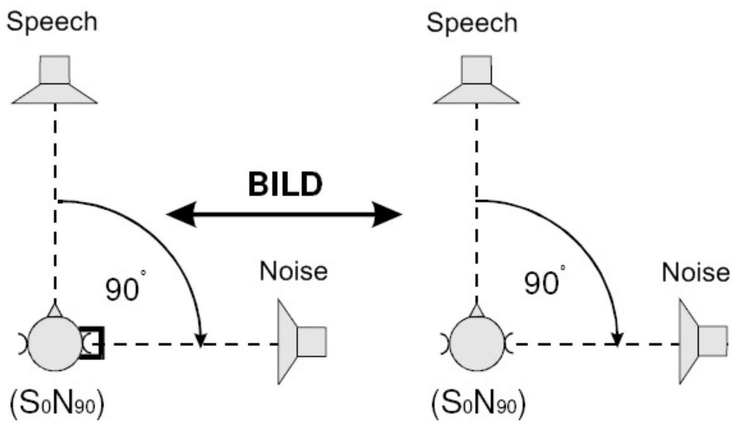


Figure 2.4: Diagrammatic representation of BILD as difference between the difference between the monaural and binaural S_0N_{90} and S_0N_{90}measurements. (source:http://hearcom.eu/prof/DiagnosingHearingLoss/AuditoryProfile/BinauralIntegration.html).

However various studies have used the above terms interchangeably. Levitt and Rabiner (1966) defined binaural intelligibility level difference (BILD) as the difference in signal level (in decibels) between two binaural conditions (S_0N_0, S_0N_{90}) for a given percent intelligibility and found it to be 6 dB for 50% intelligibility. They also found that it is only partly dependent on low-frequency interaural differences. Other studies (e.g. Blauert 1997; Johansson and Arlinger 2002) found BILD for a stationary masker to be about 4–7 dB and defined it as the difference in the speech reception threshold (SRT) in the S_0N_0 and S_0N_{90} presentation mode. As seen in figure 2.4, S_0N_0 is the condition when both speech and noise come straight ahead (0˚) while S_0N_{90} is the condition where speech comes from straight ahead and noise comes at right angle (90˚) (whether left or right side).

In summary, binaural hearing helps localization and improves intelligibility of speech in noise and the effects of this can prove advantageous to the hard of hearing.

Effects of hearing impairment on binaural masking release

Hearing impairment also affects the release from masking occurring when sound sources are spatially separated (Bronkhorst, 1999). Results obtained in a number of studies reviewed in the study by (Bronkhorst, 1999) indicate that the release from masking is, indeed, smaller for the hearing impaired than for the normal hearing and that this factor adds to the deficits discussed above: the hearing loss or speech in noise and the reduced benefit from fluctuations. The only favorable aspect is that the release from masking due to ITD or decorrelation seems to be almost intact in most hearing-impaired listeners. Only the head-shadow component is reduced (because it occurs at high frequencies and most hearing impaired have a high-frequency hearing loss). This means that their performance is less affected by an increase of the number of sources or by addition of reverberation than that of the normal hearing.

2.4.7 Localization

The accuracy with which an observer can localize an actual source has been investigated in two main ways. The observer may be asked to indicate the direction from which a sound appears to come. Among a wide range of positions, most commonly in the horizontal plane (azimuth), minimum audible angle (MAA) thus refers to the smallest difference in location between two sound sources, usually for sources located around 0° azimuth. In other words, MAA is the smallest angle (or difference in azimuth or just noticeable difference-JND) that a listener can discriminate.

Gelfand (2004) has reviewed the classic MAA studies by Mills (1958, 1963 and 1972). He found that the MAA in the horizontal plane was smallest (best) for pure tone frequencies below about 1500 Hz and above approximately 2000 Hz and was largest (poorest) between these frequencies. Thus is becomes evident that owing to the difference in physical nature of the sound, these cues are not equally effective at all frequencies. Basically, when a sinusoidal sound located to one side of the head reaches that other side of the head, hence the farther ear, it will be delayed in time and will be less intense relative to the ear reaching the nearer ear. There are thus two possible cues as to the location of the sound source; known as ITD (interaural time or phase difference) and ILD (interaural level or intensity difference). And these cues operate differentially depending on the frequency of the sound as stated above. Low frequency sounds have a wavelength which is long compared with the size of the head and thus the sounds 'bend' very well around the head. On the other hand, at high frequencies, where the wavelength is short compared to dimensions of the head, little 'bending' occurs resulting in a 'sound shadow'. Thus interaural differences in intensity are negligible at low frequencies and are more important for high frequencies. On the other hand, interaural time/phase differences become important for low frequencies and provide negligible cues for high frequencies. This is because if a tone is delayed at one ear relative to the other, a phase difference occurs which affects the relative timing of the nerve impulses at the two ears. Thus high frequency sounds whose wavelength is less than the distance between the two ears cause ambiguity regarding the sound location while the low frequency sounds due to larger wavelength can be easily identified based on the phase differences. This idea that sound localization for pure tones is based on interaural time differences at low frequencies and interaural intensity differences at high frequencies has been called the 'duplex theory' by Lord Rayleigh (1907).

Gelfand (2004) also found that the MAA was most acute (approximately 1-2 degrees) when the sound source was directly in front of the head and increased dramatically to very high values when the sources were at the side of the head. This phenomenon is often referred to as 'the cone of confusion' as illustrated in the figure 2.5 below. Thus while the duplex theory above provides a simple model for localization, it does not explain the same fully, since it becomes evident that when considering a constant elevation of the sound source that there are multiple locations which would produce identical ITDs and ILDs. If the elevation of the sound source is allowed to vary, the problem is compounded and we get a whole cone surface of points in three dimensional space that would produce the same ITDs and ILDs. The "cone of confusion" thus causes ambiguity of the ITDs and ILDs that are generated from these locations. In general, if the sound is in front, the accuracy of localization is within about 2˚ or so; if to one side, it will be up to about 7˚ and if above the typical errors range from 14˚ - 20˚ (Blauert, 1997).

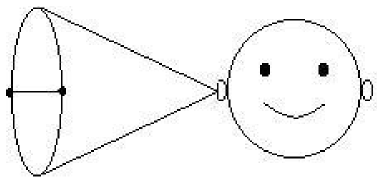


Figure 2.5: Diagram illustrating the 'cone of confusion' in localising sounds.

MAA measurement provides significant localization information relevant for everyday listening. Most studies have used physical separation of sound sources by means of a loudspeaker array. However, some studies (Wightmann and Kistler, 1989b; Besing and Koehnke, 1995) have used a virtual set-up for MAA measurements, whereby cues arising from different source locations are simulated, as described below.

2.4.7.1 Virtual localization

For virtual localization with headphones, head-related transfer functions (HRTF) are used to filter the sounds. HRTFs are usually measured with a miniature microphone placed at the entrance to the ear canal near the eardrum while sounds that vary in azimuth and elevation are being presented. Generic HRTFs can also be measured with a dummy head, so that it is not necessary to measure HRTFs for each subject. Sounds filtered with the appropriate HRTF for a certain direction and presented through headphones will be perceived as originating from that direction outside the head, as they will contain appropriate interaural time and level differences.(HEARCOM, 2005)

An advantage of virtual localization over free-field localization is that it is not necessary to have a sound-treated room with many speakers. Virtual localization behaviour can be tested in a standard audiometry room. Moreover, confounding factors like the sound-field calibration and head movements are eliminated. Further, in free-field localization testing, it is quite cumbersome, although not impossible to ensure that the sound sources and the listeners are in exactly same locations for each test session. This problem does not arise with headphone testing. The disadvantage is that, as this test is done over headphones, it cannot be done with hearing aids. Second, dummy head HRTFs are 'average HRTFs' and they are not optimal for all subjects. Therefore, the sound image resulting from filtering with these HRTFs might vary for different subjects.

2.4.8 Self-report measures

Hearing problems in elderly patients cannot be evaluated completely with conventional audiological tests in most cases. Bertoli *et al.*, (1996) reported an auditory handicap (based on psycho-social difficulties faced by them) was found in one-third of subjects with mild hearing losses and in two-thirds of subjects with greater hearing losses. Thus, diagnostic hearing assessment as well as the outcome of a hearing aid fitting can be measured in both subjective and objective terms. Subjective assessment includes the handicap caused by the hearing loss and the extent to which the hearing aid reduces impairment. In simple terms, subjective assessment is based on how the person with hearing loss feels about the loss and the communication difficulties faced by him as opposed to the objective quantification obtained from formal tests. It further involves the perceived benefit of the aid in real-life situations and can therefore be of high face validity. The most common subjective assessment is through administration of a comprehensive questionnaire which includes various everyday hearing situations including background noise, localization, sound quality, listening in a group amongst others.

Some of the commonly used questionnaires are outlined in the following table.

Table 2.3: Outline of common questionnaires.

Self report measures	Features
a) HMS (British English) (Noble and Atherley, 1970): Hearing Measure Scale	- To assess auditory disability - 42 items covering 7 areas: speech hearing, acuity for non-speech sounds, localization, emotional response, speech distortion, tinnitus, personal opinion - Validation: 27 adult males (chippers in a foundry)
b) Speech, spatial and qualities of hearing scale (Noble and Gatehouse, 2004; Gatehouse and Noble, 2004)	- To measure a range of hearing disabilities across several domains, using a self-report inventory - 50 items in 9 subscales: Hearing speech in a variety of competing contexts, the directional, distance and movement components of spatial hearing, the abilities to segregate sounds, the abilities to attend to simultaneous speech streams, qualities of hearing experience regarding ease of listening, naturalness, clarity and identifiability of different speakers, everyday sounds, musical pieces and instruments - Validation: 153 new clinic clients
c) HHIA (American English): Hearing Handicap Inventory for Adults (Newman et al., 1990; Newman et al., 1991)	- To quantify perceived handicap. Can also be used to assess benefit of hearing aids by measuring pre and post fitting of the aid - 25 items with 2 subscales concerning: emotional consequences, social and situational effects - Validation: 28 adults (29-59 years)
d) Amsterdam Inventory for Auditory disability and handicap (Kramer, 1995, 1996, 1998)	- To identify factors in hearing disability in daily life and assess its associated handicap - 30 items including detection of sounds, distinction of sounds, auditory localization, speech intelligibility in quiet and noise, intolerance of noise - Validation (Dutch version): 274 adults (16-66 years) - Translated into Dutch, English, Spanish, Swedish
e) Gothenburg Profile (Swedish, Arlinger et al., 1998; Ringdahl, 1998)	- To measure the experienced hearing disability and handicap using a self-report inventory - 20 items with 2 scales: experienced disability including being able to hear speech (5 items), being able to localize sounds (5 items), experienced handicap including impact of hearing impairment (5 items), how to performance and reaction to hearing difficulty (5 items) - Validation: 924 persons (14-91 years)

2.4.9 Listening effort

Listening is the process of receiving, constructing meaning from, and responding to spoken or nonverbal messages. However, what happens when a person cannot listen and identify a stimulus perfectly because of some sort of distortion of the signal or some background noise? There are at least two possible ways which affect the person's performance. The first and the most obvious one is reduction in speech intelligibility which leads to less amount of information available to the listener. The second is the less obvious one; the effort required or put into the listening activity. This listening effort is discussed in the following section.

An individual with hearing impairment may expend greater listening effort to maintain listening performance than normal hearing individuals (Downs, 1982). Also sometimes, when masking or distortion is not sufficient to produce errors of identification, a listener may nevertheless have to make more effort to distinguish what is said (Surprenant, 1999). Increased listening effort, and the subsequent stress create an increased working load for hearing impaired individuals in relation to normally-hearing ones, so the former tire quickly. Further, this listening fatigue is worsened due to effects like loss of environmental awareness, passive listening and finding it hard to relax in such distracting environments (Portis, 2005).

Such subjective measures give significant information about the actual communication difficulty faced by the individual as versus his objective score. For example a patient may report significant benefit from the hearing aids because he or she is expending less effort to hear in everyday listening environments, even if the clinical performance measures do not show large benefits. Thus together it can contribute tremendously towards diagnosis and hence a better intervention. In fact, Sato *et al.*, (1998) suggested subjective ratings of the easiness of speech recognition as an alternative approach to word or syllable recognition tests.

Having reviewed the relevant studies and concepts in different objectives, the next chapter describes the research methodology, subject criteria as well as detailed functioning of each test in the auditory profile designed for the present study.

Chapter Three
Research Methodology

3.1 Introduction

As discussed in the previous chapters, the test methods included in the thesis, comprise a test-battery approach with each test representing a different and potentially important domain of the peripheral auditory system. This will contribute towards explaining the two key questions put forth in the study which concern investigating the relations of speech recognition, threshold and other auditory as well as non-auditory capabilities. These questions concerned with which auditory factors are responsible for prediction of speech recognition in general across a range of hearing loss and is the variation in auditory performance across a range of hearing impairment multidimensional, or can it be approximated by a single unidimensional hearing loss construct.This perhaps is not only the most appropriate approach for the above but is also essential for outlining the various dimensions that optimally characterizes age related sensory neural hearing loss mainly affecting .cochlea.

While the last chapter outlined and discussed the different auditory and non-auditory domains, this chapter discusses the selection of actual test measures used in each of the domains. The implementation of the test battery included in the experiments is described. This is followed by the research procedures and protocols It should be noted that the various experiments carried out within the scope of the thesis were part of an EU project called HearCom and adhere to its standard. The project involved multi-centre study which was carried out across five centres in Europe including UK, Germany, Netherlands (two centres) and Sweden. While the selection of test domains was based on systematic review of the relevant studies, the selection of the actual tests to some extent was influenced by the project requirements and aims. Thus the final selection of the test measures depended on various aspects concerned with the project including ready availability of tests which were already developed and standardised by partners involved in the study and pilot studies. Some were exclusively developed for the multi-centre study.

3.2 Selection of test measures for each domain

The following sections outline the rationale for selecting the methods for use in the multicentre study. For frequency specific tests, this entailed choice of frequencies for testing. With the exception of pure tone audiometry where a full range of frequencies was included, frequencies compromised a low and a high frequency. For this purpose, the low frequency was 500 Hz and the high frequency was 3000 Hz. These were chosen to represent a low and high frequency coordinate for all the measurements. In addition to the choice of frequencies, it was necessary to choose a presentation level for a number of tests. Time restrictions prevented obtaining all measures at a wide range of presentation levels. The choice of presentation level is an important issue that may have implications for the results of the study. One approach is to choose a fixed presentation level for all participants. This must be high enough to be audible for the most impaired participant, yet not uncomfortably loud for participants with normal hearing. For some tests, this compromise may be impossible to achieve. Moreover, testing at different parts of the dynamic range for different participants may mean that results are not comparable across participants. An alternative approach was adopted for the presented study, which was to present stimuli at approximately equal loudness for all participants. This entailed obtaining a measure of, or estimating, most comfortable loudness (MCL) at each of the frequencies 500 and 3000 Hz prior to obtaining other measures. These frequencies were chosen to represent a low and high coordinate for range of frequencies and their selection was HEARCOM's decision It was also necessary to obtain MCL for a broadband signal, in order to set speech test materials and other broadband signals at approximately that level.

The following test domains were included.

3.2.1 Hearing threshold

Hearing thresholds were included as a standard measure of hearing acuity and to estimate audibility of components of speech. Pure-tone thresholds were measured using a standard clinical audiometer. Air-conduction thresholds were measured at 250, 500, 1000, 2000, 3000, 4000, 6000 and 8000 Hz and bone-conduction thresholds at 250, 500, 1000, 2000, 3000 and 4000 Hz with adequate masking of the contra-lateral ear, as per BSA audiometry procedures (2004), or equivalent.

3.2.2 Loudness perception

Loudness growth functions were measured using the procedure called ACALOS (Adaptive CAtegorical Loudness Of Sounds) by Brand and Hohmann (2002) which was partly described in chapter II (2.4.3). Also measurements for low (500 Hz), high (3000 Hz) and broadband (BB) frequencies were included. The details of the procedure used for ACALOS are as follows:

Method: Response scale and adaptive procedure

The Oldenburg-ACALOS procedure iteratively adapts the level range to the subject's responses. It is based on the constant stimuli version of the Oldenburg loudness scaling procedure.

Response Scale

The scale consists of seven main categories named – not heard, very soft, medium, loud, very loud and extremely loud and four intermediate categories. These eleven categories are converted to categorical units (CU). The conversion is as follows:

Categorical number	Categorical text	CU
1	not heard	0
2	very soft	5
3		10
4	soft	15
5		20
6	medium	25
7		30
8	loud	35
9		40
10	very loud	45
11	extremely loud	50

Figure 3.1: Response scale including the categorical number and its respective category and categorical unit (CU). The English translation of the original German scale is shown here.

Adaptive procedure in detail

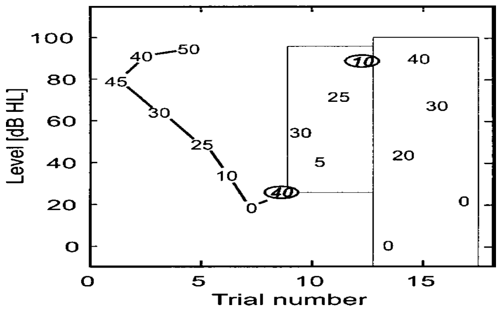


Figure 3.2: Example of a run produced by the adaptive procedure from Brand and Hohmann (2002).

The adaptive procedure also consisted of two phases, which was not obvious to the subject, because he/she rated the loudness in both phases. The auditory dynamic range of the subject was roughly estimated in the first phase. More data were collected in the second phase. In that phase, the dynamic range, in which the stimuli were presented, was re-estimated twice. The first phase started with a stimulus at 80 dB HL. When this was inaudible or too loud, the level was changed in 15 dB steps until a response between inaudible and too loud was given. Thereafter, two interleaved sequences of stimuli began. These two sequences consisted of ascending and descending tracks enabled to obtain loudness levels constituting the complete dynamic range from 0-115 dB. In the second phase it was assumed that the final levels of the two interleaved sequences in the first phase corresponded to the categorical loudness values *L*5 (very soft) and *L*50 (too loud). Thus any categorical loudness levels *L*5 -*L*50 can be estimated by linear interpolation and presented in randomized order. Two iterations were performed in this study for this estimation. During the procedure described above, the listener was protected against harmful loud stimuli by the constraint that in any case, the level was limited to maximally 5 dB above the last level which was rated as too loud before. Further the procedure also ensures that the number of inaudible stimuli are small (since they do not produce any loudness ratings) as well as even distribution of stimuli within the limits of individual auditory dynamic range to reduce bias effects.

As mentioned in the beginning of the chapter, this test is mainly included to obtain an MCL value which is used as reference starting level value for other tests described later in methodology.

3.2.2.1 Purpose of the test

The main use of this test was to obtain a suprathreshold level for each subject which can be used as starting level for other tests for left and right ears (monaurally). This was done by estimating the loudness growth functions described below for the three different types of stimuli (500, 3000 Hz, BB).

3.2.2.2 Specifications and obtaining the MCL value

One variation from the normal test procedure that was applied was that loudness functions were measured at 500 Hz, 3000 Hz and BB instead of 1000 Hz in the original test.

Signals used: For narrow band noises, a one-third-octave band of noise was used (center frequency 500 Hz and 3000 Hz) and for broadband, speech shaped ICRA1 noise was used. The duration of all signals was 2 s, sampling rate 44.1 kHz, windowed with 100-ms cos2 ramps. During each trial, the noise was presented twice with a silent interstimulus interval of 1 s duration

As stated above, the main purpose of this test was to get a reference starting level (suprathreshold) which can be used for the remaining tests in the protocol. And for the purpose of the present study, this reference level, also called as the most comfortable level, was defined as categorical unit: CU 20. (It corresponds to level 5 between the categorical labels of soft and medium, see above fig 3.3). This level was obtained for each of the three stimuli (500 Hz, 3000Hz and BB). In the test, these categorical units are objective representations of the subjective categories of loudness labels of sounds like soft, medium, loud etc and by using the following model function any level corresponding to a label and hence a categorical unit ranging from 0-50 can be obtained.

The model function estimates the best fit to the measured values on completion of a run of any measurement (500 Hz, 3000 Hz and BB) using the adaptive procedure described above.

$$F(L)=\begin{cases} 25+m_{lo}(L-L_{cut}) & \text{for } L \le L_{15} \\ bez(L, L_{cut}, L_{15}, L_{35}) & \text{for } L_{15} \le L \le L_{35} \\ 25+m_{hi}(L-L_{cut}) & \text{for } L \ge L_{35} \end{cases}$$

It consisted of two linear parts with independent slope values m_{lo} and m_{hi} . The two parts intersected at the level L_{cut} . The transition region between the levels L_{15} (soft) and L_{35} (loud) was smoothed, using a Bezier fit denoted with $bez(L, L_{cut}, L15, L35)$.

The following figure illustrates a screen shot for a typical loudness function obtained for a trial of broadband signal (BB). As seen, it constitutes the signal level in dB SPL plotted against categorical units (CU).

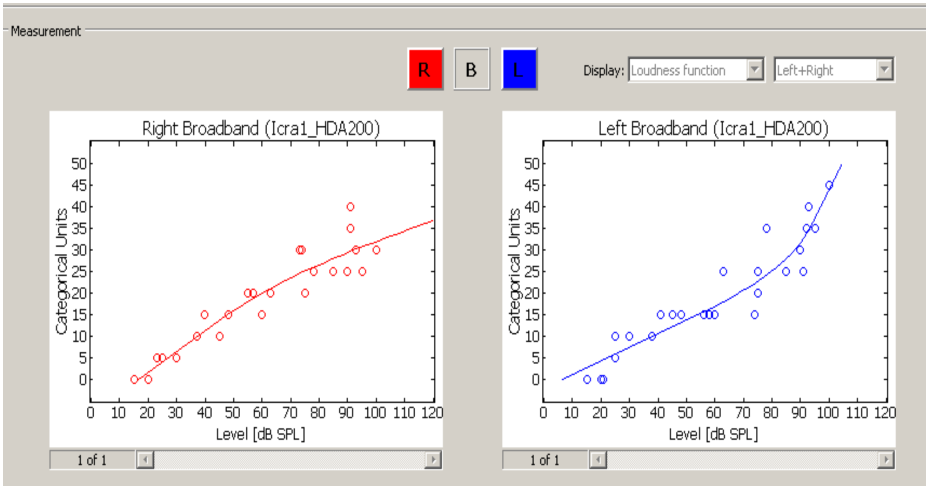


Figure 3.3: Screen shot for a typical loudness function obtained for a trial of broadband signal displaying the signal level in dB SPL plotted against categorical units (CU).

Thus the MCL values were taken as levels corresponding to CU 20 and the slope values corresponded to intersection of linear parts at CU 25 .

3.2.3 Speech recognition in noise

Speech recognition in two types of noise was measured using short meaningful sentence tests. For the English participants, the BKB sentence test was used. The BKB sentences were constructed and evaluated in 1979 by Bench, Kowal and Bamford for partially hearing children. Since then they have been commonly used as standardised speech material for audiological testing. The material consists of 20 lists with 16 sentences in each, representing a natural language sample. There are either 3 or 4 key words per sentence, with a total of 50 key words in each list. An example of a list in given in appendix III.

3.2.3.1 Purpose of the test

To measure the speech recognition threshold in noise monaurally for both ears. Measurements included:

- In quiet, diotically
- In stationary noise (ICRA-1, male-weighted version, same gender as the speaker), monaurally at both ears
- In fluctuating noise (ICRA-5_250, male-weighted version same gender as the speaker), monaurally at both ears

The ICRA noises are described below. The noise level was fixed at MCL-level (for BB noises with a maximum of 85 dB SPL) mentioned above, adaptively varying the speech level. This meant that the SPL of ICRA noise was same as the SPL of BB noise at MCL. The outcome measure is the speech recognition threshold (SRT): the signal-to-noise ratio (SNR) for 50% correct (except for the quiet condition where the outcome measure is the speech level for 50% correct).

3.2.3.2 Noise specifications

It was decided to use the same interfering noises in the SRT measurements in different languages within the multi-centre study, namely the stationary ICRA noise and the fluctuating ICRA noise that represents one interfering speaker with limited pause durations of 250 ms maximum (Dreschler *et al.,* 2001; Wagener *et al.,* 2006). Additional to the respective ICRA noises that represent a male long-term spectrum (ICRA 1, ICRA-5_250), ICRA noise with female long-term spectrum were also used (ICRA 2, ICRA-4_250) according to the speaker of the speech material. The ICRA noise has been developed for the International Collegium of Rehabilitative Audiology by the HACTES work group (Hearing Aid Clinical Test Environment Standardisation). The purpose was to establish a collection of noise signals to be used as background noise in clinical tests of hearing aids and possibly for measuring characteristics of non-linear instruments. The composed signals have well defined spectral and temporal characteristics similar to those typically found in real life speech signals and babble noise. They also have long-term average spectra and modulation characteristics like natural speech and include gender specific spectra corresponding to male and female speech in close accordance with LTASS (Byrne *et al.,* 1996) and the ANSI S3.5 (1997) standard (for the calculation of the SII).

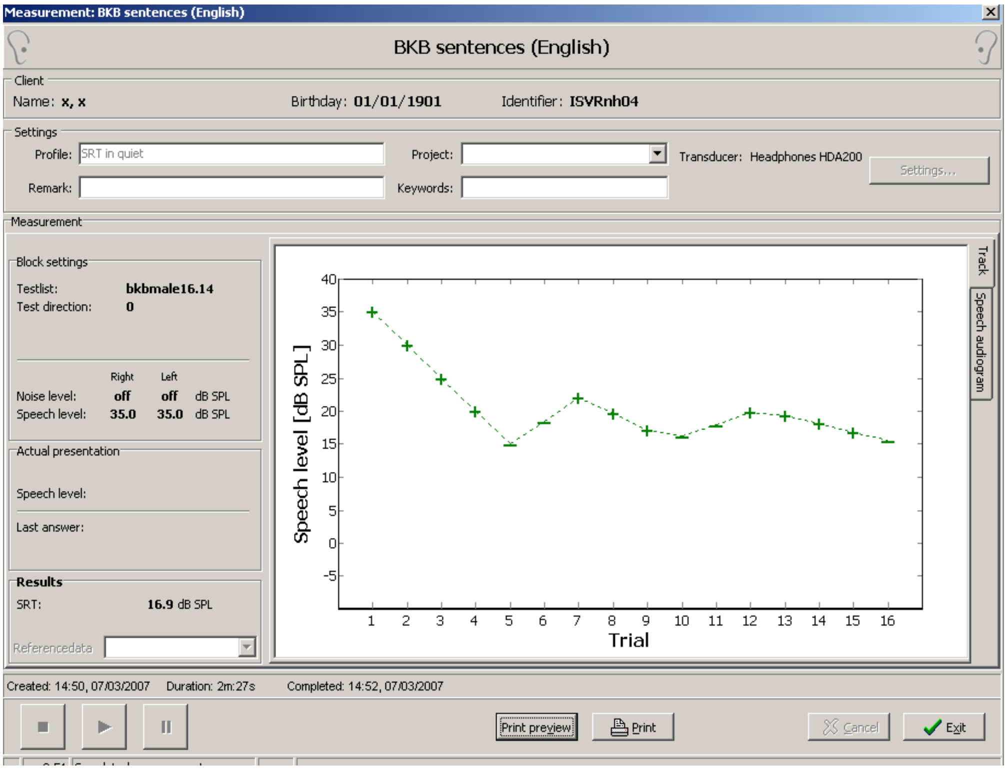


Figure 3.4: Example of an adaptive trial run for determining SRT score using BKB sentences.

3.2.3.3 Adaptive test procedure

The above figure illustrates an example of an initiated adaptive test trial or a graphical representation of the actual measurement. The symbols representing the trials have the following meaning:

+ At least 50 % of key words were recognized for this trial

–Less than 50 % of key words were recognized for this trial

Thus, for the English test version one BKB sentence list constituted one trial. The adaptive test procedure by Brand and Kollmeier (2002) was used for SRT prediction. It starts with SNR of 0 dB and then the signal level is varied (with noise level fixed at the MCL obtained from ACALOS measurements). The signal level is varied at ±5 dB steps depending on the response for first two runs and then ±0.5 dB. In other words if the word is identified correctly the signal level is reduced and if wrongly, the signal level is increased as described above. An average of these correct/wrong responses (+/– responses, see fig 3.4) at the end of the presentation and responses from 16 sentences determines the final SNR. The first two runs are not included in the calculation of the average SNR. The noise level did not exceed 85dB to avoid very high SNRs and loud noise exposure.

3.2.3.4 Different language tests

As mentioned in the beginning of the chapter, the present research constituted a multicentre study with three different centres besides the UK and hence language specific speech tests were used for each of them [Göttingen sentences (Brand and Kolleimeier, 2002) for German, Versfeld sentences (Versfeld *et al.*, 2000) for Dutch and Swedish-HINT (Hallgren, *et al.*, 2006) in Swedish].

3.2.4 Frequency and temporal resolution

The combined method of measuring frequency and temporal resolution by Larsby and Arlinger (1997) was used.

3.2.4.1 Purpose of the test

To measure frequency and temporal resolution at high (3000) and low (500) frequencies for left and right ears (monaurally).

3.2.4.2 Method

a) Signal: A Bekesy tracking technique with a pulsed test tone at a fixed frequency (500/3000 Hz) was used to determine the masked hearing thresholds. The signal/pulsed tone level is changed at a rate of 3 dB/sec (with the tone pulsed at 2.22 pulses per second with 50 ms rise/fall time, 175 ms plateau,175 ms silent interval).

b) Masking noise: The masking noise was an octave band noise mixed with a continuous white noise with a broadband spectrum level of minus 30 dB relative to the octave band noise. This noise was then modified to give the following three masked conditions:

- with no gap
- noise with spectral gap (0.5-octave wide gaps around signal frequencies)
- noise with temporal gap (10-ms silent periods symmetrically placed around the centre of the test one)

c) Measurement: Release of masking values (calculated as the difference in hearing thresholds between the condition with continuous noise and the condition with spectral or temporal gaps) are used as measures of spectral and temporal resolution. Threshold measurements always stopped automatically after nine turning points. A 'measurement' here refers to a full set of six 'threshold measurements' that were used to calculate spectral and temporal resolution values. Thus there are two repeats of each of three conditions.

3.2.5 ILD and BILD test

3.2.5.1 Binaural hearing tests

These constituted the ILD, BILD and MAA tests. All these tests required HRTFs to be incorporated into signals so that equivalent free-field calibration of all signals can be achieved and they can be presented via earphones. The details of this virtual set-up and different binaural conditions are discussed in chapter II (2.4.6).

The ILD and BILD were measured using the UK Matrix Test. The English speech sentence material was developed especially for this test using a female speaker. Development of such speech material was pre-requisite for the multicentre study since the methods developed should be uniform in order to analyse the results across different centres and languages. Development and evaluation of this material in English is included in the Appendix IV. The test was designed to be equivalent to the pre-existing tests in Swedish (Hagerman, 1982), OLSA for German (Oldenburg Sentence test, Brand and Kollemeier, 2002), NL-matrix for Dutch. The noise level was fixed and the same noises as described in section 3.2.3.2 above were used. The adaptive procedure used by Brand and Kollmeier (2002) as above was used for SRT scoring. However sentence scoring was used for the present test, whereby all five words in a sentence had to be correct.

3.2.5.2 Purpose of the test

To measure spatial hearing in terms of intelligibility level difference (ILD) and binaural intelligibility level difference (BILD) using the matrix speech material. As these tests were all conducted via headphones, virtual stimuli were used. Stimuli were filtered with generic HRTF from the appropriate directions. For the ILD and BILD test this means that the speech signal was always filtered with the HRTF for 0^{o} and the noise was filtered either with the HRTF of 0^{o} or with the HRTF of $+90^{o}$ or -90^{o}.These signals were mixed and presented dichotically, except for the monaural BILD conditions. The noise level was fixed at MCL and the speech level was varied adaptively during the test to obtain the SRT.

3.2.5.3 ILD test

For the English test, speech recognition thresholds were measured in three conditions with noise (ICRA-1, female-weighted version,):

S_0N_0: speech and noise both coming from the front (0°)

S_0N_{90}: speech coming from the front (0°) and noise coming from the right side (90^{o})

S_0N_{-90}: speech coming from the front (0°) and noise coming from the left side (-90^{o})

The ILD is defined as the SRT difference between the S_0N_0 and the S_0N_{90} or S_0N_{-90} measurement.

3.2.5.4 BILD test

To estimate the BILD, two additional, monaural, measurements were conducted.

S_0N_{90}: speech coming from the front (0°) and noise coming from the right side (90^{o}) with the right ear blocked acoustically (so both signals are directed monaurally to the left ear)

S_0N_{-90}: speech coming from the front (0°) and noise coming from the left side (-90^{o}) with the left ear blocked (both signals presented monaurally to the right ear)

The BILD represents the SRT difference between the monaural and binaural S_0N_{90} and S_0N_{-90} measurements and was averaged across −90 and +90 conditions. The relevant measure (binaural squelch) is the improvement obtained by unblocking the ear nearest the noise, so that dichotic cues are available.

3.2.6 Localization (minimum audible angle)

Localization ability was measured using a minimum audible angle test in the horizontal plane. Minimum audible angle thus refers to the smallest difference in virtual location between two sound sources that result in a different perceived location. The test used a virtual localization set-up and was developed for the project.

3.2.6.1 Purpose of the test:

To test the localization ability binaurally.
Specifically, a virtual version with binaural information introduced by applying head related transfer function mentioned in 2.4.7.1 was used in the present study.

3.2.6.2 Procedure

For each test trial a stimulus duration of 300 ms and an inter-stimulus interval of 300 ms was applied. Two stimuli (modified using HRTF) were presented consecutively from different virtual directions via earphones, symmetrically spaced around 0°. The order left first or right first was randomized. If the sounds were perceived from different angles the result was the impression of a moving sound. The task for the listener was to determine whether the sound was moving from left to right or from right to left. The test started with an angle of 32° between the two stimuli (±16°), to make sure the task was easy for the listener. In subsequent trials, the angle was reduced after two correct responses and increased after one incorrect response by means of an adaptive two-down one-up tracking procedure. In this way, a threshold value of 70.7% correct is obtained. For the first two reversals, the step size was 4°, in order to quickly reach the approximate threshold value. After two reversals the step size decreased to 2° and after four more reversals to a final value of 1°. The test continued for eight reversals after the minimum step size was reached. The MAA value was defined as the average over those last eight reversals.

Six MAA measurements were obtained for each listener to give a reliable estimate of the MAA, and to get a good impression of the variance in MAA within listeners and any possible learning effect. The final MAA estimate was the average over these six adaptive runs

3.2.6.3 Stimuli / spectral content

An important consideration for the choice of stimulus for a localization test is that the cues for sound localization (interaural level difference, ILD; interaural time difference, ITD and direction-dependent spectral filtering by the pinnae) are dependent on the frequency content of the sound. The ILD cue is mainly present in natural conditions for high frequencies. For low frequencies the large wavelengths allow the sound waves to easily diffract around the head and the ILD becomes negligible. The ITD cue is mainly important for low frequencies, as for high frequencies the wavelength becomes smaller than the difference in path length between the two ears and the ITD cue becomes ambiguous. Spectral filtering cues are useful for frequencies higher than about 4 kHz. Three sets of stimuli for the MAA localization test were used with all being presented at MCL obtained for the following three conditions (500 Hz for low pass, 3000 Hz for high pass and BB for broadband):

- broadband white noise, in which all localization cues are available
- low-pass noise, in which only ITD information is available and
- high-pass noise, in which ILD information and spectral cues are available

3.2.7. Listening effort

The listening effort method used in the present study uses the Hörtech scaling procedure developed as part of OMA(Oldenburg Multiple Applications) (explained later in 3.4). It basically uses speech material or more precisely short clips of a simple story that are mixed with noise. The English speech material was recorded specifically for this test and comprised segments of the children's story 'Irritating Irma' read by the same female speaker as for the Matrix Test. The subjective effort of hearing/understanding of speech in noise is an important point because even if speech is identified correctly, there maybe a higher load to the working memory to achieve this identification. Noise may, in effect, impose an additional "secondary task" that must be carried out whenever speech has to be understood. Though the test is subjective, the stimulus levels are controlled and hence structured. The test uses the signal-to-noise ratio of the story discourse to get a score for rating listening effort.

Thus this test measures the subjective effort required to listen to a speech/signal in non-optimal conditions. This is done by the subjects rating "how much effort is required to understand the meaning of speech?". This effort is measured on a continuous scale ranging from "no effort" to "very much effort". For computer implementation a graphical user interface was used. The subject can move the slider in 100 steps over the entire range of the scale. The slider was centered after each presentation. If the rating is finished, it has to be indicated by pressing a button and the next signal to be rated will be played. The presented signals were digitally mixed online using fluent speech with the same ICRA noises as used in other tests (either ICRA1 for continuous noise or ICRA5_250 fluctuating noise, both female weighted) at an SNR of either +/− 5 dB. All possible combinations of noises and speech are rated by the subject.

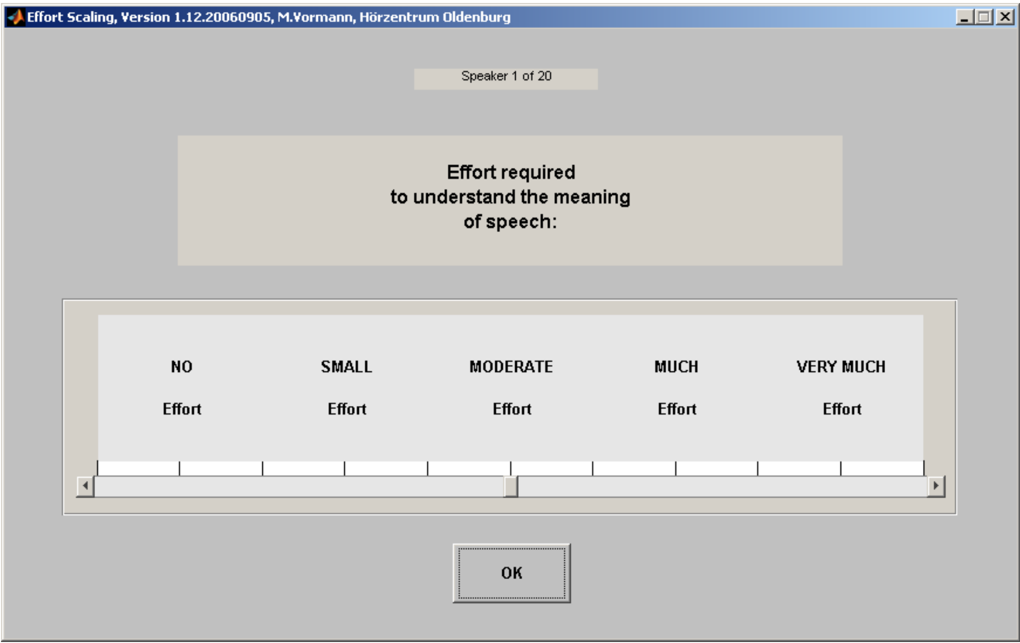


Figure 3.5: User interface picture for listening effort scale. It reveals the scales range from no effort to very much effort as scored by placement of the slider.

3.2.7.1 Scoring

As can be seen the scale extends to five categories from no to very much effort with each category having 20 points on the scale for the calculation of the total effort required with 1-20 assigned to no effort, 21-40: small effort, 41-60: moderate effort, 61-80: much effort, 81-100: very much effort. These are further averaged over 20 trials of the story with five for each of the four different SNR conditions. Thus the lower the score, the better is the performance with the best possible rating being between 1-20 for the 'no effort' scale.

3.2.8 Self-reported hearing difficulty

As a measure of self-report, the Gothenburg Profile (GP: Ringdahl, 1998) was used. This profile is a questionnaire about issues related to listening to sounds in daily life.

3.2.8.1 Purpose of the test:

To subjectively assess the subject's level of difficulty (handicap) in various everyday situations.

The GP test is basically a questionnaire answered and rated by the hearing impaired subject in order to acquire his/her psycho-social profile on various aspects of hearing loss. It includes 20 questions which describe situations in daily life where hearing is important. The full questionnaire is included in Appendix V. It was developed with content partly taken from the shortened Hearing Measurement Scale (HMS25, Erickson-Mangold *et al.*1992). The 20 items are divided into four subscales. The first subscale measures Experienced Disability in hearing speech (items 1-5) and sound localization (items 6-10). The second subscale targets the Experienced Handicap in social settings (items 11-15) and the personal reactions to the experienced handicap (items 16-20). Originally eleven scaling categories were used ("0" = no disability/handicap, "1", "2", … , "10" = maximum disability/handicap). In the present study, the following names to numbers (categories): were used, similar to eleven categories, used by Kiessling (1996).

Category “0” = never

In-between cat. “2” and “3” = rarely

Category “5” = sometimes

In-between cat. “7” and “8” = often

Category “10” = always

3.2.8.2 Response subscale and scoring

There are no “correct” responses. But rather the rating is important to assess the attitudes of subjects towards their hearing. Items in each subscale are scored and averaged separately and then a percentage score is assigned to the average which determines the ultimate score. The smaller the score the better is the performance.

3.2.9 Test of cognition

Cognition as a part of the hearing test battery is important mainly for the following reasons:

To establish optimal function of cognitive processes which underlie all sensory processes including hearing

To establish that no cognitive deficit is influencing the evaluation especially in elderly patients

This test was chosen from pilot experiments done by HEARCOM which investigated the which cognitive skills are important for speech recognition/comprehension as measured in the Hagerman speech test and in the Swedish Hearing In Noise Test (HINT). Lexical test was found to correlate well with the speech tests above and takes only few minutes to complete. Hence was chosen to be included in test battery (details in Appendix VIII

3.2.9.1 Purpose of the test:

To assess the cognitive functioning of the subjects.

Cognition was tested via the lexical-decision task which estimates the lexical skills (e.g. Bowles and Poon, 1981; Howard, 1983) of subjects. The task is to discriminate words from non-words. These word items are organised in lists of real-word / non- word combinations.

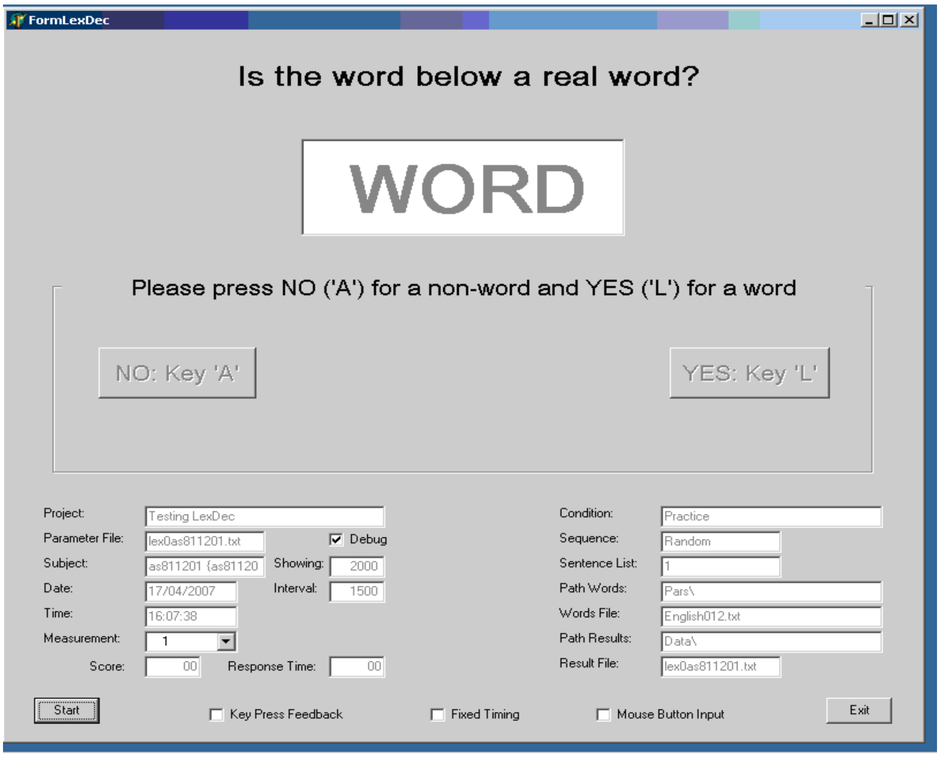


Figure 3.6: User interface picture for lexical decision test.

3.2.9.2 Task and Scoring

As can be seen, the word to be scored is visible in the box and the subjects can respond by A if it is a non-word or L if it is a real word on the keypad. Abbreviations and acronyms are considered as non-words. During the test, items are selected at random from these word pairs. The program records response times and correct scores which give a measure of the performance on the test. Thus, lower values refer to quicker response time and hence better performance. Similarly, the greater the mean correct scores, the more words are guessed correct and hence better performance.

Three measurements are carried out in total, with the first one as a practice trial containing 12 test items (words not non-words) while the next two measurements contribute towards actual scoring each containing 50 test items.

An English version of the test was developed for the present project. A list of 50 meaningful words and 50 non-meaningful words was designed. All the words were short, three lettered. The meaningful words were nouns, verbs, prepositions, determiners, adverbs, adjectives, and interrogatives while the non-words were of two types, ones that can be pronounced (e.g. shu) and ones that cannot be pronounced (e.g. dza). A complete list of the words is given in Appendix VI.

3.3 Summary of the final test domains, categories and conditions

3.3.1 Test domains

In the following table, the test domains discussed above are summarised along with actual tests included and their selection:

Table 3.1: Selection of actual tests for the auditory and non- auditory domains in the test battery.

Test no.	Test domain	Test included	selection
1	Hearing threshold	Pure tone audiometry (PTA)	routine
2	Loudness perception	ACALOS Adaptive CAtegorical Loudness Of Sounds	Standardised test available (initiated and developed by partners, Brand and Hohmann (2002)
3	Speech recognition threshold	BKB (Bench ,Kowal and Bamford) sentence test	Standardised available test in British English
4	Frequency and temporal resolution	FT test by Larsby and Arlinger	Developed & standardised by Larsby and Arlinger (1997)
5	Binaural hearing	Intelligibility level difference(ILD) and Binaural intelligibility level difference (BILD)	Developed and evaluated as part of project (included in Appendix IV)
6	Localization	Minimum Audible Angle (MAA)	Developed as part of project. at ISVR, the cross-talk cancellation version was trialed on the HearCom portal
7	Listening effort	Hörtech Effort scaling test	Available and standardised: speech story material in British English developed as part of project
8	Self-report	Gothenburg Profile	Available and developed by Ringdahl *et al*. (1998)
9	Cognition	Lexical decision test	Evaluated through a pilot study and group of other cognitive tests as part of the project

3.3.2 Test Categories

The tests can be divided into four categories depending upon the nature of testing:

Monaural: tested separately for each ear

Audiogram

Acalos

FT

BKB (SRT in noise)

Binaural:

BKB (SRT in quiet)

MAA

ILD

BILD

Subjective (self-report)

GP

Effort Scaling (listening effort)

Cognitive

Lexical decision test

3.3.3 Test Conditions

In the table below all the tests along with their measurement conditions are given:

Table 3.2: Outline of different test conditions and measurements of the test battery.

