NORTH ATLANTIC TREATY
ORGANISATION

RESEARCH AND TECHNOLOGY
ORGANISATION

AC/323(HFM-094)TP/53

www.rta.nato.int

RTO TECHNICAL REPORT **TR-HFM-094**

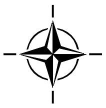
Assessment Methods for Personal Active Noise Reduction Validated in an International Round Robin

(Méthodes d'évaluation des protecteurs auditifs à atténuation active du bruit)

This Technical Report has been prepared as a result of a project on "Assessment of Personal Active Noise Reduction" for the RTO Human Factors and Medicine Panel (HFM-094/TG-028).

Published August 2004

Distribution and Availability on Back Cover



NORTH ATLANTIC TREATY
ORGANISATION

RESEARCH AND TECHNOLOGY
ORGANISATION

AC/323(HFM-094)TP/53

www.rta.nato.int

RTO TECHNICAL REPORT **TR-HFM-094**

Assessment Methods for Personal Active Noise Reduction Validated in an International Round Robin

(Méthodes d'évaluation des protecteurs auditifs à atténuation active du bruit)

This Technical Report has been prepared as a result of a project on "Assessment of Personal Active Noise Reduction" for the RTO Human Factors and Medicine Panel (HFM-094/TG-028).

by

Mr. Brian Crabtree, Dr. Sharon Abel, Canada
Dr. Arman L. Dancer, France
Dr. Karl Buck, Mr. Thomas Wessling, Germany
Dr. Herman J.M. Steeneken (Chairman), Mr. Jan Verhave, The Netherlands
Miss Susan Helen James (Secretary), Dr. Graham Rood, United Kingdom
Mr. Richard McKinley, United States of America

The Research and Technology Organisation (RTO) of NATO

RTO is the single focus in NATO for Defence Research and Technology activities. Its mission is to conduct and promote co-operative research and information exchange. The objective is to support the development and effective use of national defence research and technology and to meet the military needs of the Alliance, to maintain a technological lead, and to provide advice to NATO and national decision makers. The RTO performs its mission with the support of an extensive network of national experts. It also ensures effective co-ordination with other NATO bodies involved in R&T activities.

RTO reports both to the Military Committee of NATO and to the Conference of National Armament Directors. It comprises a Research and Technology Board (RTB) as the highest level of national representation and the Research and Technology Agency (RTA), a dedicated staff with its headquarters in Neuilly, near Paris, France. In order to facilitate contacts with the military users and other NATO activities, a small part of the RTA staff is located in NATO Headquarters in Brussels. The Brussels staff also co-ordinates RTO's co-operation with nations in Middle and Eastern Europe, to which RTO attaches particular importance especially as working together in the field of research is one of the more promising areas of co-operation.

The total spectrum of R&T activities is covered by the following 7 bodies:

AVT	Applied Vehicle Technology Panel
HFM	Human Factors and Medicine Panel
IST	Information Systems Technology Panel
NMSG	NATO Modelling and Simulation Group
SAS	Studies, Analysis and Simulation Panel
SCI	Systems Concepts and Integration Panel
SET	Sensors and Electronics Technology Panel

These bodies are made up of national representatives as well as generally recognised 'world class' scientists. They also provide a communication link to military users and other NATO bodies. RTO's scientific and technological work is carried out by Technical Teams, created for specific activities and with a specific duration. Such Technical Teams can organise workshops, symposia, field trials, lecture series and training courses. An important function of these Technical Teams is to ensure the continuity of the expert networks.

RTO builds upon earlier co-operation in defence research and technology as set-up under the Advisory Group for Aerospace Research and Development (AGARD) and the Defence Research Group (DRG). AGARD and the DRG share common roots in that they were both established at the initiative of Dr Theodore von Kármán, a leading aerospace scientist, who early on recognised the importance of scientific support for the Allied Armed Forces. RTO is capitalising on these common roots in order to provide the Alliance and the NATO nations with a strong scientific and technological basis that will guarantee a solid base for the future.

The content of this publication has been reproduced
directly from material supplied by RTO or the authors.

Published August 2004

ISBN 92-837-1121-1

Assessment Methods for Personal Active Noise Reduction Validated in an International Round Robin

(RTO-TR-HFM-094)

Executive Summary

Purpose

Methods used for the assessment of helmets and headsets equipped with Active Noise Reduction (ANR) are different from the (standardized) methods used for passive systems. ANR systems may introduce electronic noise and possess non-linear behavior at high noise levels. Therefore standard methods like comparison of subjective performance at threshold may not be valid. Alternative methods have been developed and compared in various laboratories.