Group	Test Domain	Test	Conditions measured and Details
Monaural	Audibility	Audiogram	air conduction: 250/500/1000/2000/3000/4000/ 6000/8000 Hz bone conduction: 250/500/1000/2000/3000 Hz
Monaural	Loudness perception	Adaptive Categorical Loudness of Sounds (ACALOS)	narrowband noises (500 Hz, 3000 Hz) broadband noise (male,female)
Monaural	Frequency-time resolution	FT test	500 Hz 3000 Hz
Monaural	Speech perception	Speech recognition threshold)SRT with BKB sentences	in quiet (binaural) in stationary noise (monaural) in fluctuating noise (monaural)
Binaural	Localizatio n	Minimum Audible Angle (MAA) test	broadband noise low-pass noise high-pass noise
Binaural	Binaural processing	Intelligibility Level Difference (ILD)	SRT with matrix-type sentences
		Binaural Intelligibility Level Difference (BILD)	SRT with matrix-type sentences
	Cognitive abilities	Lexical-decision Making	
Self-report	Subjective judgement	Gothenburg Profile	Questionnaire
	Listening Effort	Effort scaling	Running speech in continuous or fluctuating noise, at SNR= -5 or SNR= +5 dB

3.4 Procedure

All these tests were implemented on a common software platform called Oldenburg Multiple Applications (OMA) as shown in the figure below. All tests were performed using a PC with soundcard (RME DSP 9632 24-bit) and an external amplifier (Creek OBH-21SE) and circumaural headphones (Sennheiser HDA 200) in a sound attenuating room. Stimuli for all tests were synthesized in software or obtained from primary recordings on disc and adjusted digitally in intensity as required before replaying them via the soundcard. All subjects were tested twice; test and retest sessions and the hearing impaired subjects did the tests unaided.

3.4.1 Test set up

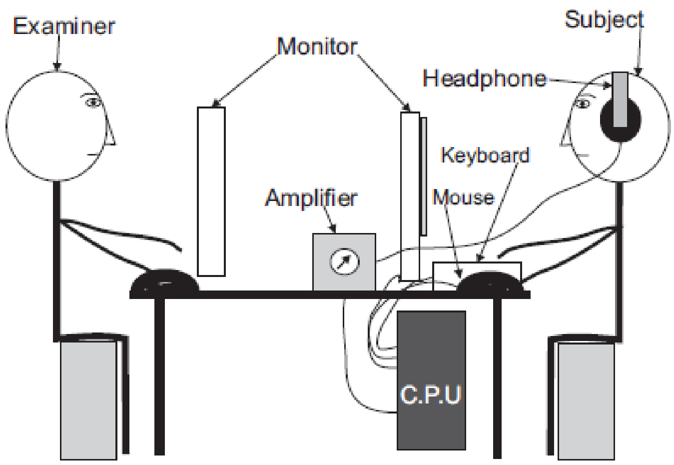


Figure 3.7: Test set up.

As seen above, the set up included a single sound treated room with examiner and subject seated across each other. Both had a separate monitor for their interface.

3.4.2 Test order and time taken

In order to minimize measurement errors caused by varying sequence effects, the order of tests was fixed, while the conditions within the tests were randomised over subjects to avoid biased average results. The test order was as shown in table 3.3:

Table 3.3: Order of tests in the test battery and time taken for each.

Test	Measurement time in minutes
ACALOS	10
Effort Scaling	5
Speech perception	15
ILD test	9
BILD test	6
Gothenburg Profile	5
FT test	15
MAA test	20
Lexical Decision Test	5
Total	90

This order ensured that ACALOS was the first test to be administered. This gave the reference starting level for all the subsequent tests. Further, the Lexical Decision Test was last to be administered, so that the subjects taking the test experience some fatigue, auditory or otherwise and could hence reveal any slowing down of cognitive processing due to previous testing.

3.5 Subject selection criteria

Subjects were selected according to the following criteria

Age between 18 and 75 years

Average hearing loss (1000,2000,4000 Hz, PTA) not more than 65-70 dB, (sensory loss only). The average is taken from the pure tone audiogram thresholds at 1000, 2000, and 4000 Hz

Maximum difference in PTA between the two ears of 30 dB

English as first language

Active and alert and able to perform the tests

Willing to spend two visits for testing (about 2 hours per visit)

No complaints of tinnitus

3.6 Ear Configurations

Each *ear* was categorized in one of the following four audiometric configurations:

- Mild flat hearing loss:
  - PTA ≤ 50 dB
  - Difference between thresholds at 500 and 4000 Hz ≤ 30 dB
- Severe flat hearing loss
  - PTA > 50 dB
  - Difference between thresholds at 500 and 4000 Hz ≤ 30 dB
- Mild sloping hearing loss
  - PTA ≤ 50 dB
  - Difference between thresholds at 500 and 4000 Hz >30 dB
- Severe sloping hearing loss
  - PTA > 50 dB
  - Difference between thresholds at 500 and 4000 Hz > 30 dB

The aim was to include at least 5 *ears* in each of the four categories for each participating centre subject to the availability.

3.6.1 Types of hearing loss

Sensorineural hearing loss was the pre-dominant type to be included defined as air bone gaps at 500 and 1000 Hz ≤ 15 dB

In case of other types: conductive or mixed hearing loss defined as air bone gaps at 500 and 1000 Hz > 15 dB

3.6.2 Ear Symmetry

Symmetry of hearing loss was defined as follows:

• symmetrical hearing loss (difference in PTA between left and right ear ≤ 10 dB)

•asymmetrical hearing loss (difference in PTA between left and right ear > 10 dB)

3.7 Additional details

In addition, there was inclusion of normal-hearing subjects (all pure-tone audiogram thresholds better than 20 dB). All subjects signed an informed-consent form and received a financial reimbursement for their time as well as travelling expenses.

Ethical Committee approval:

The hearing impaired subjects were obtained from the audiology department at Royal South Hants Hospital in Southampton. An ethical committee approval for the study was obtained for the same from Southampton University Hospital Trust. The approval ID is RHM ENT0077.

3.8 Study II

A second multicentre study was carried out following the first. This study was based on outcomes and findings of the first study and hence the description of the second study is more relevant later. Thus the methods used for the study II are described following the findings and analysis of study I.

Having discussed in detail, the test measures included in the thesis, the next chapter deals with analysis of the data collected using these tests.

Chapter Four

Results and Interpretation-I

4.1 Introduction

This chapter is constitutes the analysis, interpretation and discussion of the data gathered on the test battery described in Chapter 3.

The analysis is divided into three sections:

I. General and Descriptive Statistics
II. Group I analysis: Explanatory (aims 1-4)
III. Group II analysis: Multiple Dependant Factor (aim 5)

It should be noted that part I (General and Descriptive Statistics) concerns basic analysis of the individual tests such as difference between normal hearing and hearing impaired or between left and right ears. Thus the comparisons focus on within test differences, whereas II and III are of more interest where the focus is on relationships and interactions between the different groups of tests.

4.2 Test measures included in the data analysis

To recap, the following test measures were used.

Table 4.1: Test measures and conditions included in the data analysis.

Hearing Aspect	Test	Conditions measured & Details
Audibility	Audiogram	air conduction: .250, 500, 1000, 2000, 3000, 4000, bone conduction: 250, 500, 1000, 2000, 3000Hz
Loudness perception	Acalos	narrowband noises (500 Hz, 3000 Hz) broadband noise
Frequency and time resolution	FT test	500 Hz 3000 Hz
Speech perception	SRT with BKB sentences	in quiet (binaural) in stationary noise (monaural) in fluctuating noise (monaural)
Binaural processing	MAA test	broadband noise low-pass noise high-pass noise
	ILD	SRT with matrix sentences
	BILD	SRT with matrix sentences
Cognitive abilities	Lexical Decision Making	
Subjective judgement	Gothenburg Profile	Questionnaire
	Listening Effort	Running speech in stationary or fluctuating noise, at SNR=-5 or SNR=+5

4.2.1 Outcome measures and abbreviations used in the analysis

Listed below are different ***outcome measures*** in each test used in the analysis along with their ***abbreviations:***

Table 4.2: Outcome measures and abbreviations of the test measures and conditions included in the data analysis.

Outcome measure	Abbreviation
Per ear measurements	
Air conduction threshold at 500 Hz	AC 500
Air conduction threshold at 1000 Hz	AC 1000
Air conduction threshold at 2000 Hz	AC 2000
Air conduction threshold at 3000 Hz	AC 3000
Air conduction threshold at 4000 Hz	AC 4000
Pure tone Average of 500, 1000, 2000, 4000 Hz	PTA
Difference in thresholds between 500&4000 Hz	AGM SLOPE
Most comfortable level at 500 Hz	MCL500
Most comfortable level at 30000 Hz	MCL3000
Most comfortable level for broadband	MCLBB
Loudness level slope at 500 Hz	MCL500SLOPE
Loudness level slope at 3000 Hz	MCL3000SLOPE
Loudness level slope for broadband	MCLBBSLOPE
Masking release measures for frequency and Temporal resolution:	
Frequency resolution at 500 Hz	F500
Temporal resolution at 500 Hz	T500
Frequency resolution at 3000 Hz	F3000
Temporal resolution at 3000 Hz	T3000
Speech recognition threshold in stationary/continuous noise	SRTstat
Speech recognition threshold in fluctuating noise	SRTfluc
Binaural measurements	
Speech recognition in quiet	SRTq
Intelligibility Level difference	ILD
Binaural Intelligibility Level difference	BILD
Minimal audible angle High Low Broadband	 MAAh MAAl MAAbb
Subjective measurements	
Listening effort Continuous noise with SNR -5 dB Continuous noise with SNR +5 dB Fluctuating noise with SNR -5 dB Fluctuating noise with SNR +5 dB	 EffC5 EffCmin5 EffF5 EffFmin5
Gothenburg Profile Speech Localization Social Behaviour	 GPsp GPloc GPsoc GPbeh
Cognitive	
Lexical decision	LDT

Other abbreviations used are normal hearing (NH), hearing impaired (HI), speech recognition threshold (SRT), pure tone average (PTA), hearing threshold level (HTL). Also additional variables of asymmetry (difference between left and right ears) are used with suffix ‘diff’.

4.3 Procedure

The tests were performed in the order as listed below with test lists and frequencies randomised over subjects. Thus loudness perception was measured using the ACALOS test which gave the Most Comfortable Level (MCL) equivalent to categorical unit (CU) of 20 as explained in the previous chapter for each subject. This level (in dB) was then used as a reference starting level for the other tests following it.

Table 4.3:Test order.

ACALOS
Effort Scaling
Speech perception
ILD test
BILD test
Gothenburg Profile
F-T test
MAA test
Lexical decision making test

4.4 Subjects included

A total of 103 subjects: 30 NH and 73 HI with a wide range of hearing losses from all the centres were included in the analysis. NH: age range 19-39 years, average PTA (1, 2, 4): 5 dB HL. HI: age range 22-91 years, average PTA (1, 2, 4): 44 dB HL (range: 22-77 dB). The shaded area around the thresholds in 4.1a and b displays the approximate range or variance in dB HL for each frequency. Audiometric configurations of HI: Mild flat (38), severe flat (16), mild sloping (14), severe sloping (4); 13 subjects had asymmetric hearing loss and 60 had symmetric hearing loss as per criteria described in chapter 4. Finally, 4 ears had a mixed hearing loss and 69 ears had a purely sensorineural hearing loss. In figure 4.1 below the average thresholds from 500-4000 Hz are plotted graphically for hearing impaired (a) and normal hearing (b) subjects in the form of audiogram.

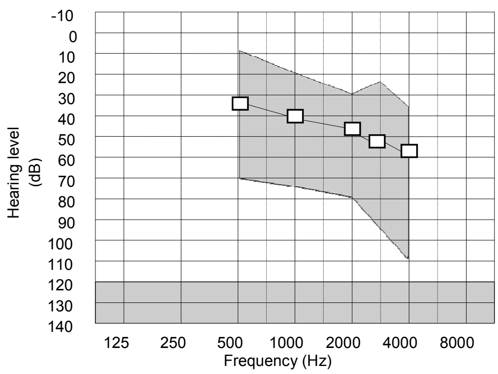


Fig 4.1a study I Thresholds: Hearing impaired.

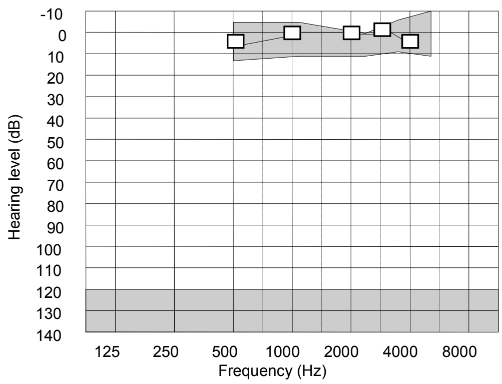


Fig 4.1b study I Thresholds: Normal hearing.

4.5 General and Descriptive Statistics

4.5.1 Descriptive statistics

The main aim of this analysis was to outline the differences between the NH and the HI groups. Useful measures would show clear differences between the two groups. This is done with box plots for the two groups for all the measures. Better ear measures (based on PTA) are included for the analysis. The Y-axis consists of the test measure of concern. Further, the box plots show the median (black bar), interquartile range (box ranges between 1st and 3rd quartile, whiskers represent highest and lowest values after exclusion of outliers). The outliers (circles, defined as any point which falls more than 1.5 times the interquartile range above the third quartile or below the first quartile) and extreme cases (stars, any point beyond the outlier) have subject numbers alongside for identification. Also the text box on the right highlights the relevant information such as frequency, type of noise, category etc for that measure. Further independent sample t-test was carried out to observe if the difference between NH and HI group scores are significant or not. They were significant for all measures for both groups ($p<0.05$). (see Appendix IX for details).

4.5.1.1 Audiogram thresholds

Figure 4.2 (a-f) below shows hearing thresholds from 500-4000 Hz for NH and HI listeners. It can be seen that in general, hearing-impaired listeners (HI) have higher/worse thresholds and steeper slope than normal-hearing listeners (NH). Moreover, there is more spread in the HI data than in the NH data.

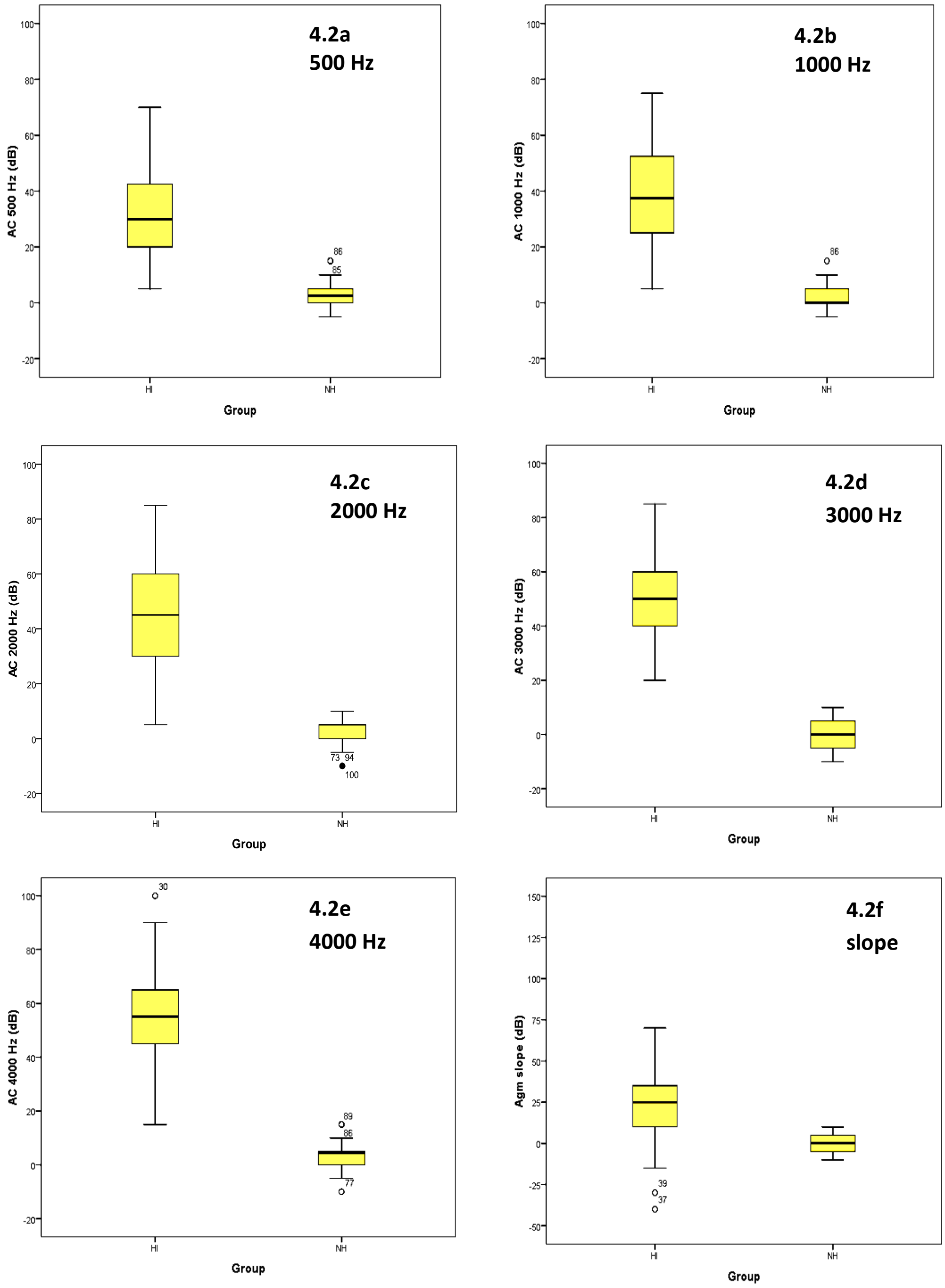
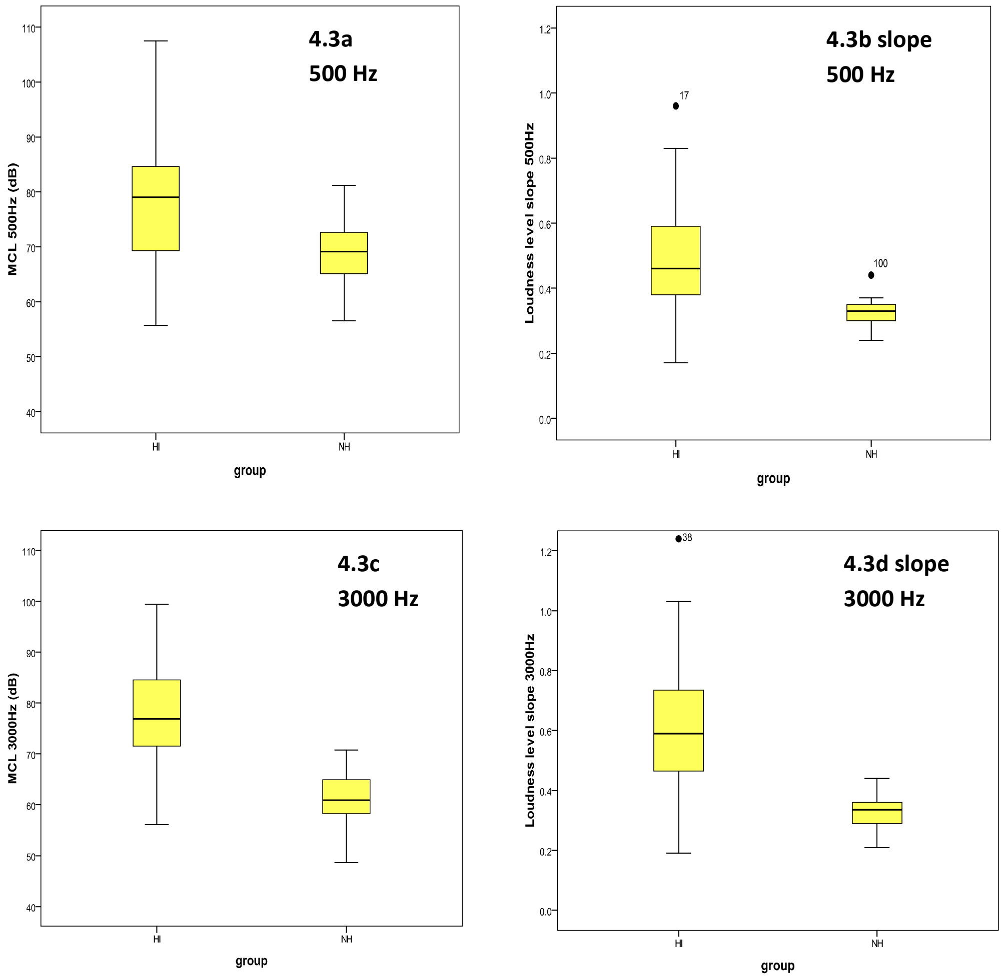
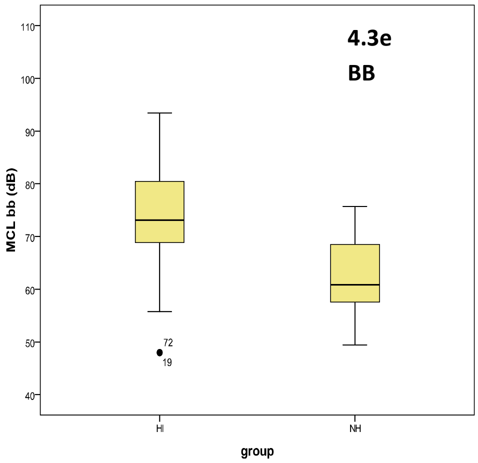


Figure 4.2: Range of Audiogram thresholds in dB HL at 500 Hz (a), 1000 Hz (b), 2000 Hz (c), 3000 Hz (d), 4000 Hz and Audiogram slope (e) for NH and HI subjects. Vertical axes show levels in dB HL. For both measures, NH subjects perform better overall (smaller values signify better performance).

4.5.1.2 ACALOS (Loudness levels)

ACALOS results are shown below in figure 4.3 (a-f). It can be seen that in general, hearing-impaired listeners (HI) have higher MCLs and steeper slopes than normal-hearing listeners (NH) with some overlap especially at MCLbb. Moreover, there is more spread in the HI data than in the NH data.





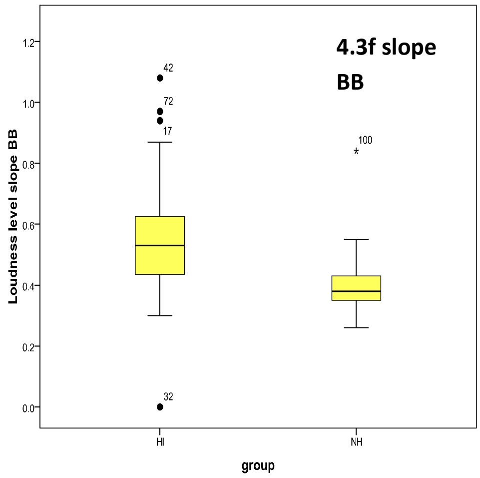


Figure 4.3: MCL at 500 Hz (a), 3000 Hz (c), BB (e) and their respective slopes 500 Hz (b), 3000 Hz (d), BB (f) for NH and HI subjects. Vertical axes show levels in dB HL for a, b, c with smaller values indicating lower levels and slope values for d, e, f with higher values indicating steeper slopes and presence of recruitment.

4.5.1.3 Frequency and temporal resolution (FT)

Figure 4.4 (a-d) shows results of the FT test for NH and HI listeners. On average, it can be seen that NH listeners have slightly better spectral and temporal resolution (greater release of masking values shown by more negative number) than HI listeners. Differences are more pronounced in the resolutions at high frequencies and less for 500 Hz.

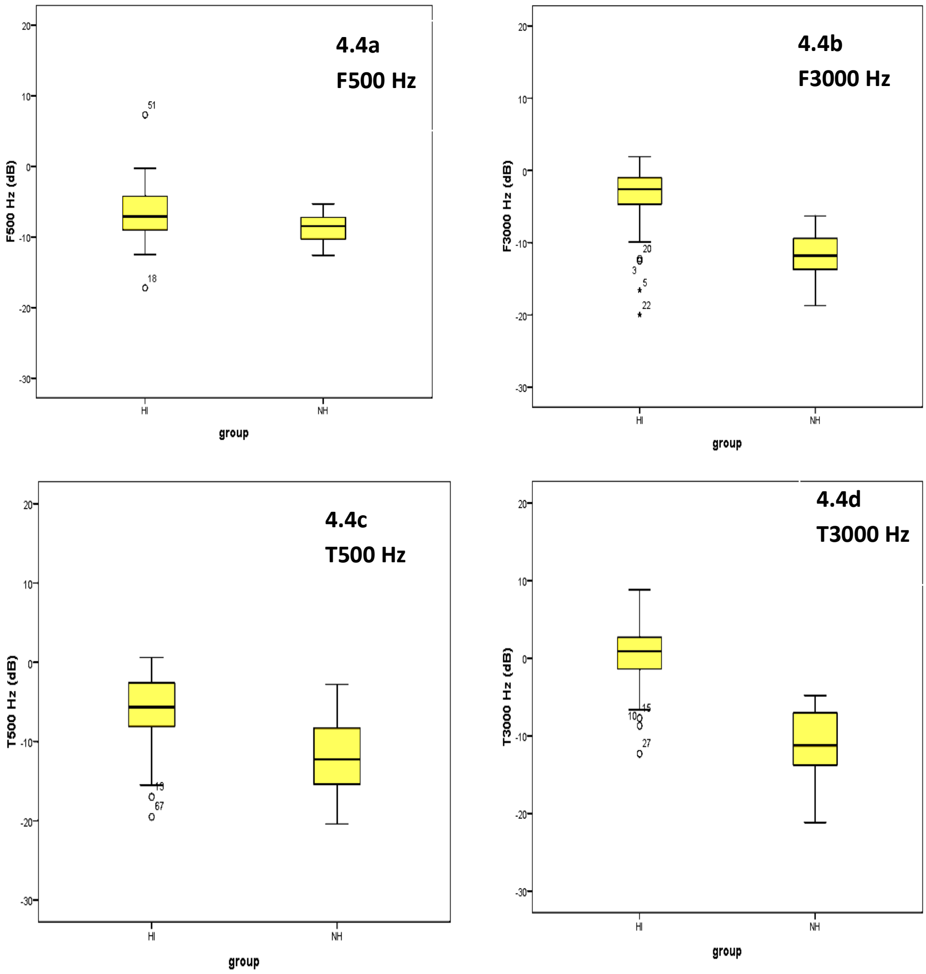


Figure 4.4: Frequency and temporal resolution for normal-hearing and hearing-impaired listeners. frequency resolution at 500 Hz (a), frequency resolution at 3000Hz (b), temporal resolution at 500 Hz (c), temporal resolution at 3000 Hz (d). Vertical axis represents resolution scores in dB with more negative values indicating better resolution.

4.5.1.4 Speech Recognition Threshold (SRT) in noise

Figure 4.5 (a-b) shows corrected SRT results in stationary and fluctuating noise. As expected, there is very little spread in the NH data, and more spread in the HI data. Additionally, differences between NH and HI listeners are larger in fluctuating noise than in stationary noise. Note that comparison between the two types of noise should recognize that results have been corrected to compensate for language differences between the various centres of the multicentre study, such that NH subjects have scores that are centred on or distributed close to zero for both noises.

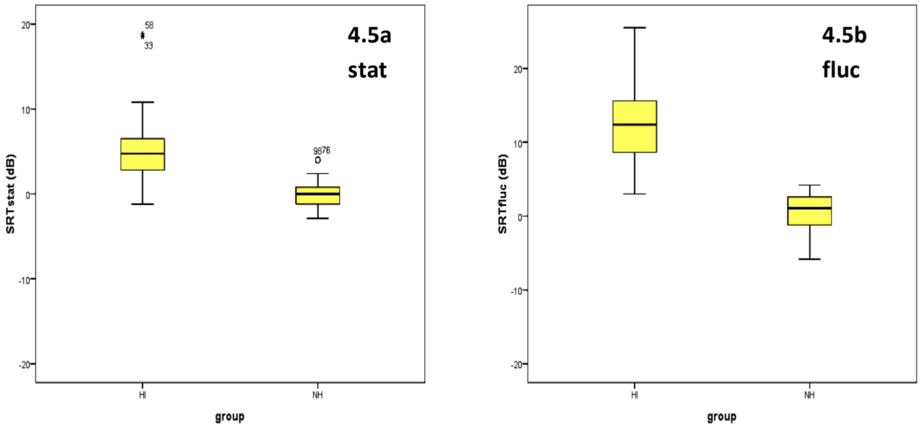


Figure 4.5: Speech recognition threshold for normal-hearing and hearing-impaired listeners, stationary noise (a), fluctuating noise (b). Vertical axis represents speech to noise score in dB, with lower/more negative SNR score indicating better performance.

4.5.1.5. Speech Recognition Threshold (SRT) in quiet

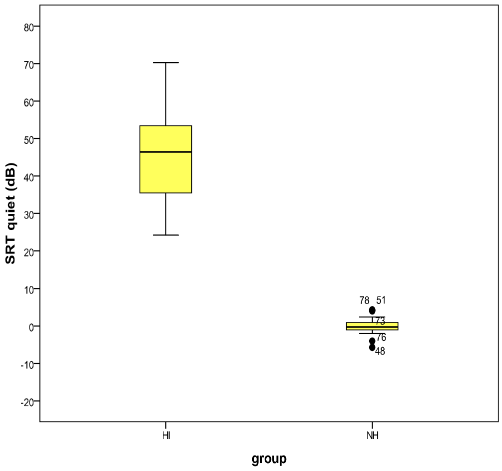


Figure 4.6: SRT in quiet (binaurally) for normal-hearing and hearing impaired listeners. Vertical axis represents the SRT in dB HL, with more negative value indicating better performance.

Results of the binaural SRT in quiet for normal-hearing and hearing-impaired listeners are shown in figure 4.5. Among HI, there is enormous spread in the results, and overall, they perform much worse than NH (higher SRT levels).

4.5.1.6 (Binaural) Intelligibility Level Difference (ILD and BILD)

Results of the binaural processing tests ILD and BILD are shown in figure 4.7(a-b). The ILD is the difference in SRT between noise and speech from straight ahead in virtual space (situation 1), and speech from straight ahead with noise from one side (situation 2). BILD is the difference in SRT between situation 2, and the same situation with ear at the ‘noise-side’ blocked. Again these are corrected for language differences as mentioned in 4.5.1.4. In both tests, more negative values refer to greater release of masking and therefore better binaural processing whereas more positive values means that the subject is unable to use binaural cues or has limited benefit from binaural processing . Only better ear results are presented here.

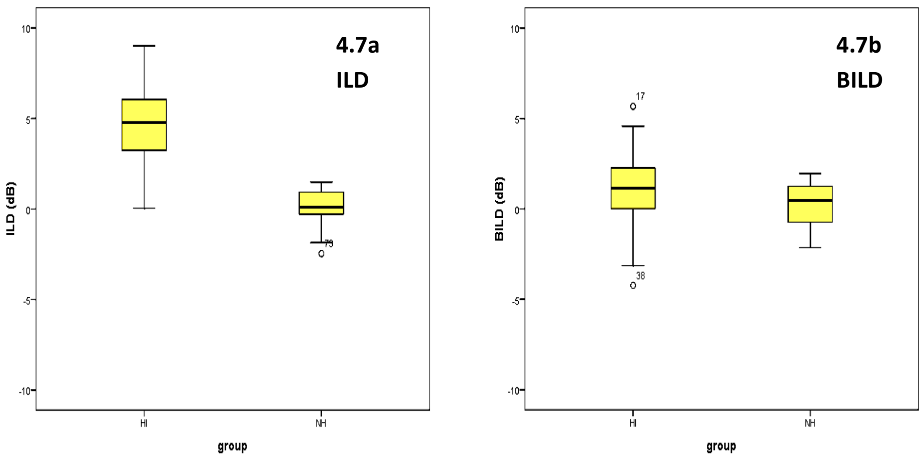


Figure 4.7: ILD (a) and BILD (b) with noise at the side of the poorer ear and hence speech was more heard at the better ear. Vertical axes show release of masking for the two conditions (more negative values refer to which means favourable SNR at the better ear and better binaural hearing and positive values mean less favourable SNR at the better ear and hence poor binaural hearing).

Hearing-impaired listeners have less benefit from spatial separation (less negative values of ILD) and binaural hearing (less negative values of BILD) than normal-hearing listeners. For ILD there is a considerable difference between NH and HI performance, and relatively little spread in the NH data. For BILD, the difference between NH and HI smaller.

4.5.1.7. Minimum audible angle (MAA)

Figure 4.8 (a-c) shows results of the MAA test for normal-hearing and hearing-impaired listeners, for three different stimulus types (low-pass noise, high-pass noise, and broadband noise). For each condition, the outcome measure of this test is the mean MAA of two measurements on one session day. In general, NH perform better than HI (smaller MAA) and spread among NH is smaller than among HI.

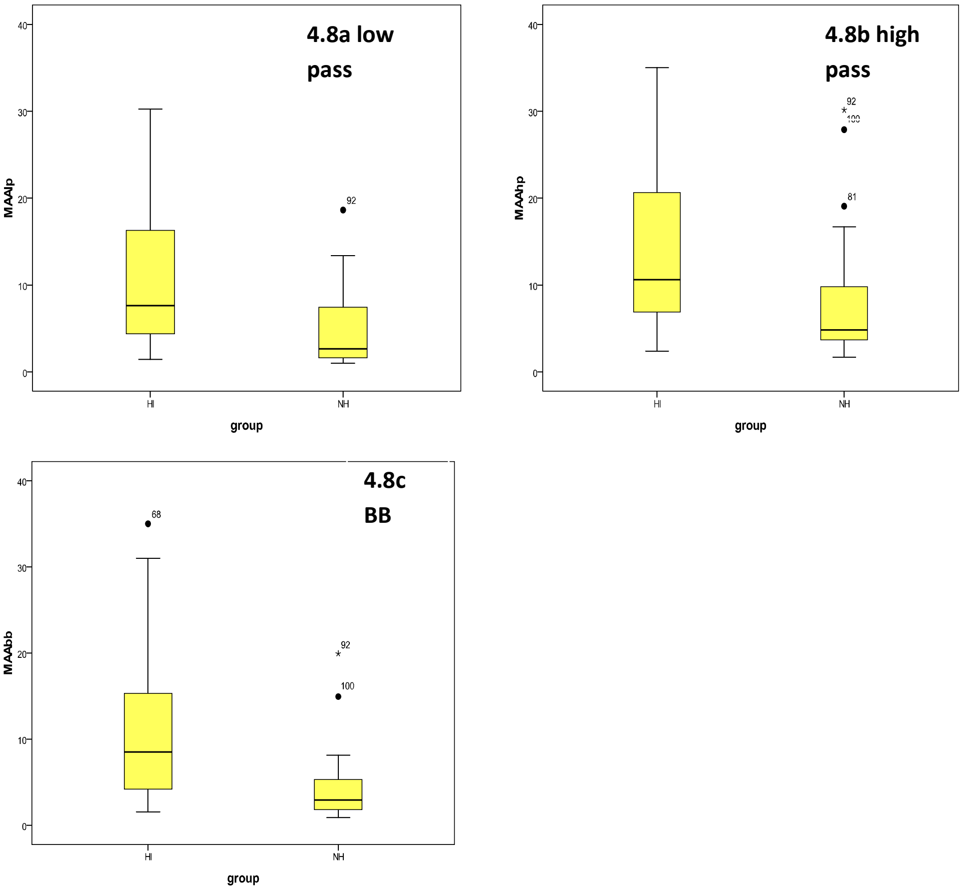


Figure 4.8: MAA results for normal-hearing and hearing-impaired listeners. The three panels show MAAs of low-pass (a), high-pass (b), or broadband (c) noise. Vertical axis shows minimum audible angle in degrees (°) with smaller value indicating better performance.

4.5.1.8 Lexical Decision Test

Figure 4.9 shows results of the cognitive test (Lexical Decision Test). On the vertical axis, percentage correct divided by response time is shown. In general, HI performs worse than NH on this test, although the task is not auditory. To some extent this is probably due to the age difference between the two groups.

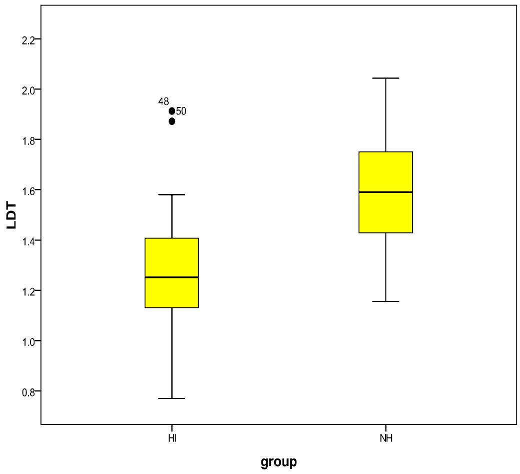


Figure 4.9: Lexical decision making test results. Vertical axis represents (%correct)/(response time), so higher values refer to quicker response time and /or better performance.

4.5.1.9. Listening Effort

Listening effort results for both NH and HI listeners are presented in Figure 4.10 (a-d). It is remarkable that in this test, in contrast to most previously described results, there is large spread in the normal-hearing data, and also normal-hearing listeners need quite some effort to understand the speech in the more difficult situations (SNR= −5). In the most difficult condition (stationary noise at SNR= −5), there is substantial overlap between NH and HI results. However, this is possibly due to the fact that the listening effort scale is subjective. NH and HI groups may respond differently to this subjective scale because the latter group have become used to having difficulties and have lower expectations.

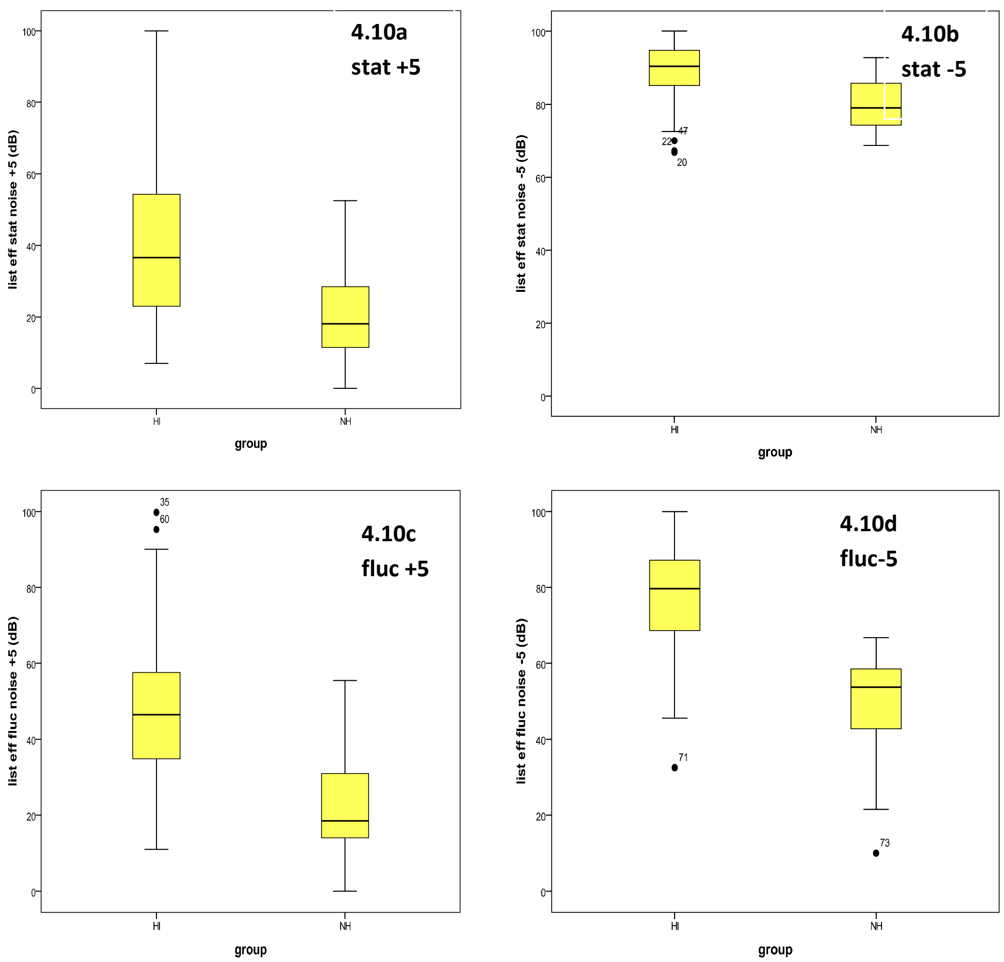


Figure 4.10: Effort required for listening to speech in presence of noise for normal-hearing and hearing-impaired listeners, in stationary noise at SNR+5 (a), in stationary noise at SNR-5 (b), in fluctuating noise at SNR-5 (c) , in fluctuating noise at SNR+5 (d). Vertical axis represents listening effort scores ranging on a scale from 0-100 with smaller values indicating better performance and hence less effort.

4.5.1.10 Gothenburg Profile

Figure 4.11 (a-d) shows results of NH and HI listeners on the Gothenburg Profile subscales: speech perception, spatial hearing, social interactions and reaction. On all four subscales there is very little spread in the normal-hearing data, and much more spread (and higher scores, so more problems) in the hearing-impaired data. Scores are mean values of answers on five questions on each topic.

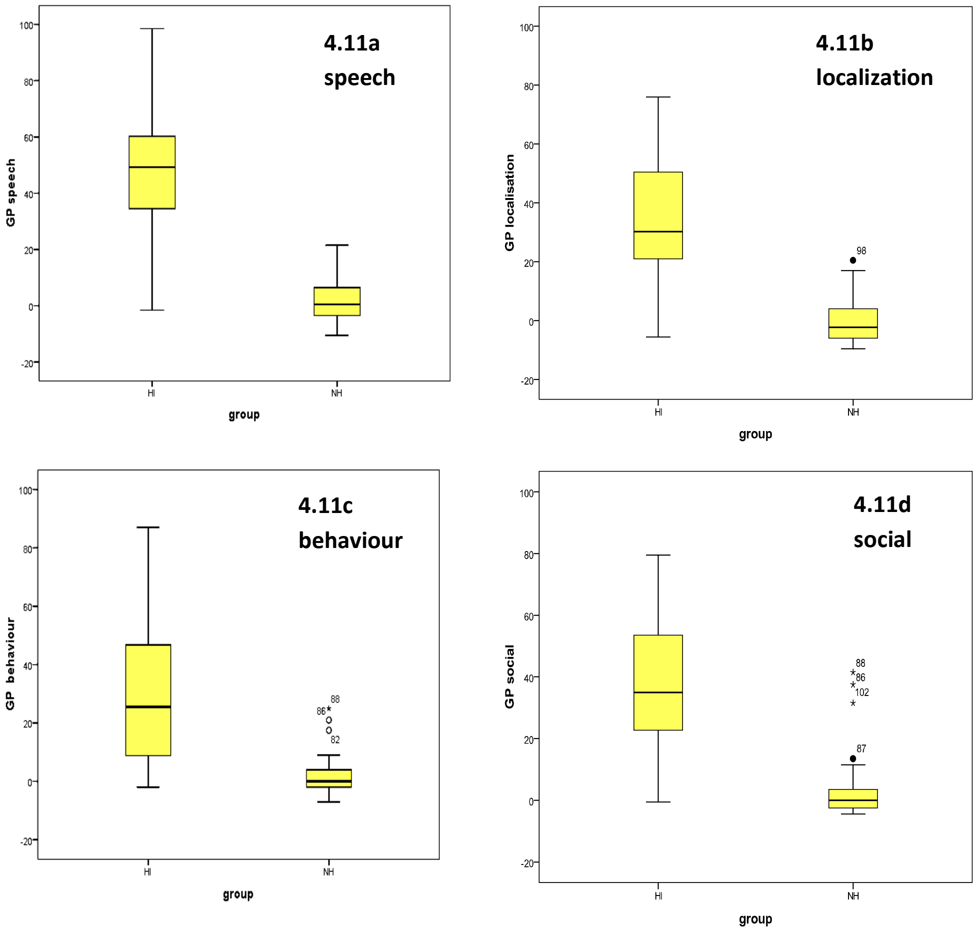


Figure 4.11: Gothenburg Profile results. The panels present scores (more negative scores refer to better hearing/less problems) of the four subscales of the questionnaire: speech perception (a), localization (b), behaviour (c) and social interactions (d).

4.5.1.11 Summary of findings from descriptive analysis

- In general, the hearing impaired group showed worse performance than the normal group as expected
- They also showed more variation than the normal group
- The scores revealed distinct absolute differences between the two groups indicating the tests are sensitive for a wide range of hearing capabilities

4.5.2 General statistics

The measures of the tests were analysed for certain basic properties like test-retest reliability, differences between left and right ears, differences between better and poorer ears.

4.5.2.1 Test-retest reliability

The test-retest reliability for the speech tests assures that for a given individual, the scores will be same at any point when measured for the same criteria, but at different point in time. Commonly, within subject SD is used to describe test-retest reliability. Further it helps ensure that correlations are not affected substantially by measurement error and hence are valid (Festen and Plomp 1983).

In addition to within subject SD, the test-retest reliability of the tests was investigated by calculating the intra-class correlation coefficient (ICC) (Shrout and Fleiss, 1979) for the total group, for HI listeners and for NH listeners. The ICC expresses the variation between subjects as a proportion of the total variation. ICC=1 when there is no variation within the subjects and all variation is attributed to differences between subjects (perfect test-retest reliability) and ICC=0 when all the variation is attributed to differences within subjects (zero variation between subjects relative to the test-retest reliability). ICCs of all outcome measures (except the audiogram) for the total group (ICC total), for HI listeners (ICCHI) and for NH listeners (ICCNH) are shown in table 4.4.

Table 4.4: Intraclass correlation coefficients for the total group (ICCtotal), hearing-impaired listeners (ICC-HI) and normal-hearing listeners (ICC-NH) and within-subject standard deviations (SDw).