Methods

The performance of a number of test methods was assessed in an international Round Robin. Five laboratories participated in this test: DRDC Canada; ISL France-Germany; TNO-HF The Netherlands; QinetiQ UK, and AFRL/HECB USA.

The passive and active sound attenuation of five headsets, all equipped with active noise reduction and an intercom, were determined. Several measuring methods were adopted including methods that are based on a human head, artificial head and artificial ear. With a human subject the maximum sound level is restricted due to a possible temporary or permanent hearing loss. Artificial ear and head based methods have no level limitation but may not be representative. This study presents a comparison of the validity of these methods.

The speech intelligibility was predicted from measurements of the Speech Transmission Index (STI) and the Modified Rhyme Test (MRT). The personal comfort and performance of the headsets were evaluated subjectively.

Results

The results on attenuation show that most measuring methods are not significantly different. Also, inter and intra subject variance was small. The measurements covered the attenuation of the passive and active performance of the systems. The results presented in this report are focused on methods to determine the passive and active attenuation and speech communication quality. Also the speech communication measurements with subjects and with an artificial head show a good correlation.

Conclusion

Various measuring methods for attenuation and speech communication quality were validated and found to be very useful for military applications.

Méthodes d’évaluation des protecteurs auditifs à atténuation active du bruit

(RTO-TR-HFM-094)

Synthèse

Objet

Les méthodes d’évaluation des casques et des serre-tête équipés d’un dispositif d’atténuation active du bruit (ANR, Active Noise Reduction) sont différentes de celles utilisées pour les protecteurs passifs. Les dispositifs ANR peuvent en effet produire des bruits parasites dûs au bruit de fond électronique. Ils peuvent aussi se comporter de façon non linéaire aux niveaux élevés. De ce fait, les méthodes d’évaluation standard (méthode subjective de mesure des seuils par exemple) ne sont pas adaptées. Des méthodes alternatives ont été mises au point et testées par différents laboratoires.

Méthode

Les performances de ces méthodes ont été mesurées et comparées par cinq laboratoires: DRDC Canada, ISL France/Germany, TNO-HF The Netherlands, QinetiQ UK, AFRL/HECB USA. Les atténuations passives et actives du bruit de cinq serre-tête équipés d’un dispositif ANR et d’un système de communication ont été mesurées. Pour ce faire on a utilisé des oreilles artificielles, des têtes artificielles et des sujets. L’utilisation de sujets est limitée par le risque pour l’audition aux niveaux de stimulation très élevés. Les oreilles et les têtes artificielles ne présentent pas cet inconvénient mais elles peuvent être non représentatives. Cette étude permet de comparer la validité des différentes méthodes. De plus, l’intelligibilité de la parole a été déterminée au moyen de la technique objective du STI (Speech Transmission Index), et l’ergonomie des serre-tête a été validée par des sujets.

Résultats

Les résultats relatifs à l’atténuation passive et à l’atténuation active des serre-tête testés ici montrent que la plupart des méthodes utilisées ne diffèrent pas de façon significative. De plus, la variabilité inter et intra-individuelle est faible. Les résultats présentés dans ce rapport concernent principalement les méthodes de détermination de l’atténuation passive et active et de l’intelligibilité de la parole.

La corrélation entre les mesures d’intelligibilité de la parole obtenues avec des sujets et les mesures obtenues avec une tête artificielle est bonne.

Conclusion

Différentes méthodes de mesure de l’atténuation des bruits et de la qualité de la communication ont été validées et leur intérêt pour des applications militaires a été démontré.

Table of Contents

Annex F – Contribution of QinetiQ, Farnborough, UK F-1

Annex G – Contribution of AFRL/HECB, Dayton, USA G-1

Preface

Military operations are often conducted under high noise conditions. In order to obtain adequate hearing protection and good speech communication hearing protectors with a communication facility are required. However, hearing protection with standard passive earmuffs is insufficient for many military applications. Noises dominated by low frequencies (propeller aircraft, diesel engines, helicopter rotor, flight deck, high performance aircraft cockpits) require increased attenuation. ANR (Active Noise Reduction) is specifically suited for reducing low frequency noise. Assessment of ANR for military applications is an important topic as the adverse military environment may degrade the performance of an ANR system (e.g. very low frequency rotor-blade noise and/or high overall acoustic intensity may overload a system).