		ICCtotal	ICCHI	ICCNH	SDw
ACALOS(dB)	MCL500	0.82	0.84	0.54	5.49
	MCL-3000	0.93	0.92	0.77	3.99
	MCLbb	0.86	0.86	0.73	5.14
(cu/dB)	SL500	0.82	0.83	0.53	0.07
	SL3000	0.86	0.84	0.69	0.09
	SLbb	0.75	0.77	0.58	0.09
GP (%)	speech	0.93	0.89	0.68	7.58
	loc	0.95	0.93	0.73	5.27
	social	0.91	0.89	0.68	7.60
	behaviour	0.95	0.95	0.61	5.69
ListEff (%)	Eff	0.90	0.86	0.79	5.00
	Eff-C-5	0.85	0.85	0.75	8.13
	Eff-C-min5	0.64	0.62	0.58	6.60
	Eff-F-5	0.89	0.86	0.76	7.35
	Eff-F-min5	0.87	0.77	0.78	7.46
MAA (degrees)	MAAbb	0.86	0.87	0.57	2.56
	MAAhp	0.87	0.88	0.78	3.20
	MAAlp	0.85	0.87	0.55	3.24
LexDec Test	LDT	0.86	0.78	0.83	0.10
FT (dB)	F500	0.73	0.77	0.47	2.04
	F3000	0.91	0.82	0.75	1.70
	T500	0.69	0.50	0.73	3.89
	T3000	0.78	0.41	0.74	3.68
BKB (dB)	SRTq	0.99	0.98	0.85	1.56
	SRTstat	0.92	0.87	0.86	1.33
	SRTfluc	0.96	0.89	0.90	1.64
Binaural(dB)	ILD	0.84	0.78	0.59	1.08
	BILD	0.66	0.68	0.52	1.14

> 0.9	excellent
0.8 - 0.9	good
0.7 - 0.8	
0.6 - 0.7	moderate
0.5 - 0.6	
< 0.5	poor

As can be seen in Table 4.4, the ICC for hearing-impaired listeners is mostly good (moderate to excellent) for nearly all variables. Only temporal resolution has poor test-retest reliability for hearing-impaired listeners. This is caused by a very small between-subject variation in temporal resolution for hearing-impaired listeners, especially at 3000 Hz. This suggests that perhaps detecting the temporal gap at 3000 Hz was too hard for HI listeners, causing a floor effect (very poor resolution, no release of masking).

For NH listeners on the other hand, most ICC values are poor to moderate, due to smaller between-subject variation. Once the test-retest reliability was established, for all the further analysis, only test values were included i.e. retest values were discarded. This is because in a clinical situation one would also have test results only, so this was considered a more realistic approach.

4.5.2.2 Differences between left and right ears

It is well established that there are asymmetries between right and left auditory function, both centrally and peripherally (Tadros et al, 2005). In young adults with normal hearing right ears tend to be more sensitive than the left to simple sounds (peripheral right-ear advantage) and for processing complex sounds such as speech (central right-ear advantage) (Tadros et al, 2005). Measuring differences between the ears gives a measure of any lateralization effects.

Differences between left and right ears of each subject in the total group were evaluated with the Wilcoxon signed-rank test (with left-right input pairs of all per-ear variables). Significant effects ($p<0.05$) were found for four variables: SRT in fluctuating noise (right ear better, mean difference 0.65 dB, $p=0.042$), ILD and BILD (left, so noise from left side better, mean differences 0.71 ($p=0.001$) and 0.20 dB ($p=0.019$) respectively) and T3000 (temporal resolution at 3000 Hz, right ear better, mean difference 0.86 dB, $p=0.043$). These effects remain significant even after excluding subjects with asymmetrical (based on PTA high > 10 dB) hearing losses.

Bonferroni's correction was then applied to the pairs above in order to decrease probability of type I errors. 16 per ear variables were compared using Wilcoxon tests. The significance value was changed to 0.003 (0.05/16). This meant that only the ILD difference remains significant (see figure 4.12). These effects might be related to the speech processing in the left hemisphere, or to other right-ear advantages discussed above. Divenyi *et al.,* (1997a) also found a lateral asymmetry favouring the right ear in their study, especially for all measures of speech understanding in presence of interference.

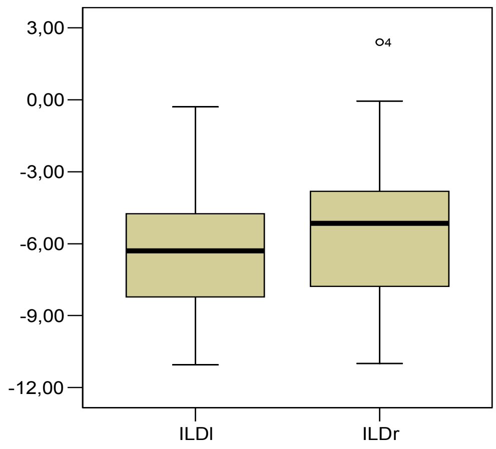


Figure 4.12: Differences between left and right ears for ILD. Vertical axes represent release of masking in dB; more negative values refer to better binaural processing. It can be seen that for ILD the overall result is better on the left side (noise from the left side).

4.5.2.3 Differences between better and poorer ears

This analysis gives a measure of asymmetry for all the subjects based on the better or poorer side of hearing. Thus for this analysis, ears of all subjects were divided in better and poorer ears (based on PTA 1, 2, 4) and differences between better and poorer ears of each subject were evaluated with Wilcoxon signed-rank tests. Of course audiogram data differed between better and poorer ears, but for the other variables, a significant effect was found only for the ILD test ($p=0.004$), where subjects overall performed better with the noise from the side of the poorer ear (see also Figure 4.13). Even in the group of subjects with asymmetrical hearing losses (defined as a difference in PTA of 1000, 2000 and 4000 Hz between ears of at least 10 dB) no other significant differences between better and poorer ears were found, but this is probably partly due to the fact that there are only 13 subjects with asymmetrical hearing loss included. However this does not remain significant after applying the Bonferroni's correction ($p<0.003$ to be significant).

4.5.2.4 Summary of findings from general analysis

- The test-retest reliability was found to range from 0.4-0.9 being moderate-good for most variables based on ICC
- Analysis of differences between better and poorer ears did not reveal any significant results
- Analysis of differences between left and right ears revealed significant results only for ILD. These effects might be related to speech processing in the left hemisphere, or to other right-ear advantages.

4.5.3 Test parts and measurements for groups I and II

The next analysis includes measurements for groups I and II. Group I mainly deals with predictions of speech recognition in noise using other auditory capabilities while group II concerns exploring the multidimensional aspect of hearing loss. The outcomes of the tests were divided into the following three parts for this purpose.

***Part 1*: Tests for each ear (or per ear tests) which are measured separately for left and right ears.**

1. Audiogram
2. ACALOS
3. FT
4. BKB (SRT in noise)

***Part 2*: Per subject tests which are measured for left and right ears combined.**

1. BKB (SRT in quiet)
2. MAA
3. ILD
4. BILD
5. Lexical decision test
6. Age

***Part 3*: Self-report/subjective tests**

1. GP
2. Effort Scaling (listening effort)

Data only for hearing impaired subjects were included since the auditory interactions of this group is of more interest here. Also only test session values for the better ear were included. Group I analysis deals with only per ear tests while group II deals with all of them together.

4.5.4 Group I analysis: different aspects explaining speech intelligibility (aims 1-4)

Before the analysis of the relations between per-ear measures, the data were checked to see if they deviated from normal distributions. This was done by performing Kolmogorov-Smirnov tests and by visual inspection. It was found that all variables except air-bone gap were distributed approximately normally. The ABGs (air-bone gap) was transformed using Blom transformation (Blom, 1958). This transforms the data to approximate normal distribution by ranking them and adjusting the distances between them. This makes them comparable with other variables. The two main statistical techniques used in the following section consist of linear regression and factor analysis which are discussed briefly in the following.

a)Multiple Linear Regression

In this context, multiple linear regression utilises a single dependent (speech intelligibility) and multiple independent variables or predictor variables. In the following section each of the regression analyses is shown by a two small tables. One (on the left side) includes the regression model summary and the best predicting model for the dependent variable, as well as the proportion of explained variance (R^2). For realistic predictions of the variation the adjusted R^2 values are taken into account and are hence highlighted in all the tables. On the right side, two values namely B (unstandardised regression coefficient) and β (standardised regression coefficient) along with the constant of the equation are given.

b) Factor Analysis:

Factor analysis is usually used as a data reduction technique so that any concept (in this case, the parameters describing the ear or auditory system) can be defined in terms of few clusters of variables (factors) from a huge corpus of variables. It also helps to observe which variables group together and which stand independently. Thus the factor loadings in each defined factor give information about the relative importance of the different variables within that factor while the grouping of different variables in the corresponding factors do so between the different factors.

Aims 1, 2 and 4 are discussed in the first section followed by 3. These aims involve prediction of SRT scores based on the set of the per ear measures. The data were subjected to linear regression. Dependent variables of the analyses were the SRTs in stationary and fluctuating noise; all other per-ear outcome measures (see above) were independent variables, except PTA high and PTA low, which are of course derived from audiogram thresholds. Variables were selected stepwise to enter or be removed from the model (enter: prob $F < 0.05$, remove: prob $F > 0.10$), using the SPSS stepwise procedure.

Results of the regression analyses are shown below. The format of presenting these findings is identical for all. For each dependent variable, the model summaries are shown with its explained variance (R^2). Also, the independent variables (predictors) that were significantly associated with the dependents and their standardized coefficients (β) and unstandardised coefficients (B) with constant for the best predicting model are shown. The tables with the heading ‘model summary’ show successive models in the stepwise procedure and the footnotes show the variables in each model. The right side tables show the B (unstandardised) and β (standardised) coefficients of the equation for the best predicting model. For all the following analysis data for HI only is included.

4.5.4.1 Findings of regression analysis (aims 1, 2, 4)

a) Stationary noise

Table 4.5: Results of the Regression analysis for prediction of SRT in stationary noise.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.562(a)	.315	**.305**	2.19710
2	.610(b)	.372	**.352**	2.12167
3	.649(c)	.422	**.394**	2.05173

a. Predictors: (Constant), AC3000
b. Predictors: (Constant), AC3000, F500
c. Predictors: (Constant), AC3000, F500, slope
d. Dependent Variable: SRTstat

		B	Beta
3	(Constant)	1.386	
	AC3000	.083	.438
	F-500	.203	.288
	slope	.033	.242

As can be seen, SRT in stationary noise is predicted by the audiogram threshold at 3000 Hz, frequency resolution at 500 Hz, and audiogram slope. This model explains nearly 40% of all variance based on the adjusted R square value (0.394). From the model summary, the additional explained variances of each predictor can be estimated. Hearing threshold at 3000 Hz explains the greatest variation accounting for over 30% while the other two variables F-500 (5%) and audiogram slope (4%) marginally account for the remaining additional 9-10% of the variation. Thus the audiogram measure was the best predictor for SRT score in stationary noise.

b) Fluctuating noise

Table 4.6: Results of the Regression analysis for prediction of SRT in fluctuating noise.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.608(a)	.369	**.360**	3.81975
2	.719(b)	.517	**.502**	3.36794
3	.749(c)	.561	**.540**	3.23807
4	.767(d)	.589	**.562**	3.15809
5	.786(e)	.618	**.587**	3.06628

a. Predictors: (Constant), AC3000
b. Predictors: (Constant), AC3000, F500
c. Predictors: (Constant), AC3000, F500, age
d. Predictors: (Constant), AC3000, F-00, age, T-000
e. Predictors: (Constant), AC3000, F500, age, T3000, SLbb
f. Dependent Variable: SRT-fluct

		B	Beta
5	(Constant)	1.987	
	AC3000	.126	.376
	F-500	.511	.403
	age	.076	.233
	T-3000	.260	.191
	SLbb	5.000	.177

As can be seen, SRT in fluctuating noise is predicted by the audiogram threshold at 3000 Hz, frequency resolution at 500 Hz, age and temporal resolution at 3000 Hz. This model explains 59% of the total variance. In terms of individual estimation, hearing threshold at 3000 Hz accounted for 36% of the variation, frequency resolution at 500 Hz a further 14% while the remaining three together accounted for almost 9% (age 4%, T-3000 2%, MCL slope BB 3%). Again the audiogram measure was the largest predictor for SRT score in fluctuating noise followed by frequency resolution at 500 Hz.

The predictions above are just estimates based on one sample of subjects. It may be possible that different estimates are obtained with other samples. Therefore it is important to not consider these findings to be particularly specific or generalisable. However, the main conclusion that high frequency threshold is important can be accepted as true since it has been observed by numerous studies before (reference table 2.2). Also frequency resolution at low frequency was the second important predictor especially for SRT in fluctuating noise.

4.5.4.2 Discussion: Aim 1

Aim 1 was to resolve the discrepancy surrounding the relation between hearing threshold level and speech recognition in hearing impaired people and hence investigate if audibility alone can explain variation in speech recognition scores.

From the above it can be seen, that among all the predictor variables, hearing threshold at 3000 Hz accounted for the greatest variation in both the kinds of noise. In this sense, the general hypothesis that hearing threshold level explains variation in speech intelligibility scores is accepted and threshold at 3000 Hz was the highest predictor. However, not all the explained variation could be attributed to audiometric variables. Significant further variation was explained by frequency resolution at 500 Hz, especially for SRT in fluctuating noise

The group A studies (Humes *et al.,* 1994; Jerger *et al.,* 1991, Divenyi *et al.,* 1997a) reveal maximum prediction of speech recognition is based on hearing threshold. In this sense, the findings from the present studies also reveal the same. However, the variation of speech recognition scores predicted by hearing threshold in these studies above range from 70-85%, which is much more than that is revealed in the present study (30-36%). Thus the results from the present study reveal only partial prediction, resembling more the findings from group B (ii) studies (Festen and Plomp, 1983; George *et al.,* 2006; van Rooij & Plomp 1990). In the study by Festen and Plomp (1983), the two main factor loadings in their PCA (Principal Component Analysis) on fifteen tests included hearing threshold and frequency resolution as seen in the present study. They explained 48% and 17% of variation respectively. Similarly George *et al.* (2006), implied that audibility is only partially responsible for the reduced masking release in speech modulation. Van Rooij and Plomp (1990) revealed two components responsible for deterioration of speech recognition scores; high frequency hearing loss accounting for 42% of variation and cognition accounting for 19%.

Thus the three studies above showed presence of other factors that could also be responsible for the SRT score variation. Two of these found were suprathreshold factors responsible for the remaining variation. These are investigated further below.

4.5.4.3 Discussion: Aim 2

Aim 2 was to observe if audibility can only in part explain the differences among speech recognition performance in hearing impaired subjects, what are the other factors (suprathreshold) responsible.

From the above it can be seen that two suprathreshold abilities are found to be predictors: frequency resolution at 500 Hz explaining 4-14% of variation, while temporal resolution is implicated to a much lesser extent. However the extent of variation for both the types of noise varies slightly. Of all the predictors for stationary noise, AC threshold at 3000 Hz accounted for the greatest extent of 30% while the additional contribution of others was relatively small (9-10%). For fluctuating noise, threshold at 3000 Hz accounted for 36% along with F-500 accounting for 14%.

These remaining variables other than threshold could be associated with the inter-individual variation in the HI subjects which is so widely seen; hence their importance should not be underestimated. Though they may have not explained a great percentage of variance once AC threshold at 3000 Hz had been included, they could well be playing an important role in deciding the SRT scores. It should further be remembered that the tests were carried out at suprathreshold level (MCL) rather than threshold level. Hence it can be said that audiogram related measures may not be actually responsible for variation in SRT scores, but they are just good predictors of the way in which suprathreshold capabilities influence speech recognition in noise. Also their role may vary depending upon the type of noise with more variation in fluctuating noise since F-500 accounted for 14% variation in fluctuating noise, but only 5% or so in stationary noise.

It is important to discuss the relative importance of frequency and temporal resolution, based on above results. Frequency resolution at 500 Hz was found to be a more important suprathreshold ability in this study since it explained 14% of variation for SRT in fluctuating noise. Festen and Plomp (1983) also found frequency resolution to be closely allied with SRT in stationary noise, explaining about 17.6 % of variation but their measure of frequency resolution was at 1000 Hz; they did not measure at any other frequencies. Temporal resolution on the other hand explained only 2%, hence is less important. This is contradictory to George *et al.,* (2006) who found temporal resolution to be more important. In fact, according to George *et al.,* (2006) frequency resolution does not qualify to be a suprathreshold deficit since it was found to be level dependent and deteriorates with presentation level even in NH subjects.

In summary, after hearing threshold, frequency resolution was found to be the main suprathreshold ability associated with variation in speech recognition scores. And as mentioned above, poor frequency resolution associated with speech intelligibility was also found in studies by Festen and Plomp (1983), Phillips *et al.,* (2000). Several other studies have also found a relationship between frequency resolution and speech recognition abilities (Dreschler and Plomp, 1985; Patterson *et al.,* 1982; Weber & Milroy, 1982). However, the present study also reveals that frequency resolution at 500 Hz is the second best predictor, after AC threshold at 3000 Hz. Such low frequency measure of frequency resolution associated with impaired speech perception is a new finding. This is not revealed by other studies. Most of those studies measured frequency resolution at high frequencies (Phillips *et al.,* 2000: 2000 Hz; Festen and Plomp, 1983, Dreschler and Plomp, 1985:1000 Hz). Partly this was because two of the studies (Phillips *et al.,* 2000; Festen and Plomp, 1983) did not include a low frequency measure. But Dreschler and Plomp (1985) included 500 Hz along with two other frequencies including 1000Hz (mid) and 2000 Hz (high) and did not find the lower frequency to be a significant predictor. The reason why spectral resolution at low frequency is important is unclear, but may relate to distinguishing vowel contrasts or semi-vowels that have cues at low frequencies. Frequency resolution at 3000 Hz may be important, but perhaps variation with frequency resolution at 3000 Hz is already accounted for by AC at 3000 Hz.

The above results can be explained both physiologically as well as based on Plomp's SRT model (1978). Any hearing loss leads to threshold elevation. However, when this loss is cochlear, the cause of this threshold elevation is loss in the sharp tuning of the basilar membrane. Along with this, higher absolute thresholds are associated with broader filters (Moore, 1995). This broadening of the filters increases the susceptibility of the nerve fibres to greater masking by noise than normal which ultimately disrupts the frequency selectivity of the signal. This explains why speech recognition in noise is affected by both threshold elevation and frequency resolution.

Similarly, Plomp's model (1978) described hearing loss for speech to be composed of two components: Attenuation (A) characterized by a reduction of the levels of both speech signal and noise, and Distortion (D), comparable with a decrease in speech-to-noise ratio. However, when both speech and noise are presented at suprathreshold level, only the D component becomes important. The results obtained in this study attempt to explain which actual measures can be attributed for this D component. Thus as shown in the present study, for a group of subjects with wide range of hearing loss when tested for various auditory capabilities the loss of the D component is attributed to frequency resolution.

4.5.4.3.1 Additional analysis

In order to support the probability of augmented role of frequency and temporal resolution, two additional analyses were carried out:

1) Partial correlations with control for audiometry/threshold measures (control variable was PTA1,2,4)

This analysis focussed on partialling out the interaction of audiogram or threshold related parameters with others especially frequency and temporal resolution shown in the table below. For this, the data was subjected to partial correlations . Any correlations that remain significant after this reveal the influence of measures (other than threshold measures) that influence the speech recognition in both types of noise. Thus it serves to support, establish and confirm their exclusive influence over speech recognition in noise i.e., without taking into account the influence of threshold measures which is already established as the primary predictor.

2) Linear Regression to predict the SRT scores using only frequency and temporal resolution measures

1] The following two tables show correlations before (table 4.7a) and after (table 4.7b) partialling out threshold measures

Table 4.7a: Correlations between per-ear measurements for HI. Each cell displays Pearson's correlation coefficient with its significance (p) in bottom row. Significant correlations are marked green (p<0.01) and yellow (p<0.05) in top row. Suffix b for better ear values.

		Audiogram						Acalos						FT				SRT	
		AC500b	AC3000b	PTAlb	PTAhb	slopeb	nABG	MCL500b	MCL3000b	MCLbbb	SL500b	SL3000b	SLbbb	F500b	T500b	F3000b	T3000b	SRTstatb	SRTfluctb
Audiogram	AC500b	1.000																	
	AC3000b	0.436 0.000	1.000																
	PTAlb	0.904 0.000	0.607 0.000	1.000															
	PTAhb	0.704 0.000	0.821 0.000	0.903 0.000	1.000														
	slopeb	-0.599 0.000	0.342 0.003	-0.406 0.000	0.025 0.832	1.000													
	nABG	0.225 0.057	-0.078 0.514	0.096 0.421	-0.017 0.887	-0.263 0.026	1.000												
Acalos	MCL500b	0.608 0.000	0.184 0.122	0.553 0.000	0.419 0.000	-0.418 0.000	0.245 0.038	1.000											
	MCL3000b	0.478 0.000	0.628 0.000	0.555 0.000	0.638 0.000	0.057 0.636	0.105 0.382	0.592 0.000	1.000										
	MCLbbb	0.608 0.000	0.382 0.001	0.663 0.000	0.614 0.000	-0.249 0.036	0.098 0.415	0.649 0.000	0.561 0.000	1.000									
	SL500b	0.501 0.000	0.156 0.190	0.523 0.000	0.401 0.000	-0.354 0.002	-0.148 0.216	-0.032 0.790	-0.034 0.779	0.294 0.013	1.000								
	SL3000b	0.009 0.942	0.421 0.000	0.194 0.103	0.299 0.011	0.198 0.096	-0.144 0.229	-0.244 0.039	0.009 0.938	-0.116 0.333	0.225 0.058	1.000							
	SLbbb	0.395 0.001	0.147 0.218	0.353 0.002	0.266 0.024	-0.237 0.045	-0.064 0.592	0.101 0.401	0.020 0.865	-0.074 0.537	0.510 0.000	0.206 0.082	1.000						
FT	F500b	0.301 0.012	0.209 0.084	0.370 0.002	0.347 0.003	-0.107 0.382	0.041 0.740	0.119 0.330	0.260 0.031	0.365 0.002	0.338 0.004	0.056 0.645	0.069 0.573	1.000					
	T500b	0.331 0.006	0.148 0.224	0.236 0.051	0.149 0.221	-0.216 0.075	0.087 0.477	0.133 0.277	0.195 0.108	0.256 0.035	0.228 0.059	-0.103 0.401	0.172 0.157	0.371 0.002	1.000				
	F3000b	-0.188 0.121	0.426 0.000	-0.053 0.665	0.164 0.178	0.477 0.000	-0.345 0.004	-0.127 0.297	0.313 0.009	0.041 0.742	-0.117 0.338	0.230 0.058	-0.124 0.311	0.104 0.395	-0.014 0.909	1.000			
	T3000b	-0.122 0.319	0.251 0.038	-0.021 0.863	0.109 0.372	0.290 0.016	-0.242 0.045	-0.292 0.015	-0.078 0.526	-0.099 0.421	-0.016 0.894	0.210 0.083	-0.065 0.598	0.070 0.568	-0.046 0.708	0.432 0.000	1.000		
SRT	SRTstatb	0.231 0.051	0.588 0.000	0.324 0.006	0.515 0.000	0.349 0.003	0.070 0.558	0.020 0.870	0.372 0.001	0.257 0.030	0.034 0.775	0.224 0.058	0.148 0.214	0.373 0.002	0.250 0.038	0.262 0.030	0.155 0.205	1.000	
	SRTfluctb	0.372 0.001	0.608 0.000	0.490 0.000	0.623 0.000	0.194 0.102	-0.037 0.757	0.010 0.936	0.332 0.004	0.285 0.016	0.239 0.043	0.297 0.011	0.229 0.053	0.503 0.000	0.238 0.049	0.262 0.030	0.332 0.005	0.824 0.000	1.000

Table 4.7b: Partial correlations between per-ear measurements for HI (control variable: PTA (1, 2, 4)). Significant correlations are marked green (p<0.01) and yellow (p<0.05). Suffix b for better ear values.

		Acalos						FT				SRT	
		MCL500b	MCL3000b	MCLbbb	SL500b	SL3000b	SLbbb	F500b	T500b	F3000b	T3000b	SRTstatb	SRTfluctb
Acalos	MCL500b	1.000 .											
	MCL3000b	0.464 0.000	1.000 .										
	MCLbbb	0.546 0.000	0.278 0.020	1.000 .									
	SL500b	-0.240 0.044	-0.411 0.000	0.066 0.588	1.000 .								
	SL3000b	-0.426 0.000	-0.247 0.038	-0.399 0.001	0.120 0.320	1.000 .							
	SLbbb	-0.012 0.918	-0.201 0.093	-0.313 0.008	0.457 0.000	0.138 0.252	1.000 .						
FT	F500b	-0.031 0.802	0.053 0.670	0.205 0.097	0.231 0.058	-0.053 0.667	-0.026 0.834	1.000 .					
	T500b	0.078 0.525	0.131 0.285	0.211 0.087	0.186 0.129	-0.156 0.203	0.139 0.258	0.344 0.004	1.000 .				
	F3000b	-0.219 0.073	0.274 0.024	-0.077 0.534	-0.202 0.098	0.192 0.117	-0.176 0.151	0.051 0.681	-0.039 0.749	1.000 .			
	T3000b	-0.374 0.002	-0.192 0.116	-0.212 0.085	-0.066 0.593	0.187 0.127	-0.098 0.428	0.034 0.781	-0.063 0.609	0.422 0.000	1.000 .		
SRT	SRTstatb	-0.251 0.034	0.066 0.583	-0.087 0.473	-0.219 0.066	0.086 0.476	0.014 0.908	0.241 0.048	0.205 0.094	0.210 0.086	0.116 0.348	1.000 .	
	SRTfluctb	-0.354 0.002	-0.110 0.362	-0.159 0.188	-0.015 0.902	0.148 0.218	0.084 0.486	0.391 0.001	0.188 0.125	0.207 0.090	0.340 0.005	0.751 0.000	1.000 .

As can be seen, the correspondence between various measures reduced materially, which undoubtedly reveals the influence of the threshold measures. Even for SRT measures, correlations reduced from the order of 0.3-0.5 to 0.2-0.3. However, the fact that F-500 and T-3000 still remained significant is of more importance here. This further supports the hypothesis 2 that along with hearing threshold, suprathreshold abilities also play a role in SRT scores; in this case the ability is frequency resolution.

2] SRT scores using only frequency and temporal resolution measures as predicted by linear regression

Summary of the regression models with their explained variance (R^2) for both the types of noise is given below.

Table 4.8: Results of the regression analysis for prediction of SRT in stationary noise (left) and fluctuating noise (right) with only frequency and temporal resolution measures as independent variables.

Model	R	R Square	Adjusted R Square
1	.324[a]	.105	**.091**
2	.435[b]	.189	**.164**

a. Predictors: (Constant), F500

b. Predictors: (Constant), F500, F3000

c. Dependent Variable: SRT stat

Model	R	R Square	Adjusted R Square
1	.503[a]	.253	**.242**
2	.585[b]	.342	**.322**

a. Predictors: (Constant), F500

b. Predictors: (Constant), F500, T3000

c. Dependent Variable: SRT-fluc

4.5.4.3.2

a Findings from additional analysis

As can be seen, for the stationary noise the explained variance increased from 5% (table 4.5) to 9% (table 4.8) for frequency resolution at 500 Hz with an additional 7% from frequency resolution at 3000 Hz. Similarly for the fluctuating noise the explained variance increased from 14% (table 4.6) to 24% (table 4.8) for frequency resolution at 500 Hz and a further 1% or so for temporal resolution at 3000 Hz.

b Discussion regarding the additional analysis and other predictor variables

From the additional analysis above, it becomes evident that though hearing threshold measures were the best predictors, frequency and temporal resolution also play an important role in influencing speech recognition scores in noise. This can be understood since it is the auditory resolution rather than auditory sensitivity that is expected to be more important at the presentation levels used in the study. However, threshold does seem to have some importance, perhaps as an indirect predictor that is not reflected in the resolution measures used here.

c. Other factors: Age

Age was found to be a minor predictor for SRT in fluctuating noise explaining about 4% of variation in combination with other factors (see table 4.6). Other studies that have found age as a factor explaining speech recognition include Divenyi *et al.*, (1997a), George *et al.,* (2007), Jerger *et al.,* (1993), Dubno *et al.,* (1984). Most of these studies allocated it to a factor in conjunction with other variables such as temporal resolution (George *et al.,* 2007) or for a specific test like SSI (Jerger *et al.,* 1993).

4.5.4.4 Discussion (Aim 4):

Aim 4 was to investigate if there are different factors that underlie speech recognition for the two different types of noise.

From the above it can be seen, that threshold at 3000 Hz and frequency resolution at 500 Hz are common predictors for both the types of noise while additional predictors like age and T-3000 are present for fluctuating noise. The first two were also found to be responsible for masking release in George *et al.,* (2006). The latter two can be thought to be the variables differentiating the performance in the two types of noise. Another inference could be that at any given time, hearing sensitivity and frequency resolution are important for speech recognition; however in fluctuating backgrounds along with these two, the individual's temporal resolution and age also become significant. This is understandable because it is known that hearing impaired listeners show less masking release in fluctuating noise unlike the normal hearing listeners who show a significant improvement in their SNR when compared to stationary noise. The hearing impaired group are unable to take advantage of the information present in the gaps of the fluctuating backgrounds.

Also SRT in fluctuating noise is perhaps a better measure of performance in the hearing impaired since it explains greater variance in terms of percentage (58%) as compared o stationary noise (39%). This could further imply that that fluctuating backgrounds reveal more information about the processes in the auditory system than stationary noise and hence should be included amongst measures to test the hearing impaired people which is not done in routine clinical practice. The importance of fluctuating noise has been discussed in Chapter 1(1.4.1 III). Also not many studies except George *etal.*

(2006) have studied speech recognition performance in fluctuating noise. Further, it can also be said that the prediction of SRT scores depends to some extent on type of noise used as there were differences in the few parameters that influenced them individually.

Finally, regression analysis for the two types of noise was repeated after excluding the four cases with mixed hearing loss to observe if the results varied for sensory neural hearing loss. However, this did not alter the results and SRT stat was still predicted by hearing threshold and SRT fluc a by combination of hearing threshold and frequency resolution.

Hence the hypothesis that for the two types of noise (stationary and fluctuating), different factors are responsible for speech recognition scores, is accepted for now since additional factors like temporal resolution and age were found to predict fluctuating noise. But their contribution was quite small and hence their repeatability has to be confirmed which will be done in chapter 5.

4.5.4.5 (Aim 3): To investigate factors predicting speech recognition in noise with differing magnitude of hearing loss.

a) Analysis for aim three

To investigate the influence of hearing loss magnitude, data was split up into three groups with the first group I consisting of PTA (1000, 2000, 4000 Hz) <= 40 dB (mild HL), the second one (group II) with PTA between 41 and 55 dB (moderate HL) and the third (group III) with PTA above 55 dB (moderate-severe). This sort of analysis helped characterize different ranges of hearing loss and the variables affected in each. The analysis would also reveal if there are any factors or trends specific to any group. Group I had 31 cases, group II had 26 and group III had 15.

b) Findings of aim three:

In the table below the findings for the three groups are summarised with their variances (R^2) for the two types of noises.

Table 4.9: Results of the regression analysis for prediction of SRT in stationary noise and fluctuating noise in the three groups of hearing loss. Variables included in the final regression model are shown with the percentage of variance explained.

Group of HL/Type of noise	Stationary noise	Fluctuating noise
Mild	A'GM slope (28%)	Age (28%)
	T500 (13%)	T3000 (12%)
Moderate	MCL slopes (40%)	F500 (32%)
	F500 (10%)	T3000 (23%)
		MCL slope (14%)
		Age (11%)
Severe	-----------	F500 (36%)

c) Discussion: Aim 3

It should be noted that the analysis is more exploratory and hence the included subjects were merely divided based on their hearing loss and hence there was no uniformity in terms of the number of cases, their age, audiogram shape etc. This limits the precision and robustness of the findings, but it would nevertheless be interesting to observe if and how the predictive variables change across groups of hearing loss when studied as a function of types of noise.

As can be seen, different predictive variables are present in different groups of hearing loss. Several other interesting observations can be made. Audiogram measures are present only in the mild group for stationary noise. Presumably, threshold at 3000 Hz explains predominantly variation between groups rather than within groups including variation between NH and HI. On the same lines, auditory resolution measures explained greater variance in fluctuating noise in general. Also, the variance explained by auditory resolution (whether frequency or temporal) increases in fluctuating noise and decreases in stationary noise as the degree of hearing loss increases. More on this aim is discussed later in the Chapter V.

4.6 Factor analysis with per ear variable

Factor analysis was performed to reduce data within the set of ear-dependent outcome measures for hearing-impaired listeners. The result will be a few parameters characterizing the better ear. All per-ear variables were included in the analysis, except PTA low, PTA high and SRT stat, because their correlations with other variables were too high, making the determinant of the correlation matrix too small. Factors are reported if their Eigenvalues were above 1. For extraction the principal component method was used, and a Varimax rotation was applied to ease interpretation.

4.6.1 Findings from Factor analysis

Results of the analysis are shown in the rotated component matrix, in Table 4.10. This table shows factor loadings for all variables except values below 0.4 are suppressed; values above 0.7 are printed bold. The determinant of the correlation is 0.00003, KMO (sampling adequacy) = 0.605 and Bartlett's test significance < 0.001 which means that the data were suitable for factor analysis. The total variance explained by these four factors was 67 %.

Table 4.10: Results of the Factor analysis: Rotated component matrix for per ear measures explained by four distinct sets of factor loadings.

	Component			
	1	2	3	4
slope	**0.779**			
F3000	**0.746**			
T3000	0.612			
MCL3000		**0.884**		
MCL500	-0.437	**0.789**		
MCLBB		**0.729**		0.414
SLbb			**0.829**	
SL500			**0.774**	
SL3000	0.542		0.492	
F500				**0.811**
T500				**0.701**
AC3000	0.609	0.624		
ABG	-0.460			
SRTfluct	0.525			0.502
AC500		0.636	0.535	
Explained				
variance:	20.3	19.8	14	12.8
Interpretation:	high-freq processing	audibility	recruitment	low-freq processing

4.6.2 Results and discussion of factor analysis of per ear variables

The first factor contains mainly variables at 3000 Hz (F and T resolution, threshold, audiogram slope), and is probably related to high frequency processing. The second factor contains all MCL values from ACALOS, and audiogram thresholds. It can be identified as being an 'audibility' factor and is a result of the experimental methodology. In other words, though MCL measures represent suprathreshold domain here they are grouped with threshold measures. This is because MCL was used to determine the presentation level that was adapted to the subject's individual hearing level, which led to the intrinsic correlation between the two measures. Factor 3 consists of slopes from ACALOS. So this factor appears to be associated with recruitment. In the 4th factor, finally, both spectral and temporal resolutions at 500 Hz have high loadings, so this factor is related to low-frequency resolution. It can be noted that the audiogram is represented in most of four factors, but never has high factor loadings.

Regarding the FT test, factors cluster frequency (factor 2 for 3000 Hz and factor 4 for 500 Hz) rather than properties (spectral and temporal resolution). Moreover, the different ACALOS outcome measures (MCL and slope) represent two different factors: audibility, and loudness recruitment. Finally, the SRT in fluctuating noise is present in the same factors as spectral and temporal resolution. This grouping helps to support findings from aim two which endorses the importance of suprathreshold factors in speech recognition in noise. Various studies have revealed similar factors; high frequency processing represented by auditory sensitivity and auditory resolution was also reported in study by Lutman (1987). Similarly, an audibility factor was also shown by Humes *et al.*(1994) while high frequency associations were present in the study by Divenyi *et al.* (1997 c).

4.7 Summary of findings from group I analysis

- For speech recognition in stationary noise, hearing threshold at 3000 Hz was the largest predictor explaining the most variation accounting for over 30% while the other two variables F500 (5%) and audiogram slope (4%) accounted for the rest of 9-10% of the variation. The overall model explained 40% of variation in the speech test scores.
- For speech recognition in fluctuating noise, again, hearing threshold at 3000 Hz was the largest predictor explaining the maximum variation accounting for over 36% followed by F500 explaining 14% of the variation while the remaining three together accounted for almost 8% (age 4%,T-3000 2%,MCL slope BB 2%). The overall model explains 59% of the total variance.
- Thus in general threshold measurements proved to be associated with the greatest variation in speech recognition scores in noise followed by F500 (especially for fluctuating noise).
- Additional analysis assessing exclusive roles of suprathreshold measurements like auditory resolution (both frequency and temporal) included examining correlations partialled for threshold measures and linear regression assessing predictions using resolution measures only. The former revealed that after controlling for audiogram measures, correlations between SRT in noise and F500, T3000 reduced (from 03-0.5 to 0.2-0.3) but remained significant. The latter revealed 4-10% overall increase in variation of SRT explained by frequency and temporal resolution measures which further supports the importance of frequency and temporal resolution in influencing the speech perception in noise.
- Low frequency auditory resolution as one of the important predictors for speech recognition in noise (indicated by F500) was a novel finding compared to other studies (Phillips et al, 2000; Festen and Plomp, 1983; Dreschler and Plomp, 1985) which indicated higher frequencies.
- Temporal resolution at 3000 Hz and age helped differentiate the performance between the two types of noise while frequency resolution at 500 Hz and hearing threshold at 3000 Hz were common predictors

- In different groups of hearing loss, SRT in noise (generalising findings from both stationary and fluctuating noise) could be predicted by audiogram and auditory resolution measures in mild losses and only auditory resolution measures for moderate and higher losses. This indicates decreased influence of hearing sensitivity and increased influence of auditory resolution with increase in magnitude of hearing loss. In other words a shift from threshold related to suprathreshold processing and the fact that factors other than threshold come into play when the degree of loss is more. However, the finding may only be for the present group of subjects and cannot be reasonably generalised. This will be discussed more in chapter V.
- It was observed also that when the type of noise or the magnitude of hearing loss differed, the subsequent factors influencing speech recognition also differed to some extent. Thus predictive variables for speech recognition were governed by two factors: type of noise and magnitude of hearing loss.
- Factor analysis led to test measures grouping into four separate groups including high frequency processing, audibility, recruitment and low frequency processing. Together, they explained 67 % of variation.
- Also in factor analysis predictions thus tended to be based not just on test domain (audibility, recruitment) but frequency of stimuli (high /low) which indicates the importance of testing auditory performance at different frequencies. Also the grouping of frequency and temporal resolution measures with SRT measures highlights their association.

4.8 Group II analysis: Exploring multidimensionality (aim 5)

Group I analysis focussed on per ear variables including threshold, loudness, speech recognition and frequency and temporal resolution. Now, group II analysis will include the relations of these per ear variables with the per subject variables and subjective variables as well as relations within each of the latter groups individually.

This analysis includes all the test measures including per ear, binaural, subjective and cognitive as opposed to group I which focussed on only per ear variables. This analysis aims to explore all the tests and investigate their interrelations. The division of this section is as follows:

- Per subject variables: SRT in quiet, ILD, BILD, Lexical Decision Test, MAA, age
- Subjective variables: GP and listening effort (and their relations with per ear and per subject variables)

4.8.1 Per subject measures

Per subject variables (SRT in quiet, ILD, BILD, Lexical decision, MAA, age) were tested for normality both by Kolmogorov-Smirnov tests and by visual inspection. All variables except the MAA variables were distributed approximately normally. The MAA variables were transformed using the Blom method.

4.8.1.1 Factor analysis

Factor analysis was performed to reduce data within the set of ‘binaural’ outcome-measures for hearing-impaired listeners. The result will be a few parameters characterizing the listener.

Results of the analysis are shown in the rotated component matrix, see table 4.11. Values below 0.4 are suppressed and above 0.7 are printed bold. The determinant of the correlation is 0.058, KMO (sampling adequacy) = 0.687 and Bartlett’s test which means that the data was suitable for factor analysis.

Table 4.11: Results of the Factor analysis: Rotated component matrix for per subject measures explained by three distinct factor loadings.

	Component		
	1	2	3
age			**0.811**
LexDec			**-0.805**
SRTq		**0.715**	
ILD		**0.725**	
BILD		0.691	0.442
MAAb	**0.856**		
MAAlp	**0.862**		
MAAhp	**0.846**		
Explained variance:	29.79%	20.584	19.631
Interpretation	MAA	speech	age/cogn.

The interpretation of this factor analysis was fairly straightforward. The first factor clearly incorporates all three MAA parameters (binaural processing) explaining almost 30% of variance, the second factor contains all speech-reception related measures explaining 21% of variation and the third factor is determined by age and cognition explaining almost 20% of variation. Together, the three factors explained nearly 70% of variation. It can be seen from above that the localization and speech-spatial measures stood independently. Spatial separation of speech and localization of speech were expected to group together since both require use of binaural cues. However, the fact that they have not and ILD grouped with SRT in quiet perhaps indicates that ILD/BILD measures are more dependent on the absolute ability to recognise speech than the binaural advantage. Independent existence of spatial separation of speech and noise measures was also seen in the study by Divenyi *et al.* (1997 c). Age and cognition formed a non-auditory factor cluster. The negative sign indicates an inverse relation between them which means that as age increases lexical scores decrease.

Further, regression analysis examined the relationship between the measures of spatial hearing (ILD/BILD and MAA) in per subject variables and other measures: better-ear hearing threshold levels, spectral/temporal resolution, loudness tolerance; asymmetry of hearing threshold levels, spectral/temporal resolution, and loudness tolerance.

4.8.1.2 a Findings of Regression analysis (ILD/BILD)

As seen below, ILD can be predicted by hearing threshold at 3000 Hz, measures of asymmetry (taken as difference between left and right ears, indicated by suffix 'dif') for loudness level (BB), frequency resolution at 500 Hz, air conduction threshold at 3000 Hz, loudness levels slope (SL)(3000 Hz,BB). BILD can be predicted from MCL at 500 Hz.

Table 4.12: Results of the regression analysis for prediction of ILD with per ear measures as independent variables.

ILD

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.417(a)	.174	**.161**	1.91490
2	.568(b)	.323	**.302**	1.74695
3	.627(c)	.393	**.364**	1.66711
4	.673(d)	.453	**.417**	1.59615
5	.699(e)	.489	**.447**	1.55505
6	.727(f)	.529	**.481**	1.50601

		B	Beta
6	(Constant)	-.580	
	AC3000	.065	.444
	MCL-BBdif	.073	.305
	F-500dif	.135	.280
	AC3000dif	.033	.233
	SL3000dif	2.752	.234
	SLb	2.569	.208

a. Predictors: (Constant), AC3000
b. Predictors: (Constant), AC3000, MCL-BBdif
c. Predictors: (Constant), AC3000, MCL-BBdif, F500dif
d. Predictors: (Constant), AC3000, MCL-BBdif, F-00dif, AC3000dif
e. Predictors: (Constant), AC3000, MCL-BBdif, F500dif, AC3000dif, SL3000dif
f. Predictors: (Constant), AC3000, MCL-BBdif, F500dif, AC3000dif, SL3000dif, SLbb
g. Dependent Variable: ILD

Table 4.13: Results of the regression analysis for prediction of BILD with per ear measures as independent variables.

BILD

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.445(a)	.198	**.186**	1.65388

		B	Beta
1	(Constant)	-4.530	
	MCL-500	.074	.445

a. Predictors: (Constant), MCL500
b. Dependent Variable: BILD

4.8.1.2 b Discussion:

The above analyses for BILD demonstrate the importance of low-frequency hearing in the processing of spatial hearing cues which in turn implies the significance of interaural time differences while resolving the issues of binaural hearing. It is known that binaural release from masking and other forms of binaural processing that could potentially improve performance have their greatest effect at lower frequencies (Humes and Roberts, 1990). For ILD the hearing threshold and asymmetry measures predict the variation among all measures entered as independents. Further, the presence of positive coefficients reveals that binaural performance decreases with increasing ear asymmetry since more positive values on ILD scale indicates more impairment.

4.8.1.3a Findings of regression analysis (MAA)

Table 4.14: Results of the regression analysis for prediction of MAA (broadband condition) with per ear measures as independent variables.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.327(a)	.107	**.093**	.81213

b. Dependent Variable: MAAbb

		B	Beta
1	(Constant)	-1.675	
	MCL-3000	.025	.327

a. Predictors: (Constant), MCL3000

Table 4.15: Results of the regression analysis for prediction of MAA (low pass condition) with per ear measures as independent variables.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.362(a)	.131	**.116**	.83934
2	.442(b)	.196	**.166**	.81500

		B	Beta
2	(Constant)	-.453	
	AC500	.020	.368
	MCL-BBdif	.026	.254

a. Predictors: (Constant), AC500
b. Predictors: (Constant), AC500, MCLBBdif
c. Dependent Variable: MAAlp

4.8.1.3b Discussion (MAA)

As seen above only MAA in BB and low pass conditions could be predicted while the high pass condition could not be predicted. Also the extent of prediction was very limited especially for the low pass condition. Thus MAA measure showed limited association with other measures. This was also seen in factor analysis where they stood independently.

4.8.1.4 Self report/subjective measures

The normality of the distributions of the subjective (Listening effort and GP) measures was investigated by Kolmogorov-Smirnov tests and visual inspection. The Gothenburg-Profile variables were transformed using Blom transformation because of their skewed distributions. Distributions of the listening-effort results were approximately normal.

a) Gothenburg Profile

The regression analysis in this section, involved the Gothenburg Profile subscales as dependent variables. For the *Speech* subscale, SRT in quiet explained 38% of the variance. For the *Localization* subscale, SRT in quiet, ILD and age explained 41% of the variance. For the *Social/Relation* subscale, SRT in fluctuating noise, F-3000, MCL-3000 explained 23% of the variance. For the *Perform (behaviour)* subscale, SRT in quiet and SRT in stationary explained 20% of the variance. (See table 4.16-4.19)

Table 4.16: Results of the regression analysis for prediction of GP subscale: speech with per ear measures and per subject as independent variables.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.626(a)	.392	**.381**	.58931

		B	Beta
1	(Constant)	-.582	
	SRTq	.037	.626

a. Predictors: (Constant), SRTq
b. Dependent Variable: speech

Table 4.17: Results of the regression analysis for prediction of GP subscale: localization with per ear measures and per subject as independent variables.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.581(a)	.337	**.326**	.63392
2	.631(b)	.398	**.377**	.60957
3	.666(c)	.444	**.414**	.59088

		B	Beta
3	(Constant)	-.136	
	SRTq	.029	.481
	ILD	.101	.274
	age	-.011	-.215

a. Predictors: (Constant), SRTq
b. Predictors: (Constant), SRTq, ILD
c. Predictors: (Constant), SRTq, ILD, age
d. Dependent Variable: loc

Table 4.18: Results of the regression analysis for prediction of GP subscale: social with per ear measures and per subject as independent variables.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.384(a)	.148	**.133**	.72790
2	.454(b)	.206	**.178**	.70887
3	.519(c)	.270	**.231**	.68578

		B	Beta
3	(Constant)	-2.112	
	SRT-fluc	.062	.376
	F-3000	-.061	-.317
	MCL-3000	.020	.277

a. Predictors: (Constant), SRTfluc

b. Predictors: (Constant), SRT-fluc, F3000
c. Predictors: (Constant), SRT-fluc, F3000, MCL-000
d. Dependent Variable: social

Table 4.19: Results of the regression analysis for prediction of GP subscale: behaviour with per ear measures and per subject as independent variables.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.379(a)	.143	**.129**	.78546
2	.451(b)	.204	**.176**	.76388

		B	Beta
2	(Constant)	-.459	
	SRTq	.018	.271
	SRTstat	.065	.268

a. Predictors: (Constant), SRTq
b. Predictors: (Constant), SRTq, SRTstatb
c. Dependent Variable: behav

As can be seen, the speech perception and spatial hearing subscales are most predictable. This is understandable since perhaps both of them are directly dependent on the extent of hearing disability and hence are easier to self-assess. There could be a lot of individual variation in the other two (social and behaviour) depending upon an individual's perception of the disability as well as relevance of a number of other psycho-social factors including the individual's level of social activity, age etc. The localization subscale was partially predicted by ILD unlike in factor analysis where they stood apart (MAA and ILD). Partly this could be due to the correlation between SRTq and ILD. Further, overall, SRTs rather than hearing thresholds were found to be the most important predictors for all subscales, which may reflect the focus of the questionnaire items on speech tasks.

b) Listening Effort

The regression analysis in next section involved the Listening Effort subscales as dependent variables with per subject measures and per ear measures as independent variables.