To assess performance test methods for ANR in military applications a Round Robin study was initiated. The study included:

(a) Experiments at five laboratories using the same assessment methods for ANR in military environments,

(b) Comparison of the results and selection from all tests a compact set of assessment methods that encompass the specifications for military use,

(c) Proposals for standardization of selected methods within NATO and also for civil applications.

This report details the results of the study. These results will also be presented in a Lecture Series in 2004 (LS 244) organized by the RTO and disseminated to international standardization organizations (ISO and CEN).

Préface

Les opérations militaires sont souvent conduites dans des environnements très bruyants. Pour protéger l'audition et conserver une bonne intelligibilité de la parole, il est nécessaire d'utiliser des protecteurs auditifs munis d'un dispositif de communication. La protection auditive passive fournie par les serre-tête classiques est souvent insuffisante. Les bruits de basse fréquence (moteurs diesel, rotors d'hélicoptères, moteurs d'avions...) nécessitent une atténuation plus importante. L'atténuation active du bruit (ANR) est particulièrement adaptée pour les basses fréquences. Il est important de pouvoir évaluer cette technique (ANR) pour des applications militaires dont les environnements sonores particuliers peuvent entraîner une réduction inattendue des performances (par exemple: un bruit de très basse fréquence produit par les pales d'un rotor peut excéder les capacités d'un dispositif conçu pour une application civile, et bloquer son fonctionnement).

Une étude visant à déterminer l'intérêt et la validité des méthodes d'évaluation des dispositifs à atténuation active du bruit a été réalisée. Cette étude comprend :

(a) des expérimentations réalisées dans cinq laboratoires, selon les mêmes procédures, dans des environnements sonores spécifiquement militaires,

(b) une comparaison des résultats obtenus et la sélection d'un ensemble de méthodes couvrant les besoins des applications militaires,

(c) une évaluation des résultats et une standardisation des méthodes retenues pour une utilisation militaire (OTAN) et, si possible, également pour une utilisation civile.

Les résultats de l'étude sont contenus dans le rapport. Ces résultats seront présentés en 2004 au cours d'une série de conférences (LS 244) organisées par le RTO. Ils seront également transmis aux organismes internationaux.

Foreword

Adequate hearing protection and efficient speech communication is recognized as a critical capability in most military applications such as vehicle and aircraft operations, command and control, and in the battlefield. Advanced hearing protectors are required for a range of environmental conditions, especially those with extremely high levels of low frequency noise. Passive hearing protectors equipped with an additional active noise reduction system may offer sufficient sound attenuation and suitable speech communication capabilities for these harsh noise environments.

Application oriented assessment is required to guarantee optimal performance. Hence, a study was conducted to assess and select assessment methods for active noise reduction systems.

The study was organized as a Round Robin test where various laboratories performed the same test with the same test material. The reproducibility of the various methods can thus be determined. The laboratories involved in this study were: DRDC Canada; ISL France/Germany; TNO-HF The Netherlands; QinetiQ UK; AFRL/HECB USA and formed special Task Group (HFM-TG028).

The results of the study were presented in this report. The relevant assessment methods for sound attenuation and speech communication are described (chapter 3). The test set-up according to the Round Robin paradigm and the required variables are described in chapter 4. The experimental results and the statistical analysis are detailed and discussed in chapter 5. The conclusions and recommendations are presented in chapter 6.

A lecture series on this topic will be conducted in 2004.

Human Factors and Medicine Panel

CHAIRMAN

Dr. Robert ANGUS
Director General
Defence R & D Canada – Suffield
PO Box 4000 – Station Main
Medicine Hat, Alberta T1A 8K6
CANADA

DEPUTY/VICE-CHAIRMAN

Dr. Jean-Michel CLERE
Chef du Département Sciences Médicales et
Facteurs Humains DGA/DSP/STTC/SH
8, Boulevard Victor
F-00303 Armées
FRANCE

PANEL EXECUTIVE

Col. Carel E.M. BANSE, MA
BP 25
92201 Neuilly-sur-Seine
FRANCE
Tel: +33 1 55 61 22 60/62 Fax: +33 1 55 61 22 98
Email: bansec@rta.nato.int or pelatd@rta.nato.int

MEMBERSHIP OF TASK GROUP HFM-094/TG-028

CHAIRMAN

Dr. H.J.M. Steeneken
TNO Human Factors
Kampweg 5
P.O. Box 23
3769 ZG Soesterberg
THE NETHERLANDS

SECRETARY

Miss Susan Helen James
QinetiQ
Cody Technical Park
Room 2001/A6. Ively Road
Farnborough, Hants, GU14 0LX
UNITED KINGDOM

MEMBERS

Dr. Sharon M. Abel
DRDC Toronto
(Defence Research Development Canada)
1133 Sheppard Ave. West
Toronto, Ontario M3M 3B9
CANADA

Dr. A.J. Brammer
Institute for Microstructural Sciences
National Research Council
Montreal Road
Ottawa, Ontario K1A OR6
CANADA

Dr. Karl Buck
Institut de St. Louis
P.O. Box 34, Rue du Général-Cassagnou
68301 Saint-Louis
FRANCE

Mr. Brian Crabtree PEng
DRDC Toronto
(Defence Research Development Canada)
1133 Sheppard Ave. West
Toronto, Ontario M3M 3B9
CANADA

Dr. Armand L. Dancer
Institut de St. Louis
P.O. Box 34, Rue du Général-Cassagnou
68301 SAINT-LOUIS
FRANCE

Mr. Thomas Wessling
Wehrtechnische Dienststelle für Waffen
und Munition
Postfach 1764
49716 Meppen
GERMANY

Mr. J.A. Verhave
TNO Human Factors
Kampweg 5
P.O. Box 23
3769 ZG Soesterberg
THE NETHERLANDS

Dr. Graham M. Rood
QinetiQ, Cody Technical Park
Room 2001/A6. Ively Road
Farnborough, Hants GU14 0LX
UNITED KINGDOM

Mr. Richard McKinley
AFRL/HECB
Wright Patterson AFB
2255H Street
Wright Patterson AFB
OH 45433-7022
UNITED STATES

Chapter 1 – INTRODUCTION

For many military applications hearing protection with standard passive earmuffs is insufficient. For noises dominated by low frequencies (diesel engines, helicopter rotor) increased attenuation is required. Active Noise Reduction (ANR) is specifically suited for noise reduction at low frequencies. Assessment of ANR for military applications is an important topic as the adverse military environment may cause unexpected reductions in ANR performance (e.g. very high intensity noise and/or very low frequency rotor-blade noise may overload systems.).

The proliferation of personal ANR in various countries has been accompanied by a diverse range of assessment methods and performance criteria. At the AGARD/RTO conference in Copenhagen (1997) a proposal was made to select and develop standard assessment methods focused on the military applications of ANR (Steeneken and Verhave, [27]). This led to an initiative by TNO-HF to set up a test consortium with five internationally recognized research groups working in the field of military ANR applications. These are: DRDC – Canada, Institute of Saint Louis – France/Germany, TNO-Human Factors – the Netherlands, QinetiQ – United Kingdom and AFRL/HECB Wright Patterson AFB – USA.

In 1999 this group compiled an overview of test methods for ANR systems in military applications and also initiated a *Round Robin test* in which a number of selected ANR systems were assessed in all participating laboratories [31]. This study started in 2000 and measurements were finished in 2002.

In order to disseminate results, a special Task Group on ANR assessment under the NATO-RTO Human Factors and Medicine Panel was initiated. Task group HFM-TG028 was formed in 2001 and this report discusses preferred assessment methods for military application of ANR systems.

Chapter 2 – OBJECTIVES

The aim of the study was to select or develop tests for evaluation of ANR systems that are used under adverse military conditions. For this purpose benchmark tests were identified to be performed at all participating laboratories. Additionally, each laboratory had the freedom to perform other tests as long as these were non-destructive or harmful to the systems being tested. In this project the following activities were performed.

(1) Selection of test methods to be included in the benchmark and assessment trials to be performed at all test labs (completed in 1999);

(2) Selection of systems to be included in the benchmark tests and supply by all laboratories of one system to be tested in the Round Robin and one spare (completed in December 1999);

(3) Performing the selected tests on the systems supplied (these tests were completed at all sites in 2002);

(4) Evaluation of the individual laboratory results of the benchmark tests and evaluation of the laboratory specific tests. This analysis is designed to result in the selection of useful tests based on the requirements for military use and on the reproducibility between laboratories. This is the major research goal and was completed in autumn 2003;

(5) The results of the experiments and the lessons learned will be published in an RTO Technical Report. Also an RTO lecture series is scheduled for 2004;

(6) Dissemination in the civil world and standardization is foreseen in ISO and CEN.