Findings:

Table 4.20: Results of the regression analysis for prediction of Listening effort in stationary (continuous) noise at SNR +5, with per ear measures and per subject as independent variables.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.424(a)	.180	**.165**	18.87408
2	.514(b)	.264	**.238**	18.02857

a. Predictors: (Constant), SRTfluct
b. Predictors: (Constant), SRTfluct, T3000
c. Dependent Variable: Eff-C-5

		B	Beta
2	(Constant)	12.253	
	SRT-fluct	2.278	.526
	T-3000	-1.817	-.309

Table 4.21: Results of the regression analysis for prediction of Listening effort in stationary (continuous) noise at SNR -5, with per ear measures and per subject as independent variables.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.304(a)	.093	**.077**	7.61147

a. Predictors: (Constant), SRTstat
b. Dependent Variable: Eff-C-min5

		B	Beta
1	(Constant)	85.494	
	SRTstat	.694	.304

Table 4.22: Results of the regression analysis for prediction of Listening effort in fluctuating noise at SNR +5, with per ear measures and per subject as independent variables.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.457(a)	.209	**.195**	17.23800

a. Predictors: (Constant), SRTq
b. Dependent Variable: Eff-F-5

a. P

		B	Beta
1	(Constant)	27.60 7	
	SRTq	.688	.457

Table 4.23: Results of the regression analysis for prediction of Listening effort in fluctuating noise at SNR -5, with per ear measures and per subject as independent variables.

Model	R	R Square	Adjusted R Square	Std. Error the Estima
1	.503(a)	.253	**.240**	11.842
2	.560(b)	.313	**.289**	11.455

		B	Std. Error	Beta
2	(Constant)	53.364	5.405	
	SRTq	.381	.136	.358
	AC3000	.272	.122	.285

b. Predictors: (Constant), SRTq
c. Predictors: (Constant), SRTq, AC3000b
d. Dependent Variable: Eff-F-min5

Discussion

As can be seen, effort in fluctuating noise is more predictable than in stationary noise and like GP results, SRTs are the most important predictors. It is known that increased listening effort, and the subsequent stress, creates an increased working load for hearing impaired individuals in relation to normally-hearing ones, so the former tire quickly. So this effort is perhaps better predicted by fluctuating noise than stationary especially in hearing impaired people. However only a small percentage of variances were found to be explained.

4.9 Summary of findings from group II analysis

- Factor analysis for per subject measurements revealed localization and speech-spatial measures are independent of other measures.
- Age and cognition were grouped together and were inversely related.
- The importance of low-frequency hearing and measures of asymmetry along with hearing sensitivity in the processing of spatial hearing cues is demonstrated by binaural hearing measures.
- SRT measures were the best predictors for self-rated hearing disability.
- The four listening effort subscales conditions were predicted best by SRT in fluctuating noise, with overall limited predictions. The same was seen with MAA measures where limited predictions were observed.

4.10 Aim 5: To explore the multi-dimensionality of hearing loss.

Both group I and group II analysis explored the interrelations between the three groups of tests mainly through regression and factor analysis which are both based on correlations. The regression analysis including the different dependent variables usually revealed more than one predictor variable, indicating a significant amount of inter-dependence and any measure (SRT, ILD etc) was usually predicted by more than one other measure. This implies that the predictions are multi-dimensional and the correlations are distributed over the measures. Though hearing sensitivity dominates predictions for most of the measures it does not in any way singly qualify to characterize all the different processes in the auditory system. The relative importance of speech recognition, frequency resolution and subjective/ self-rated hearing disability measures is evident from above. Thus the hypothesis that hearing disability is multidimensional is accepted. Alternatively the hypothesis that measures of hearing disability are highly correlated and hence describe a single underlying dimension is rejected.

However, aspects such as how the specific measures combine in this multidimensionality or are the findings in the present study repeatable, still remain to be answered. This is covered in the next chapter. It also includes the findings using the SII model approach to complement those from regression analysis.

4.11 Need for study II

This section deals with the shortcomings of the first study which led to designing the second study. Execution of a second study was important for several reasons. The main findings of the present study which include threshold measurements at high frequency followed by frequency resolution at low frequency to be the two main factors predicting the speech recognition in noise, agree closely with studies of group B (ii), while they differ from two other groups. Further, though high frequency hearing threshold as an important predictor is agreed by many studies, frequency resolution (as opposed to temporal resolution, cognition or age found by various studies) and particularly at low frequency (as versus high frequency revealed by some studies) are less readily supported. All these findings need to be verified to establish whether or not they are specific to one set of subjects. Thus, a second study was essential to increase the reliability, to observe if the findings can be generalised to a larger group as well as confirm the findings from the present one along with comparing them to other studies.
Further, an extent of multidimensionality of hearing loss is established in the first study. A second one can help to confirm auditory or other domains that characterize this multidimensionality.

4.11.1 Limitations of the study I

Most tests were carried out with stimuli set at the most comfortable loudness (MCL) measured using ACALOS. Due to this a certain amount of structural correlation was built into the data, glimpses of which were apparent when MCL/loudness measures were found to be appearing as predictors in most of the other measures. This approach introduces two types of structural correlation: one deterministic and the other statistical. The deterministic correlation arises because test outcomes (e.g. frequency resolution measure) may depend on the stimulus presentation level and so the test outcome could appear to be correlated with MCL even though no correlation would be shown when tested at the same stimulus level in all participants. The statistical correlation occurs because random measurement errors in determining MCL are perpetuated through all the outcome measures through dependence of test outcomes on presentation level. These were some of the disadvantages of using MCL as presentation level. However, the approach of basing stimulus levels on MCL has the advantage of setting a comfortable level for all participants, regardless of hearing impairment. Use of a fixed level for all participants may mean that it is uncomfortably loud for some participants with normal hearing and/or too quiet for some hearing-impaired participants. Also it efficiently served the purpose of providing a level that is above threshold (suprathreshold) which was a core requirement for measuring the suprathreshold deficits and hence observing the relation with threshold.

Chapter Five

Results and Interpretation-II: Comparison with study I

5.1 Introduction, method and premise for study II

A second multicentre study was essential for numerous reasons as outlined in Chapter IV (section 4.11). A few of the reasons were to overcome or reduce some of disadvantages regarding the starting stimulus levels, to be able to confirm the main findings of the study I and list the characterization of the multidimensionality of hearing loss. Further a comparison between the two studies will also give an insight into varied aspects that need to be considered when exploring the relations between threshold, suprathreshold and other subjective auditory tests.

In order to overcome some of the disadvantages of using MCL as a reference stimulus level another method to define a reference level was required. An alternative hearing-threshold dependent level method was used. This method was hoped to reduce the statistical dependence caused by use of MCL as reference level. With this approach, there would still be a level effect on many test outcomes, but it would be a systematic one as well as more predictable and possibly corrected for. It was further debatable whether fixed/adaptive noise levels should be used for testing the suprathreshold capabilities level for the second study. Use of fixed noise level would mean all the participants would be compared at the same level, leading to more easily interpreted results as well as further reducing the structural dependence of tests. However it has the disadvantage of measuring all NH listeners at a potentially loud level, while measuring many HI at levels that may not be large enough and produce audibility problems. Also this would mean including only the milder hearing impaired which would interfere with the study design of including subjects with a wide range of hearing loss. Furthermore, adjusting stimulus levels according to hearing loss reduces the influence of audibility and increases the potential visibility of supra-threshold effects. On balance, it was decided to make measurements using stimulus levels adjusted according to hearing threshold levels.

Hence the presentation level was calculated approximately according to the one third gain formula (60 dB+ PTA/3). This formula was considered a good compromise between sufficiently loud level for all listeners to be considered suprathreshold and not reaching too harmful or too loud levels for some others which was possible with use of other formulas such as a ½ gain rule. The presentation level according to the approximate ⅓ gain-formula was as follows:

Table 5.1 : PTA and the corresponding presentation level according to 1/3 gain formula.

PTA level	Presentation level
PTA <= 5 dB	60 dB
5dB < PTA <= 15 dB	63 dB
15dB < PTA <= 25 dB	67 dB
25dB < PTA <= 35 dB	70 dB
35dB < PTA <= 45 dB	73 dB
45dB < PTA <= 55 dB	77 dB
55dB < PTA <= 65 dB	80 dB
65dB < PTA <= 75 dB	83 dB
75dB < PTA <= 85 dB	87 dB
85dB < PTA <= 95 dB	90 dB
95dB < PTA <=105 dB	93 dB

Also certain measures used in study I revealed results that were of limited value. Their correspondence with other variables was also low. These included listening effort and MAA. So they were eliminated in study II. Further MCL bb was also not included to reduce measurement time. In study I it was required for SRT measurements, but this was not the case in II due to use of ⅓ gain-formula for presentation level.

5.2 Test measures included in the data analysis

Table 5.2: Test measures and conditions included in the data analysis.

Hearing Aspect	Test	Conditions measured & Details
Audibility	Audiogram	air conduction: 250,500,1000, 2000, 3000, 4000, 6000 bone conduction: 250-3000Hz
Loudness perception	Acalos	narrowband noises(500 Hz, 3000 Hz)
Frequency-time resolution	FT test	500 Hz 3000 Hz
Speech perception	SRT with BKB sentences	in quiet (binaural) in stationary noise (monaural) in fluctuating noise (monaural)
Binaural hearing	ILD	SRT with matrix-type sentences
	BILD	SRT with matrix-type sentences
Cognitive abilities	Lexical Decision Making	
Subjective judgement	Gothenburg Profile	Questionnaire

The outcome measures and abbreviations are the same as used in the analysis for chapter IV except the loudness level measure Lcut. It is obtained by measuring the intersection of two linear parts of the loudness growth function with independent slope (for details refer chapter III, 3.2.2.2) and approximately corresponds to the slope for CU 20. Other abbreviations used are normal hearing (NH), hearing impaired (HI), Speech recognition threshold (SRT), pure tone average (PTA), Hearing threshold level (HTL). Also additional variables of asymmetry (difference between left and right ears) are used with suffix ‘diff’.

5.3 Structure and focus of the chapter

This chapter thus constitutes the analysis, interpretation and discussion of the data gathered on the second multicentre study. It includes the same test battery and same test order as described in Chapter III with elimination of listening effort, MAA and BB measures as mentioned above along with use of modified presentation levels. The analysis follows the same pattern as chapter 4 which includes general and descriptive analysis to begin with followed by an objective-focussed analysis and interpretation.

However, the main focus of this chapter is to outline the findings of the data presented as part of study II and compare it with study I and other studies in the literature. Thus all the findings are discussed in the same light. The box plots in the descriptive statistics for normal and hearing impaired for study II for the different measures are displayed along with those of study I. Similarly, the same principles of regression and factor analysis of study I are applied here, and are presented in the form of summarized tables which include the predictive variables (regression analysis) and factor loadings (factor analysis) of both the studies. Further, the comparative findings of the two studies are presented and discussed in two parts:

Part I: Includes a very brief discussion of findings from regression and factor analysis of the two studies in their original forms. This was essential in order to recognise the performance of the tests using the same measures (with some eliminations), but on a different group of subjects and using a different method of presentation level.

However, following part I, it was realised that though both the studies used the same measures and tests, there were a few subtle but important differences between the two. Some of the significant ones included frequencies considered for PTA for the two studies, inclusion of measurement of asymmetry etc. Thus in order that they are more comparable across various parameters, the similarities and differences between the two studies are outlined.. Thus a few aspects were modified in study I as per the criteria of study II and analysis were repeated for all measures. This third type of analysis was referred to as modified study I. They are discussed in detail in Appendix VII

Part II: Includes discussion of findings from regression and factor analysis of study I (original and modified) and study II together.

It becomes evident that the above two parts compare the two studies with similar and different parameters. Study I (original) and II used different methods as well as some different test variables. Study I (modified) and study II used different methods but the same test variables. It was attempted to investigate the present sets of data in all possible ways in order to explore their relations on the selected tests, methods and subjects. These in turn are discussed with studies from the literature as chapter IV.

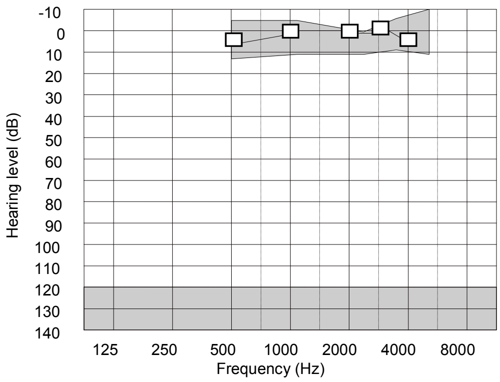
Part III: Includes outline and discussion of the important results deduced from parts I and II. As will be seen below, parts I and II elaborate on a significant number of analyses and different results. Part III on the other hand focuses in detail only on the important, new and interesting results which are an outcome of the various comparisons. Following this, the data from study II is subjected to analysis using a Speech Intelligibility Index model and its findings are briefly discussed. Finally, the various objectives and hypothesis put forth in the thesis (ref chapter I, section: 1.6) are outlined as per the findings from both the studies followed by the outlining the multiple dimensions of hearing loss, again based on the findings from both the studies.

5.4 Subjects included in Studies I and II

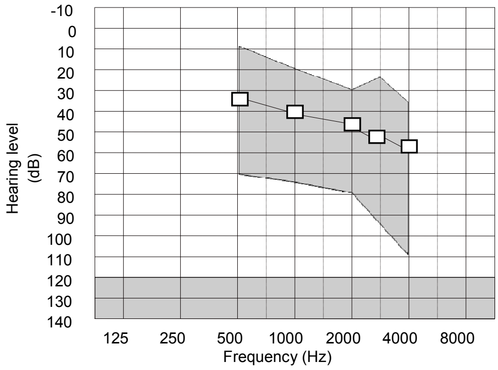
Table 5.3: Details of subjects included in the studies I and II.

Study I	**Study II**
Total number of subjects: 103	Total number of subjects: 128
NH: 30	NH: 26
HI: 73	HI: 102
Age range: NH: 19-39, HI:22-91 years	Age range: NH: 20-40, HI:22-82 years
NH: average PTA (500,1,2,4): 2 dB (range: -10 to 15 dB).	NH: average PTA (500,1,2,4): 1.9 dB (range: -5 to 8.7 dB).
HI: PTA (500,1,2,4): 43 dB (range: 5-100 dB), slope (4000-500):25 dB	HI: average PTA (500,1,2,4): 39.3 dB (range: 10-68 dB), slope (4000-500):30 dB
Audiometric configurations: Mild flat : 38 severe flat: 16 mild sloping: 14 severe sloping: 4 mixed hearing loss 4 sensory neural hearing loss 69	Audiometric configurations: Mild flat : 25 severe flat: 7 mild sloping: 62 severe sloping: 8 mixed hearing loss 0 sensory neural hearing loss 102

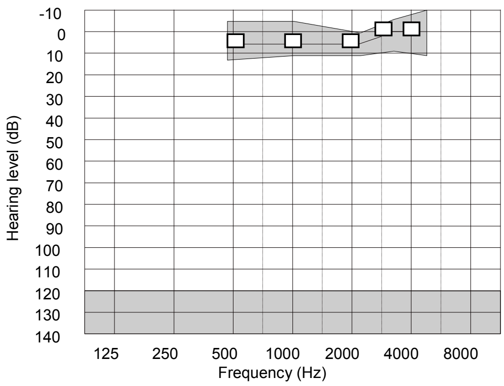
Study I   Thresholds : Normal hearing



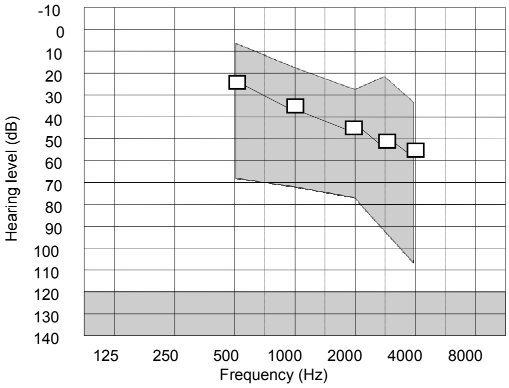
Study I Thresholds : Hearing impaired



Study II   Thresholds : Normal hearing



Study II Thresholds : Hearing impaired



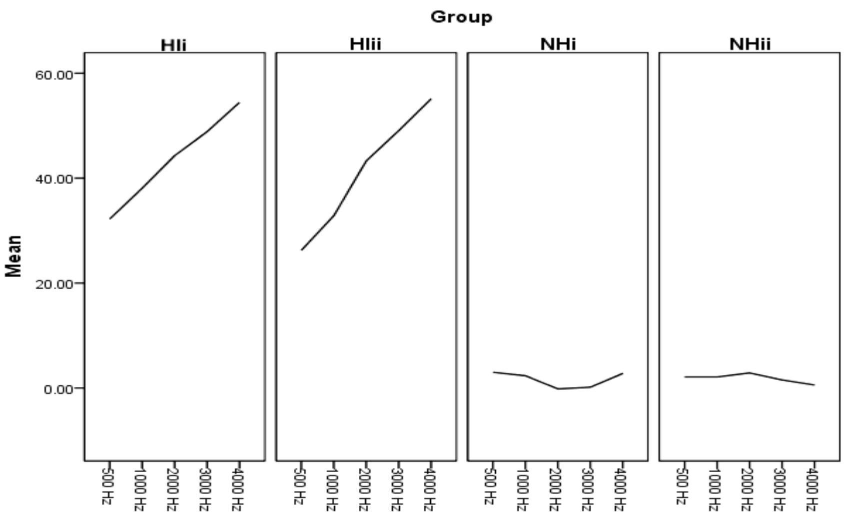


Figure 5.1: Averaged audiogram thresholds at 500- 4000 Hz for study I and II for normal hearing and hearing impaired (upper) and combined (lower).

As can be seen from figure 5.1 (lower panel) the thresholds of hearing impaired in study II show a bigger range when seen on graph as well as slightly more steep slope than study I. The shaded area around the thresholds in the upper panel displays the approximate range for each frequency. The number of subjects in study II are more than that of study I. However, the slope is higher for study II than that for study I. This is also evident from the different audiometric configurations where study I has more of flat audiograms and II has more of sloping.

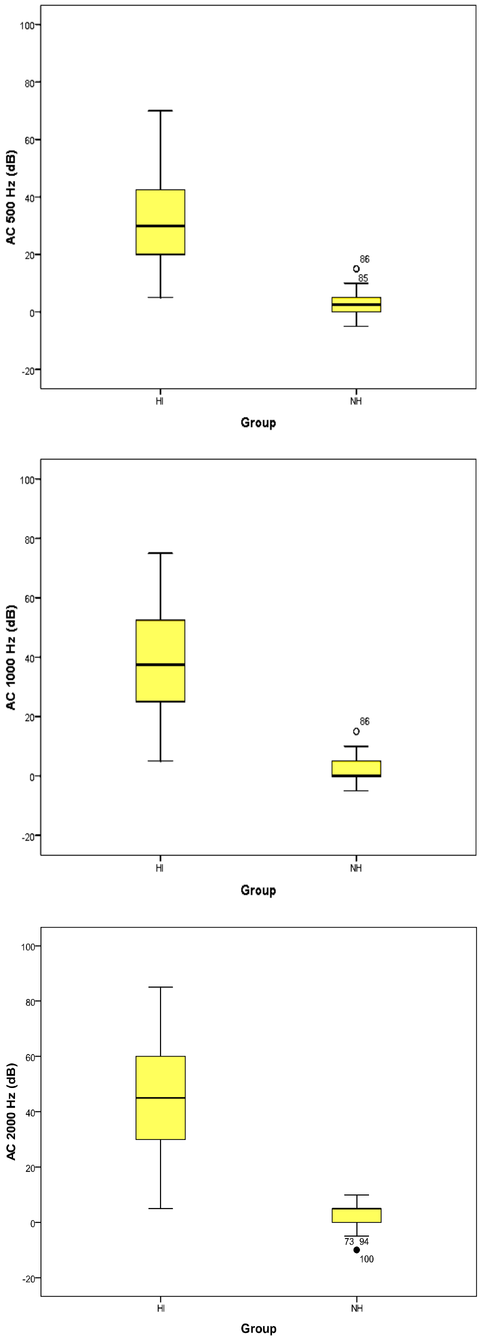
5.5 General and Descriptive Statistics

This analysis outlines the differences between the NH and the HI groups. This is done with box plots for the two groups for all the measures. Better ear measures are included for the analysis. NH listeners are shown in the right box plot and HI in the left box plot in all the graphs. The Y-axis consists of the test measure of concern. Further, the box plots show the median (black bar), interquartile range (box ranges between 1st and 3rd quartile) whiskers represent highest and lowest values after exclusion of outliers. The outliers (circles, defined as any point which falls more than 1.5 times the interquartile range above the third quartile or below the first quartile), and extreme cases (stars, any point beyond the outlier) have subject numbers alongside for identification. It should be noted that this part of the analysis focuses only on the difference between normal hearing and hearing impaired for study II and its comparison with study I as stated above. Other comparisons like right/left, better/worse, test-retest reliability have already been covered in the chapter IV and are not included here. Further independent sample t-test was carried out to observe if the difference between NH and HI group scores are significant or not. They were significant for all measures for both groups and both studies ($p<0.05$). (see Appendix IX for details).

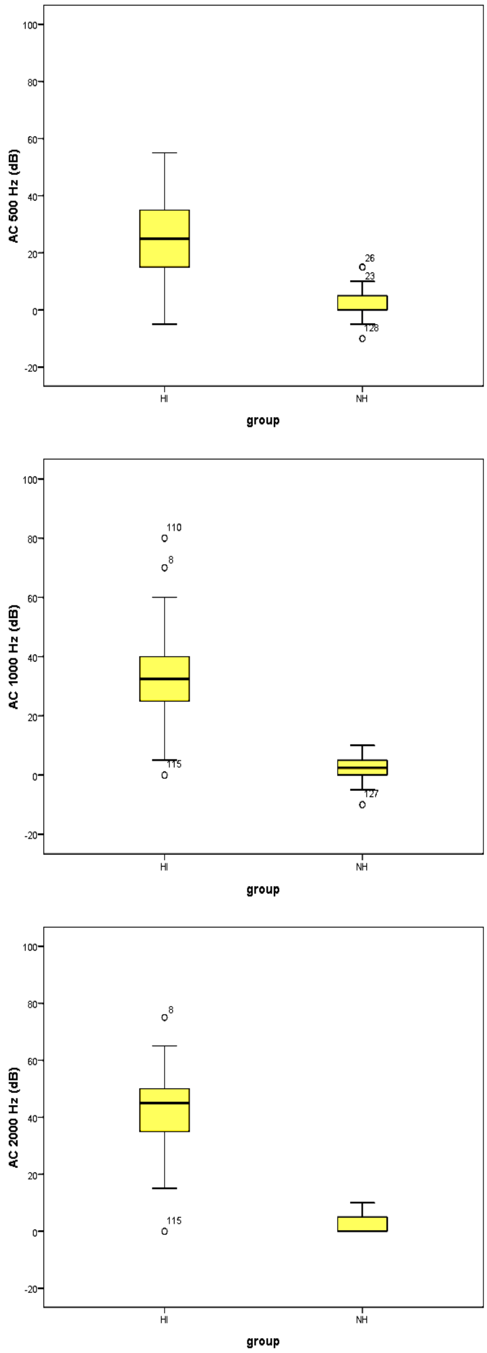
5.5.1 Audiogram thresholds and measures

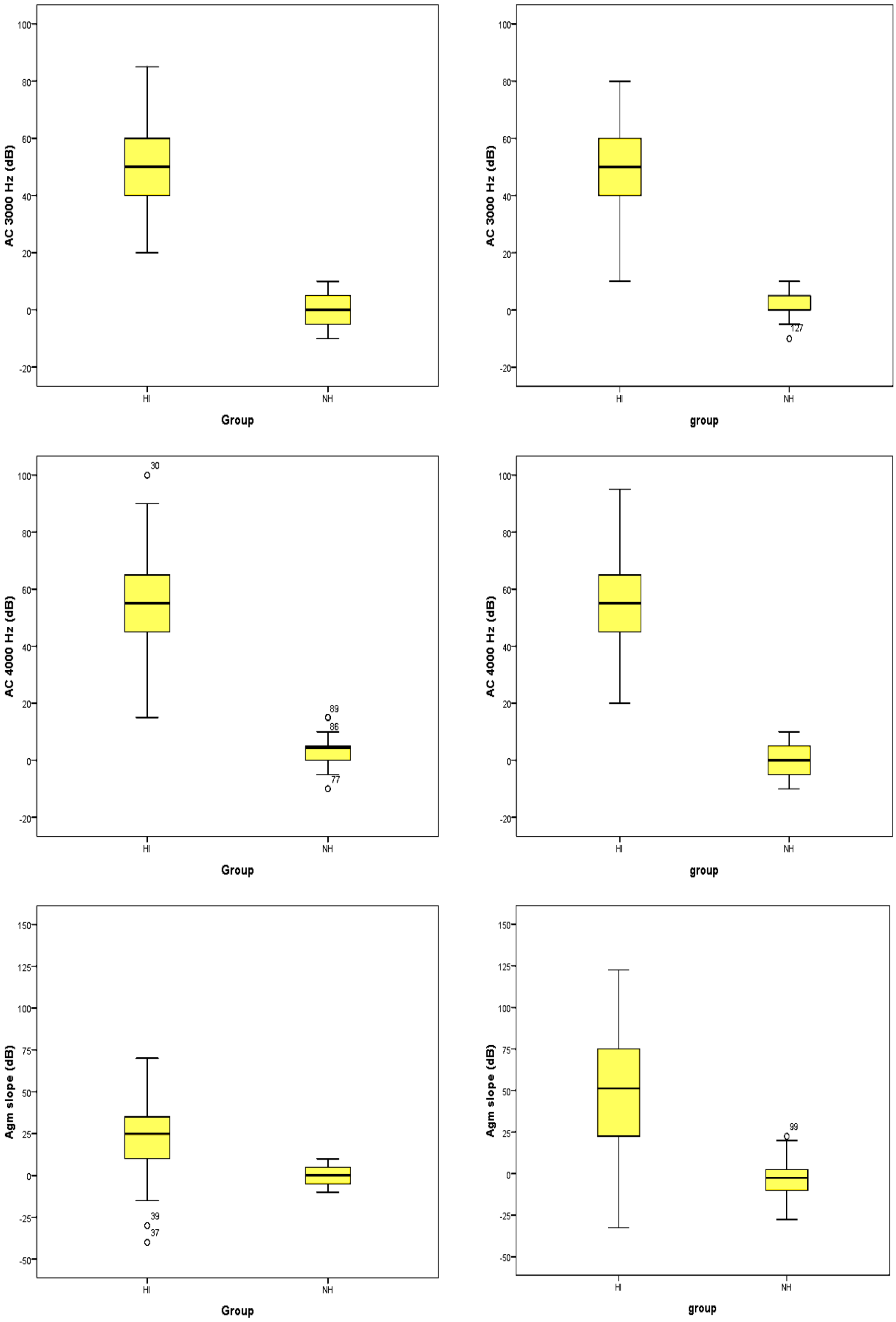
Figure 5.2 below shows hearing thresholds at 500-4000 Hz for NH and HI listeners along with audiogram slope (defined as difference in thresholds between 4000 and 500 Hz) in study I and II along with reference starting level (study II) obtained according to table 5.2.

Study I



Study II




AC 3000 Hz (dB)
Group
group
AC 4000 Hz (dB)
Agm slope (dB)
HI
NH
30
89
86
77
127
39
37
99

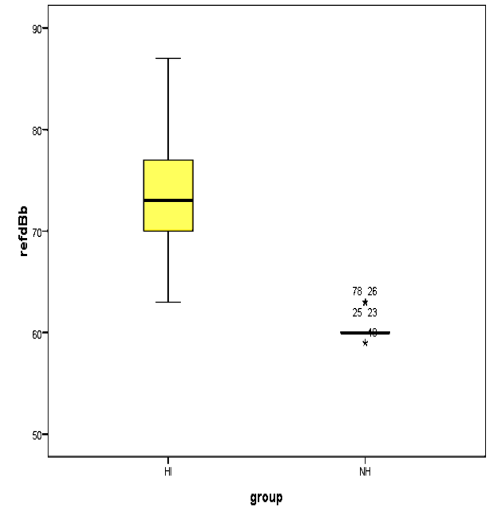
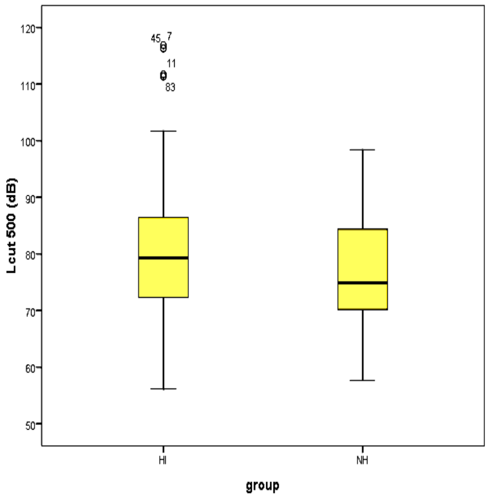


Figure 5.2: Range of Audiogram thresholds at 500- 4000 Hz, audiogram slope and ref dB levels for NH and HI subjects in study I (left column) and study II (right column) Vertical axes show levels in dB HL. In both the studies, NH subjects have better thresholds overall (smaller values signify better performance).

Overall, there is no significant difference in the plots between the two studies and across all groups with NH listeners performing better in both studies and HI revealing more spread. However, on a closer look, across the different frequencies in HI groups the spread of levels is further upwards (beyond 40 dB) for study I than study II. And this is more pronounced for lower frequencies (500, 1000 Hz) and to a lesser extent for higher frequencies (3000 Hz) with almost none for 4000 Hz. Further, for 2000 Hz, the HI group for study I reveals much more spread than study II. On the other hand, for audiogram slope, the spread is more for study II than study I for HI subjects. This is due to the differences in audiometric configurations discussed in table 5.3 .It can be seen that mild sloping group is much larger in study II than study I.

5.5.2 ACALOS (Loudness levels)

ACALOS results in the form of Lcut measures are shown below in figure 5.2. It can be seen that in general, hearing-impaired listeners have steeper slopes than normal-hearing listeners although there is substantial overlap. Only Lcut values (Intersection of two linear slopes) for study II are shown here, since with use of 1/3 gain levels, other MCL-related measures are of less significance.



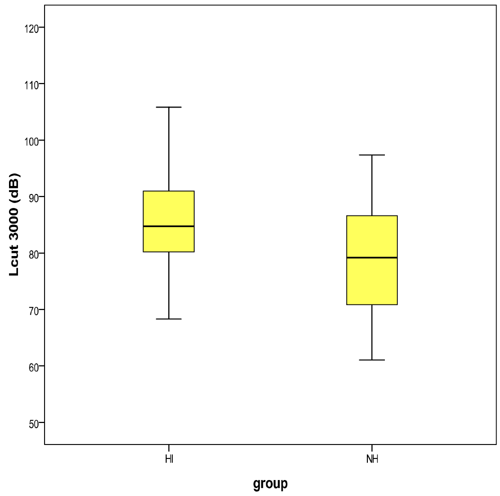
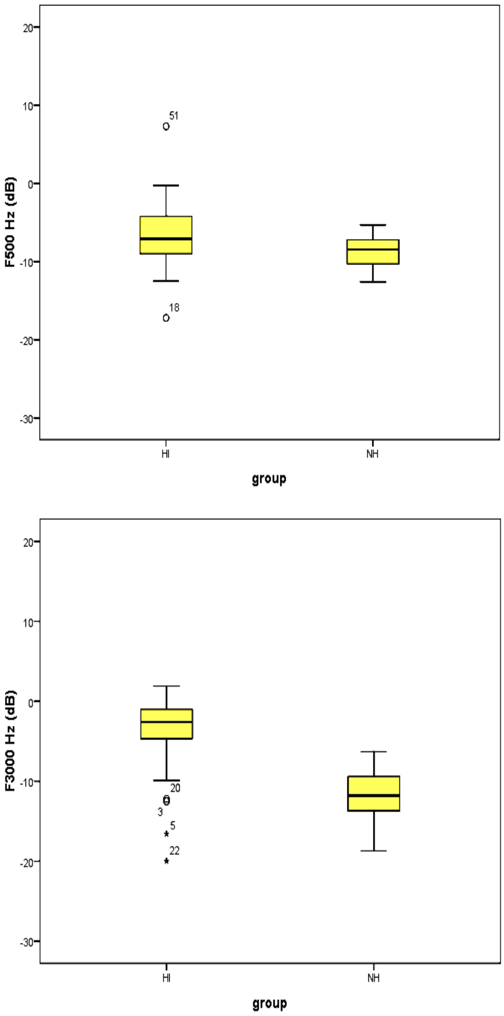


Figure 5.3: Lcut values (intersection of two linear slopes) at 500 Hz and 3000 Hz for NH and HI subjects in study II. Vertical axes show levels in dB SPL with smaller values indicating better performance and slope values with higher values indicating steeper slopes and presence of recruitment.

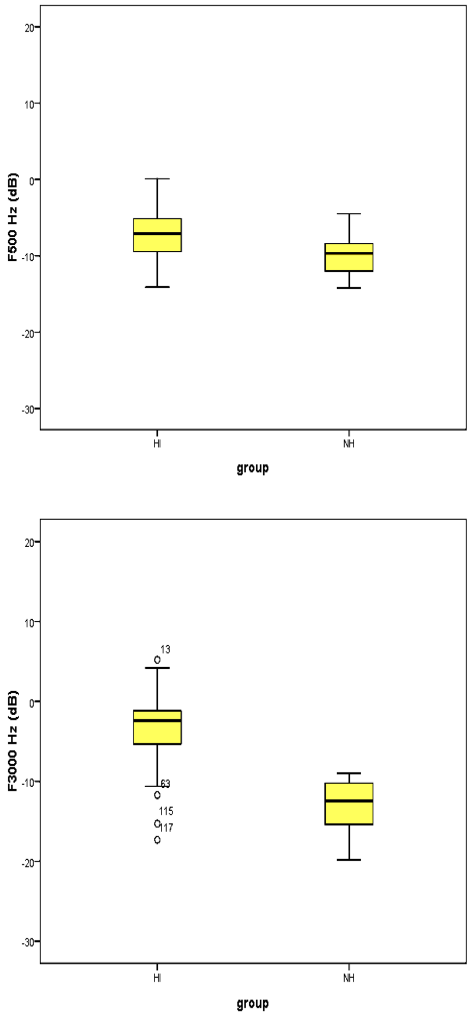
5.5.3 Frequency and Temporal resolution (FT)

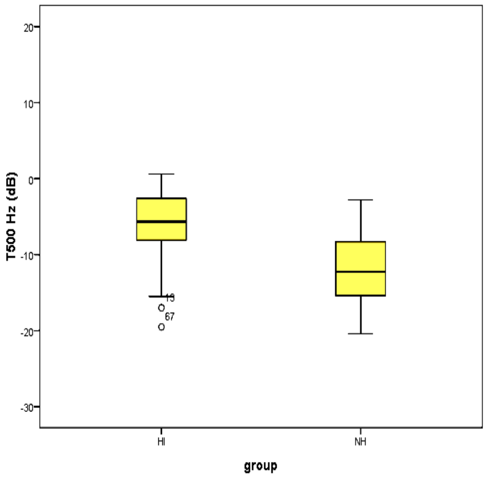
Figure 5.4 shows results of the FT test for NH and HI listeners for the two studies. On average, it can be seen that NH listeners have better spectral and temporal resolution (greater release of masking values shown by more negative number) than HI listeners. There is an equal overall spread in both groups, with slightly more for the HI group. The scores are more or less similar across the two studies for both groups. The difference between the scores is more pronounced for the high frequency in both studies and is less for low frequency. This is related to the fact that the subjects' thresholds were higher for high frequencies. The ceiling effect seen for the HI seen for T3000 measure is discussed below.

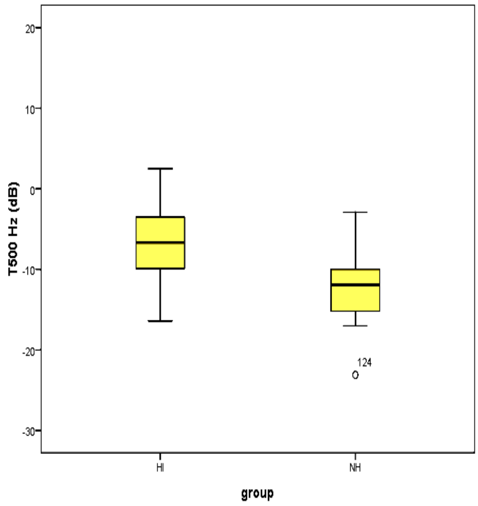
Study I

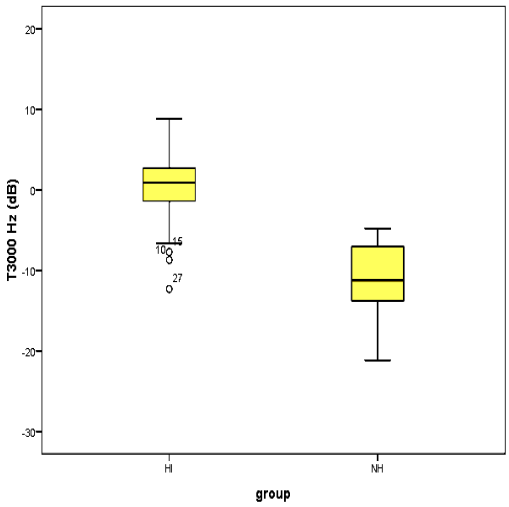


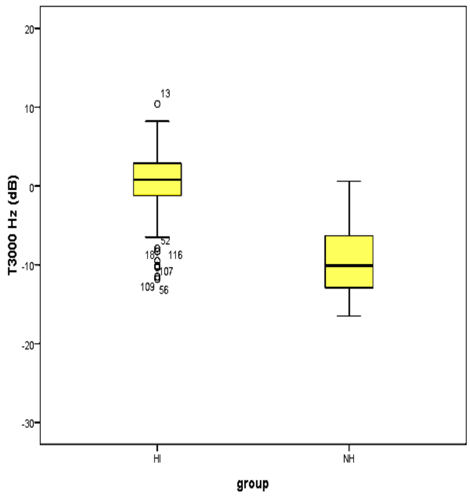
Study II




T500 Hz (dB)
20
10
0
-10
-20
-30
15
67
HI
NH
group


T500 Hz (dB)
20
10
0
-10
-20
-30
124
HI
NH
group


T3000 Hz (dB)
20
10
0
-10
-20
-30
15
10
27
HI
NH
group


T3000 Hz (dB)
20
10
0
-10
-20
-30
13
52
18
116
107
109
56
HI
NH
group

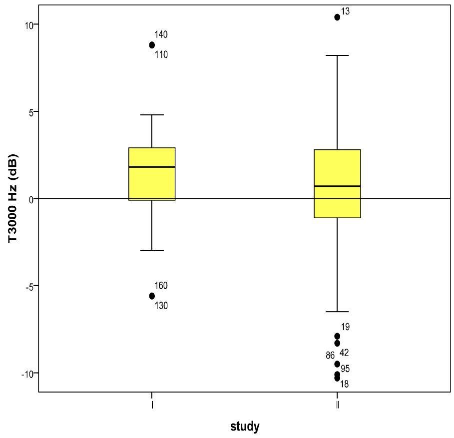


Figure 5.4: Frequency and temporal resolution for NH and HI 500 Hz and 3000Hz in study I (left column) and study II (right column). Vertical axis represents resolution scores in dB with smaller or more negative values indicating better resolution. The last panel compares the T3000 scores of both studies for evidence of a ceiling effect.

As seen above the scores for temporal resolution at 3000 Hz reveal a ceiling effect as discussed in chapter IV (see section 4.5.1.3). This effect is also evident in the study II, with considerable extent of scores lying above the marked line (i.e. above zero which means more positive and hence worse) but perhaps to a lesser extent, mainly because there is considerable spread below the marked line as well. This is not in case in study I. This ceiling effect reflects that the test was difficult for many HI listeners.

5.5.4 Speech Recognition Threshold (SRT) in noise

Study I Study II

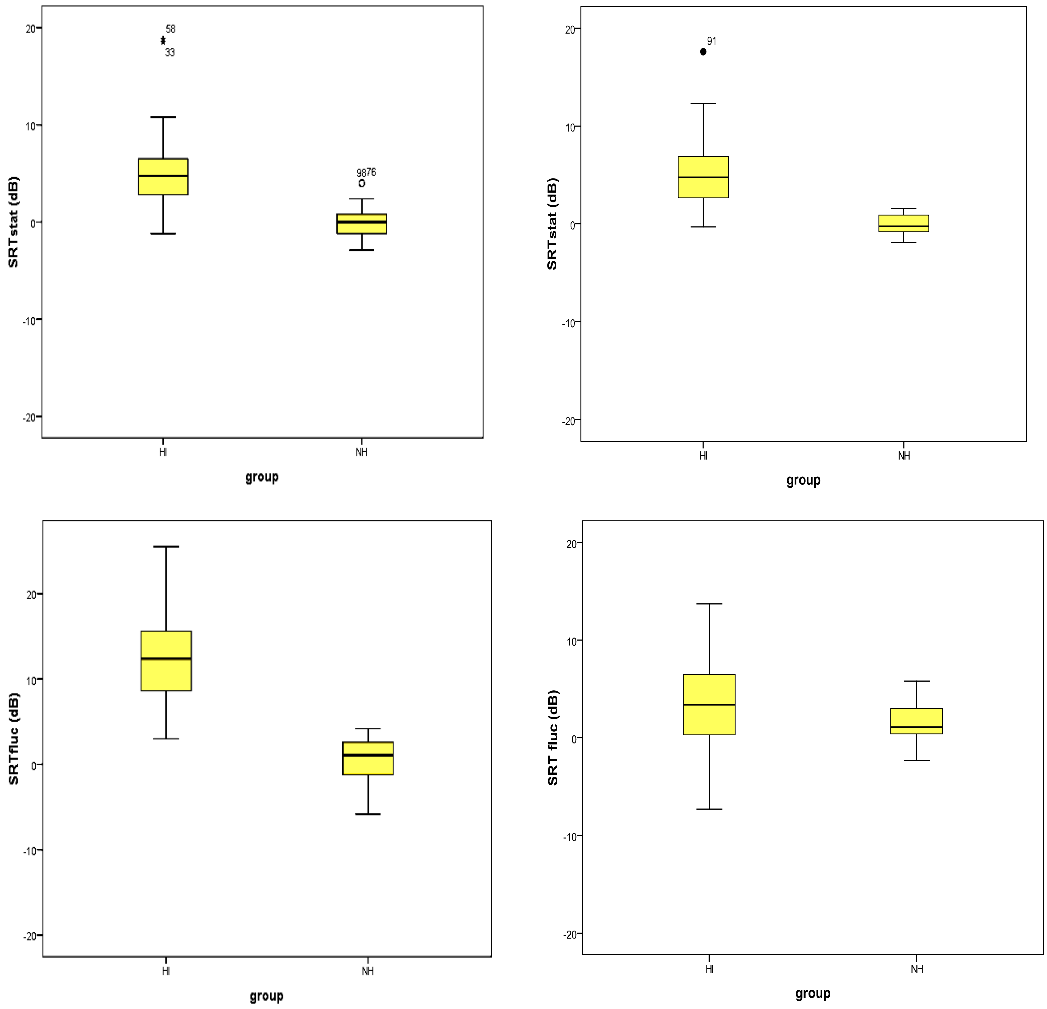


Figure 5.5: Speech recognition threshold for NH and HI listeners in stationary noise and fluctuating noise in study I (left column) and study II (right column). Vertical axis represents speech to noise score in dB, with lower/more negative SNR score indicating better performance.

Figure 5.5 shows corrected SRT results in stationary and fluctuating noise for both the studies. The scores for NH listeners centre around 0 since they have been corrected for language differences. As expected, there is more spread in the HI data. Additionally, differences between NH and HI listeners are larger in fluctuating noise than in stationary noise. Further, the performance scores for both the groups across the studies are similar for the stationary noise, but for fluctuating noise the scores for both the groups are higher (worse) in study I than II and the difference is more pronounced for the HI group. This could be related to the difference in the hearing threshold configuration of the subjects in the two studies as seen in figure 5.1 above.