Chapter 3 – ASSESSMENT METHODS

Methods used for the assessment of the attenuation of helmets and headsets equipped with ANR are different from the (standardized) methods used for passive hearing protection devices. ANR systems may introduce audible electronic noise and possess non-linear behaviour at high noise levels. Therefore international standardized methods such as comparison of subjective performance at threshold are not applicable (see: EN 352-1, EN 24869-1, ISO 4869-2). As ANR based hearing protectors are often equipped with an electro-acoustic feedback system, it is possible to use these systems for intercom applications. Therefore, assessment of speech communication is included.

ANR systems may give some discomfort to the user and may show instability that can result in a hazardous noise. This requires subjective assessment of users under representative conditions.

For all these topics many test methods were developed in the past. In this chapter we describe briefly the selected methods that were used in the Round Robin test.

3.1 ATTENUATION

The attenuation of a personal hearing protector, equipped with ANR, can be determined subjectively or objectively.

Subjective assessment can be based on matching of two, narrow band, noise signals provided with and without the ANR system activated (see Annex E). A subject can adjust the level of one of the noise signals. The task is to match the two levels for equal loudness (subjective perception of the level). The level difference at equal loudness between the two noise signals represents the contribution of the ANR. This method requires external control of the ANR system. This is not always an option for ANR systems. The advantage of the method is that it may be representative of the attenuation at the position of the eardrum.

We compared this method extensively with two objective methods and concluded that the use of a test microphone in the concha close to the entrance of the earcanal (MIRE, microphone in real ear, see figures 1A, 1B) provided results similar to the results obtained with the subjective level matching method (see Annex E and [35]). Therefore, we used the MIRE-method for the majority of the attenuation measurements (see also [3, 7, 15, 22, 27, 29, 32, 34]). However the MIRE-method requires subjects. In order to get information on the subject dependency (inter subject variability) and on reproducibility (intra subject variability), tests were performed with a number of subjects (typically five), and tests were repeated (typically five replications).

To prevent noise-induced hearing loss, a subject cannot be exposed for a long time to noise levels higher than an 8 hour, energy equivalent sound level (Leq) of 80-85 dBA. Therefore, tests were designed for an artificial ear (or head) in order to assess systems safely at higher levels (MIArtE, microphone in artificial ear, Figure 2). This method may use an artificial ear with a test microphone inside the ear canal with a representative cavity.

The MIRE-method can also be used with an artificial head when a microphone is placed in the concha in a similar position as with subjects (see Figures 1, 2). A simpler method is to use a microphone mounted in a flat plate coupler [3, 6, 31].

Figure 1: Microphone Configuration for the MIRE-Method.

Figure 2: Artificial Ear Mounted in an Artificial Head (MIArtE).

Three objective methods were compared: the artificial ear method with the HMSII.3 (artificial head), the MIArtH method with the same artificial head, and the simple flat plate coupler that consists of a microphone mounted in a dummy head with flat sides.

The objective attenuation was determined in 1/3 octave frequency bands by subtracting the 1/3 octave spectra obtained with the test microphone, with and without the hearing protector placed on the head. The preferred frequency range was 12.5 Hz to 20 kHz. (Note: Most acoustic test facilities cannot provide a diffuse acoustic field at frequencies below 50 Hz and generation of sufficient sound pressure levels in a large volume test chamber at low frequencies is difficult and can be expensive).

The dynamic range of the system (maximum level of the acoustical noise outside the hearing protector to the minimum noise level under the hearing protector) determines the range of attenuation values that can be obtained accurately. With the use of an active system the minimum noise level will increase due to the electronic self-noise of the system. The noise measured under an earcup mounted on the head of a subject may additionally introduce physiological noise. A typical environment for measurements performed with subjects or with a dummy head is shown in figure 3.

Figure 3: Artificial head placed in a high noise room, designed for generation of a diffuse sound field (max. 120 dBA). Two of the five high power loudspeakers (back view) are visible at the left. In the center of the wall a sound absorption module is placed. The small loudspeakers are not used (these are available for hearing threshold experiments).

Attenuation at High (Impulsive) Noise Levels

In the military environment, crews are regularly exposed to munitions noise and hearing protectors should therefore also be evaluated under these conditions. The peak levels needed for a realistic evaluation (150 dB to 190 dB) of such an exposure are created using explosive charges. This technique allows for the production of Friedlander waves with different levels and duration in the free field (see Annex D).

In order to avoid overload of the microphones, due to the amplification of the Transfer Function of the Open Ear (TFOE), it has been necessary to use a procedure for the insertion loss (IL) measurement that does not expose unprotected “artificial ears” to impulse noise with peak pressure levels higher than 160 dB. The procedure used is based on knowledge of the TFOE, which can be determined with lower levels. Once this function is known, the IL can be calculated:

$$\mathbf{IL(f) = L_{free\ field}(f) - L_{protected\ ear}(f) + TFOE(f)}$$

The pressure time histories in the free field and at the ear underneath the earcup are measured simultaneously using the set up shown in figures 4 and 5 and converted to 1/3 octave-band levels. Then the insertion loss is calculated using the above formula.

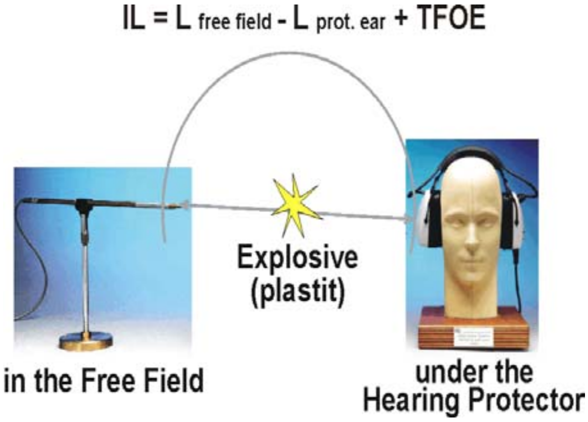


Figure 4: Set-up for the IL Measurement using Explosive Charges.

Figure 5: Experimental Set-up for Impulse Measurement at ISL.

As the signals created by explosive charges are highly reproducible, this method can also be used to show the influence of the ANR system on the time signal. If the pressure time history underneath the earcup is measured with the ANR system switched on-off, the difference between these two measurements represents the pressure signal produced by the ANR system.

The use of explosive charges as an impulse noise source associated with the above measurement method, allows the evaluation of the effectiveness of ANR hearing protectors for all noise levels that can be found in a military environment. It also allows the limits of the electro-acoustic system and its behaviour when overloaded to be determined.

At high continuous (stationary) noise levels a simple flat plate coupler in a small volume enclosure can be used. Such a system consists of a microphone mounted in a dummy head with flat sides. However, this type of fixture can introduce errors due to the increase in the trapped volume due to lack of a pinna or pinna simulator and the flat sides artificially reducing the probability and size of acoustic leak when compared to real heads. The dummy head can be placed in a small volume that allows generation of high noise levels, although the resulting sound field is not diffuse. With the system shown in figure 6 levels up to 130 dBA can be achieved [34].

Figure 6: Dummy Head with "Flat-Plate" Microphone in a Volume Suitable for Relatively High Noise Levels (130 dBA).

3.2 SPEECH COMMUNICATION

The speech communication quality of an ANR system mainly depends on the method used for the injection of the speech signal in the system. Some systems make use of an advanced method that prevents suppression of the speech due to ANR. However, some systems inject the speech signal directly into the ANR loop and make use of a correction filter to decrease distortion.

As the design of the ANR system, the speech injection method, and the suppression of background noise determine the speech communication quality, it is important to assess the speech intelligibility in representative conditions.

This assessment can be done with subjective measurement (by making use of standardized speech intelligibility test materials with live talkers and listeners) or predicted by objective methods such as the Speech Transmission Index (STI).

Standardized subjective test materials normally use simple test words (CVC words, Consonant-Vowel-Consonant) or rhyme words (Modified Rhyme Test, MRT, comparison of similar words with a different initial consonant or central vowel).

Modified Rhyme Test

The Modified Rhyme Test (MRT, [9]) is a standardised subjective speech intelligibility test described in ANSI S3.2 to determine the response to consonants and vowels. For this purpose test words are presented from which subjects select the most likely word. This type of test was performed at AFRL/HECB (see Annex G).

The experimental design included five signal to noise ratios (SNRs), five ANR headsets, ten listeners, and two talkers. One male and one female talker was used. A total of 50 trials were conducted. Five listeners participated per session. The recorded MRT speech stimuli were presented via a PC sound card. For all conditions the ANR system was “ON”. A research intercommunication system was used with a 10 kHz bandwidth. The AFRL/HECB facility (Figure 7) was used to generate the ambient sound field, to generate the MRT stimuli, and collect the subject responses to the stimuli.

Figure 7: High Noise Listening Room at AFRL/HECB. The room is used for MRT listening tests.