5.5.5 Speech Recognition Threshold (SRT) in quiet

Study I Study II

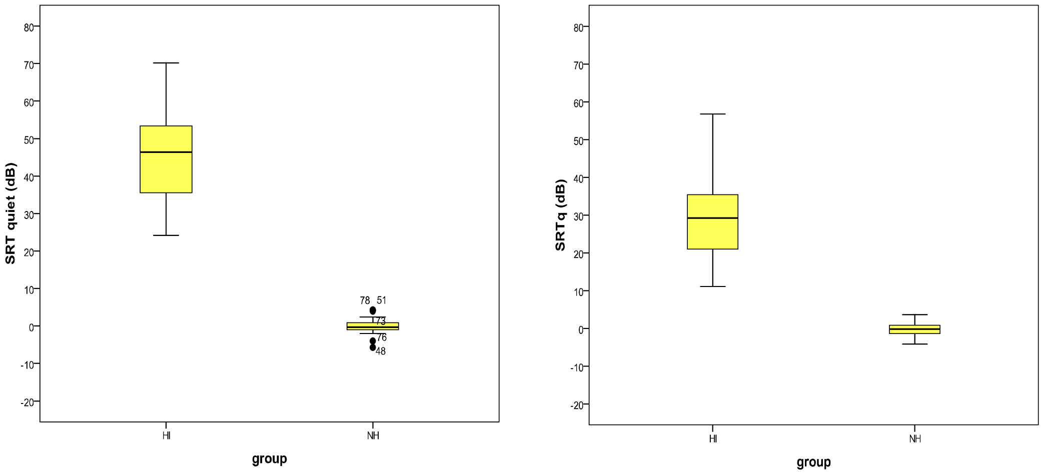


Figure 5.6: SRT in quiet (binaural) for NH and HI in study I (left column) and study II (right column). Vertical axis represents the score in dB, with smaller/more negative value indicating better performance.

Results of the binaural SRT in quiet for normal-hearing and hearing-impaired listeners are shown in Figure 5.5 for both studies. The scores for NH listeners centre around 0 since they have been corrected for language differences. Among HI, there is significant spread in the results, and overall, they require higher presentation levels than NH as expected. Also between the two studies, study I has higher average than that of study II. This could be related to differences in overall hearing range of HI in the two studies. The range for study II was 10-68 dB while study I it was 5-100 dB. This difference is also reflected here where study I has higher average due to some subjects having thresholds ranging as high as 100 dB.

5.5.6 (Binaural) Intelligibility Level Difference (ILD and BILD)

Results of the binaural processing tests ILD and BILD are shown in figure 5.6. The ILD is the difference in SRT between noise and speech from straight ahead in virtual space (situation 1), and speech from straight ahead with noise from one side (situation 2). BILD is the difference in SRT between situation 2, and the same situation with ear at the 'noise-side' blocked. In both tests, absolute values are calculated, hence more negative values refer to less release of masking and therefore worse binaural processing as seen for the HI group. The scores for NH listeners centre around 0 since they have been corrected for language differences.

Study I

Study II

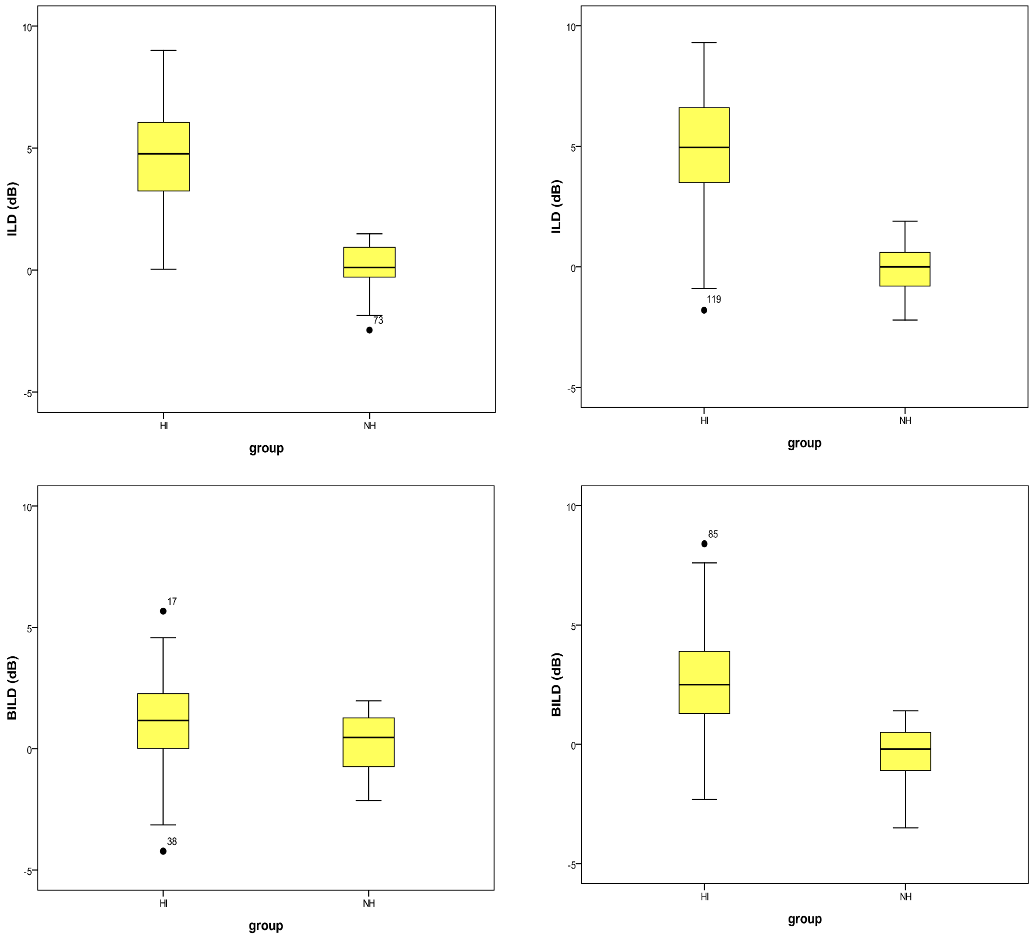


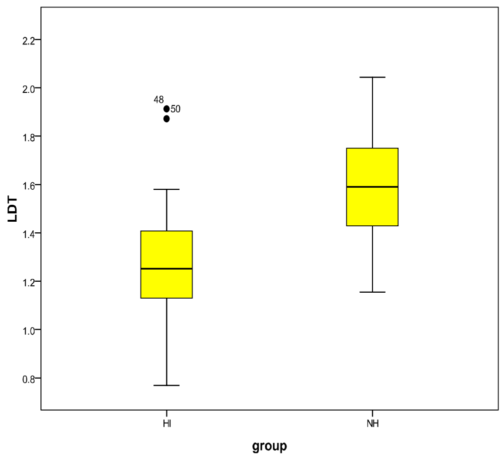
Figure 5.7: ILD and BILD with noise at the side of the poorer ear in study I (left column) and study II (right column). Vertical axes show absolute values for release of masking for both conditions (more negative values refer to better binaural hearing).

Hearing-impaired listeners have less benefit from spatial separation and binaural hearing than normal-hearing listeners. Thus across both studies HI group have worse (more positive) scores with more variation and spread of scores. Between the studies the scores are more or less comparable, except with some difference in HI where there is relatively less variation in study I for both measures.

5.5.7 Lexical Decision Test

Figure 5.8 shows results of the cognitive test (Lexical Decision Test) calculated as percentage of correct score/response time. Thus higher values refer to better performance and lower values refer to poor performance. Thus in general for both studies, HI perform poorer than NH group. To some extent, the differences between the two groups can be attributed to the age difference between them. Between the two studies the variation of scores as well as difference between the two groups is more pronounced for study I than study II. This could be related to differences in age ranges for the two studies. Study I had participants ranging as high as 91 years while for study II it was 82 years. The presence of relatively more elderly participants in study I could have resulted in more variation.

Study I



Study II

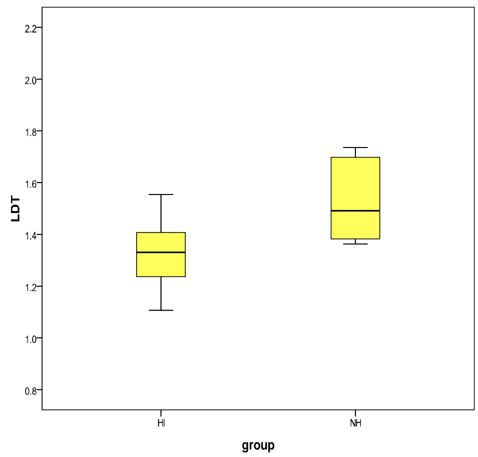
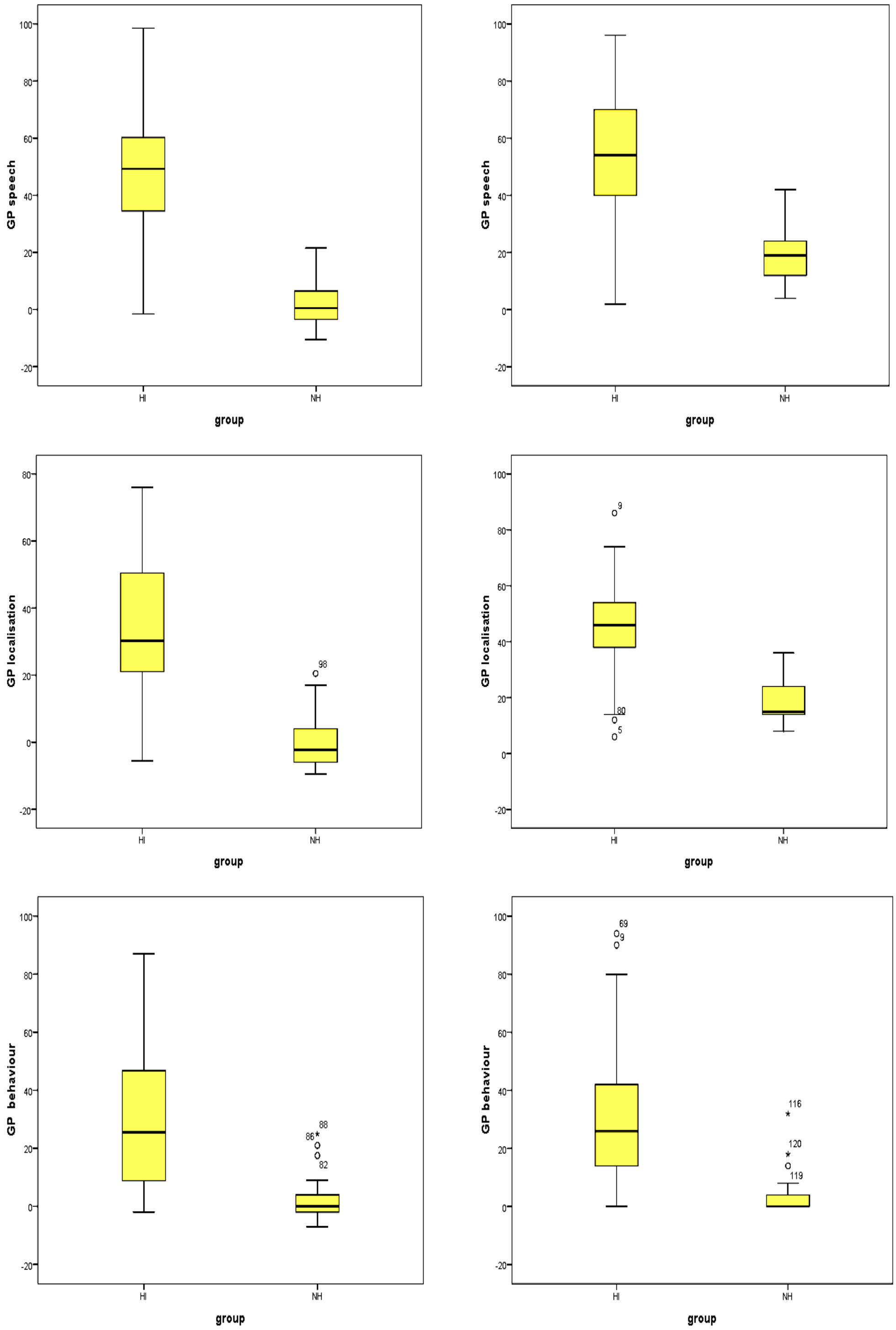


Figure 5.8: Lexical decision making test results in study I (left column) and study II (right column). Vertical axis in represents score (%correct)/ response time, so higher values refer to better performance seen for NH.

5.5.8 Gothenburg Profile

Figure 5.9 shows results of NH and HI listeners on the Gothenburg Profile subscales: speech perception, spatial hearing, social interactions and behaviour (reaction); for the two studies. In general, lower or more negative values indicate better performance. Across the two studies, the performance of the two groups is similar. In both studies, overall the spread of scores is more and performance is worse (higher scores, so more problems) for HI group than NH. The performance is as expected with NH showing better scores along with little spread. There are some small differences between the different subscales for two studies. For speech and localization the spread of scores in NH group for the two studies is slightly more than the other two groups. This could simply because to some extent NH listeners are also susceptible to difficulties in hearing speech as well as locating sounds in the presence of background noise which meant that these two categories can reveal some variation. However for behaviour and social subscales, the difference between the NH and HI groups across the two studies is greater. This again can be expected since the psycho-social attitudes of HI groups are likely to more negatively affected than NH, hence less variation for NH group.

Study I Study II



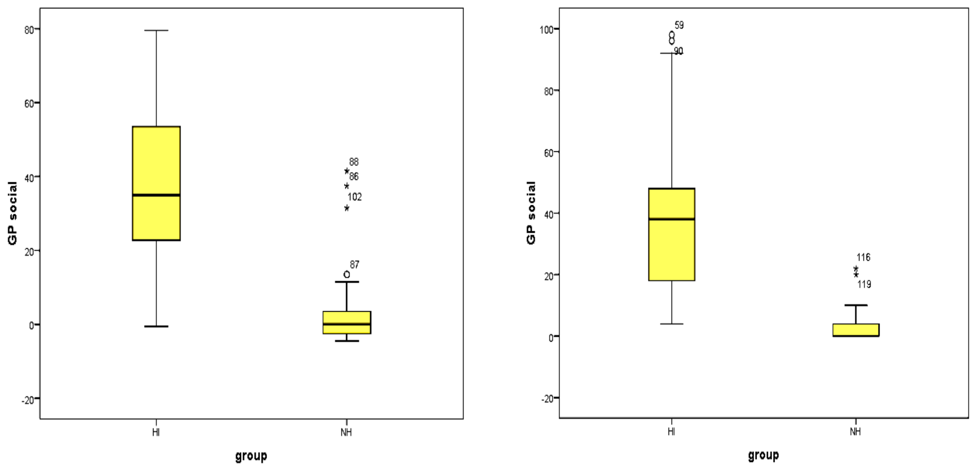


Figure 5.9: Gothenburg Profile results in study I (left column) and study II (right column). The panels present scores (more negative scores refer to better hearing/less problems) for the four subscales of the questionnaire: speech perception, spatial localization, social interactions and behaviour

5.5.9 Summary of findings from descriptive analysis

Performance on most of the tests was similar and comparable to study I with few exceptions for HI group in SRTfluc and audiogram slope for study II higher than in study I. This was attributed to differences in the audiogram configurations. It was the reverse for SRTquiet where I was higher than II. This was attributed to the overall range of threshold measures being higher in study I. Finally there was more and higher variation within HI and NH groups for study I than II. This was attributed to the relative age range differences for the two studies. Study I had more elder (higher in age) subjects than study II and hence perhaps greater variation.

- In general, the hearing impaired group showed worse performance than the normal group as expected
- They also showed more variation than the normal group
- The scores revealed distinct absolute differences between the two groups in study II like study I indicating the tests are sensitive for a wide range of hearing capabilities

5.6 PART I

Per ear measures for study II

Before the analysis of the relations between per-ear measures, the data were checked to see if they deviated from normal distributions. This was done by performing Kolmogorov-Smirnov tests and by visual inspection. The following variables were not distributed normally: AC2000, refdB, SRTstat, F3000, T3000. They were transformed using Blom transformation (Blom, 1958). This transforms the data to approximately normal distribution by ranking them and adjusting the distances between them. This makes them comparable with other variables. Again like study I, only HI data was included since the focus of the study was to score their performance and hence investigate the pattern in majority of clinical population.

Per subject measures for study II

Per subject variables (SRT in quiet, ILD, BILD, Lexical decision, age) and subjective (GP) were tested for normality both by Kolmogorov-Smirnov tests and by visual inspection. All variables except GP (social), GP (reaction), lexical score and response time were distributed approximately normally. They were transformed using the Blom method.

Other details of study II analysis

Like study I, in order to investigate the influence of hearing loss magnitude, data was split up into three groups with the group I consisting of PTA defined as average of 500, 1000, 2000 and 4000 Hz: PTA<= 40 dB (mild HL), the second one (group II) with PTA between 41and 55 dB (moderate HL) and the third (group III) with PTA above 55 dB (moderate-severe). This sort of analysis helped characterize different ranges of hearing loss and the variables affected in each. The analysis would also reveal if there are any factors or trends specific to any group as well as if it was similar or different to study I. Group I had 60 cases, group II had 33 and group III had 9.

Summarised results for all measures of study II as compared to study I are given in the following table.

Table 5.4: Summarised results for all measures of study I and II.

Test measures/ dependant	Predictors Study I	Individual variation (%)	Total % of variation R^2	Predictors Study II	Individual variation (%)	Total % of variation R^2	Comments/changes in study II as compared to I
SRT stat	AC 3000	30%	39%	AC 4000	35%	40%	Frequency of threshold measure changed from 3000Hz to 4000 Hz, however, this still remains to be in high frequency domain. Slight decrease (3%) in variation explained by F500. AC1000 replaced A'gm slope, but both are threshold related.
	F500	5%		F500	2%		
	A'gm slope	4%		AC 1000	3%		
SRT fluc	AC 3000	36%	58%	AC 4000	20%	38%	Frequency of threshold measure changed from 3000Hz to 4000 Hz, however, this still remains to be in high frequency domain. slight increase (4%) in variation explained byF500. T3000 and MCL BB were not present. Overall variation reduced
	F500	14%		F500	18%		
	Age	4%					
	T3000	2%					
	MCL SL BB	2%					
SRT stat (Mild)	A'gm slope	28%	41%	AC4000	37%	41%	Change in group measure from audiogram slope to AC 4000 Hz, however, both belong to same domain of threshold related measures. Similarly T500 to F500
	T500	13%		F500	4%		
SRT fluc (Mild)	Age	28%	40%	AC3000	11%	26%	Change in group measures to audiogram and frequency resolution.
	T3000	12%		AC4000	4%		
				F500	11%		
SRT stat (moderate)	MCL SL 3000	32%	50%	AC4000	15%	15%	Significant decrease in variation explained by the model. Group measures changed from MCL and F500 to threshold, but overall predominantly were related to audibility in both studies.
	F500	10%					
	MCL SL BB	8%					
SRT fluc (Moderate)	F500	32%	80%	AC4000	10%	15%	Significant decrease in variation explained by the model . Primary predictor (frequency resolution) remained unchanged
	T3000	23%		F500	15%		
	MCL SL 3000	14%					
	AGE	11%					
SRT stat (Mod-sev)	----	----	-----	AC 4000	76%	76%	AC4000 Hz explained the variation in study II only
SRT fluc (Mod-sev)	F500		36%	F500	88%	88%	Only extent of variation increased
ILD	AC 3000	16	48%	AC 1000diff	17%	48%	The primary predictor for both studies was audibility measure.
	MCL BB diff	14		T500	10%		
	F500 diff	6		AC 2000	6%		
	AC 3000 diff	4		Lcut3000diff	3%		
	SL3K diff	4		AC 500	12%		
	MCL SL BB diff	4					
BILD	MCL 500	19%	19%	Lcut500	3%	13%	Low frequency measures in both studies
				ILD	11%		
GP speech	SRTq	38%	38%	SRTfluc	3%	15%	For speech and social subscales, SRT measures were predominantly responsible for the variation in GP scores in both studies while for the other two, they varied.
				slope	4%		
				AC4000diff	8%		
GP loc	SRTq	32%	41%	F500diff	4%	21%	
	ILD	5%		T3000	4%		
	Age	4%		BILD	3%		
				F500	10%		
GP soc	SRT fluc	13%	23%	BILD	4%	13%	
	F3000	4%		SRTfluc	9%		
	MCL 3000	6%					
GP reaction	SRTq	13%	18%	T500	4%	4%	
	SRTstat	5%					

5.7 Summary of findings from comparison of study I and study II

1) Similar to findings from study I, for speech recognition in stationary noise, hearing threshold at 4000 Hz was the largest predictor explaining the maximum variation accounting for over 35%, while the other two variables F500 (3%) and hearing threshold at 1000 Hz (2%) account for the rest. The overall model explained 40% of variation in the speech test scores. The only difference was frequency of hearing threshold in study I was 3000 Hz.

2) Similar to findings from study I, for speech recognition in fluctuating noise, hearing threshold at 4000 Hz was the largest predictor explaining the maximum variation accounting for over 20% followed by F500 explaining 18% of the variation. The overall model explains 38% of the total variance. The only difference was frequency of hearing threshold in study I was 3000 Hz and other variables like age and temporal resolution were not present.

3) Thus like study I, findings from study II also resemble findings from group A studies (Humes *et al.,* 1994; Jerger *et al.,* 1991, Divenyi *et al.,* 1997a) in the sense that threshold measures were responsible for large variation in speech recognition scores. However, frequency resolution also contributed to the prediction and hence again like study I, the findings resemble group B (ii) studies (Festen and Plomp, 1983; George *et al.,* 2006; van Rooij & Plomp 1990).

4) Both the studies revealed low frequency auditory resolution (indicated by F500) as a predictive variable which was a novel finding since most other studies (Phillips *et al.,* 2000; Festen and Plomp, 1983; Dreschler and Plomp, 1985) that showed frequency resolution as a contributing factor revealed a high frequency measure, one of the reasons being perhaps non-inclusion of low frequency measure for frequency resolution.

5) Unlike study I, temporal resolution at 3000 Hz and age did not predict speech recognition in fluctuating noise. The difference is attributed to the different approaches used in the two studies with regards to reference starting level. However, even in the study I these two accounted for minimal variation only. Thus unlike findings from study I, findings from study II reveal that speech recognition in the two types of noise in fact can be predicted by similar variables (which are threshold and frequency resolution in this case).

6) In different groups of hearing loss, SRT in noise (generalising findings from both stationary and fluctuating noise) could be predicted by audiogram and auditory resolution measures in mild losses and only auditory resolution measures for moderate and higher losses. This is opposite of what would be expected, since at the presentation levels used it would be expected that mild hearing losses would be unaffected by threshold effects. In study II this was seen for fluctuating noise. This indicates decreased influence of hearing sensitivity and increased influence of auditory resolution with increase in magnitude of hearing loss. In other words a shift from threshold related to suprathreshold processing and the fact that factors other than threshold come into play when the degree of loss is more is revealed here. This again was similar to that of study I and was found to be especially true for fluctuating noise in the present study where the variation explained by frequency resolution increases as degree of hearing loss increases. However, the trends seen here cannot be considered stable due to such small and non-uniform number of cases. The findings from the two studies are similar to that of Pavlovic (1984). His interpretations revealed that supra-threshold distortions were unimportant for mild/moderate hearing loss as compared to more severe losses, which was also seen in the present study. However Lutman (1987) showed the opposite findings where auditory resolution was important for mild hearing losses and not so much for more severe losses. More on this is discussed in 5.9.

7) Similar to study I predictive variables for speech recognition were governed by magnitude of hearing loss. However, unlike study I, the different predicting variables did not change for two types of noise. Study II mainly included only threshold and frequency resolution with their extents varying for both stationary and fluctuating noise.

8) As in study I, the importance of low-frequency hearing and measures of asymmetry along with hearing sensitivity in the processing of spatial hearing cues is demonstrated by binaural hearing measures in this study as well. Study by Humes and Roberts (1990) has also revealed that binaural release from masking that could potentially improve performance has the greatest effect at lower frequencies.

9) Unlike study I, where SRT measures were the best predictors for self-rated hearing disability, study II did not reveal predominance of any particular group across all subscales. Further, the percentage of variation explained by the predictors in study II was relatively less compared to study I. This emphasizes that any study is just an estimate and not necessarily reliable.

From above it can be seen, the primary predictors of most of the group measures did not change in the two studies. Thus overall both studies agreed fundamentally except the few methodological differences which could have led to any minor discrepancies. And these minor changes in variables do not occur consistently across studies and are probably chance findings, hence not meaningful. However, the general stability of the main findings suggests that they are robust (as opposed to chance findings.)

5.8 PART II

Summarised results for all measures of study II as compared to study I along with modified study I are given in the following table. It consists of findings as per criteria discussed in 5.3 and appendix VII.

Table 5.5: Summarised results for all measures of study I, modified study I and study II.

Test measures/ dependant	Predictors Study I	Individual variation (%)	Total % of variation R^2	Predictors Study II	Individual variation (%)	Total % of variation R^2	Predictors Study I (modified)	Individual variation (%)	Total % of variation R^2
SRT stat	AC 3000	30%	39%	AC 4000	35%	40%	AC 4000	33%	43%
	F500	5%		F500	2%		F500	7%	
	A'gm slope	4%		AC 1000	3%		Age	3%	
SRT fluc	AC 3000	36%	58%	AC 4000	20%	38%	AC 3000	36%	56%
	F500	14%		F500	18%		F500	14%	
	Age	4%					Age	4%	
	T3000	2%					AC 4000	2%	
	MCL SL BB	2%							
SRT stat (Mild)	A'gm slope	28%	41%	AC4000	37%	41%	AC 4000	43%	51%
	T500	13%		F500	4%		T500	8%	
SRT fluc (Mild)	Age	28%	40%	AC3000	11%	26%	AC4000	38%	47%
	T3000	12%		AC4000	4%		age	9%	
				F500	11%				
SRT stat (moderate)	MCL SL 3000	32%	50%	AC4000	15%	15%	AC 3000	26%	26%
	F500	10%							
	MCL SL BB	8%							
SRT fluc (Moderate)	F500	32%	80%	AC4000	10%	25%	AC 3000	30%	59%
	T3000	23%		F500	15%		Age	19%	
	MCL SL 3000	14%					F3000	10%	
	AGE	11%							
SRT stat (Mod-sev)	----	----	-----	AC 4000	76%	76%	AC 500	28%	28%
SRT fluc (Mod-sev)	F500		36%	F500	88%	88%	F500	53%	53%
ILD	AC 3000	16	48%	AC 1000diff	17%	43%	SRTfluc diff	24%	54%
	MCL BB diff	14		T500	10%		AC4000	14	
							MCL BB diff	4%	
	F500 diff	6		AC 2000	6%				
	AC 3000 diff	4		refdB	7%		AC 500	4%	
							MCL BB	5	
	SL3K diff	4		Lcut3000diff	3%		SRT fluc	3%	
	MCL SL BB diff	4							
BILD	MCL 500	19%	19%	AC 500	12%	15%	MCL500	18%	18%
				Lcut500	3%				
GP speech	SRTq	38%	38%	ILD	11%	18%	SRTq	38%	38%
				SRTfluc	3%				
				slope	4%				
GP loc	SRTq	32%	41%	AC4000diff	8%	19%	SRTq	32%	38%
	ILD	5%		F500diff	4%		ILDdiff	6%	
	Age	4%		T3000	4%				
				BILD	3%				
GP soc	SRT fluc	13%	23%	F500	10%	14%	SRT fluc	13%	33%
	F3000	4%		BILD	4%		AC500diff	8%	
	MCL 3000	6%					F3000	6%	
							MCL3000	6%	
GP reaction	SRTq	13%	18%	SRTfluc	9%	13%	SRTq	13%	23%
	SRTstat	5%		T500	4%		SRTstat	5%	
							AC500diff	5%	

5.9 Summarized Comparisons and discussion study I (original and modified) and II

Speech recognition in the two types of noise, (objective 1, 2, 4):

SRT (stationary noise): As can be seen, in all studies high frequency hearing threshold (3000/4000 Hz) explained the highest variation in SRT scores ranging from 30-35% followed by frequency resolution at 500 Hz ranging from 2-7%. Beyond these two, different variables such as age and audiogram slope were seen to predict stationary noise. They were not consistent and hence can be chance findings. Overall, it can be said that the findings did not vary much across studies.

SRT (fluctuating noise): Again, as seen, in all studies high frequency hearing threshold (3000/4000 Hz) explained great variation in SRT scores ranging from 38-58% followed by frequency resolution at 500 Hz ranging from 14-18%. Study I (both) revealed presence of other variables such as age, temporal resolution and MCL slope. Age and temporal resolution as discussed above have been known to be associated with speech recognition in noise while presence of MCL measure is probably the direct consequence of method used which utilised MCL-related level as the reference starting level. However, they explained only 2-4% of variation across studies and can be considered to be of less significance and hence it becomes evident that even for fluctuating noise, the main findings did vary much across studies.

The findings concerning hearing threshold and frequency resolution found in the present study confirm those of other studies such as Humes and Roberts, (1990), Jerger *et al.,* (1991); Humes *et al.*, (1994); Divenyi *et al.,* (1997a) (hearing threshold) and Festen and Plomp, (1983); Lutman (1987); Dreschler and Plomp (1985) (frequency resolution).

Speech recognition in different groups of hearing loss (objective 3):

The results from table 5.5 suggest the regressions within hearing loss groups are not stable with the subject numbers available. Hence in order to increase the number of subjects in each group (and hence the reliability of the results) data was combined for two studies and regressions were repeated with threshold and FT measures as independents. Data consisting of 175 HI (study I+II) was divided into two groups this time; mild hearing loss (below 40 dB, n=94) and moderate (above 40 dB, n=81). This is discussed below with help of table 5.6 below where combined results for the two studies is displayed.

Table 5.6: Findings from regression analysis for different groups of hearing loss combined for study I and II

Test measures/ dependant	Predictors Study I + II combined	Individual variation (%)	Total variation	Number of cases
SRT stat (Mild: below 40 dB)	**AC3000**	**30%**	**39%**	**94**
	AC4000	6%		
	F500	3%		
SRT fluc (Mild below 40 dB)	**AC4000**	**23%**	**48%**	**94**
	Slope	14%		
	AC500	6%		
	T500	2%		
	AC3000	3%		
SRT stat (moderate: above 40 dB)	AC4000	22%	32%	81
	F500	**10%**		
SRT fluc (Moderate: above 40 dB)	**F500**	**15%**	**26%**	**81**
	AC1000	5%		
	age	6%		

In the table it can be seen that for mild group threshold measures are main predictors for both noises, while in moderate group, for stationary noise: the prediction for frequency resolution increases from 3 to 10 % while in fluctuating noise it takes over as the main predictor. Thus a weaker trend seen in table 5.5 becomes more apparent here and it can be seen that as the degree of hearing loss increases speech recognition in both types of noise is predicted more by frequency resolution and less by auditory sensitivity measures. Scatter plots for groups of hearing losses are further plotted in part III.

In clinical terms, the presence of this trend could mean that the speech recognition in milder losses is subject to change mainly due to insufficiencies in threshold-related factors, while as the loss increases, it tends to alter due to insufficient threshold, and /or supra-threshold and resolution factors. Of course the groups have presence of the other factor as well; it is just the relative extent that varies. Physiologically, this can be related to more hair cells being damaged as the degree of hearing loss increases. This would explain the role of thresholds to depict speech recognition. Alternatively, it is suggested that, in mild losses, speech recognition in noise is affected due to problems in the peripheral auditory system, but as the degree of loss increases, the damage perhaps also includes some processes beyond the periphery even though threshold and frequency resolution are fundamentally peripheral. This suggestion rises from the fact that speech redundancies are often processed beyond the cochlea. Also according to some studies analysis of specific signal attributes such as frequency, intensity and duration is partly done by the central auditory system (Albeck et al., 1992; van Rooij & Plomp, 1990). Of course more in depth study and analysis is required to establish the same. Also, it is known that the primary role of outer hair cells is to actively influence the mechanics of cochlea, so as to produce high sensitivity and sharp tuning (Moore, 1986) and hence ultimately increase frequency discrimination.. Thus as more number of outer hair cells are damaged it is more likely to affect suprathreshold abilities like frequency and temporal resolution. Speech recognition in noise for different groups of hearing loss has been studied by Lutman (1987) and Pavlovic (1984). The former used fixed presentation levels while the latter used predictions based on SII. The present study on the other hand utilised adaptive presentation levels. Hence the trends seen here have not been observed before as well as are novel in this context.

Objective 5

Binaural measures: ILD measures in two out of three analyses were predicted primarily by threshold measures and by SRT measures in one. Of course occurrence of one particular measure as a predictor was seen to a lesser extent. However, measures of asymmetry for different group measures surfaced in all three which is an important finding. Further for BILD, all showed low frequency measure predictors indicating their importance. Humes and Roberts (1990) also revealed that improved performance due to binaural release from masking has its greatest effect at lower frequencies.

Subjective measures: GP subscales in all studies were predicted mainly by SRT measures with the exception of social and localization measures in study II.

Lexical decision and age: In both studies they showed a negative correlation revealing that as age increases cognitive processing could be affected,

Factor analysis studies I&II:

Table 5.7: Summary of results of the factor analysis for per ear measures in study I original and modified and study II).

Study I original	Variables	Variation (%)	Study II	Variables	Variation (%)	Study I modified	Variables	Variation (%)
High frequency processing	Audiogram slope, T3000 (Includes SRT)	20%	Low-mid frequency processing	AC500, AC1000, ref dB, F500	26%	Audibility (includes SRT)	AC500-AC4000, PTA, SRTstat, SRTfluc	26%
Audibility	MCL 500,3000, BB	20%	High frequency processing (Includes SRT)	AC4000, AC3000, SRTstat, SRTfluc	25%	High frequency processing	F3000, T3000	19%
Recruitment	Loudness level slopes at 500,BB	14%	Loudness tolerance	Lcut500 Lcut3000	13%	Recruitment	MCL slope 500,MCL slope BB	12%
Low frequency processing	F500, T500	13%				Low frequency processing	F500, T500	11%
Total % of variation		67%	Total % of variation		64%	Total % of variation		68%

As seen in the table above high frequency processing, low frequency processing, audibility and recruitment were the four factors that were observed in all the three studies. It was the order (and hence the factor that explained the highest variation and highest factor loading) that varied in all. Also the actual groups of measure in each factor were not necessarily the same but were rather named depending on the pattern of the highest factor loading. For example the audibility factor included MCL measures in study I (original), but AC thresholds in study II or low frequency processing included auditory resolution and /or threshold measures. However, some common observations can still be seen. SRT measures grouped with high frequency threshold in all three analyses. This further supports the main findings discussed above. The factors in all three studies were grouped according to the frequency (high/low) of the group measure rather than the measure itself which highlights its importance. Results of factor analysis by various studies have revealed similar measures high frequency processing represented by auditory sensitivity and auditory resolution (Lutman, 1987); audibility factor (Humes *et al.* 1994) and high frequency associations (Divenyi *et al.* 1997 c).

5.10 Summary

The sections above dealt with comparing the two studies with same and different parameters. Study I (original) and II somewhat used different methods as well as some different test variables. Study I (modified) and study II used somewhat different methods but the same test variables. It was attempted to investigate the present sets of data in all possible ways in order to explore their relations on the selected tests. As it becomes evident that even with the use of different subjects, methods used or modifications, the main findings and/or general trends did not vary greatly. This is an important finding which increases the credibility of the results as well as the tests themselves. It also suggests that speech recognition has been repeatedly explained by threshold and frequency resolution.

Of course ILD and GP measures showed some variation but this can be understood since the methods and subjects used were different. These two tests were perhaps most sensitive to these factors since for ILD, the predictive variables will vary depending upon the audiogram configurations, asymmetry between the two ears and hence localization ability while the GP measures were susceptible to subjective bias that is always involved with any self-report tests.

Thus while chapter IV was dedicated to outline, interpret and discuss the findings in the thesis, the present chapter served to compare, contrast and confirm the findings by using the same tests with certain modifications and eliminations on a different group of subjects.

As can be seen, the primary predictors for each of the test measures were quite similar in both the studies. Only the extent of their variation differed for some. This can be expected because, in study I the measures were MCL-controlled while in study II they were threshold-controlled. In the second study in fact some of the interrelations became more evident than what were seen in the first especially for the different groups of hearing loss. On the other hand, presence of threshold measurements dominated most predictions like the MCL-measures in the first study. Thus this internal dependence of the variables is unavoidable to some extent in such studies. However, the important finding here is that it did not change the results to a great extent. This implies that at any given suprathreshold level the SRT predictors as well as for other measures would possibly be more or less same with a few small variations. Thus the findings from the studies complement and support each other. Also while study I itself the established the different aspects of hearing loss, the other helped to further characterize these multiple aspects.

Part III

5.11 Outline and discussion of the important findings deduced from interpretation and analysis presented in parts I and II

The most important finding revealed in the two studies is that hearing threshold and frequency resolution are responsible for the majority of the explained variation in speech recognition in noise especially for fluctuating noise. It thus becomes essential to explore the relation between these three variables in detail for the two studies. The following figure thus reveals a matrix of scatter plots for the three variables in both the studies followed by their actual correlations.

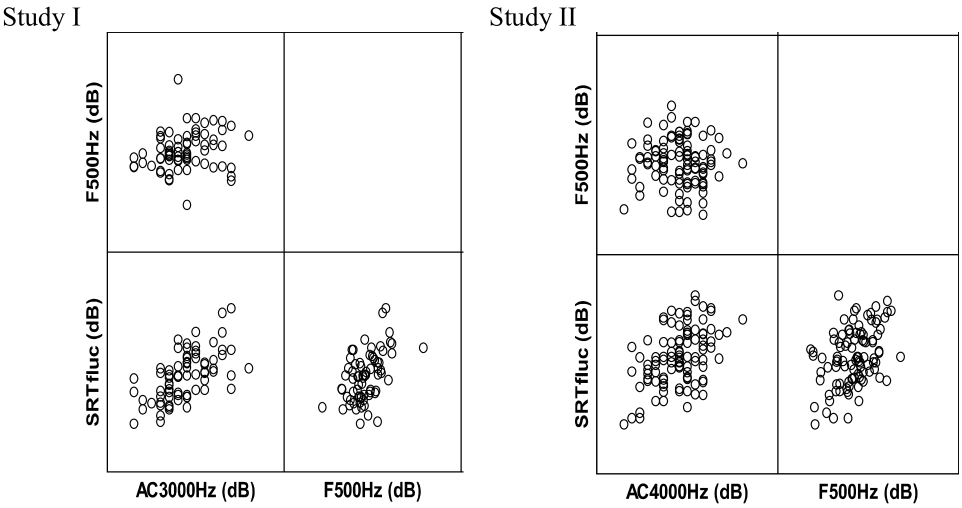


Figure 5.10: Scatter plot matrix of SRTfluc, AC3000/4000 and F500 where n=72 (Study I) and n=102 (Study II)

Table 5.8: Correlation between SRT in fluctuating noise, frequency resolution and hearing threshold in studies I and II.

Study I	F500	AC3000	SRTfluc
SRTfluc	.503**	.608**	1

Study II	F500	AC4000	SRT fluc
SRT fluc	.408**	.467**	1

As can be seen there is considerable spread across the plots for both the studies revealing significant variation of performance in the hearing impaired. This further becomes evident by the correlation tables (significant at 0.01 level which is represented by stars) displayed below. The correlation coefficient is significantly more in study I for both variables (threshold and frequency resolution) than in study II. This could be related to the different subjects, test methods etc. Correlations between AC 3000/4000 and F500 for both studies were not significant.

The above figure gives simple correlations between the three variables and hence general outlook of the results. In order to observe the relations more specifically, partial correlation controlling for audiogram measures followed by partial regression and residual plots was performed. These are discussed below.

Partial correlations for the two studies are given in table 5.9 below. As can be seen, in both studies frequency resolution measures remain significant after controlling for audiogram measure (pure tone average 500-4000 Hz) which helps confirm its role in predicting speech recognition in noise.

Table 5.9: Partial correlations between SRT and FT measures controlling for audiogram measure (control variable: PTA (5,1,2,4) in study I and II.

			Study I				**Study II**			
			F500	T500	F3000	T3000	F500	T500	F3000	T3000
SRT	SRTstat (dB)	Correlation	**0.241**	0.205	0.210	0.119	0.081	-0.106	**0.348**	0.145
		Significance (2-tailed)	**0.048**	0.094	0.086	0.348	0.443	0.311	**0.001**	0.167
	SRTfluc (dB)	Correlation	**0.391**	0.188	0.207	**0.340**	**0.292**	0.085	**0.286**	0.140
		Significance (2-tailed)	**0.001**	0.125	0.090	**0.005**	**0.004**	0.417	**0.005**	0.181

Secondly, partial regression plots are further given below in figure 5.11. They essentially help investigate the relation between the dependent and independent variables and will serve to complement the findings from partial correlations above. Partial regression plots give the strength of marginal relationship between the independents in the full model. Thus it will reveal the relation of SRT fluc with AC3000/4000 while partialling for F500 or vice versa. The partial regression plots are generated for each of the predictive/independent variables of SRTfluc (dependent variable) in the full model during the regression analysis as seen part II above (table 5.4 and 5.5).

Study II Study I

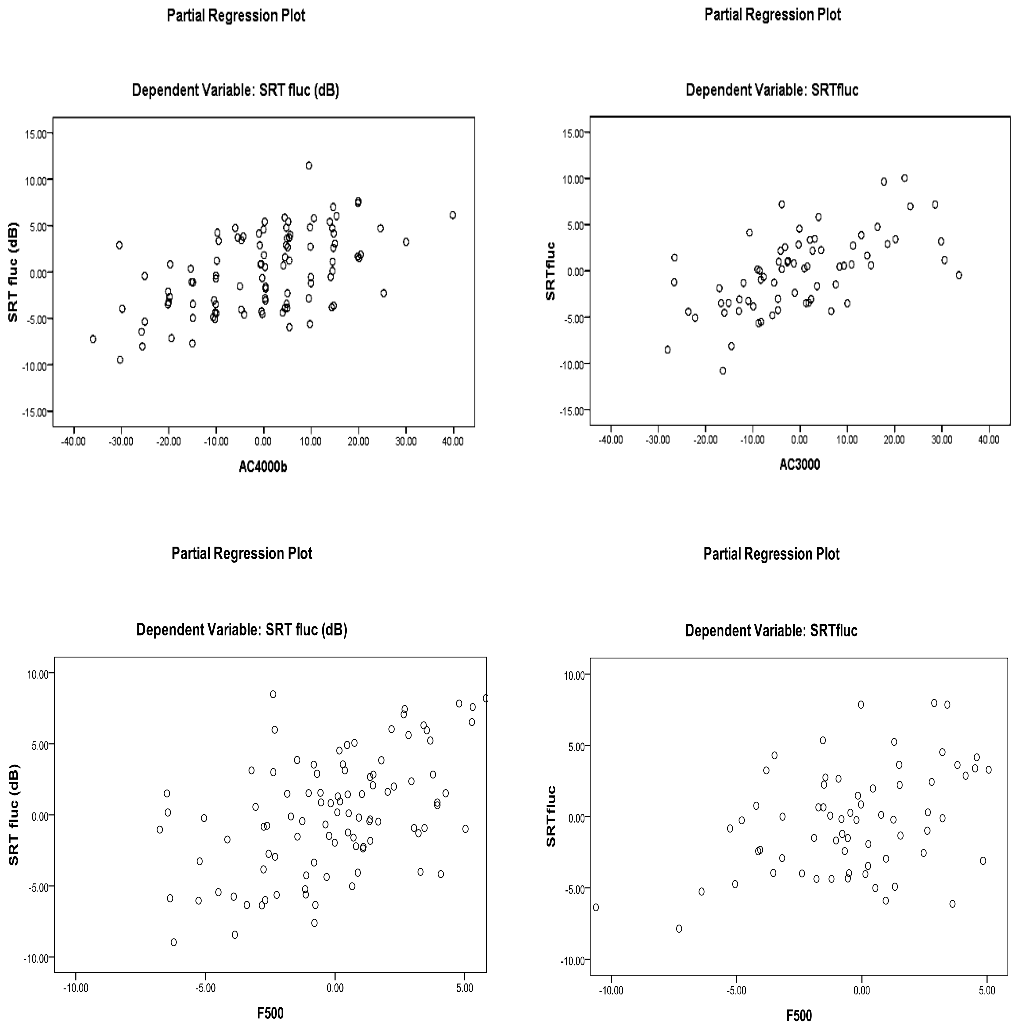


Figure 5.11: Partial regression plots for AC3000/AC4000Hz (dB) and F500Hz (dB) for study I and study II where n=72 (study I) and n=102 (study II)

From above it can be seen, the relation of SRTfluc to F500 reveals more spread and variation than AC3000/4000. Of course AC3000/4000 also reveals considerable variation but is less random and more correlations than that of F500. The area under ellipse serves to highlight this correlation. Thus speech recognition in noise is best predicted by threshold measurements and secondarily by frequency resolution. Thus the detailed investigation of the relation of three variables helps confirm our findings from the regression analysis for the two studies.

In the next figure (5.12), scatter plots for different groups of hearing loss for speech recognition in fluctuating noise across the two studies are given for mild (upper) and moderate hearing losses (lower). It can be observed that, the spread of scores for SRTfluc is more linear for threshold in mild group and more variation (and hence relatively more spread of scores) for frequency resolution and vice-versa for greater degree of hearing loss.