MRT Maximum Likelihood

The MRT maximum likelihood method was used at DRDC ([8, 9, 36], Annex C). The MRT test material consists of 50 word sets containing six rhyming words each, for a total vocabulary of 300 words. In this implementation, the 50 word sets or “screens” selected at random are presented on a computer monitor as a 2 x 3 matrix, as one of these words selected at random is presented as an audio stimulus within a carrier phrase. The level of the stimulus is chosen to produce a desired SNR with respect to a continuous noise masker, and the result of the run (all 50 screens) is the percentage discrimination proficiency achieved at that SNR.

The Maximum Likelihood Procedure or MLP when embedded in the MRT provides a means of varying the stimulus level with respect to the masking noise. The procedure adaptively finds the signal level necessary for the subject to achieve a given discrimination proficiency, in this case, 60 per cent. It assumes that the shape of the MRT psychometric function is already known, and in effect estimates the position of the function along the SNR axis. As the run progresses, the routine selects the levels having the maximum likelihood of achieving the target proficiency as based on the response history within the run, thus it quickly converges to the associated level. In comparative ANR system assessments, lower stimulus levels are associated with better performance in a given noise field.

Objective Intelligibility Assessment

A standardized objective method to predict speech intelligibility is provided by the STI (Speech Transmission Index [10, 28]), and by a revised version STIr ([33]; IEC 60268-16).

The STI is obtained by applying a specific speech-like test signal at the audio input and by analysis of this transmitted test signal through the same measuring microphone as used with the MIRE attenuation measurements with the same subject.

In general the STI is measured as a function of the environmental noise level for a condition with and without the ANR system (see Annex E). The STI methods accounts for: noise, band-pass limiting, non-linear distortion, and temporal distortion (reverberation, echoes, AGC).

The objective was to measure the speech intelligibility versus SNR and to compare the STI predictions with the MRT scores for the five different ANR headsets. The approach was to fix the speech presentation levels at 75 dB SPL and to use five different ambient noise levels 90 dB, 95 dB, 100 dB, 105 dB, and 110 dB to achieve a range of SNR's at the listeners ears.

3.3 SUBJECTIVE ASSESSMENT

The comfort and performance of, an ANR system may be assessed additionally to the physical specifications. This may include stability, noise from the system, acoustical performance, acceptability of using the system for a long period, etc. The subjects were asked to score their impressions on a subjective scale. For example, a five point scale may be used with a range: excellent, good, fair, poor, bad.

Stability

Verification that no instability or oscillations are detected during the donning and doffing of the headset. A cautionary note: Instability may result in high signal levels. Subjects need to be protected against a high noise dose, for instance by using an earplug. It has to be verified that closure of the earcanal does not affect the acoustical conditions which may influence the results. However, it is rare that closure of the earcanal affects ANR performance for circumaural earmuff type ANR systems.

Acoustical Performance

The subjective appreciation of the acoustical signals (noise and, if applicable, the speech signal), is determined during representative usage.

Acceptance

The subjective appreciation of wearing the system is determined. This may include judgment of weight, pressure of the earmuff on both sides of the head, ease of placing the system on the head, and the use in combination with other systems such as a helmet, gas mask, oxygen mask, spectacles, etc.

Chapter 4 – ROUND ROBIN VARIABLES

A Round Robin test implies that several laboratories perform the same tests with the same systems. Such an experimental design provides information on the reproducibility of the test methods included in the Round Robin. Each participant was asked to supply one ANR system and a spare system of the same type, one for inclusion in the test and the spare system in case of failure of the primary system.

In order to get a representative sample of ANR systems, systems were selected that included a wide performance range. The systems included in the tests are listed in Table 1. Photographs of the systems are shown in Annex A.

Table 1: Type Specification of the ANR Systems used for the Round Robin

System	Provided by
Bose Pru-57C56	AFRL/HECB
In-house system	QinetiQ
Telemit	ISL
Telex type ANR-tm	DRDC
In-house system	TNO-HF

There are various assessment methods (described in chapter 3), and it was suggested that all laboratories assess the attenuation of the systems with passive, and active attenuation. The common experiments included measuring methods that use a subject in order to consider the fit of the earmuff to the human head. Some countries noise regulations do not allow use of human subjects. In this case an artificial head or artificial ear had to be used. In some countries the use of high noise levels with subjects, with a possible risk of hearing damage, is not allowed. For this reason participants were encouraged to use subjects as well as artificial heads or artificial ears in order to compare results of the different methods. Table 2 provides an overview of the assessment methods that were used in the five participating laboratories. Basically, most laboratories performed tests with the MIRE-method and three laboratories with artificial replacements of a subject. Also the application of high noise levels and impulsive noise was addressed. Communication quality was tested at three sites including subjective and objective intelligibility tests. The human factors issues were examined in two laboratories. The range of tests presented above allows for a robust analysis when all individual results are compared.