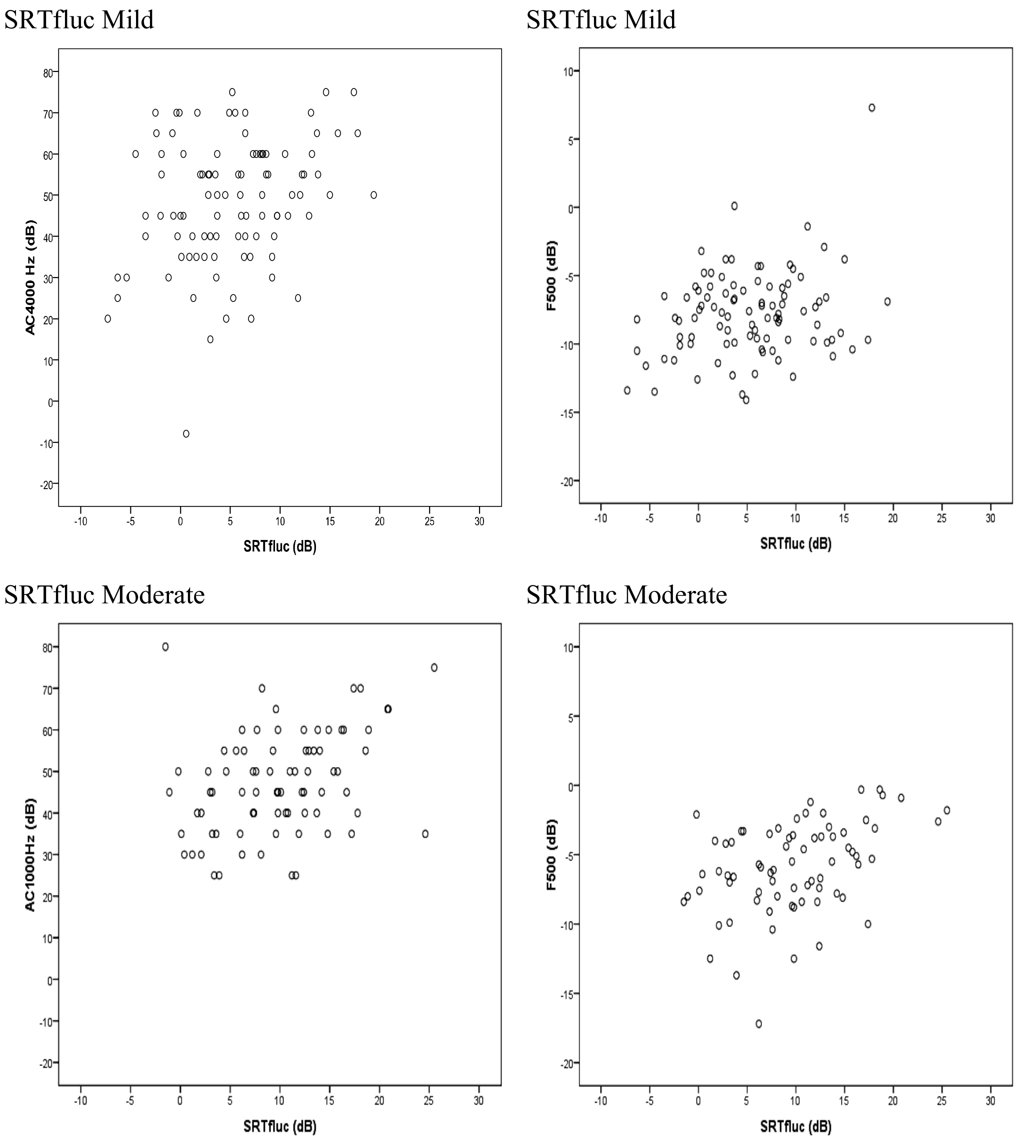


Figure 5.12: Scatter plots for different groups of hearing loss for speech recognition in fluctuating noise across the two studies for mild (upper); n=94 and moderate hearing losses (lower); n=81.

Finally, graphs are plotted for other measures like GP, ILD and BILD with the measure that explained the highest variation. These are all referred from results displayed in table 5.5.

Gothenburg Profile

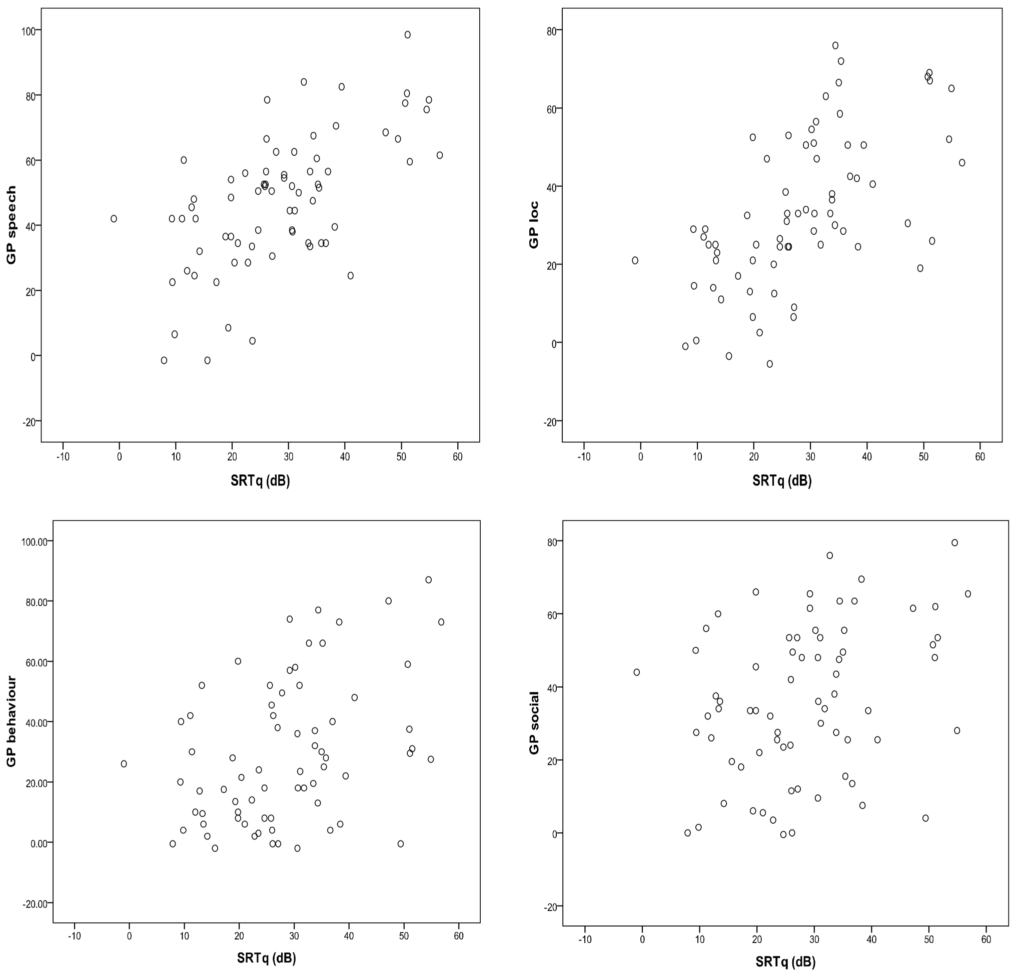


Figure 5.13: Scatter plot of SRT in quiet against the four subscales of Gothenburg Profile. where n=72

Speech recognition in quiet seemed to predict GP measures among all the other independent variables as seen from table 5.5. Hence its scatter against the four subscales of Gothenburg Profile are given above. It can be seen that the variation of speech and localization subscales is less than that of behaviour and social scale. This has been discussed in 5.5.8.

ILD

In general it can be seen from table 5.5 that measures of asymmetry (whether of speech recognition or threshold) predicted ILD . Their scatter plots are given below.

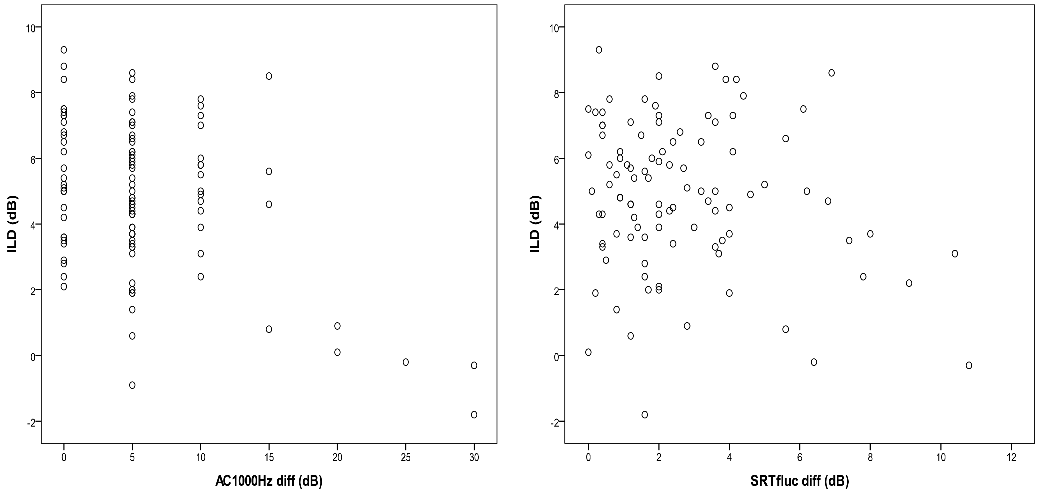


Figure 5.14 : Scatter plot of ILD and measures of asymmetry where n=102 (left panel) and n=72 (right panel)

As can be seen, though there is considerable spread , but beyond 15dB (left) a negative trend is seen wherein the binaural processing (ILD) decreases as the asymmetry increases. This can be understood since between the two ears the better ear will have better ILD scores than worse ear and the analysis includes only better data. Which means the the more better ILD on better ear, the less it is on the worse ear and hence the diff between the two(which is measure of asymmetry) is more.

BILD

Low frequency measure (AC500/MCL500) best predicted BILD as seen in table 5.5. This is displayed below.

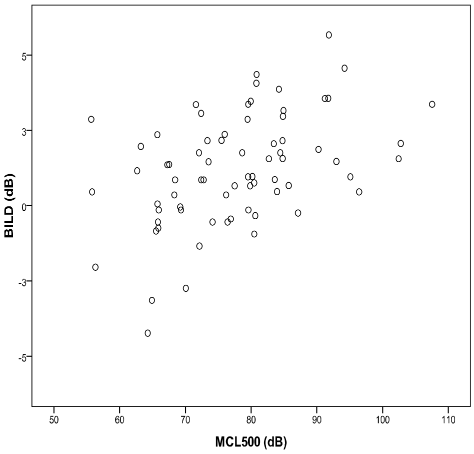


Figure 5.15: Scatter plot of BILD and MCL500 where n=72

As can be seen, though some variation is seen a linear positive trend is visible which means as the MCL measure increases (or becomes worse), BILD also increases and becomes positive (and hence worse).

Speech Intelligibility Index (SII) Analysis

In the sections above SRT was used to quantify subjects' ability to perceive speech in the presence of noise. Though this gave a very good estimate of the ability above, but to some extent it did not take into account the inter-subject audiogram and presentation level differences which were present in both studies. A predictor of speech intelligibility performance known as the Speech Intelligibility Index or SII (ANSI S3.5-1997) on the other hand is able to handle these differences since its model already accounts for audiogram and spectrum differences. This measure gives an estimate of the amount of speech information available in a certain condition using the individual's audiogram/threshold levels and signal and masker levels as inputs. The concept of SII and its relevance to SRT has already been discussed in first two chapters. Most studies have used either one or the other approach. Pavlovic (1984) has used the SII approach for his predictions whereas others have measured speech recognition. A third way is to incorporate both the approaches to observe if findings from each can help complement each other. A similar approach was used by George *et al.,* (2006) and will also be included in the present research. Their study investigated SII by analysing speech signals, using a modified version of model by Rhebergen & Versfeld (2005), as well as differences between stationary and fluctuating noise. Their calculations showed that the hearing impaired group revealed a larger (or exaggerated) SII than normal hearing subjects. They concluded that an SII value of about 0.3 indicates that the listener has no auditory deficits (as seen in normal hearing subjects), apart from possibly an elevated threshold, (as seen in their simulated hearing loss subjects) while an increased SII can serve an indication of a supra-threshold deficit (as seen in their hearing impaired subjects).

In the present study, the calculated SII is transformed into an SNR value. This SNR is referred to as SNR-SII and estimate of this ability can be directly compared to the SRTn namely, SRTstat and SRTfluc. It was hypothesized that if audibility is the only important property, SNR-SII will correlate well SRTstat and SRTfluc. The details of SII calculations are as follows:
SII was calculated as per ANSI (S3.5-1997) using the MATLAB application (available on SII website, http://www.sii.to/). The one-third octave method was used and speech and noise spectrum levels were calculated for 18 frequency bands using the thresholds and an index of band importance function.

The spectrum of the matrix test stationary noise was weighted according to the frequency response of the Sennheiser HDA-200 earphones on the IEC 318 coupler, to estimate the coupler levels and summated to get the total coupler SPL. Further the coupler-to-freefield correction to estimate levels in freefield at the position of the listener's head (as required for SII) were already incorporated in the test software during the SRTn measurement and hence were not applied. These band levels are converted to spectrum levels (dB/Hz) relative to the overall coupler SPL used for calibration. Noise spectrum levels were same as the speech. Band importance functions for general speech (ANSI S3.5-1997) were applied.

These speech and noise spectrum levels were then transformed to SNR values equivalent to the individual SII values. This was necessary since the SRTn was obtained using an adaptive procedure. For this transformation, the noise levels were kept constant and the speech levels were varied adaptively similar to SRTn measurements. The adaptive procedure varied between -20 to 20 dB SNR till the SII reached 0.2. The SNR equivalent to 0.2 SII was taken as the SNR-SII value required to achieve 50% score (similar to SRTn for stationary and fluctuating noise). According to ANSI S3.5-1969 a value of 0.2 corresponds to 50% SRT score for sentences when calculated with first presentation to listeners. And hence this was chosen for SII calculations since SRTn was also obtained using the first presentation to listeners.

Thus SRT-SII values were obtained for study II for left and right ears as for stationary and fluctuating noise as well as with and without including the SII distortion factor. Correlation between corresponding SRTn and SNR-SII across all measurements was found to be around 0.4, which was much lower than expected. Inclusion of SII distortion factor made essentially no difference, as it only affects presentation levels higher than used in present study. SNR-SII also did not predict variation in SRTn (in both stationary and fluctuating) when used as an independent variable in regression analysis.

Thus SII analysis proved to be less useful in determining the role of threshold in prediction of speech recognition. It was hypothesized that if SNR-SII will correlate well with SRTstat and SRTfluc, audibility can be considered to be exclusively responsible for variation in speech recognition in noise. The fact that it does not, seems to suggest otherwise. It can also be argued that factors other than threshold could be involved. This suggestion in a way complements our main findings which reveal that along with threshold, frequency resolution also helps predict speech recognition.

To a certain extent this could also be due to limitations of SII model which relies on the principle that speech components throughout the 30-dB dynamic range are equally important. However, loudness recruitment makes the ear highly linear. Also, for speech in noise, the speech peaks convey the important information and losing the lower intensities (especially at high frequencies) probably makes little difference. Finally, for high frequencies, the LTASS (Long-term Average Speech Spectrum) can be misleading - the average speech spectrum is quite low but it is composed mainly of short-duration high-intensity bursts of energy. A simple masking model like SII suggests these are inaudible when there is either masking noise or hearing loss at high frequencies, but in fact the high-intensity bursts are easily audible. For a more thorough analysis, the use of ESII by Rhebergen & Versfeld (2005) would add value. However, considering the scope of the current study, as well as the fact that SII analysis was used merely to complement the findings from regression analysis, which was the main focus, the use of E-SII was not considered obligatory. Having said that, the use of E-SII would be an interesting extension of the current study, and could be considered for future work.

5.12 Hypothesis/objectives/findings: Based on combined findings from the three analyses

Table 5.10: Summary of the hypotheses put forth in chapter I.

Hypothesis	Accepted/ rejected	Explanation
Auditory sensitivity alone cannot explain or predict the variation in speech recognition performance	Accepted	Speech recognition performance was predicted mainly by a combination of hearing threshold at high frequency and frequency resolution at low frequency
Besides audibility, certain measurable suprathreshold factors can explain the variation in speech recognition scores	Accepted	Besides audibility, variation in speech recognition was predicted by frequency resolution at 500 Hz mainly for fluctuating noise
Different factors will affect the speech intelligibility as the degree of hearing loss differs	Accepted	In mild hearing losses, speech intelligibility was predicted by hearing threshold followed by frequency resolution, in moderate hearing losses this was reversed where it was predicted mainly by frequency resolution.
The performance of hearing impaired people for stationary and fluctuating noise depends on different factors for the two types of noise.	Rejected	Speech recognition in the two types of noise was commonly predicted by hearing threshold level and frequency resolution. The additional variables such as temporal resolution and age did not count for any significant variation and were not consistent across studies.
Measures of hearing capability are not highly correlated and can be understood to be multidimensional	Accepted	Most of the measures of hearing capability included in test battery showed considerable variation. Further measures like speech recognition, hearing threshold, frequency and temporal resolution showed considerable interdependence while others like measures of localization, loudness perception, cognition were independent. Also the predictions of the different measures included auditory domains beyond hearing sensitivity measures.

5.13 Key questions investigated in the thesis

Lastly, the thesis aimed to answer two important questions mentioned in the first chapter:

- Which auditory factors are responsible for prediction of speech recognition in general across a range of hearing loss?

  Hearing threshold at high frequencies and frequency resolution at 500 Hz are the best predictors of speech recognition in general across a wide range of hearing loss

- Is the variation in auditory performance across a range of hearing impairment multidimensional or can it be approximated by a single unidimensional hearing loss construct?

  Variation in auditory performance across a range of hearing impairment is multidimensional and the measures mainly responsible for this variation include high frequency hearing threshold, speech recognition in stationary and fluctuating noise, frequency resolution at 500 Hz. Self-report measures (GP), lexical decision and binaural measures also add to characterize this multidimensionality further.

These are summarized in the following section

5.14 Outline of multiple dimensions of hearing loss

Study I (original and modified)

Test Domain	Hearing Sensitivity	Speech Perception	Auditory Resolution	Loudness Perception	Binaural Hearing	Localization	Listening Effort	Self Report of Hearing disability	Cognition
Type	Threshold	Supra-threshold	Supra-threshold	Supra-threshold	Supra-threshold	Supra-threshold	Supra-threshold	subjective	Cognitive /lexical access
Test variable	Hearing Threshold Level	Speech recognition in noise	Frequency & Temporal Resolution	Loudness Level Measure-ments	ILD BILD	MAA	Effort scaling	GP	Lexical decision

MAA & Listening effort did not show much correspondence with other measures, MCL BB not found to be useful measure due to change in experimental method. Hence all three were eliminated.

⬇

Study II

Test Domain	Hearing Sensitivity	Speech Perception	Auditory Resolution	Loudness Perception	Binaural Hearing	Self Report of Hearing disability	Cognition
Type	Threshold	Supra-threshold	Supra-threshold	Supra-threshold	Supra-threshold	subjective	Cognitive /lexical access
Test variable	Hearing Threshold Level	Speech recognition in noise	Frequency & Temporal Resolution	Loudness Level Measurements (except BB)	ILD BILD	GP	Lexical decision

Final outcome measures responsible for characterizing multidimensionality in hearing loss [based on results from studies I (original and modified) &II]

⬇

Test Domain	Hearing Sensitivity	Speech Perception	Auditory Resolution	Binaural Hearing	Self Report of Hearing disability	Cognition
Type	Threshold	Supra-threshold	Supra-threshold	Supra-threshold	subjective	Cognitive /lexical access
Test variable	Hearing Threshold Level	Speech recognition in noise	Frequency & Temporal Resolution	ILD BILD	GP	Lexical decision
Test measures/ Subscales	AC & BC 250-8000 Hz	Stationary fluctuating noise	F500	0,0 0,90r 0,90l	speech localization social behaviour	reaction time
Reason	(Emphasis On high frequency since evident in both regression & factor analysis in all studies)	(Main measure in concern)	Consistent predictor in all Studies)	(Not significantly of value on its own, but could be important when binaural issues are to be investigated, also highlighted importance of low frequency for localization)	(collectively, showed correlation with both SRT & HTL, helps to assess individual perception of hearing loss)	(Correlated With age In both studies)

In Summary, based on results of studies I and II, for two separate group of subjects with a wide range of hearing loss

⬇

Multidimensionality of hearing loss characterized by

⬇

High Frequency Hearing Sensitivity **Speech Recognition in Stationary &Fluctuating Noise**
Low Frequency Auditory Resolution **Binaural Hearing**
Self-report of hearing ability **Cognition**

Chapter Six

Summary and Discussion

6.1 General discussion and overall comparison with other studies

This thesis is mainly concerned with exploring interrelations among threshold and a range of suprathreshold auditory capabilities along with the possible interaction with subjective ratings. The above motivation was further strengthened by the fact that though several studies have attempted similar work in the past, the empirical evidence is nevertheless inconclusive in various respects. And these very respects served to form the key objectives discussed in the thesis. As recalled from the first two chapters the main shortcomings from the previous literature included:

Discrepancy regarding which auditory capabilities (threshold or suprathreshold) helps predict speech recognition in noise

Comment [M1]: Delete this word

Limited focus on multidimensionality of hearing loss leading to restriction of auditory domains in the test battery

6.1.1 Discrepancy regarding auditory capabilities (threshold or suprathreshold) influencing speech recognition in noise

The results from both multicentre studies in the present work showed high frequency hearing threshold (3000/4000 Hz) followed by frequency resolution at 500 Hz to be the best predictors of speech recognition in noise. Thus a combination of threshold and suprathreshold abilities was found to be responsible for variation of speech recognition in noise. Their relative importance in the different types of noise is discussed previously. It should however be noted that the tests carried out were at suprathreshold level as opposed to threshold which would mean the effects of audibility would be reduced as audibility of signal was ensured. This is an important consideration as it affects the interpretation of the main findings; the findings should be looked upon as correlations rather than causations. This is especially true for relation between hearing threshold level and speech recognition in this context because all the measures tested were at a level where any effects caused by lack of audibility were minimized or at least reduced. When this consideration is combined with the knowledge that hearing threshold levels are the best predictors of speech recognition performance, it must be recognised that the prediction does not necessarily imply direct causation. Reduced performance may be correlated with poorer hearing threshold levels, but it does not necessarily occur through lack of audibility of speech components. Presumably, reduced performance on the speech recognition task occurs because of a variety of supra-threshold deficits (Plomp's D parameter) that happen to be predicted well by hearing threshold level. In this case the main supra-threshold deficit was found to be frequency resolution. At the same time, it can be counter-argued that despite efforts to remove the effects of audibility in investigating the predictions of speech recognition in noise, the measures of audibility were nevertheless present. This implies that while studying the relations of speech recognition in noise, the role of threshold measures (whether causative or not) cannot be underestimated. Taken together, according to this study, when considering speech recognition in noise (whether for the purpose of clinical diagnosis or while fitting the hearing aid), it is hearing threshold level and frequency resolution among the huge array of auditory domains that are most crucial over others. Whether their relative influence varies perhaps depends on the individual subject and his hearing loss, but the present findings at least help place individual performances in a more specific framework especially in the context of such varied auditory capabilities.

The graphs below help illustrate this specificity by observing the scatter of scores for the three important measures in the discussion above, namely SRT (in fluctuating noise), HTL and F500. Their correlation (r) with each other was overall 0.4 (significant at 0.01 level). While in chapter V this relation was discussed more with high frequency context, here it is discussed more generally by using the average threshold level for a broader perspective.

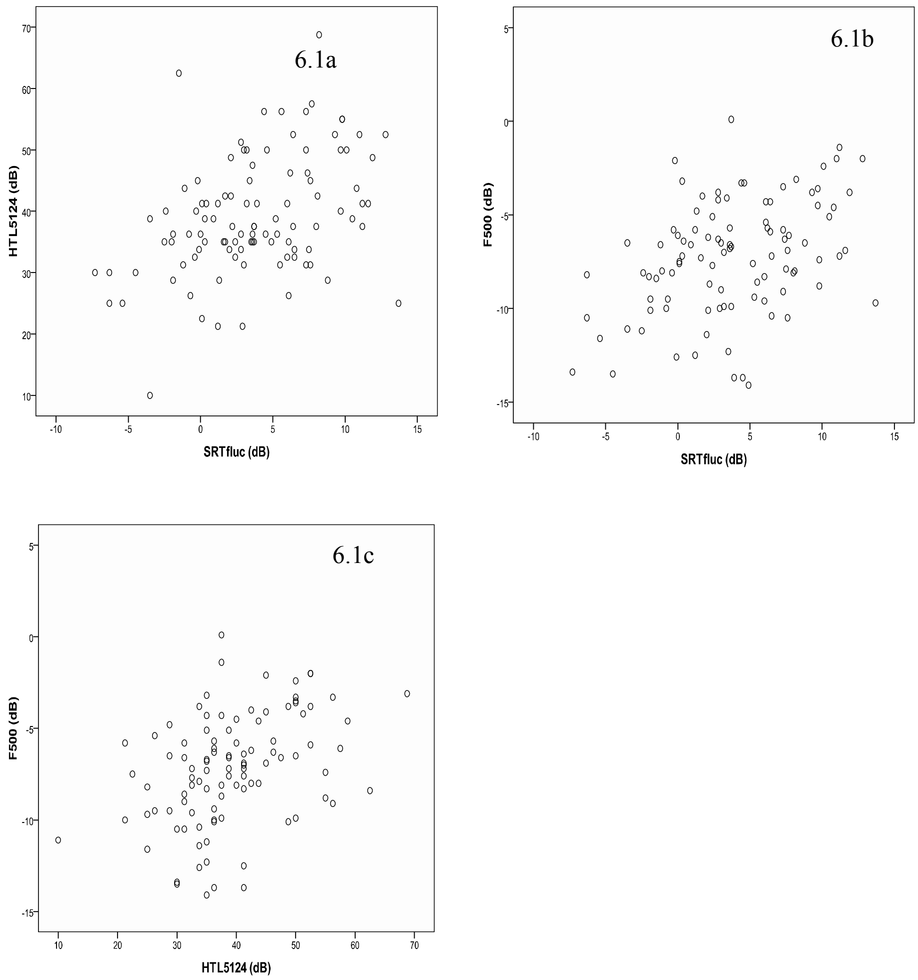


Figure 6.1: a) scatter plot of speech recognition in fluctuating noise (x-axis) and hearing threshold level (500, 1000, 2000, 4000 Hz) (y-axis); b) scatter plot of speech recognition in fluctuating noise (x-axis) and average frequency resolution at 500 Hz (y-axis), c) scatter plot of hearing threshold level (500, 1000, 2000, 4000 Hz), (x-axis) and average frequency resolution at 500 Hz (y-axis).

As can be seen in fig 6.1, there is considerable scatter across the three dimensions, but at the same time a linear positive trend is also observed which is relatively greater for HTL than F500. Such graphs thus reveal and help understand the relative influence discussed above and the trends seen across a wide range of hearing impaired subjects.

Further, the findings agree with most previous studies (Lutman, 1987; van Rooij and Plomp, 1990; Jerger *et al.,* 1991; Humes *et al.,* 1994; Divenyi *et al.,* 1997; George *et al.,* 2006) that have been discussed in the thesis (refer to Table 2.2). There is a general consensus among these studies that speech recognition in noise can be predicted fairly well by hearing threshold levels, particularly at 2000-4000 Hz. One reason why high frequency thresholds predict speech recognition performance may be simply the influence of audibility of the important cues at these frequencies which are needed to recognise speech in the presence of background noise. Further it is known that high frequency components give clarity to speech which gets compromised when there is interfering noise.

The studies differ when considering whether the prediction can be improved by considering alternative measures such as frequency (Festen and Plomp, 1983) and temporal resolution (George *et al.,* 2006) or both (Lutman, 1987). In fact only two studies (Dreschler and Plomp, 1985; Festen and Plomp, 1983) did not acknowledge the influence of threshold measurements on speech recognition. They found that both frequency and temporal resolution (Dreschler and Plomp, 1985) or just frequency resolution (Festen and Plomp, 1983) are responsible for the explained variation in speech recognition scores and are independent of audiogram measures. However, the study by Dreschler and Plomp (1985) included younger subjects aged 13-20 years which could be main reason for this difference. The present study too to some extent sought to demonstrate that variance in speech recognition in noise performance could be explained by measures other than the audiogram: the methodology aimed to enable the non-audiometric variables to show their importance (by measuring the performance at suprathreshold rather than threshold level). However, the fact that it revealed audiometric variables to be more important seemed to suggest the opposite. But at the same time, it must be remembered that frequency resolution was found to be predicting speech recognition secondary to hearing threshold. This has been demonstrated by regression analysis (chapter IV). More importantly partial correlations controlling for audiometric variables were studied in chapter IV (table 4.8) and chapter V (table 5.8); they showed significant correlations of the order of 0.2-0.4 between speech recognition in fluctuating noise and frequency resolution. Similarly residual plots in chapter V (figure 5.10) serve to support the same. And lastly the SII analysis did not prove particularly useful to show the exclusive role of hearing thresholds in predicting speech recognition in noise which seems to suggest involvement of other measures.

Hence both threshold and suprathreshold (frequency resolution) factors are important overall for speech recognition in noise. And the results in this study reveal their relative importance.

6.1.1.1 Comparison of other factors influencing speech recognition with previous literature

Besides hearing threshold level and frequency resolution, a few other factors discussed below influenced speech recognition to a lesser degree.

a) Temporal resolution: This was found to influence speech recognition in three studies (Lutman, 1987; Dreschler and Plomp, 1985; George *et al.,* 2006). The finding also existed in the present study, for fluctuating noise; however to a much lesser degree when compared to hearing threshold or frequency resolution and was not seen in study II. The reason behind the lower contributing value of temporal resolution is not completely clear. However in study I, a ceiling effect was found for this measure when its scores were considered against speech recognition, shown in the graph below.

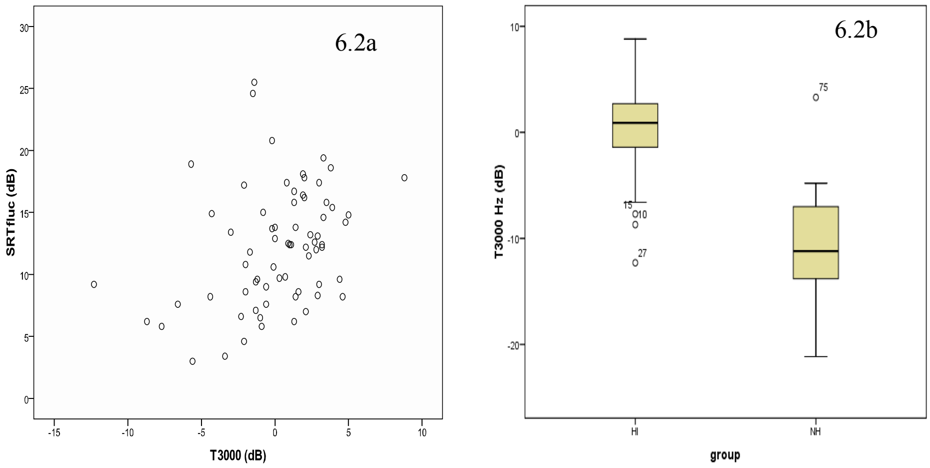


Figure 6.2: Scatter plot of SRT in fluctuating noise versus temporal resolution at 3000 Hz in hearing impaired listeners (a), ceiling effect in study I where they scored worse than expected (b).

As can be seen, the HI scored more badly than expected. This would mean that the test was perhaps difficult for the present set of listeners which in turn affected their correlation with SRT.

b) Cognition: A very minimal influence of cognitive abilities on speech recognition has been shown by studies (van Rooij and Plomp, 1990; Jerger *et al.,* 1991; Humes *et al.,* 1994). In fact Akeroyd (2008) summarised the results from 20 studies observing cognition and speech recognition in noise. Overall results demonstrated that no one cognitive test always gave a significant result, with some measures (e.g. working memory) being more effective than others (IQ). He concluded that there is a link, but it is secondary to the predictive effects of hearing loss, and it is somewhat mixed across studies. In the present study two cognitive measures when combined (response time and percentage of correct responses) as reflected by lexical decision test did not show any significant influence on other test measures. One of the main reasons for this could be the type of test chosen. It was too short and perhaps did not place the required additional cognitive load. However, as mentioned in Chapter III, the test of lexical decision was chosen from a battery of tests as part of the project based on pilot experiments, mainly based on the clinical feasibility and time efficiency. It was further administered at the end to ensure that the subjects showed some fatigue; auditory or otherwise. It also served the purpose of ruling out the presence of any impairment other than hearing loss. But it should be remembered that the results of cognitive tests depend on the type of subjects to a great extent. It is highly unlikely for subjects in the study who are physically mobile and otherwise active and available for prolonged psycho-acoustical testing will reveal any material cognitive dysfunction. Nevertheless, in both studies cognitive measures were inversely related to age, so perhaps there is some slowing down with age, but this may not necessarily be reflected on their hearing abilities.

c) Age: Age as a contributing factor for speech recognition in noise has been revealed by various studies (Divenyi *et al.,* 1997a; George *et al.,* 2006), while some (van Rooij and Plomp, 1990; Lutman, 1987, Jerger *et al.,* 1991;) reveal no effect. In the present research, only study I showed a minimal contribution of age for speech recognition in fluctuating noise while study II did not (refer to Table 4.25). Age and hearing loss are intrinsically related and it is very difficult to separate the effects of one from other (Divenyi *et al.,* 1997a). This could be one of the main reasons for age to not converge as a significant unique factor in the present study. It nevertheless correlated with cognition as mentioned above as also seen in the study by van Rooij and Plomp (1990) and its importance for influencing both hearing sensitivity and speech perception is well established.

d) Frequency of stimuli

Frequency of the test stimuli included to study the auditory domains is as important as the domains themselves since they reveal more specific information about the various deficits. These gain more importance while studying the older hearing impaired subjects since their typical audiograms reveal frequency specific loss. While the present study shows agreement with most other studies (Lutman, 1987; van Rooij and Plomp, 1990; Jerger *et al.,* 1991; Humes *et al.*1994) in terms of high frequency hearing loss being a major contributing factor for speech recognition, it differs from others in terms of the low frequency measure of frequency resolution found to be an additional factor. Most other studies (Festen and Plomp, 1983; Lutman 1987; Dreschler and Plomp, 1985) that found frequency resolution as a contributing factor have revealed so at high frequency. In this sense, the finding from the present study is a novel one. Reasons why the low frequency measure (500 Hz) of frequency resolution has been shown in both the studies to be a contributing factor for speech recognition as opposed to the high frequency measure are unclear. It could also be related to distinguishing vowel contrasts or semi-vowels that have cues at low frequencies. The information that humans require to distinguish between vowels can be represented purely quantitatively by the frequency content of the vowel sounds called formants. A formant is a concentration of acoustic energy around a particular frequency in the speech wave. (Wood, 2005). And it is known that vowels have formants which concentrate on low frequencies and consonants have energy at high frequencies. Benade (1976) suggests the following ranges of frequencies for the formants of a male voice: 1st formant 150-850 Hz, 2nd formant 500-2500 Hz, 3rd formant 1500-3500 Hz, 4th formant 2500-4800 Hz. Most often the two first formants, f_1 and f_2, are enough to distinguish between the vowel (Ladefoged, 2001) and as can be seen they both have frequencies concentrated at low regions.

6.1.2 Limited focus on multidimensionality of hearing loss leading to restriction of auditory domains in the test battery

One of the features of the present study is the structure of the test battery in terms of exploring all the auditory and non-auditory domains related to hearing loss while limiting redundancy. As with most other studies, hearing sensitivity, speech perception, frequency resolution, temporal resolution and cognition were included. However, certain additional measures which were not included by most other studies were outlined in the first chapter. These included loudness perception, localization, binaural hearing, subjective assessment.

a) Loudness perception: The loudness levels obtained in the present study were used to calculate the most comfortable level which was used as the reference starting level for each subject in study I. Further, measures of this at both high and low frequency provided frequency specific information and the slopes obtained were used as a measure of recruitment. In study I, the measures were distributed in factor analysis while in study II, they stood independently. Also in study I, slopes of the measures grouped separately as a recruitment factor. Thus overall loudness perception correlated with other measures as well as showed independent existence. This difference between study I and study II may be also due to difference in setting stimulus levels, allowing ACALOS to be somewhat independent in study II.

b) Localization: Minimal audible angle measures showed minimal correlations with other measures and stood independently in the present study. Difficulties of localization are known to be present in people with hearing impairment. The fact that they are not correlated with other measures like speech recognition or hearing threshold in the study could mean that they may not necessarily be influenced by other auditory capabilities. Also it did not correlate with self-report measures in Gothenburg Profile.

c) Binaural Hearing: Taken together from both the studies ILD was influenced by measures of asymmetry (whether hearing thresholds or SRT) while BILD was influenced by low frequency measures indicating the importance of low frequency hearing while using binaural cues.

d) Subjective measurements: Two subjective assessments were included in the test battery: Listening effort and Gothenburg Profile. Measures of Listening effort could be predicted to some extent by speech recognition in fluctuating noise which implies that some hearing impaired subjects require more listening effort in fluctuating background than stationary. However, overall they showed limited correlations with other measures. One of the main reasons for this could be that though an increased listening effort in noisy conditions is known to be present in hearing impaired listeners, it may not be directly related to their speech recognition performance or hearing loss. This means, the effort required is differential and there is a lot of variance in the majority of subjects. Furthermore, such effort is known to be affected by other psycho-social factors like alertness, individual susceptibility to noise etc. The Gothenburg Profile on the other hand overall showed the influence of both speech recognition and hearing threshold. This is an important inference since it relates the objective-subjective aspects in the test battery; not present in any of the previous studies discussed here. The prediction of variation in different subscales ranged from 18-62 % (ref table:5.5) which meant there was considerable variation and hence could be chance findings in some cases. However, it is still worthwhile to note that the GP subscales showed some correlation with the two most important auditory capabilities in the study which meant that subjective assessments are possibly to some extent based on the objective performance or vice-versa, an implication that would be of significant value for rehabilitation of the clinical population.

e) Multidimensionality: From above it can be observed that some of the additional measures in the test battery interrelated with others, while some did not. Thus some of these aspects showed influence on others while some stood independently. Overall though a certain level of correlation was evident, so was the considerable variation. However, each of them contributed some unique information about a different aspect of hearing loss. All these facts serve to highlight the multidimensional nature of hearing loss as a sensory entity. Though this is an obvious implication, empirical evidence focussing on this aspect has been very limited in the past. Studies in the past have no doubt implied this before but not many have made efforts to systematically observe and verify the actual auditory and non-auditory capabilities that outline this multidimensionality. Methods applied in the present study have focussed on investigating all the different auditory capabilities in order to get a comprehensive, yet specific view of the clinical manifestations associated with the peripheral auditory system. This is unlike others in the literature which seemed to have included only one or few measures besides speech recognition and hearing threshold; often leading to an insufficient depiction.

Further the fact that only a limited proportion (approx 60%) of variation is captured in predictions of speech recognition scores helps to support the aspect of multidimensionality further. The remaining 40% is still unexplained. Thus even if measurement uncertainty/error is considered, it still indicates considerable involvement of other aspects. Factor analysis in both studies revealed several factor loadings as opposed to just one, and they were predominated by different auditory measures ranging from high frequency processing to recruitment. All these facts further serve to illustrate this multidimensionality.

6.2 Influence of signal levels used for measurements

Three important issues concerning the signal levels are discussed here.

6.2.1 Adaptive or fixed level of signal presentation: This has been discussed above. The adaptive method allows one to test subjects with greater degree of hearing loss without reaching intolerable levels.

6.2.2 The influence of signal levels on statistics or results: This too has been discussed above. In study I, due to use of MCL as the reference signal level, an intrinsic correlation between the MCL measures and others was evident. To reduce this, a slightly altered method was devised which made use of the 1/3-gain formula based on threshold levels described previously. Both these methods ensure a good balance between excessive sound levels and being sufficiently above threshold to assess any suprathreshold difficulties. It should be remembered that including methods that measure subjects in a uniform way supersedes any other experimental criteria. Further, the influence of measurement levels can be avoided only by use of fixed presentation levels (Lutman, 1987; Divenyi et al, 1997) unlike here, where adaptive levels were used which were essential for suprathreshold comparisons (that would make the signal audible for a wide range of hearing impaired subjects) as desired in the present study. In fact with use of both the methods, there was little difference in the key findings. Only a small amount of variation with some measures was observed.

6.2.3 Frequency resolution and level dependence: The issue of frequency resolution and its possibility of being influenced by signal levels has been discussed before in chapter II. According to Festen and Plomp (1983) a precautionary measure to reduce this influence is to perform measurements in a narrow range of sound-pressure levels of the maskers. A similar approach was also used in the present research, wherein the subsequent signal levels chosen were within the narrow range of the thresholds in both the studies. Thus in study I, this consisted of MCL which is usually 20 dB above the threshold while in study II this was via 1/3-gain rule. In either of the cases, the levels were high enough to monitor any suprathreshold deficits such as frequency resolution. However, many studies in the past have shown that auditory filters broaden with increased levels. However, Rosen *et al.* (1998) revealed that models with filter parameters depending on probe level fit the data much better than masker-dependent models. Thus when the probe level is varied (as versus the masker level), auditory filter shapes are less prone to changes due to level differences. This was also applied in the present study where the probe level (a pulsed tone of 500/3000 Hz) was varied and the masker level (octave band noise) was fixed. Also, Wightman and Raz (1980) found a difference of only 2 dB between high and low signal levels for their thresholds measuring frequency selectivity and hence they concluded that deficits in selectivity exhibited by hearing-impaired listeners is not an artefact due to stimulus level. Of course, a confounding effect of this dual influence may have occurred in some subjects with higher degree of hearing loss, but there were few subjects with an average hearing loss exceeding 65 dB compared to the total sample of 72 in study I and 102 in study II. Additionally, the levels rarely reached very high or intolerable levels due to adaptive measurements.

6.3 Influence of statistical methods

The statistical methods used in the present studies, multiple linear regression and factor analysis, are primarily descriptive and exploratory. Predominantly, they explore covariance in the available data set and are also dependent on idiosyncrasies present in the pool of participants who took part in the study. Due to these limitations a certain level of caution is required in interpreting the various findings as well as generalising the findings to clinical populations. To some extent, this can be taken care of by including a large number of subjects. Taken together from both studies, the number of hearing impaired subjects was 173 in the present research, which is quite large by conventional standards in the field. However, the number of measures included were also quite large. In such cases, it is more appropriate to place greater importance on variables that explain substantial percentages than variables that explain only small incremental percentages of variance. The interpretation of results in the present study has adhered to this. Further, stepwise regression is susceptible to strong correlations between two variables. Due to this it may reveal only one predictor variable which has the highest correlation with the dependent variable or include another which has an intrinsic correlation with predictor variable and not the dependant variable. But, if the aim is to derive predictive estimation of one variable based on group of variables like in the present study, stepwise regression is the most commonly used procedure.

6.4 Other key findings

6.4.1 Right-left symmetry

Right ear advantage for various auditory tasks has been established by a number of studies before. This advantage was seen for ILD. These effects might be related to the speech processing in the left hemisphere, or to other right-ear advantages. The findings serve to support other studies (e.g. Divenyi *et al.,* 1997c) in the literature, however the effect is small and is not of clinical significance for the present studies.

6.4.2 Fluctuating noise

As mentioned before, predictions of speech recognition scores for this type of noise have been limited with perhaps only one study (George *et al.,* 2006) covering the aspect in detail. From both the studies, it was seen that speech recognition in fluctuating noise was predicted mainly by hearing threshold followed by frequency resolution as well as temporal resolution and age to a lesser extent. Hearing threshold, temporal resolution and age were also found as factors by George *et al.,* (2006), however, the explained variances differed. In fact their study did not find frequency resolution as a predicting factor. As mentioned before, it could also simply be related to differences in type of tests, subjects etc or the fact that the test at 3000 Hz in the present study was difficult for many participants and hence did not surface. Nevertheless, these findings related to fluctuating noise are important since fluctuating backgrounds are common in everyday life and the fact that the findings are relatively novel compared to predictions in stationary noise which has been studied by the majority of studies. It should be stressed that the majority of previous research focussing on speech in stationary noise is therefore limited. Conclusions based solely on such research may not be generalisable to everyday situations where noise is not stationary.

6.4.3 Groups of hearing loss

Collectively, from both studies auditory sensitivity measures explained greater variance in mild hearing loss group and auditory resolution measures in greater degree of hearing loss. Thus, a common trend across groups, types of noise and measures reflected that the variance explained by auditory resolution (in this case frequency) increases and that explained by audiogram measures decreases as the degree of hearing loss increases. Its clinical and physiological implications have been discussed previously. These findings differ from other studies like Lutman (1987) which found auditory resolution measures to be influencing speech recognition measures for mild group while auditory sensitivity was found to be influential for moderate groups. The main reason for this difference could be because the present study shows comparisons of two types of noise as versus one. Division of groups according to hearing loss was also different in both the studies. However, Lutman's study includes predictions for overlapping groups such as normal + mild and mild+ moderate unlike the present, where there is no overlap. More importantly, Lutman study used fixed presentation levels while the present one used adaptive. However, the findings from the present study are consistent with the interpretation of Pavlovic (1984). He applied the AI (articulation index) procedure to audiograms of normal and HI subjects to predict SRT in noise. He concluded that suprathreshold distortion factors were present in addition to loss of audibility, for the more severe losses. In other words, supra-threshold distortions were less important for mild/moderate hearing loss as compared to more severe losses, which was also seen in the present study.

6.5 Clinical implications

The tests used in the present study can be directly applied clinically as a measurement tool in order to observe the performance of hearing impaired population on various auditory capabilities. Their value is limited to a certain extent by the finding that clinically essential measures can be predicted by hearing threshold or speech recognition with an addition of frequency resolution. It nevertheless may prove useful to obtain an individual auditory profile that can be used clinically to monitor the diversity of results seen in the people with hearing loss. In fact, based on the results of the two multicentre studies, the tests in the profile could be divided into the following three phases. These phases could be used to place a particular patient in a broad framework by observing his/her performance of speech recognition and hearing threshold as well as investigate a specific auditory capability such as frequency resolution. Again, based on the several predictions in the two studies, they could be used as a hierarchical protocol for clinical diagnosis.

Phase I (Basic):

Auditory capability	**Actual test/measure**
Hearing Threshold	Pure tone audiogram 250-8000 Hz
Speech Recognition in noise	SRT in stationary and fluctuating noise
Subjective	Gothenburg Profile

Hearing threshold and speech recognition are the two most common, frequent and essential measures used clinically. Again as per the predictions from the present studies, SRT can be predicted for most part based on the individual's HTL to a great extent. Also subjective measures such as questionnaires are usually available in individual clinics and can be used to provide additional information of the perception of hearing loss by the patient as well for hearing aid intervention. They would thus be in phase I where all subjects would require to be tested.

Alternatively, it can also be deduced that for the purpose of clinical assessment and hence intervention, the performance of many patients can be sufficiently predicted using the phase I measures. Thus for most patients with predominantly age-related hearing loss whose performance more or less 'fits' within these measures, a straightforward intervention process can be initiated. However, due to considerable diversity of the patients evident in clinics, there would be quite a few whose performance cannot be predicted using only these three. These could be patients with additional problems like recruitment or asymmetrical loss or simply with performance scores that could be affected due to an additional deficient measure. In the present study, frequency resolution was found to be predicting performance scores following hearing threshold and hence is included in the advanced phase. Temporal resolution is also found to be responsible for predicting the scores by some studies and hence would be helpful here. Similarly binaural hearing and loudness level measurements serve to provide additional information as mentioned above and hence these too are included in the advanced phase.

Phase II (Advanced):

Auditory capability	**Actual test/measure**
Frequency Resolution	Frequency Resolution at 500 Hz
Temporal Resolution	Temporal Resolution at 3000 Hz
Binaural hearing	ILD/BILD
Loudness level measurements	MCL and loudness slopes at 500 and 3000 Hz

Thus in general, phase II would be warranted if more specific information about the subjects hearing loss is required as well as if sufficient information regarding the subject is not obtained. Further as mentioned above, if the performance cannot be predicted by first three measures, these additional measures (frequency and temporal resolution) can serve as basis to provide more insight as to perhaps explaining the diversity or variation seen in clinical population. Though the measures in phase II will not directly impact the treatment plan, information obtained from them can aid in aspects like counselling or hearing aid adjustments etc. Also as stated above, in cases where the loss is asymmetrical, the binaural measures would provide valuable information. Similarly, loudness measures can be useful for dynamic range, or where there is a possibility of recruitment etc.

Phase III (Detailed):

Auditory capability	**Actual test/measure**
Cognition	Lexical Decision

A majority of patients would be more or less assessed in phase II. However, in a small number of patients, especially in elderly it may be needed to ascertain their cognitive-sensory functions. The lexical decision test can be used here. It can also be used in second language users or even in others as a 'ruling out' test whereby involvement of any (cognitive) sensory difficulties other than hearing can be estimated.

The phases can serve as guidance for clinical protocol. The main suggestion is that through the findings from the present study, there is availability of such a clinical tool and that each of the tests in the auditory profile help characterize the individual. In fact, the tests in each phase can be used individually or separately as well as with inclusion/exclusion of phases dependent on the performance of a particular patient, his/her needs. Each patient is unique and should be treated accordingly, while the above tests help establish an orderly pattern of achieving the same.

The diagrammatic representation of the above protocol is as follows:

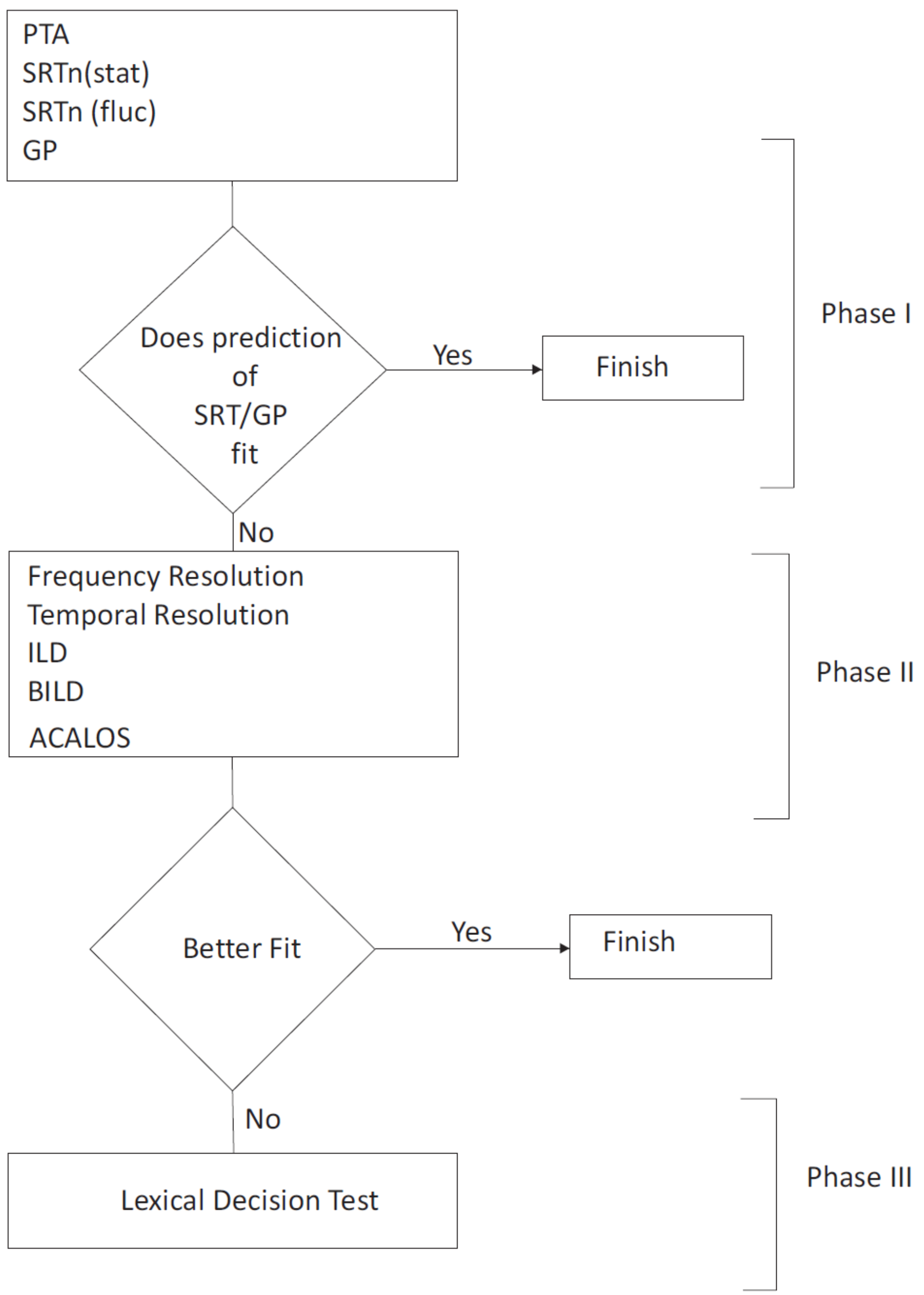


Figure 6.3: Clinical protocol based on findings of study I and II.

Chapter Seven

Conclusions and Contribution

7.1 Conclusions

Broadly, speech recognition performance in noise can be predicted mainly by hearing threshold level and frequency resolution. Specifically the high frequency hearing threshold (3000Hz/4000Hz) which has been used in many other studies to summarise hearing impairment, is a good predictor. Prediction of speech recognition performance in noise is improved by the addition of a measure of spectral resolution at low frequency (e.g. 500 Hz), especially for speech in fluctuating noise. Speech Intelligibility Index proved to be of limited value in depicting the exclusive role of thresholds in prediction of speech recognition in noise.

Frequency resolution appears to be the ability influencing the 'D' parameter which is consistent with Plomp's speech reception model (1978) and high frequency hearing threshold for 'A' parameter. However, prediction from thresholds does not necessarily mean that the relationship arises from lack of audibility because Plomp's D parameter is statistically correlated with threshold elevation. The prediction by a frequency resolution term confirms the importance of Plomp's 'D' parameter. These findings are important from research perspective since they contribute to understanding of the cocktail party effect which is concerned with suprathreshold deficits as opposed to threshold levels.

Relationships described above vary according to the degree of hearing loss. For mild hearing losses threshold related factors are important while for severe losses suprathreshold factors (frequency resolution) gain more importance.

Thus it can be concluded that for clinical purpose, measurement of hearing threshold, SRT in fluctuating noise and frequency resolution at 500 Hz are sufficient to broadly characterize an individual's peripheral auditory status.

However, if the purpose is to obtain a comprehensive auditory profile binaural measures, cognitive measures and self-report measures should be included

Tests like the ones used here were able to differentiate between the normal hearing and hearing impaired participants as well as between different degrees of hearing loss and hence were found to be sensitive and specific. Most of them were devised exclusively for the present studies. They revealed good test-retest repeatability as well as giving consistent results for the two studies while investigating their correlations. Thus it can be concluded that the tests used in the present studies can be used together as a tool for assessing peripheral auditory system for both clinical and research purposes.

Hearing loss is multidimensional. This is an obvious but important conclusion. This multidimensionality is characterized by measures of hearing threshold, speech recognition, auditory resolution (frequency and temporal) and cognition. Additionally self-report measures helped bridge the gap between subjective-objective measurements while binaural measures highlighted the importance of low frequency cues when listening to speech in the presence of interference in a binaural environment.

For future research it is suggested that studies should focus on these multiple dimensions as well as hearing loss with differing severities and incorporate measuring all the different auditory capabilities rather than inclusion of few redundant ones. It is required to go beyond the relation of speech recognition and threshold/suprathreshold measures in order to get a complete depiction of the peripheral auditory system. This is important even when most of them are correlated or could be predicted by audiogram measures since quite a few stood independently which implies that there is considerable individual variability.

7.2 Contribution and novel findings

Auditory frequency resolution measured at low frequency is a significant predictor of speech recognition in both stationary and fluctuating noise. This is a new finding, not revealed in any study before. Most studies (Festen and Plomp, 1983; Dreschler and Plomp, 1985; Lutman, 1987; Philips *et al.,* 2000) which established the relation between the two capabilities revealed a high frequency measure.

For clinical purposes, in order to predict speech recognition, high frequency hearing threshold is sufficient with an added measure of frequency resolution. The finding that high frequency hearing threshold is a predictor of variation in speech recognition has been highlighted by studies before. However, the interesting and novel part about this whole approach was that in spite of all attempts to prove otherwise, it was consistently seen in two studies that hearing threshold is ultimately responsible for variation in speech recognition in both the types of noise. In order to reduce the effects of audibility, all the tests were carried out at suprathreshold level; inclusion of a wide range of test battery was ensured so that interrelations among various measures other than just audibility were explored.

The two studies have overcome substantial methodological limitations in previous studies. The key advances include: large samples of participants (231 including both studies), measurement of auditory capabilities at both low and high frequencies, inclusion of diverse auditory and non-auditory domains for testing, measurements in both right and left ears independently, inclusion of a subjects with a wide range of hearing loss and age and ensuring audibility of speech across the range of hearing impairment.

The present study has also demonstrated multidimensional nature of hearing and identified factors characterizing the same.

The present study extended existing knowledge by also addressing issues of binaural hearing. The results emphasise the importance of low frequency hearing, including low frequency spectral resolution, on binaural speech recognition in noise. The other (rather obvious) finding is that binaural performance decreases with increasing ear asymmetry.

The present study also extended existing knowledge by examining predictors of self-reported hearing abilities. The results indicated that they are influenced mainly by both hearing threshold and speech recognition in noise, although there was considerable unexplained variation in the subscales. This nevertheless attempts to quantify the subjective-objective relationship, rarely attempted before.

The present study also observed the performance of subjects as a function of degree of hearing loss, type of noise and type of auditory measures using adaptive method of measurement. Such diverse and collective analysis in one study was rarely evident in the literature. In doing so it was revealed that as the degree of hearing loss increases speech recognition is increasingly influenced by frequency resolution and decreasingly by hearing threshold. Such a trend is also new though few studies like Lutman (1987) and Pavlovic (1983) have attempted similar investigation.

Overall, the research methods used in the present study contribute towards:

- Understanding of the cocktail party effect by outlining the factors responsible for the deficient speech recognition.
- Outlining the multiple dimensions and the actual measures that characterize hearing loss.

References

Akeroyd M.A, (2008), Are individual differences in speech reception related to individual differences incognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int J Audiol.* 47 (Suppl. 2):53-71

Albeck Y, Nebenzahl I & Lewis A, (1992), Temporal structure model of binaural masking level difference. *J. Acoust. Soc. Am.*92:1389-93

Allen J.B, Hall J.L and Jeng P.S, (1990), Loudness growth in ½ octave bands (LGOB)- A procedure for assessment of loudness .*J. Acoust Soc Am.*88;745-53

Allen P, (2000), Audiology Diagnosis. In Acoustics and Psychoacoustics by Roeser *et al.,* 2000 pp 169. Thieme medical publishers Inc., New York

Andersson U, (2001), Cognitive deafness. Department of Behavioural Science, Linköping University, Sweden.

Andersson U, Lyxell B, (1998), Phonological deterioration in adults with an acquired severe hearing impairment. *Scand Audiol Suppl* 49:93-100.

ANSI S3.5 (1997), "American National Standard Methods for the Calculation of the Speech Intelligibility Index," American National Standards Institute, New York.

Arlinger S, Billermark E, Oberg M, Lunner T, Hellgren J, (1998), Clinical trial of a digital hearing aid. *Scand Audiol* 27:51-61.

Bacon S.P, Opie J.M, Montoya D.Y, (1998), The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. *J Speech Lang Hear Res* 41:549-563.

Benade, Arthur H . (1976) Fundamentals of Musical Acoustics, Oxford University Press

Bench J, Kowal A, Bamford J, (1979), The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *Br J Audiol* .13:108-112.

Bertoli S, Probst R, Jordan P, (1996), Hearing handicap--an addition to audiometric hearing loss. Results of an exploratory study of auditory communication disorders in the elderly. 44(7):376-84

Bernstein JGW (2010) in 'Why do hearing impaired subjects fail to benefit from masker fluctuations' in Neurophysiological bases of auditory perception. Springer.

Besing J.M, Koehnke J, (1995), A test of virtual auditory localization. *Ear Hear* 16:220-229.

Blauert J, (1997), *Spatial hearing: the psychophysics of human sound localization*. MIT Press, Cambridge.

Blom G, (1958), *Statistical Elements and Transformed Beta Variables.* Wiley, New York.

Bowles N.L, Poon L.W, (1981), "The effect of age on speed of lexical access." Exp Aging Res. 7, 417-425.

Brand T and Kollmeier B, (2002), Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests. *J. Acoust. Soc. Am.* 111(6):2801-10

Brand T and Hohmann V, (2002), An adaptive procedure for categorical loudness scaling. *J. Acoust. Soc. Am.* 112 (4); 1597-1604

Bronkhorst A.W, and Plomp R, (1992), Effect of multiple speech like maskers on binaural speech recognition in normal and impaired hearing. *J Acoust Soc Am.*92:3132-39

Bronkhorst A.W. (2000), The Cocktail Party Phenomenon: A Review of Research on Speech Intelligibility in Multiple-Talker Conditions. *Acta Acoustica* 86:117-128

British Society of Audiology, (2004), Recommended Procedure for pure tone air and bone conduction threshold audiometry with and without the use of masking and determination of uncomfortable loudness levels. www.thebsa.org.uk

Byrne D *et al.,* (1996), : http://www.icra.nu/Prod_Noise.html

Cox R.M, Alexander G.C, Taylor I.M and Gray G.A, (1997), The contour test of loudness perception. *Ear Hear.* 18 (5); 389-400

Cameron S, Dillion H, Newall P, (2006). The Listening in Spatialized Noise test: Normative data for children. *Int J Audiol* ; 45:99-108

Davis A.C, (1989), The Prevalence of Hearing Impairment and Reported Hearing disability among Adults in Great Britain. *International Journal of Epidemiology.*18 (4) :911-17

Dirks D.D, Morgan D.E and Dubno J.R, (1982), A procedure for quantifying the effects of noise on speech recognition. *J. Speech Hear Disord.* 47: 114–123.

Divenyi P.L and Haupt K.M, (1997b), Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. II. Correlation Analysis. *Ear Hear*. 18:100–113

Divenyi P.L and Haupt K.M, (1997c), Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. III. Factor representation. *Ear Hear.* 18: 189–201.

Divenyi P.L and Haupt K.M, (1997a), Audiological correlates of speech understanding under non-optimal conditions in elderly listeners with mild-to-moderate hearing loss. I. Age and laterality effects. Ear and Hearing, 18, 42-61

Downs D.W, (1982), Effects of hearing aid use on speech discrimination and listening effort. *J Speech Hear Disor*. 47;189-193

Dreschler W.A and Plomp R, (1980), Relations between psychophysical data and speech perception for hearing-impaired subjects I *J. Acoust. Soc. Am* 68(6):1608-15

Dreschler W.A and Plomp R, (1985), Relations between psychophysical data and speech perception for hearing-impaired subjects IIa *J. Acoust. Soc. Am* 78(4):1261-70

Dreschler W.A, Verschuure H, Ludvigsen C, Westermann (2001), ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. International Collegium for Rehabilitative Audiology. *Audiology* 40 (3), 148-157.

Dubno J.R and Schaefer A.B, (1992), Comparison of frequency selectivity and consonant recognition among hearing-impaired and masked normal-hearing listeners *J. Acoust. Soc. Am* 91, 2110-2121.

Dubno J.R, Horwitz A.R, Ahlstrom J.B, (2002), Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing. *J Acoust Soc Am* 111:2897-2907

Duquesnoy A.J & Plomp R, (1980), Effect of reverberation and noise on the intelligibility of sentences in cases of presbyacusis. *J. Acoust. Soc. Am* 68, 537-544.

Duquesnoy A.J, (1983), The intelligibility of sentences in quiet and in noise in aged listeners. *J. Acoust. Soc. Am* .74, 1136-1144

Eddins D.A, (1999), Amplitude-modulation detection at low- and high-audio frequencies. *J Acoust Soc Am* 105:829-837.

Eddins D.A, (2001), Measurement of auditory temporal processing using modified masking period patterns. *J Acoust Soc Am* 109:1550-1558.

Eisenberg L.S, Dirks D.D, Bell T.S, (1995), Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing. *J Speech Hear Res* 38:222-233.

Elberling C, (1999), Loudness scaling revisited. *J. Am. Acad. Audiol.* 10:248–260

Elberling C and Nielson C, (1993), The dynamics of speech and the auditory range in sensorineural hearing impairment in Recent Developments in hearing instrument technology-15th Danavox Syposium

Eriksson-Mangold M, Hallberg LR-M, Ringdahl A, Erlandsson SI, (1992), The development of a shortened Hearing MeasurementScale: the HMS-25. *J Audiol Mcd* 1:161-75.

Evans E.F, (1977), Frequency selectivity at high signal levels of single units in cochlear nerve and nucleus. In: E.F. Evans and J.P. Wilson, Editors, *Psychophysics and Physiology of Hearing*, Academic Press, London (1977), pp. 185–192

Festen J.M, (1993), Contributions of comodulation masking release and temporal resolution to the speech-reception threshold masked by an interfering voice. *J Acoust Soc Am* 94:1295-1300.

Festen J.M and Plomp R, (1981), Relations between auditory functions in normal hearing. *J. Acoust. Soc. Am*. 70: 356-69.

Festen J.M and Plomp R, (1983), Relations between auditory functions in impaired hearing. *J. Acoust. Soc. Am*. 73: 652–662.

Festen J.M and Plomp R, (1990), Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J Acoust Soc Am* 88:1725-1736.

Fletcher H, (1940), Auditory patterns.*Rev.Mod.Phys*.12, 47-65.

Florentine M, Buus S, Scharf B and Zwicker E, (1980), Frequency selectivity in normal hearing and hearing impaired observers. *J Speech Hear Res* 23: 457-61

Freyman R.L, Helfer K.S, McCall D.D and Clifton R.K, (1999), The role of perceived patial separation in the unmasking of speech. *J. Acoust. Soc. Am.* 106: 3578–88

French N.R and Steinberg J.C, (1947), Factors Governing the intelligibility of speech sounds. *J. Acoust. Soc. Am.*19: 90-119

Gatehouse S, Naylor G, Elberling C, (2003), Benefits from hearing aids in relation to the interaction between the user and the environment. *Int J Audiol* 42 Suppl 1:S77-S85.

Gatehouse S, Noble W, (2004), The Speech, Spatial and Qualities of Hearing Scale (SSQ). *Int J Audiol* 43:85-99.

Gelfand S.A, Piper N & Silman S, (1986), Consonant recognition in quiet and in noise with aging among normal hearing subjects. *J. Acoust. Soc. Am.* 80, 1589-1598.

Gelfand SA, (2004), Hearing: An introduction to psychological and physiological Acoustics. 4th edition. New York: Marcel Dekker.

George E.L.J, Festen J.M and Hougast T, (2006), Factors affecting masking release for speech in modulated noise for normal hearing and hearing impaired listeners. *J Acoust Soc of Am.*120:2295-2311

Glasberg B.R and Moore B.C.J, (1989), Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech, *Scand audio Suppl.*32, 1-25.

Glasberg B.R and Moore, B.C.J, (1992), Effects of envelope fluctuations on gap detection. *Hear res* 64(1): 81-92

Gustafson HA and Arlinger SD, (1993), Masking of speech by amplitude modulated noise. *J Acoust Soc of Am.*95:518-529

Hagerman B, (2002), Speech recognition threshold in slightly and fully modulated noise for hearing-impaired subjects. *Int J Audiol* 41:321-329.

Hagerman B, Kinnefors C, (1995), Efficient adaptive methods for measuring speech reception threshold in quiet and in noise. Scand Audiol 24:71-77.

Hagerman, B. 1982. Sentences for testing speech intelligibility in noise. *Scand Audiol* 11: 79-87

Hallgren M (2005) Hearing and cognition in speech comprehension. PhD thesis (published),Linkoping University

Hallgren M, Larsby B, Arlinger S, (2006), A Swedish version of the hearing in noise test (HINT) for measurement of speech recognition. *Int J Aud.* 45:227-/237

Hawkins J.E and Stevens S.S, (1950), The masking of pure tones and speech by white noise. *J. Acoust. Soc. Am*. 22: 6–13.

HEARCOM (2005) Implementation of a preliminary test set for auditory impairments (D-2-1), unpublished

Hirsh I J, (1950), The Relation between Localization and Intelligibility *J. Acous. Soc. Am*. 22(2); 196-200

Howard D.V, (1983), "The effects of aging and degree of association on the semantic priming of lexical decisions." Exp Aging Res. 9, 145-151.

Humes L.E and Roberts L, (1990), Speech-recognition difficulties of the hearing-impaired elderly: The contributions of audibility *J. Speech Hear. Res*. 33:726-735.

Humes L.E, Watson B.U, Christensen L.A, Cokely C.G, Halling D.C, Lidia L, (1994), Factors associated with individual differences in clinical measures of speech recognition among the elderly. *JSHR.* 37:465-474

Irwin R.J and McCauley S.F, (1987), Relations among temporal acuity, hearing loss and the perception of speech distorted by noise and reverberation. *J. Acoust. Soc. Am.*81:1557-65

Jerger J, Jerger S, Oliver T and Pirozzolo F, (1989), Speech understanding in the Elderly. *Ear Hear* 10:79-89

Jerger J, Jerger S and Pirozzolo F, (1991), Correlational analysis of speech audiometric scores, hearing loss, age and cognitive abilities in the elderly. *Ear and Hearing.*12:103-109

Johansson M.S.K and Arlinger S.D, (2002), Binaural masking level difference for speech signals in noise. *Int J Audiol* 41(5) 279-84

Keidser G, Seymour J, Dillon H, Grant F, Byrne D, (1999), An efficient, adaptive method of measuring loudness growth functions. *Scand Audiol,* 28:3-14

Kießling, J: Scaling methods for the selection, fitting and evaluation of hearing aids. In: Kollmeier, B. (ed.): Psychacoustics, speech and hearing aids. World Scientific, Singapore 1996

Killion M, Niquette P, Gudmundsen G, Revit L, Banerjee S, (2004), Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners.116(4); 2395-2405

Killion M.C and Niquette P.A, (2000), What can the pure-tone audiogram tell us about a patient's SNR loss? *Hear. J.* 53: 46–53.

Killion, M.C, (1997), SNR loss: I can hear what people say, but I can't understand them. *Hearing Review.*4 (12): 8-14.

Kramer S.E, Kapteyn T.S, Festen J.M, (1998), The self-reported handicapping effect of hearing disabilities. Audiology 37:302-312.

Kramer S.E, Kapteyn T.S, Festen J.M, Tobi H, (1995), Factors in subjective hearing disability. Audiology 34:311-320.

Kramer S.E, Kapteyn T.S, Festen J.M, Tobi H, (1996), The relationships between self-reported hearing disability and measures of auditory disability. *Audiology* 35:277-287

Kryter K.R, (1962), Methods for calculation and use of Articulation Index. *J. Acoust. Soc. Am*. 34:1689-97

Ladefoged, Peter (2001) http://en.wikipedia.org/wiki/Formant

Larsby B, Arlinger S, (1997), A method for evaluating temporal, spectral and combined temporal-spectral resolution of hearing. *Scand Audiol* 27:3-12.

Larsby B, Arlinger S, (1999), Auditory temporal and spectral resolution in normal and impaired hearing. *J Am Acad Audiol* 10:198-210.

Launer S, Hohmann V, and Kollmeier B, (1996), Modeling loudness growth and loudness summation in hearing-impaired listeners'' in *Modeling Sensorineural Hearing Loss*, edited by W. Jesteadt ~Erlbaum, Hillsdale, NJ. 175–185.

Leeuv A.R and Dreschler W.A., (1994), Frequency-resolution measurements with notched noises for clinical purposes. *Ear and Hear.* 15:240-55

Levitt H and Rabiner LR (1966) Binaural Masking-Level Differences and Improvements in Intelligibility (A) *J Acoust Soc Am.*39 (6): 1232

Ludvigsen C and Kuk F, (2001). New Solutions for Age-Old Hearing Aid Problems.(issue stories) *Hearing Review* November 2001.

Lunner T (2003) Cognitive function in relation to hearing aid use. *Int J Audiol* 42 Suppl 1:S49-S58.

Lutman M.E and Clark J, (1986), Speech identification under simulated hearing aid frequency response characteristics in relation to sensitivity, frequency resolution and temporal resolution , *J. Acoust. Soc. Am* 80: 1030-1040.

Lutman M.E, (1987), Psycho-acoustical characterization of sensorineural hearing loss in the UK population. unpublished abstract: 325

Lutman M.E,(2008) 'Acoustical analysis of ILD and BILD signals HearCom draft report, unpublished

Lyregaard P.E, (1982), Frequency selectivity and speech intelligibility in noise. *Scand. Audiol.* Suppl. 15: 113–122.

Lyxell B, Andersson U, Borg E, Ohlsson I.S, (2003), Working-memory capacity and phonological processing in deafened adults and individuals with a severe hearing impairment. *Int J Audiol* 42 Suppl 1:S86-S89.

MacKeith NW and Coles RRA, (2007). Binaural advantages in hearing of speech. *The J Laryngol and Otol* 85(3):213-32

Mills A.W, (1958), .In chapter 13, Minimum Audible Angle, 408. Hearing: An introduction to psychological and physiological Acoustics. 4th edition, Gelfand S. New York: Marcel Dekker.

Mills A.W, (1963), .In chapter 13, Minimum Audible Angle, 408. Hearing: An introduction to psychological and physiological Acoustics. 4th edition, Gelfand S. New York: Marcel Dekker

Mills A.W, (1972), .In chapter 13, Minimum Audible Angle, 408. Hearing: An introduction to psychological and physiological Acoustics. 4th edition, Gelfand S. New York: Marcel Dekker

Moore B.C.J, (1989), How much do we gain by gain control in hearing aids. *Acta Otolaryngol*

Moore, B.C.J, (1986), Temporal resolution measured by discrimination of stimuli in Temporal Resolution of the Auditory system in 'An introduction to Psychology of hearing (Third ed. B.C.J Moore), Academic Press, London and New York

Moore, B.C.J, (1986), Psychophysical tuning curves in Estimating the shape of the auditory filter in 'An introduction to Psychology of hearing (Third ed. B.C.J Moore), Academic Press, London and New York

Moore, B.C.J, (1995), Perceptual Consequences of Cochlear Hearing Loss and their Implications for the Design of Hearing Aids. *Ear & Hear* 17:133-60

Nelson P.B, Jin S.H, Carney A.E, Nelson D.A, (2003), Understanding speech in modulated interference: cochlear implant users and normal-hearing listeners. *J Acoust Soc Am* 113:961-968.

Nelson P.B, Hornsby B, Moore B.C.J, Trine T, (2007), Understanding Individual differences in Aided Speech Recognition and Hearing Aid Success. *The Hearing Review.* Issue: June

Newman C.W, Weinstein B.E, Jacobson G.P, Hug G.A, (1990), The Hearing Handicap Inventory for Adults: psychometric adequacy and audiometric correlates. *Ear Hear* 11:430-433.

Newman C.W, Weinstein B.E, Jacobson G.P, Hug G.A, (1991), Test-retest reliability of the hearing handicap inventory for adults. *Ear Hear* 12:355-357.

Noble W, Atherley G.R.C, (1970), The hearing measure scale: a questionnaire for the assessment of auditory disability. *J Aud Res* 10:229-250.

Noble W, Gatehouse S, (2004), Interaural asymmetry of hearing loss, Speech, Spatial and Qualities of Hearing Scale (SSQ) disabilities, and handicap. *Int J Audiol* 43:100-114.

Noordhoek I.M, Houtgast T, Festen J.M, (2001), Relations between intelligibility of narrow-band speech and auditory functions, both in the 1-kHz frequency region. *J Acoust Soc Am.* 109:1197-1212.

Olsen W.O, Noffsinger D, Carhart R, (1976), Masking level differences encountered in clinical populations. *Audiology* 15; 287–301.

Patterson R.D and Moore B.C.J, (1986), Auditory filters and Excitation patterns as representations of frequency resolution. In Frequency Selectivity in Hearing (ed. B.C.J Moore), Academic Press, London and New York

Patterson R.D, Nimmo-Smith L, Weber D.L, and Milroy R, (1982), The deterioration of hearing with age: Frequency selectivity the critical ratio, the audiogram and speech threshold. *J. Acoust. Soc. Am* 72, 1788-1803.

Patterson R.D, (1976), Auditory filter shapes, derived with noise stimuli. *J. Acoust. Soc. Am*.59: 640-54

Pavlovic C.V, (1984), Use of articulation index for assessing residual auditory function in listeners with sensorineural hearing impairment. *J. Acoust. Soc. Am.*75(4):1253-58

Peters R.W, Moore B.C.J, and Baer T, (1998), Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people *J. Acoust. Soc. Am*. 103, 577–587.

Philips S.L, Gordon-Salant S, Fitzgibbons P.J, Yeni-Komshian G, (2000), Frequency and temporal resolution in elderly listeners with good and poor word recognition.*JSLHR,*43:217-28

Pichora-Fuller M.K, (2003), Processing speed and timing in aging adults: psychoacoustics, speech perception, and comprehension. *Int J Audiol* 42 Suppl 1:S59-S67.

Pick G.R, Evans E.F, Wilson J.P, (1977), Frequency resolution in patients with hearing loss of cochlear origin; in Evans EF, Wilson JP (eds), Psychophysics and Physiology of hearing. New York, Academic Press, 273-81

Plomp R, Mimpen A.M, (1979), Improving the reliability of testing the speech reception threshold for sentences. *Audiology* 18:43-52.

Plomp R.A, (1978), Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J. Acoust. Soc. Am*.63(2):533-49

Portis T, (2005), In Cognitive issues with hearing aid or Why brain matters? Edwards & Hafter, AAA 2006

Poth E.A, Grose J.H, (1992), Masking level differences for tones and speech in elderly listeners with relatively normal audiograms (A). *J. Acoust. Soc. Am*.92(4) 2296

Rasmussen A.N, Olsen S.Q, Borgkvist B.V and Nielsen L.H, (1998), Long term test-retest reliability of category scaling in normal-hearing subjects using pure-tone stimuli. *Scand Audiol*. 27:161-67

Rickets T.A and Bentler R.A, (1996), The effect of test signal type and bandwidth on the categorical scaling of loudness. *J. Acoust Soc Am.*99: 2281-87

Rhebergen K.S and Versfeld N.J, (2005), A Speech Intelligibility Index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners. *J. Acoust Soc Am.*117:2181-92

Ringdahl A, Eriksson-Mangold M, Andersson G, (1998), Psychometric evaluation of the Gothenburg Profile for measurement of experienced hearing disability and handicap: applications with new hearing aid candidates and experienced hearing aid users. *Br J Audiol* 32:375-385.

Rosen S, Baker R.J, Darling A, (1998), Auditory filter nonlinearity at 2 kHz in normal hearing listeners. *J Acoust Soc Am* 103:2539-50

Ross M, (2006), Are Binaural Hearing Aids Better? The Hearing Journal (March-April)

Rummel R.J, (1975), Applied factor analysis, John Wiley Sons, New York

Sato H, Yoshino H and Nagatomo M, (1998).Relationship between speech transmission index and easiness of speech perception in reverberatory fields. *J Acoust Soc Am.*103 (pt.2); 2999

Schorn K and Zwicker E, (1990), Frequency selectivity and Temporal Resolution in patients with various inner ear disorders. *Audiology* 29:8-20

Shrout P.E and Fleiss J.L, (1979), Intraclass correlations: Uses in assessing rater reliability. *Psycho Bullet*. 2:420-28

Studebaker G.A, and Scherbecoe R.L, (1988), Magnitude estimations of the intelligibility and quality of speech in noise. *Ear Hear.* 9; 259–267.

Summers V, Molis M.R, (2004), Speech recognition in fluctuating and continuous maskers: effects of hearing loss and presentation level. *J Speech Lang Hear Res* 47:245-256.

Supin A.Y, Popov V.V, Milekhina O.N, Tarakanov M.B, (1997), Frequency-temporal resolution of hearing measured by rippled noise. *Hear Res* 108:17-27.

Surprenant A, (1999), The Effect of noise on memory for spoken syllables. *Int J of Psychology.*34(516);328-33

Tadros S.F, Frisina S.T, Mapes F, Kim S, Frisina R, (2005), Loss of Peripheral Right-Ear Advantage in Age-Related Hearing Loss *Audiol Neurootol* 10:44-52

Takahashi G.A, Bacon S.P, (1992), Modulation detection, modulation masking, and speech understanding in noise in the elderly. *J Speech Hear Res* 35:1410-142

Ter Keurs, M., Festen, J. M., and Pierap, R. (1992). Effect of spectral envelope smearing on speech perception, *J. Acoust. Soc. Am* 91:2 872-2880.

Thompson B, (1984), Canonical correlation analysis. London: Sage Publications

Tyler R.S, Summerfield Q, Wood E.J and Fernandes M.A, (1982), Psychoacoustic and phonetic temporal processing in normal and hearing-impaired listeners. *J. Acoust. Soc. Am*. 72 (7) 53-760.

van Rooij J C.G.M and Plomp R, (1990), Auditive and cognitive factors in speech perception by elderly listeners II: Multivariate analyses. *J. Acoust. Soc. Am*. 88: 2611-2624.

van Rooij J C.G.M and Plomp R, (1991), Auditive and cognitive factors in speech perception by elderly listeners III: Additional Data and Final discussion. *J. Acoust. Soc. Am* 91 (2): 1028-33

van Rooij J C.G.M, Plomp R and Orlebeke J.F, (1989), Auditive and cognitive factors in speech perception by elderly listeners I: Development of test battery. *J. Acoust. Soc. Am* 86 (4):1294-1309

van Rooij J.C.G.M & Plomp R (1990), Auditive and cognitive factors in speech perception by elderly listeners. II: Multivariate analyses. *J. Acoust. Soc. Am.* 88, 2611-2624.

Versfeld N.J, Daalder L, Festen J.M, Houtgast T, (2000), Method for the selection of sentence materials for efficient measurement of the speech reception threshold. *J Acoust Soc Am* 107:1671-1684.

Wagener K, Josvassen J.L, Ardenkjaer R, (2003), Design, optimization and evaluation of a Danish sentence test in noise. *Int J Audiol* 42:10-17.

Wagener K.C, Brand T, Kollmeier B, (2006), The role of silent intervals for sentence intelligibility in fluctuating noise in hearing-impaired listeners. *Inter J Audiol* 45(1), 26-33.

Wightman F.L and Kistler D.J, (1989b), Headphone simulation of free-field listening II: Psychophysical validation. *J Acoust Soc Am.*, 85(2), 868-878.

Wightman F.L and Raz I, (1980), Frequency selectivity as a function of sound pressure level. *J Acoust Soc Am.Suppl 1* 67 (S8)

Wilson R.H, Moncrieff D.W, Townsend E.A, Pillion A.L, (2003), Development of a 500 hertz masking-level difference protocol for clinic use. *J Am Acad Audiol.* 14; 1–8.

Wood S, (2005) http://person2.sol.lu.se/SidneyWood/praate/whatform.html

Zwicker E, Schorn K, (1978), Psychoacoustical tuning curves in audiology. *Audiology* 17 (2):120-40.

Appendices:

APPENDIX I

Outline and main findings of the classified studies discussed in chapter I and II

Class	Major outcome	Authors	Study	Domains covered	Main findings	Other specifications
A mainly	*Speech intelligibility can be predicted mainly based on a person's hearing threshold.*	Humes LE, Watson BU, Christensen LA, Cokely CG, Halling DC, Lidia L. (1994).	Studied factors associated with individual differences speech recognition among the elderly	Threshold measurement, auditory discrimination, speech recognition in quiet and noise, cognition	Principal component analysis revealed hearing loss as the single largest factor for differences in speech recognition (70-75%)	
A mainly	*Speech intelligibility can be predicted mainly based on a person's hearing threshold.*	Jerger J,Jerger S and Pirozzolo F.(1991)	Observed correlations between speech recognition,HL,age&cognition	Threshold ,speech recognition, cognition	Speech recognition was attempted to be predicted including all meaaures.Degree of hearing loss accounted for maximum variation, while cognitive scores accounted for very little variance for monotic test procedures. For dichotic tests both degree of hearing loss & cognition (speed of mental processing)accounted for significant variance	Age accounted for unique variance only in SSI
A mainly	*Speech intelligibility can be predicted mainly based on a person's hearing threshold.*	Divenyi, P. L., & Haupt, K. M. (1997a)	Investigated the age and laterality effects on audiological correlates of speech understanding	Threshold measurement, speech recognition in quiet and noise, perception of spectrally or temporally	Hearing loss is major factor differentiating auditory performance in elderly, age affects auditory & spatial	

				distorted speech and auditory resolution of frequency, time, and space	resolution, right ear advantage found for central auditory processing &left ear for auditory resolution	
A mainly	*Speech intelligibility can be predicted mainly based on a person's hearing threshold*	Humes and Robers 1990	The role that sensorineural hearing loss plays in speech recognition difficulties of the hearing impaired elderly is examined	Hearing threshold, speech recognition in noise	The primary determiner of speech recognition performance in the elderly hearing impaired subjects was their threshold elevation	
B(i) mainly	*Speech intelligibility may or may not be predicted based on a person's hearing threshold*.	Killion 1997, Killion et al., 2004	Observed relation between hearing threshold and speech recognition	Hearing threshold, speech recognition	Wide range of SNR loss is seen in persons with similar pure tone hearing losses.	
B (i)mainly	*Speech intelligibility may or may not be predicted based on a person's hearing threshold*.	Nelson et al, 2007	Observed relation between hearing threshold and speech recognition	Hearing threshold, speech recognition	Considerable variability exists among listeners with hearing loss and this could affect their success with amplification	
B(i) mainly	*Speech intelligibility may or may not be predicted based on a person's hearing threshold*.	Pavlovic CV (1984).	Used SII to predict speech recognition in NH & HI	Threshold and speech recognition in noise	Good SII predictions for normal and less impaired, but not for greater HL.Disproportionate loss in speech discrimination compared to that predicted on SII,suprathreshold deficit is frequency specific	
B(ii) mainly	*Speech intelligibility can be only partly predicted by a person's hearing threshold.*	George etal, 2006	Measured SRT in several amplitude modulated noises to determine cause of reduced benefit form	Pure tone audiometry,Speech recognition, frequency And temporal resolution	Reduced masking release can only partly be accounted by reduced audibility, rather temporal resolution and age are	Reduced spectral resolution does not qualify as an actual supra-threshold

			masker modulations		main factors governing masking release	deficit
B(ii) mainly	*Speech intelligibility can be only partly predicted by a person's hearing threshold.*	Festen, JM and Plomp R (1983)	Observed relations between auditory functions in impaired hearing.	Auditory threshold, frequency resolution, temporal resolution, speech recognition in quiet & noise	Principal component analysis revealed frequency resolution to form a cluster and independent of hearing loss. Speech recognition in noise is closely allied to frequency resolution, whereas speech recognition in quiet is governed by hearing loss	All tests were performed at 1000Hz.
B(ii) mainly	*Speech intelligibility can be only partly predicted by a person's hearing threshold.*	van Rooij &Plomp 1990 (II)	Observed relations between Auditive, cognitive and speech perception in the elderly population	Auditory threshold, frequency resolution, temporal resolution, speech recognition in noise, cognition	The results show that the deterioration of speech perception in the elderly consists of two statistically independent components ,progressive high-frequency hearing loss with age that accounts for approximately two-thirds of the systematic variance of the tests of speech perception and cognition accounting for one-third.	
B(ii) mainly	*Speech intelligibility can be only partly predicted by a person's hearing threshold.*	Lutman ME (1987)	Examined the relation amongst a wide variety of psycho-acoustical abilities along with the confounding effects of age	Auditory threshold, frequency resolution, temporal resolution, intensity resolution, temporal (integration, suppression, distortion and adaptation) and speech	Results from factor analysis revealed 7 identifiable factors with the first /largest factor representing low-mid frequency sensitivity and the second factor representing high frequency (4000 Hz) sensitivity, frequency	Speech recognition in noise was predicted by variables in factor 2 and was not related to age once psycho-acoustical variables had been accounted

				recognition in noise,	resolution and temporal resolution	for. Prediction for different groups was also studied.
B(ii) mainly	*Speech intelligibility can be only partly predicted by a person's hearing threshold.*	Dreschler and Plomp (1985)	Observed relations between psycho-physical data with a extensive battery of tone/phoneme/speech perception for adolescents	Auditory threshold, frequency resolution, temporal resolution, speech/tone/phoneme perception	Results indicate that hearing loss for speech is related to both, the frequency resolving power and temporal processing by the ear.	Phoneme perception parameters proved to be related to filtered speech thresholds than unfiltered and hence their role of bridging function between tone and speech perception is limited
other		Festen, JM and Plomp R (1981)	observed relations between auditory functions in normal hearing.	Auditory threshold, frequency resolution, temporal resolution, speech recognition in quiet & noise	Correlations obtained between- steepness of shallow edge of tuning curve & width of auditory filter (simultaneous masking),inverse between width of tuning curve& temporal window, positive between width of auditory filter and strength of cubic difference tone	All tests were performed at 1000Hz.Low correlations were not caused by poor test reliability
other		Divenyi PL and Haupt KM. (1997b)	Investigated correlations between audiological correlates of speech understanding	Threshold measurement, speech recognition in quiet and noise, perception of spectrally or temporally distorted speech and auditory resolution of frequency, time, and space	Auditory sensitivity measures showed persistent correlation with all except auditory resolution for frequency, time and space. After controlling for sensitivity, perceptual segregation was related to speech	It was important to remove the effects of hearing loss to uncover other factors

					intelligibility	
other		Divenyi PL and Haupt K. M. (1997c)	Studied factor representation for audiological correlates of speech understanding	Threshold measurement, speech recognition in quiet and noise, perception of spectrally or temporally distorted speech and auditory resolution of frequency, time, and space	Three factors were extracted including interference (general susceptibility to noise, spatial separation),high frequency hearing and basic auditory function	Hearing loss is a component of different factors. speech in reverberation should be included in clinical testing, since it was found to be independent of other measures
other		van Rooij &Plomp 1990 (I)- Development of test battery	Observed relations between Auditive, cognitive and speech perception in the elderly population	Auditory threshold, frequency resolution, temporal resolution, speech recognition in noise, cognition	In young listeners, individual differences in speech perception are small with low correlations, in elderly they overlap considerably between phoneme, spondee and sentence	Performance in the elderly is only partially correlated with age.
other		van Rooij &Plomp 1990 (III)- Additional data and final discussion	Observed relations between Auditive, cognitive and speech perception in the elderly population	Auditory threshold, frequency resolution, temporal resolution, speech recognition in noise, cognition	Age differences with respect to speech perception are most likely due to differences in auditive factors	

APPENDIX II

List of the commonly used speech recognition tests in different languages

Language	Test	Authors	Year	Number of sentence /sentence lists	Noise type	Speech/Noise presentation Level	Adaptive Procedure used	Scoring method	SRT in noise	SRT in quiet	Slope
Canadian French	HINT for adult Canadian Francophone populations	*Vaillancourt et al.,*	2005	240 sentences(12 lists /20 sentences)	Speech shaped noise	Speech (sentence)presentation level 65 dB(A)	+/- 2 dB steps is used depending upon the previous response	Single word responses, scores were calculated in percent correct	Mean SRT in noise was -3.0 dB (A) in noise front condition and -11.4 dB (A) in noise side condition (SD=1.1 dB)	Mean SRT in quiet was 16.4 dB (A) (SD=2.2dB)	10.3%/dB
Cantonese	Chinese HINT	*Wong et al.*	2005	240 sentences(two separate lists with 12 sentences/20lists and 20 sentences/12 lists	Speech shaped noise	noise fixed at 65dBA with the level of speech varied, for quiet condition :20dB A	+/- 2 dB steps for both quiet and noise conditions	Single word responses, scores were calculated in percent correct	Mean SRT's: Noise front:-3.9 dB,Noise right:-10.6 dB,Noise left: -10.5 dB	Mean SRT's: In quiet: 19.4 dBA	----
Swedish	Swedish version of HINT	*Hallgren ,Larsby, Arlinger*	2006	250 sentences, 25 lists/10 sentences	Speech shaped noise	noise level fixed at 65 dB SPL	+/- 2 dB steps is used depending upon the previous response	Whole sentence scoring single word scoring	The mean S/N ratio at threshold: -3.0 dB,with a SD of 1.1 dB	---	17.9%/dB at its steepest (calculated between -4 and -2S/N) for whole sentence scoring and 15.4dB/dB (calculated between -6 and -4 S/N) for word scoring

American English	Speech in noise sentence test (SPIN test*)*	*Kalikow et al.,*	*1977*	200 sentences 8 lists / 25 sentences with high (PH)and low(PL) predictabilities	twelve talker babble	Speech and noise were presented at 80dB SPL (0dB SNR)	Non adaptive	Single word responses, scores were calculated in percent correct	scores : 88% (PH)and 38% (PL)	--	14%/dB (PH)and 3% 8%/dB(PL)
American English	Standardization of Speech in noise sentence test (SPIN test)	Bilger et al.	1984	the original 10 lists /25 sentences ,with high(PH) or low(PL) predictability	twelve talker babble	Speech presentation level: 8 dB SNR /50 dB above individual threshold	Non adaptive	Single word responses, scores were calculated in percent correct	Average scores: PH-43.6% (list SD =1.0),PL: 22.9% (list SD=2.8), Overall (PH and PL): 33.3% correct (list SD = 1.6)	--	
American English	Hearing in noise sentence test (HINT sentences*)*	*Nilsson et al*	*1994*	25 lists / 10 sentences,plus 3 practice lists / 12 sentences	Speech shaped noise	Sentences presented in quiet and for sentences in noise: noise presentation level-72dB(A)	+/- 2 dB steps is used depending upon the previous response	Single word responses, scores were calculated in percent correct	Mean SRT in noise : -2.92 dB, SD - 0.78dB	Mean SRT in quiet :23.9dB(A) ,SD :3.5dB	
Dutch	Dutch speech reception test (Plomp sentences)	*Plomp and Mimpen*	1979	10 lists/13 sentences	Speech shaped noise	Sentences presented in quiet and for sentences in noise: speech fixed, presentation level-50 dB SPL(Plomp); fixed noise level - 65 dB(A)(Smoorenburg)	Non adaptive		Mean monaural SRT in noise : -5.9 dB, SD:0.9 dB & free-field presentation, the mean SRT in quiet: -19 dB (Plomp); male speaker, headphones, the mean monaural SRT in noise : -5.1(SD=1.8dB) Smoorenburg(1992)	Mean monaural SRT in quiet :19dB(A), Plomp; free-field presentation, mean SRT in quiet was -15.8 dB (SD=2.3), Smoorenburg (1992)	Slope 20% /db(Plomp)

Dutch	*Alternative Dutch Speech Reception test, VU98 sentences*	*Versfeld et al.,*	*2000*	10 lists /13 sentences	Speech shaped noise	Sentences presented in quiet and for sentences in noise: noise presentation level-72dB(A)	+/- 2 dB steps is used depending upon the previous response	Whole sentences are scored	Mean monaural SRT in noise : -4.1 dB, SD – 1.1dB	-----	15% per dB for both male & female speakers
British English	BKB sentences	Bench et al.,	1979	21 lists / 16 sentences				Keyword scoring			
British English	*IHR ASL sentences*	*Macleod and Summerfield*	1990	10 lists /15 sentences	Lo w pass White noise, Speech shaped noise (Moore et al.,2001and Alcantara et al.,2003)	sentences in noise: noise presentation level-60dB(A) with audio(A) and audio-visual (AV)modes;60,65 & 70 dB SPL (Moore et al.,2001and Alcantara et al.,2003)	+/- 2 dB steps is used depending upon the previous response	Keyword scoring	Mean monaural SRT in noise : - A:-16.8 dB (SD=1.2 dB),AV: -23.2 dB (SD=2.3 dB); 60dB: -11(SD=0.3),75 dB: -10(SD=1.2)& 65 dB SPL: -8.6 (SD=1.0) (Moore et al.,2001and Alcantara et al.,2003)	------	9.9 and 7.4% for A and AV condition respectively
Swedish	*Closed set Swedish sentences test*	*Hagerman,*	,1984	11 lists / 10 sentences	Speech shaped noise	Sentences presented in quiet and for sentences in noise: speech presentation level-55 dB SPL	----	Closed speech test	SRT in noise:-8.1 dB (SD:0.44dB)	SRT in quiet :21.6 dB SPL(SD 2.5dB).	20 and 25%dB per sentence, respectively for HI&NH

Swedish	Adaptive Swedish test	Hagerman and Kinnefors	1995	11 lists /10 sentences	Speech shaped noise	Sentences presented in quiet and for sentences in noise: noise presentation level: -8dBSNR	+/-1,2 or 3dB depending upon on the number of correct words	Single word responses, scores were calculated in percent correct	mean SRT in noise is -7.8dB, SD=0.6	mean SRT in quiet : 21 (SD of 1.6)dB	----
German	*Gottingen sentences*	*Kollmeier and Wesselkamp*	*1997*	20 lists /10 sentences	Speech shaped noise	for sentences in noise: noise presentation level: 65dB SPL	------	Single word responses, scores were calculated in percent correct	average SRT: average SRT -6.2dB (SD 0.3 dB)	---	average slope :19%/dB
German	*Oldenburg sentences (closed set)*	*Wagener et al.*	*1999a; 1999b, 1999 c*	10 lists / 10 sentences	Speech shaped noise	sentences in noise: noise presentation level: 65dB SPL	Closed set	Single word responses,s cores were calculated in percent correct	average SRT of the lists : -7.1(SD=0.2)dB, learning effect : 0.3dB per list	---	average slope :17%/dB

German	*Gottingen ,Oldenburg sentences (adaptive test)*	Brand and Kollmeier	2002	---	Speech shaped noise	sentences in noise: noise presentation level: 65dB SPL (NH), medium loudness for the HI listeners	Step size and direction depended on word score of the previous sentence, converging at 20 and 80% correct	Single word responses, scores were calculated in percent correct	Gottingen sentences (NH): SRT's -5.5 (SD=0.8),HI: -0.8 (SD=1.1)dB; Oldenburg sentences- SRTs-(NH)-6.1 (SD=0.6) dB;HI: -2.4 (SD=0.8)	----	Gottingen sentences converging at 20 and 80% correct:slopes were 16(SD=2.6) and 13 (SD=3.1)%/dB for NH and HI respectively Oldenburg sentences converging at 20 and 80% correct:slopes were 16(SD=3.4) and 14(SD=4.1)5/dB for NH and HI,respectively
Danish	*Dantale II sentences*	*Wagener etal.*	*2003*	16 lists /10 sentences	Speech shaped noise	sentences in noise: noise presentation level: 65dB SPL	Step size and direction depended on word score of the previous sentence	Single word responses,s cores were calculated in percent correct	SRTs in noise : -8.43 dB(SD=0.95 across subjects) SRTs in noise : -8.38(SD=0.16 across lists)dB	---	Slopes:13.2(SD=1.9)%/dB
American English	Quick SIN test	Killion etal.,	2004	12 lists, /6 sentences	four-talker babble	noise presentation level: 70dB HL	Fixed SNR's	Key word scoring	SRTs in noise : 1.9 dB,(NH),too diverse for HI	---	estimate of SNR loss accurate to +/-2.7 dB at the 95% confidence level

American English	Original SIN	*Etymotic research*	*1993*	9 lists / 40 sentences	four-talker babble	Sentences presentedat two levels (83 and 53dB SPL)	Four SNRs;15,10, 5,0 dB	Key word scoring	---	----	----
American English	BKB-SIN	*Fikret-Pasa*	*1993*	16 lists / 10 sentences	four-talker babble	Sentences presented at SNR of +21,+18,+15,+12 ,+9,+3,0,-3 and - 6 dB	Fixed SNR	Key word scoring	Normal hearing adults : - 2.5 dB Normal hearing children: 0.5dB(ranging between 0 to 1 dB) Cochlear Implant adults :These were divided into three categories Best performers : 11.9dB Mid performers :9.0 dB Poor performers :6.5 dB	---	---
American English	*Connected speech test*	*Cox et.al*	*1987*	48 passages of conversationally produced connected speech	six-talker speech babble	passages presented with speech level fixed at 61 dB A	Fixed SNR (-3,-8)	Key word scoring from passage	group mean score : 60.1 rau		Slopes: 12%/dB

APPENDIX III

An example of a list of BKB sentences with words in capitals being key words

The CLOWN had a FUNNY FACE.
The CAR ENGINE's RUNNING.
SHE CUT with her KNIFE.
CHILDREN LIKE STRAWBERRIES.
The HOUSE had NINE ROOMS.
THEY're BUYING some BREAD.
The GREEN TOMATOES are SMALL.
HE PLAYED with his TRAIN.
The POSTMAN SHUT the GATE.
THEY're LOOKING AT the CLOCK.
The BAG BUMPS on the GROUND.
The BOY DID a HANDSTAND.

APPENDIX IV

Development and Evaluation of English Matrix Test

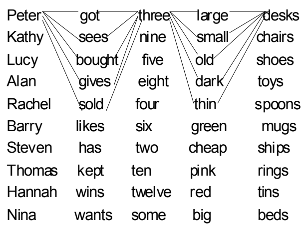
Selection of the speech material:

The material used for the test as the base list consisted ten sentences with five words each. It was devised and recorded by Stuart Hall (2006) as a part of his Master's dissertation. The syntactic structure of all sentences is identical: *Name verb numeral adjective object*. This structure has been previously used in other languages, like the Swedish Hagerman sentences (Hagerman, 1982), the German Oldenburg Sentence test and the Danish Dantale II test (Wagener et al, 2003) and is now being adapted for English. The base list approximates the mean phoneme distribution of the English language. The test sentences are generated by choosing randomly one of the ten alternatives for each part of the sentence. Consequently each test list consists of the same word material. There are 100,000 possible permutations (sentences) using this approach. The English word matrix (base material) is as follows:

Name	**Verb**	**Numeral**	**Adjective**	**Object**
Peter	got	three	large	desks
Kathy	sees	nine	small	chairs
Lucy	bought	five	old	shoes
Alan	gives	eight	dark	toys
Rachel	sold	four	thin	spoons
Barry	likes	six	green	mugs
Steven	has	two	cheap	ships
Thomas	kept	ten	pink	rings
Hannah	wins	twelve	red	tins
Nina	wants	some	big	beds

Recording of the speech material

A female speaker with a neutral southern accent read out sentences in a large recording studio having minimal reverberation. She was instructed to maintain the same speed and pronunciation throughout. The sentences were recoded onto digital audio tape at a sample rate of 44.1 kHz then transferred to digital waveform files (.wav). In order that the final sentences had a naturally spoken pattern, one hundred sentences were recorded in a manner which meant that each word in a given column would be recorded in combination with all the words from the following column. This would also facilitate use of correct co-articulation between words while cutting the material. The method of combination is demonstrated below:



The sentences were recorded for three takes to provide enough recorded material for cutting. This editing and cutting was done using the Adobe Audition programme which visualizes and edits time waveforms and short time spectra.

Editing the material :

The three recordings of the sentences were played several times and the best material from the three in terms of speed of delivery, rhythm, pronunciation etc. was selected as the final recording to be cut. The silence at beginning and end of the selected sentences was then almost entirely deleted. A small duration of silence (15 ms) was left. The recorded files were further edited by equalising the overall rms levels of the sentences.

Cutting the material

The sentence to be cut was played several times on Adobe Audition. It was then viewed in both waveform and spectrogram to identify the point at which the cut should be placed. As the words were chosen at random for generation (with the constraint of correct co-articulation), it was important that all words we cut are in the same way. So the cutting was performed in a way that the cutting point cuts the file into two parts, so nothing is duplicated. For example, when cutting the pair of words ‘got three’ (see figure below), the waveform corresponding to ‘three’ is identified and cut closely at the beginning, allowing about 15 ms before the start of the frication that commences the /th/ phoneme. The reminder of the word pair is ‘got’ plus any co-articulation between the two words. In this way, the co-articulation is attached to the end of words. Note that all the possible combinations of the last two words were actually spoken in the recordings. However following validation it was necessary to independently change the levels of each of the last two words so they had to be cut at a later stage.

Peter got three large desks

Peter ~ ✂ *got* **….**
…. ✂ **got ~** ✂ *three ….*
…. ✂ **three ~** ✂ *large* **….**
…. ✂ **large desks**

The co-articulation effects are taken into account to achieve a natural intonation. Only the utterances with the correct co-articulation to the following word in the final sentence are used, i.e. the boldface words, to generate the sentence: Peter got three large desks. The co-articulation part is indicated by ~, the cutting place by ✂.

Thus each file was cut into single words so each particular word can be addressed in the optimization. The cuttings were done in the zero crossings of the gross waveforms which means that each file starts with 0° phase and ends with 360° phase. This could be done accurately by zooming the view of the file to the selected cutting point and adjusting its location to the nearest zero crossing point. The cutting of the hundred sentences with five words lead to five hundred sound files, each stored as (.wav) format. However, during the generation only ten files from last column (objects) were required for formation of sentences and the best example were chosen, based on listening. Thus overall, four hundred and ten individually cut sound files were obtained.

<u>*Labelling the files:*</u>

Each file had to be labelled separately so as to track the word followed by its respective co-articulation. The base sentences were hence allotted numbers from 0-9 while the individual words had letters a-e as shown below:

Index	a	b	c	d	e
0	Peter	got	three	large	desks

1	Kathy	sees	nine	small	chairs
2	Lucy	bought	five	old	shoes
3	Alan	gives	eight	dark	toys
4	Rachel	sold	four	thin	spoons
5	Barry	likes	six	green	mugs
6	Steven	has	two	cheap	ships
7	Thomas	kept	ten	pink	rings
8	Hannah	wins	twelve	red	tins
9	Nina	wants	some	big	beds

This labelling enables identification if any word in the matrix followed by its co-articulation without too much difficulty. For e.g.: Peter sees five thin ships
(Peter-a0, sees-b1, five-c2, thin-d4, ships-e6 which is the following word)
Thus the sound file Peter here which is followed by sees is named as a0b1 sees as b1c2 and so on.

Generation of sentences for optimisation:

The individually cut and labelled sound files were ready to be combined to generate sentences which would provide material for optimisation measurements. However, before the final generation of sentences the words were spliced together, the waveforms were tapered and overlapped to ensure a smooth transition. The overlap was about 20 ms. The generation of the new sentences was performed using a specially written program which randomly generated a new list of ten sentences by combining the cut wave files. In constructing the sentences a word in a given column is selected to produce the correct co-articulation for the following word, regardless of the previous word. The sentence generating program produced a single waveform file for each new sentence of each of the twenty lists of ten sentences made up of five original wav files. These were labelled 0101.wav for the first sentence of list one, 0102.wav for the second sentence of list 1 and so on. These sentences (two hundred in all), were then used by Hall (2006) to test fourteen normal hearing subjects using non-adaptive procedures. The data obtained was subjected to statistical analysis. This was essential to obtain an optimised perceptual homogeneity. Optimisation basically standardises the speech material, making all sentences equally intelligible and hence ensuring that each one would give the same result as other when randomly selected.

<u>*Data Analysis*</u>

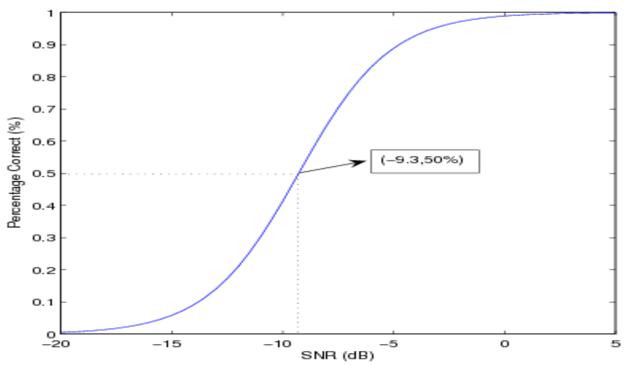
Data was analysed using SPSS software.

- The analysis of Hall's original data was aimed to obtain a psychometric function of % correct level (dB) vs. SNR and hence estimate the SNR for 50% correct recognition of each individual word as well as for the whole corpus. The values obtained could be used to adjust individual words and hence ensure equal intelligibility for all sentences.

-The method used for obtaining each of the above value is logistic regression.

1. Overall data value & graph is given below.

(Psychometric function of Mean correct values in dB vs SNR)



The above graph reveals that at SNR value of -9.3dB all of the sentence material is 50 % intelligible. Thus,

-9.3 dB SNR = avg Reference value (range -3 to -13 dB SNR)

2. Table below displays threshold values of SNR values individually for each word, as well the optimisation values to be actually applied to each file along with their slope.

Optimisation values in dB for Matrix sound files along with the SNR's and slope.

Word	constant	B	p(0.5) dB	Rounded (SNR) dB	reference dB	correction dB	slope
Peter	6.394	0.708	-9.031	-9	-9.3	-0.3	0.18
Kathy	4.006	0.524	-7.645	-7	-9.3	-2.3	0.13
Lucy	8.269	0.728	-11.358	-11	-9.3	1.7	0.18
Alan	6.227	0.487	-12.786	-12	-9.3	2.7	0.12
Rachel	7.004	0.599	-11.692	-11	-9.3	1.7	0.15
Barry	5.261	0.583	-9.027	-9	-9.3	-0.3	0.15
Steven	7.512	0.769	9.768	-9	-9.3	-0.3	0.19
Thomas	6.229	0.614	-10.144	-10	-9.3	0.7	0.15
Hannah	1.166	0.244	-4.778	-4	-9.3	-5.3	0.06
Nina	5.227	0.64	-8.167	-8	-9.3	-1.3	0.16
got	6.635	0.792	-8.453	-8	-9.3	-1.3	0.2
sees	5.876	0.574	-10.263	-10	-9.3	0.7	0.14
bought	6.108	0.662	-9.226	-9	-9.3	-0.3	0.17
gives	4.944	0.584	-8.465	-8	-9.3	-1.3	0.15
sold	4.72	0.523	-9.024	-9	-9.3	-0.3	0.13
likes	6.668	0.615	-10.842	-11	-9.3	1.7	0.15
has	9.874	0.909	-10.862	-11	-9.3	1.7	0.23
kept	4.848	0.567	-8.55	-8	-9.3	-1.3	0.14
wins	3.26	0.559	-5.442	-5	-9.3	-4.3	0.14
wants	5.4	0.693	-7.792	-7	-9.3	-2.3	0.17
three	4.21	0.488	-8.627	-8	-9.3	-1.3	0.12
nine	6.41	0.623	-10.288	-10	-9.3	0.7	0.16
five	3.795	0.391	-9.705	-9	-9.3	-0.3	0.1
eight	10.096	0.913	-11.058	-11	-9.3	1.7	0.23
four	7.101	0.671	10.582	-10	-9.3	0.7	0.17
six	7.269	0.635	-11.447	-11	-9.3	1.7	0.16
two	7.287	0.71	-10.263	-10	-9.3	0.7	0.18
ten	6.927	0.675	-11.083	-11	-9.3	1.7	0.17
twelve	6.634	0.722	-9.188	-9	-9.3	-0.3	0.18
large	8.342	0.877	-9.511	-9	-9.3	-0.3	0.22
small	4.354	0.508	-8.507	-8	-9.3	-1.3	0.13
old	3.004	0.349	-8.607	-8	-9.3	-1.3	0.09
dark	5.24	0.529	-9.905	-10	-9.3	0.7	0.13
thin	0.926	0.75	-1.234	-1	-9.3	-8.3	0.19
green	4.983	0.599	-8.318	-8	-9.3	-1.3	0.15
cheap	6.952	0.684	-10.163	-10	-9.3	0.7	0.17
pink	3.069	0.559	-5.49	-5	-9.3	-4.3	0.14
red	6.866	0.805	-8.529	-8	-9.3	-1.3	0.2
big	3.494	0.582	-6.003	-6	-9.3	-3.3	0.15
desks	7.865	0.789	-9.968	-10	-9.3	0.7	0.2
chairs	7.76	0.748	-10.374	-10	-9.3	0.7	0.19
shoes	4.804	0.436	11.018	-11	-9.3	1.7	0.11
toys	4.135	0.468	-8.835	-8	-9.3	-1.3	0.12
spoons	4.524	0.507	-8.923	-9	-9.3	-0.3	0.13
mugs	5.309	0.564	-9.413	-9	-9.3	-0.3	0.14
ships	3.078	0.275	-11.192	-11	-9.3	1.7	0.07
rings	4.827	0.603	-8.004	-8	-9.3	-1.3	0.15
tins	5.889	0.653	-9.018	-9	-9.3	-0.3	0.16
beds	3.886	0.534	-7.227	-7	-9.3	-2.3	0.13

These values were pooled from the ten examples of each word (except for object words). It was sufficient to apply a single correction factor to all sound files of the same word (e.g. all 'Peters') since there was no significant difference between all samples of sound files of the same word. The above correction factors were finally applied to sound files using the Adobe Audition program. The matrix material was thus standardised.

Final Editing: Before the sentences were actually generated, they underwent one more editing session. Here the sentences were once again played for qualitative analysis. It was observed that some words, especially beginning with stop sounds (/b/), sibilants (/s/ /sh/) sounded unnatural in certain combinations. Hence it was necessary to include additional silence gaps at the beginning/end of some words. Moreover, further editing included in some instances tapering the ends, smoothing the wave so that the sentences sound natural.

Generation of sentences: The final step was to generate a set of standardised sentences for the main study. This generation was performed using the same specially written program used during the generation of sentences for optimisation. The program also produced a text file for each of the lists of ten sentences describing the contents of each sentence. Twenty such lists each containing ten sentences were produced. Thus, two hundred sentences overall were generated for HearCom. These text files were then pasted onto a new text file combining all the text files from all lists. The text files required to be in a specific format , in order that they are compatible with the OMA test platform. Twenty such lists each containing ten sentences were produced.

Normative data:15 normal hearing individuals were aged between 18-30 yrs (average: 24years) were tested for their left and right ears in noise (generated from the long term average spectrum of the material and in quiet) with one test list for each ear and condition. In noise the score was obtained adaptively with noise level fixed and speech level varied in same way described to obtain SRT in the studies above. Their average and SD across the two ears are as follows: stationary noise (-8.4 dB,0.86 SD) and quiet (19.1 dB, 1.2 SD)

APPENDIX V

Gothenburg Profile (full questionnaire)

The Gothenburg Profile consists of the following questions:

Subscale 1: Speech intelligibility and localization.

1. How often does it occur that you cannot hear conversation when speaking to one person at home?
2. How often does it occur that you cannot hear conversation in a group at home?
3. How often does it occur that you cannot hear the speaker at a meeting, if you are well positioned?
4. How often does it occur that you cannot hear the newsreader on the TV, when the volume is not turned up?
5. How often does it occur that you cannot hear the newsreader on the radio?
6. How often does it occur that you cannot localize different sound of traffic?
7. How often does it occur that you turn your head to the wrong direction when someone is calling out to you?
8. How often does it occur that you are surprised because cars have come closer to you than you thought?
9. How often does it occur that you cannot hear when someone is opening a door behind you?
10. How often does it occur that you cannot (only by hearing) decide if the water is boiling in a pan?

Subscale 2: Experienced handicap, relation to others

11. How often do you find hearing problems an obstacle for your social life?
12. How often does it occur that you avoid social gathering because it is too hard to follow a conversation?
13. How often does it occur that you feel that people are ignoring you just because of your hearing difficulties?
14. How often do you feel that people find it hard to talk to you?
15. How often does it occur that you have a feeling of being excluded from things because of your hearing difficulties?
16. How often does it occur that you are reluctant to meet new people due to your hearing difficulties?
17. How often does it occur, if you are sitting quietly in a group of people, that you are afraid of saying something foolish?
18. How often does it occur that your self confidence is affected because you are having hearing difficulties?
19. How often does it occur that your poor hearing makes you feel inadequate?
20. How often does it occur that you feel sad or angry if you cannot join in a conversation?

APPENDIX VI
Complete list of the words used in development of lexical decision test

Word	Part of speech	word frequency	Non words	Can pronounce/ not
FAN	noun	48	SHU	P
RUG	noun	14	DZA	NOP
GAP	noun	45	KAS	P
KEY	noun	76	ZVU	NOP
GET	verb	2210	CAG	P
FUN	noun	34	TPI	NOP
JOB	noun	326	POB	P
SUM	noun	56	WTE	NOP
ASK	noun	610	NAR	P
SKY	noun	56	BLT	NOP
BUY	verb	262	DAR	P
SIT	verb	301	CLD	NOP
CUP	noun	134	LIK	P
FEE	noun	58	TRW	NOP
GUN	noun	55	MOL	P
JAM	noun	10	HWO	NOP
PET	noun	19	HUS	P
LEG	noun	118	EDN	NOP
CUT	noun	58	DUR	P
DOG	noun	124	TKU	NOP
PAN	noun	15	MUB	P
BAT	noun	13	EHJ	NOP
WAY	noun	1148	TIV	P
END	noun	458	QFY	NOP
EYE	noun	392	ROX	P
USE	noun	328	NVE	NOP
CAR	noun	353	POY	P
EAR	noun	59	FGI	NOP
PUB	noun	51	DEG	P
WAR	noun	297	BTU	NOP
FOR	preposition	8412	DIF	P
BUT	preposition	22	HTI	NOP
HIS	determiner	4285	JIK	P
OUR	determiner	950	JBO	NOP
NOW	adverb	1382	GIS	P
OUT	adverb	1542	MPI	NOP
HOW	interrogative	1016	SIK	P
WHY	interrogative	509	LNU	NOP
FAR	adverb	310	BAS	P
OLD	adjective	648	DGE	NOP
BIG	adjective	338	GUL	P
NEW	adjective	1154	UGT	NOP
HOT	adjective	94	BON	P
DRY	adjective	56	CKE	NOP
CAN	verb	2672	GAK	P
SAY	verb	3344	MGI	NOP

TRY	verb	552	LAF	P
RUN	verb	406	QHU	NOP
PAY	verb	381	HAR	P
MAP	noun	56	XVI	NOP

APPENDIX VII

Similarities and differences of studies I and II

In order to compare the two studies, it becomes essential to outline the various similarities and differences between the two studies.

Table 5.21: Similarities and differences between the studies I and II

No.	Study I	Study II
1.	AC thresholds of 500 and 3000 Hz only included	AC thresholds of 500-4000 Hz included
2.	PTA was defined as average of 1000,2000 and 4000 Hz	PTA was defined as average of 500,1000,2000 and 4000 Hz
3.	Included MAA, Listening effort and MCL BB measurements	These were eliminated
4.	MCL-defined reference value. Hence six variables related to loudness measurements were included (MCL500, MCL3000,MCLbb and slopes for the three)	1/3 gain rule-defined reference value. Hence variables included refdB and only two variables related to loudness measurements (loudness slopes only for 500 and 3000 Hz since BB measures were eliminated)
5.	Better ear, test session values used for data analysis	Better ear, test session values used for data analysis
6.	Blom transforms used for not normally distributed data	Blom transforms used for not normally distributed data
7.	Measures of asymmetry not included in regression as variables for prediction of GP variables	Measures of asymmetry included in regression as variables for prediction of GP variables
8	ILD and BILD measures did not include AC1000,2000,4000 Hz and their measures asymmetries as their independent variables.	Includes them.
9	GP measures (subscales) not included in FA per subject analysis	GP measures (subscales) included in FA per subject analysis

From above it becomes evident that, the two studies are different on six out of eight aspects. Thus 5 and 6 are similar in both studies while they differ in 1-4 and 7-8. Also 3 and 4 are discussed before and the reasons for the changes are justified. But to compare both, 1,2 ,7 and 8 need to altered in either of the studies. This would result in more appropriate comparison of the two studies. Thus these aspects (1,2,7,8) were modified in study I as per the criteria of study II and regression analysis were repeated for all measures (SRTstat, SRTfluc, SRT for both types of noises and in the three groups of hearing loss, ILD, BILD, GP measures). Thus analysis for study I was repeated by:

- Addition of AC thresholds of 1000, 2000 and 4000 Hz as variables from the raw data.
- PTA was redefined as average of 500-4000 Hz and was added as a variable. Thus the subgroups were also divided based on this.
- Measures of asymmetry (defined as difference in threshold between the left and right ears) were included in regression as variables for prediction of GP variables
- Factor analysis was repeated including GP subscales.
- Addition of AC thresholds of 1000, 2000 and 4000 Hz and their asymmetry measures for ILD/BILD regression.

APPENDIX VIII

Cognitive tests:

The purpose of the experiments, conducted is: To study which cognitive skills are important for speech recognition/comprehension as measured in the Hagerman speech test and in the Swedish Hearing In Noise Test (HINT). Protocol : In the experiment the following tests were used:

1) Speech recognition was measured with the **Hagerman speech test**. The speech signal was fixed at 70 dB SPL (C-weighted). The noise level was adjusted adaptively in an interleaved method to reach 50 % and 80 % correct responses. The S/N values for 50% and 80% correct responses, respectively, were calculated and used as outcome measures.

2) **HINT** (Hearing In Noise Test): Speech recognition was measured with the Swedish HINT sentences (Hällgren et al, manuscript). The speech level was fixed at 70 dB SPL (C-weighted). The noise signal was adjusted adaptively in two different procedures, to reach 50% correct sentences and to reach 60% correctly recognized keywords. The S/N values for the two procedures were calculated and used as outcome measures.

3) **Vigilance test (CVC):** Letters were presented on a computer screen in front of the test subjects one at a time. Every second letter is a consonant and every other a vowel. The letters were presented with an inter-stimulus interval of one (fast) or two (slow) seconds. The subject's task was to press the space-bar on the computer every time three letters in a sequence of consonant-vowel-consonant (C-V-C) constituted a real word. The test items were twenty real C-V-C words for each inter-stimulus interval, i.e. a total of forty. Outcome measures were the numbers of existing words recognized in the conditions CVC d'slow and CVC d'fast, to be combined to CVC d'average.

4) **Reading span test:** The subject's task was to comprehend sentences and to recall either the initial or the final words of a presented sequence of sentences in correct serial order. The words in each sentence are presented in a word-by-word fashion. Half of the sentences are absurd and half are normal sentences. The subjects' task was to respond "yes" (for a normal sentence) or "no" (for an absurd sentence) after the presentation of each sentence. After a sequence of sentences the test leader indicated that the subject should recall either the initial or the final word of each presented sentence in the sequence. Outcome measure was the total number of correctly identified initial and final words.

5) **Lexical decision making:** The subject's task was to judge whether a string of three letters constituted a real Swedish word or not. Half of the non-words sounded like a real word when pronounced, whereas the remaining half did not . Stimuli were presented on the screen and the subjects pressed predefined keys, one for "no" and one for "yes". Both number of correct answers and reaction times were measured.

6) **Physical matching:** The subject's task was to judge whether two simultaneously presented letters had the same physical shape (e.g., A - A) or not (A - a). Stimuli were presented on the screen and the subjects pressed predefined keys, one for "no" and one for "yes". Both number of correct answers and reaction times were measured.

7) **Wordspan:** Test of serial record of monosyllabic words (3 letters). The subject's task was to repeat a series of words presented one by one on the computer screen. After a sequence of words (3-8) the test leader indicated that the subject should start to recall all words presented since last recall. A sequence of n words was repeated three times. Outcome measure was the total number of correctly recalled words.

8) **Test of verbal ability:** The verbal ability was assessed by giving the subjects a paper with twenty-seven groups of words, each group including 5 words. The subject's task was to choose and underline the two words among the five which were mutually opposite. The subjects were allowed a maximum of four and a half minutes to solve the task. Outcome measure was the number of correctly identified opposite pairs.

Among the cognitive tests especially the reading span test and the test of lexical access were found to correlate significantly with the S/N in the Hagerman test and in the HINT test. And since lexical test was more time efficient clinically, it was chosen to be selected in the test battery over reading span.

APPENDIX IX

Independent Samples Test

STUDY I (NH=30, HI= 73)		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
AC 500 Hz (dB)	Equal variances assumed	23.907	.000	-9.587	98	.000	-29.21429	3.04739	-35.26173	-23.16684
	Equal variances not assumed			-13.288	95.440	.000	-29.21429	2.19848	-33.57856	-24.85001
AC 1000 Hz (dB)	Equal variances assumed	33.973	.000	-10.968	98	.000	-35.73810	3.25849	-42.20446	-29.27173
	Equal variances not assumed			-15.575	91.245	.000	-35.73810	2.29465	-40.29597	-31.18022
AC 2000 Hz (dB)	Equal variances assumed	39.211	.000	-13.773	98	.000	-44.45238	3.22756	-50.85736	-38.04740

	Equal variances not assumed			-19.643	90.291	.000	-44.45238	2.26297	-48.94797	-39.95679
AC 3000 Hz (dB)	Equal variances assumed	16.876	.000	-17.926	98	.000	-48.69048	2.71614	-54.08057	-43.30038
	Equal variances not assumed			-24.302	97.541	.000	-48.69048	2.00358	-52.66676	-44.71420
AC 4000 Hz (dB)	Equal variances assumed	14.863	.000	-16.966	98	.000	-51.62857	3.04309	-57.66749	-45.58965
	Equal variances not assumed			-23.395	96.086	.000	-51.62857	2.20682	-56.00902	-47.24813
Agm slope (dB)	Equal variances assumed	28.001	.000	-6.211	98	.000	-22.41429	3.60909	-29.57641	-15.25216
	Equal variances not assumed			-8.813	91.395	.000	-22.41429	2.54336	-27.46605	-17.36252

Independent Samples Test

STUDY I (NH=30, HI= 73)		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
F500 Hz (dB)	Equal variances assumed	6.404	.013	-2.611	97	.010	-1.90609	.73012	-3.35516	-.45701
	Equal variances not assumed			-3.264	92.654	.002	-1.90609	.58397	-3.06579	-.74639
F3000 Hz (dB)	Equal variances assumed	.109	.741	-10.397	97	.000	-8.73565	.84019	-10.40319	-7.06811
	Equal variances not assumed			-11.197	66.072	.000	-8.73565	.78017	-10.29328	-7.17802
T500 Hz (dB)	Equal variances assumed	1.041	.310	-5.465	97	.000	-6.02899	1.10316	-8.21846	-3.83951
	Equal variances not assumed			-4.814	42.640	.000	-6.02899	1.25236	-8.55522	-3.50275

T3000 Hz (dB)	Equal variances assumed	6.831	.010	-12.441	97	.000	-11.18978	.89943	-12.97490	-9.40467
	Equal variances not assumed			-10.661	40.651	.000	-11.18978	1.04962	-13.31008	-9.06949

Independent Samples Test

STUDY I (NH=30, HI= 73)		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
MCL500	Equal variances assumed	10.561	.002	-4.151	99	.000	-8.90785	2.14598	-13.16595	-4.64976
	Equal variances not assumed			-5.266	94.471	.000	-8.90785	1.69151	-12.26616	-5.54955

MCL3000	Equal variances assumed	9.775	.002	-8.095	99	.000	-16.66803	2.05909	-20.75371	-12.58236
	Equal variances not assumed			-10.753	98.795	.000	-16.66803	1.55013	-19.74390	-13.59217
MCLbb	Equal variances assumed	2.066	.154	-6.573	99	.000	-11.77848	1.79184	-15.33389	-8.22308
	Equal variances not assumed			-7.570	76.916	.000	-11.77848	1.55587	-14.87667	-8.68030
SL500	Equal variances assumed	33.903	.000	-5.683	99	.000	-.16166	.02845	-.21810	-.10521
	Equal variances not assumed			-8.198	90.377	.000	-.16166	.01972	-.20083	-.12249
SL3000	Equal variances assumed	24.528	.000	-7.193	99	.000	-.29026	.04035	-.37033	-.21019
	Equal variances not assumed			-10.508	87.181	.000	-.29026	.02762	-.34516	-.23536
SLbb	Equal variances assumed	5.624	.020	-4.810	99	.000	-.15077	.03134	-.21296	-.08858

	Equal variances not assumed			-5.579	78.266	.000	-.15077	.02702	-.20457	-.09697

Independent Samples Test

STUDY I (NH=30, HI= 73)		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
GP speech	Equal variances assumed	18.320	.000	-11.595	99	.000	-44.63310	3.84919	-52.27074	-36.99546
	Equal variances not assumed			-16.033	97.612	.000	-44.63310	2.78387	-50.15788	-39.10832
GP localisation	Equal variances assumed	16.590	.000	-8.767	99	.000	-32.65164	3.72441	-40.04169	-25.26160

	Equal variances not assumed			-11.675	98.876	.000	-32.65164	2.79662	-38.20083	-27.10246
GP behaviour	Equal variances assumed	32.309	.000	-6.046	99	.000	-26.67347	4.41143	-35.42670	-17.92025
	Equal variances not assumed			-8.570	94.117	.000	-26.67347	3.11252	-32.85335	-20.49360
GP social	Equal variances assumed	16.956	.000	-7.790	99	.000	-31.79014	4.08112	-39.88798	-23.69230
	Equal variances not assumed			-9.624	90.223	.000	-31.79014	3.30321	-38.35232	-25.22796
SRTquiet (dB)	Equal variances assumed	36.095	.000	-12.154	99	.000	-28.52164	2.34678	-33.17817	-23.86512
	Equal variances not assumed			-18.378	76.743	.000	-28.52164	1.55198	-31.61219	-25.43110
SRTstat (dB)	Equal variances assumed	6.997	.009	-8.172	99	.000	-4.88601	.59793	-6.07243	-3.69959
	Equal variances not assumed			-10.275	93.136	.000	-4.88601	.47554	-5.83032	-3.94170

SRTfluc (dB)	Equal variances assumed	11.755	.001	-13.142	99	.000	-11.66845	.88790	-13.43024	-9.90667
	Equal variances not assumed			-16.482	92.738	.000	-11.66845	.70795	-13.07434	-10.26256

Independent Samples Test

STUDY I (NH=30, HI= 73)		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
ILD (dB)	Equal variances assumed	15.398	.000	-11.178	99	.000	-4.51561	.40398	-5.31720	-3.71402
	Equal variances not assumed			-14.328	95.845	.000	-4.51561	.31516	-5.14122	-3.89000
BILD (dB)	Equal variances assumed	.129	.720	-2.959	99	.004	-1.43744	.48574	-2.40126	-.47363

	Equal variances not assumed			-2.453	38.605	.019	-1.43744	.58588	-2.62287	-.25201
LDT (lexical Decision)	Equal variances assumed	.083	.774	6.555	99	.000	.27945	.04263	.19486	.36404
	Equal variances not assumed			6.982	63.321	.000	.27945	.04003	.19947	.35943

Independent Samples Test

STUDY I (NH=30, HI= 73)		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
EffC5	Equal variances assumed	9.978	.002	-2.914	99	.004	-13.20035	4.53071	-22.19027	-4.21044
	Equal variances not assumed			-3.539	87.112	.001	-13.20035	3.73045	-20.61488	-5.78582

EffCmin5	Equal variances assumed	3.516	.064	-3.027	99	.003	-5.56268	1.83775	-9.20918	-1.91618
	Equal variances not assumed			-3.318	67.973	.001	-5.56268	1.67665	-8.90840	-2.21695
EffF5	Equal variances assumed	9.499	.003	-3.457	99	.001	-15.73685	4.55234	-24.76969	-6.70402
	Equal variances not assumed			-4.281	90.646	.000	-15.73685	3.67566	-23.03849	-8.43522
EffFmin5	Equal variances assumed	3.791	.054	-4.126	99	.000	-16.60927	4.02532	-24.59639	-8.62215
	Equal variances not assumed			-4.769	77.593	.000	-16.60927	3.48281	-23.54359	-9.67496
MAAbb	Equal variances assumed	13.777	.000	-2.716	99	.008	-4.27043	1.57258	-7.39077	-1.15009
	Equal variances not assumed			-3.471	95.493	.001	-4.27043	1.23025	-6.71262	-1.82824
MAAhp	Equal variances assumed	8.499	.004	-2.538	99	.013	-5.09907	2.00879	-9.08495	-1.11320

	Equal variances not assumed			-2.883	74.310	.005	-5.09907	1.76859	-8.62282	-1.57532
MAAlp	Equal variances assumed	9.283	.003	-2.360	99	.020	-3.88962	1.64825	-7.16011	-.61912
	Equal variances not assumed			-2.964	92.958	.004	-3.88962	1.31237	-6.49575	-1.28348

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
AC 500 Hz (dB)	Equal variances assumed	12.598	.001	-8.779	114	.000	-24.21795	2.75849	-29.68249	-18.75341
	Equal variances not assumed			-12.623	88.314	.000	-24.21795	1.91863	-28.03063	-20.40527

AC 1000 Hz (dB)	Equal variances assumed	13.385	.000	-10.629	114	.000	-30.60684	2.87943	-36.31097	-24.90271
	Equal variances not assumed			-16.643	107.341	.000	-30.60684	1.83906	-34.25242	-26.96126
AC 2000 Hz (dB)	Equal variances assumed	22.067	.000	-15.856	114	.000	-39.83761	2.51248	-44.81481	-34.86040
	Equal variances not assumed			-26.566	113.956	.000	-39.83761	1.49955	-42.80821	-36.86700
AC 3000 Hz (dB)	Equal variances assumed	17.506	.000	-17.150	114	.000	-46.46154	2.70916	-51.82837	-41.09471
	Equal variances not assumed			-27.101	108.987	.000	-46.46154	1.71436	-49.85935	-43.06373
AC 4000 Hz (dB)	Equal variances assumed	20.545	.000	-18.772	114	.000	-52.70085	2.80738	-58.26224	-47.13947
	Equal variances not assumed			-29.439	107.638	.000	-52.70085	1.79016	-56.24940	-49.15231
F500 Hz (dB)	Equal variances assumed	.816	.368	-4.281	114	.000	-2.74726	.64173	-4.01852	-1.47601

	Equal variances not assumed			-4.671	46.668	.000	-2.74726	.58812	-3.93064	-1.56389
T500 Hz (dB)	Equal variances assumed	.001	.978	-5.243	114	.000	-4.85983	.92694	-6.69609	-3.02357
	Equal variances not assumed			-4.943	37.477	.000	-4.85983	.98311	-6.85095	-2.86871
F3000 Hz (dB)	Equal variances assumed	.628	.430	-11.571	114	.000	-9.87368	.85333	-11.56412	-8.18323
	Equal variances not assumed			-13.249	51.067	.000	-9.87368	.74523	-11.36974	-8.37761
T3000 Hz (dB)	Equal variances assumed	.014	.907	-8.230	114	.000	-9.44162	1.14716	-11.71414	-7.16911
	Equal variances not assumed			-8.872	45.685	.000	-9.44162	1.06424	-11.58423	-7.29901
SRTstat (dB)	Equal variances assumed	14.092	.000	-7.856	114	.000	-4.80530	.61166	-6.01699	-3.59360
	Equal variances not assumed			-12.834	113.181	.000	-4.80530	.37443	-5.54709	-4.06351

SRTfluc (dB)	Equal variances assumed	13.290	.000	-12.457	114	.000	-11.45940	.91993	-13.28178	-9.63703
	Equal variances not assumed			-17.172	79.031	.000	-11.45940	.66732	-12.78766	-10.13114
Lcut 500 (dB)	Equal variances assumed	.122	.727	-.990	114	.005	-2.57596	2.60095	-7.72842	2.57651
	Equal variances not assumed			-1.060	45.151	.005	-2.57596	2.42964	-7.46905	2.31713
Lcut 3000 (dB)	Equal variances assumed	1.473	.227	-3.748	114	.000	-7.30228	1.94852	-11.16229	-3.44227
	Equal variances not assumed			-3.476	36.726	.001	-7.30228	2.10070	-11.55976	-3.04479
slope	Equal variances assumed	25.738	.000	-6.678	114	.000	-47.33974	7.08909	-61.38317	-33.29632
	Equal variances not assumed			-10.695	111.036	.000	-47.33974	4.42643	-56.11098	-38.56851
ILD	Equal variances assumed	5.076	.026	7.446	114	.000	3.48222	.46769	2.55572	4.40872

	Equal variances not assumed			9.596	66.455	.000	3.48222	.36287	2.75781	4.20663
BILD	Equal variances assumed	.261	.610	2.623	114	.010	1.22009	.46514	.29865	2.14152
	Equal variances not assumed			2.268	33.945	.030	1.22009	.53796	.12675	2.31342
LDT (lexical Decision test)	Equal variances assumed	1.569	.213	5.487	114	.000	.02975	.00542	.01901	.04050
	Equal variances not assumed			6.094	48.182	.000	.02975	.00488	.01994	.03957
GP speech	Equal variances assumed	2.297	.132	-10.814	114	.000	-.64816	.05994	-.76690	-.52943
	Equal variances not assumed			-12.742	54.127	.000	-.64816	.05087	-.75014	-.54619
GP localisation	Equal variances assumed	1.028	.313	-5.129	114	.000	-.23558	.04593	-.32657	-.14459
	Equal variances not assumed			-5.240	41.898	.000	-.23558	.04495	-.32631	-.14485

GP behaviour	Equal variances assumed	3.999	.048	-10.611	114	.000	-.71874	.06773	-.85292	-.58456
	Equal variances not assumed			-13.801	67.972	.000	-.71874	.05208	-.82266	-.61481
GP social	Equal variances assumed	4.784	.031	-7.881	114	.000	-.57488	.07294	-.71939	-.43038
	Equal variances not assumed			-10.164	66.550	.000	-.57488	.05656	-.68780	-.46197
SRTquiet (dB)	Equal variances assumed	29.296	.000	-10.807	114	.000	-24.7881	2.2936	-29.3318	-20.2444
	Equal variances not assumed			-17.884	113.881	.000	-24.7881	1.3860	-27.5339	-22.0424

References for appendices

Bench J, Kowal A, Bamford J, (1979), The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. Br J Audiol 13:108-112.

Bilger R.C, Nuetzel J.M, Rabinowitz W.M, Rzeczkowski C, (1984), Standardization of a test of speech perception in noise. J. Speech Hear. Res. 27:32-48.

Brand T, Kollmeier B, (2002), Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests. J Acoust Soc Am 111:2801-2810.

Cox R.M, Alexander G.C and Gilmore C.A, (1987). Development of the connected speech test (CST). *Ear and Hearing*, 8 (suppl): 119S-126S .

Fikret-Pasa. (1993), Speech in noise test (BKB-SIN). http://www.etymotic.com/pdf/bkbsin-user-manual.pdf.

Hagerman B, (1984), Clinical measurements of speech reception threshold in noise. Scand Audiol 13:57-63.

Hagerman B, Kinnefors C, (1995), Efficient adaptive methods for measuring speech reception threshold in quiet and in noise. Scand Audiol 24:71-77.

Hall S.J, The development of a new English sentence in noise test and an English number recognition test. University of Southampton, M.Sc. dissertation, 2006.

Hallgren M, Larsby B, Arlinger S, (2006), A Swedish version of the hearing in noise test (HINT) for measurement of speech recognition. *Int J Aud.* 45:227- 237

Kalikow D.N, Stevens K.N and Elliott L.L, (1977), Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. J Acoust Soc Am 61:1337-1351.

Killion M, Niquette P, Gudmundsen G, Revit L, Banerjee S,(2004), Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners.116(4);2395-2405

Kollmeier B, Wesselkamp M, (1997), Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment. Journal of the Acoustical Society of America 102:2412-2421.

MacLeod A, Summerfield Q, (1990), A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise: rationale, evaluation, and recommendations for use. Br J Audiol 24:29-43.

Nilsson M, Soli S.D, Sullivan J.A, (1994), Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. J Acoust Soc Am 95:1085-1099.

Plomp R, Mimpen A.M, (1979), Improving the reliability of testing the speech reception threshold for sentences. Audiology 18:43-52.

Speech in noise test (SIN). (1993). http://www.etymotic.com/pdf/bkbsin-user-manual.pdf.

Vaillancourt V, Mayer C, Laroche C, Basque C, Nali M, Eriks-Brophy A, Soli S, Giguère C, (2004), Adaptation of the HINT (Hearing In Noise Test) for adult francophone populations. Int J Audiol 44(6):358-69

Versfeld N.J, Daalder L, Festen J.M, Houtgast T, (2000), Method for the selection of sentence materials for efficient measurement of the speech reception threshold. J Acoust Soc Am 107:1671-1684.

Wagener K, Brand T, Kollmeier B, (1999a), Development and evaluation of a German sentence test part I: Design of the Oldenburg sentence test. Z Audiol 38:4-15.

Wagener K, Brand T, Kollmeier B, (1999b), Development and evaluation of a German sentence test part II: Optimization of the Oldenburg sentence test. Z Audiol 38:44-56.

Wagener K, Brand T, Kollmeier B, (1999c), Development and evaluation of a German sentence test part III: Evaluation of the Oldenburg sentence test. Z Audiol 38:86-95.

Wong L.L, Soli S.D, (2005), Development of the Cantonese Hearing In Noise Test (CHINT). Ear Hear 26:276-289.