In order to facilitate the comparison of the results of the five laboratories and to simplify analysis methods, ANOVA (analysis of variance), linear regression analysis, and correlation coefficients were computed. The results were compiled in a standardized spreadsheet consisting of the original raw data and a common method for obtaining mean values of the results.

Table 2: Overview of the Measurements Performed by Each Participant of the Round Robin Test

Round Robin Measurements	DRDC	ISL	TNO-HF	QinetiQ	HECB
Attenuation					
MIRE inter subject var.	5 subj.		5 subj.	6 subj.	10 subj.
MIRE intra subject var.	3 replica		2 subj. 5 replica	3 replica	3 replica
MIArtE (artificial head)			X		
MIArtE (artificial head high level stationary)	120 dB 32-125Hz		126 dBA wide band		
MIArtE (flat plate coupler)	X				
MIArtE (artificial Ear)	3 replica	10 replica	X	90-115 dB 63-250Hz	
MIArtE (artificial Ear) Impulse		3 replica			
Self-noise	X	X	X	X	
Speech communication					
Frequency Response	X		7 octaves	X	
Subjective MRT, MLH (60% score Lnoise)	5 subj.				
Subjective MRT					X
STI MIRE (function Lnoise)	5 subj.		5 subj.		
STI artificial head (function Lnoise)			X		
Human Factors					
Comfort stability comments/questionnaire	X		X		

Chapter 5 – EXPERIMENTAL RESULTS AND DISCUSSION

5.1 ATTENUATION MEASUREMENTS

Four of the five participating laboratories provided attenuation results based on the MIRE-method for five ANR systems (see Annex A) and 5-10 subjects (Table 2). The data were presented with the specific spreadsheet to obtain a standardized data presentation. This spreadsheet provided the attenuation curves for the total attenuation, passive attenuation and the ANR contribution. The validity of the results was verified with respect to the self noise spectrum of the systems and measuring set-up (see 5.1.5). From these basic results a selection was made for the various statistical analyses and conversions to insertion loss (see 5.1.1 and 5.1.2).

The measurements with artificial heads or flat plate were performed by four of the participating laboratories. In total seven methods were evaluated. The results were collected with the same generic spreadsheet as used for the MIRE-method results. The evaluation of these results is described in section 5.1.3.

Some of the attenuation measurements were performed at a high noise level with stationary and/or impulsive noise. These results are described in section 5.1.4.

5.1.1 MIRE-Method, Inter Subject Variation

The fit of a headset may differ from person-to-person. This implies that leakage as well as trapped volume underneath the earcup may occur and that both the passive and active attenuation may vary. Hence, the spread of the attenuation values obtained with different subjects is a measure for the “goodness of the fit” and the inter-individual differences in morphology. In figure 8 an example of the mean attenuation as a function of frequency is given for two measured conditions (passive and total insertion loss) and the calculated attenuation contributed by the ANR system. The vertical bars represent the standard deviation based on the number of measurements (5 to 10 different subjects). It is clear that the standard deviation values are small for the active attenuation and larger for passive and total attenuation. This may be explained by the small sensitivities of the ANR to acoustic impedance changes. For the passive attenuation, leakage is one of the important parameters while for an ANR system the effect of leakage is smaller.

The main aim of this study was to assess the reproducibility of a measuring method applied at various measuring sites. For this purpose the results of the participating laboratories were compared using an ANOVA (analysis of variance). The variables for this analysis were: 4 laboratories, 5 headsets, left-right earcup, 5 subjects (used as replica), and “the attenuation curve”. This last variable consists of 33 1/3 octave-band levels over the frequency range from 12.5 Hz to 20 kHz. This makes the analysis complicated as 33 different analyses have to be performed which are not directly related to the application of a hearing protector. Therefore, the A-weighted IL (insertion loss[1]) was considered. This describes the decrease of the level of a specific noise by application of a hearing protector. As the IL-A (A-weighted insertion loss) depends on the type of noise; six different military noises that cover a range of frequency spectra and military environments were used. The spectra of these noises are given in Annex B together with their A-weighted level.

---

[1] Insertion loss (IL): The mean algebraic difference in decibels between the A-weighted sound pressure level, measured by the microphone in a specified sound field under specified conditions, with the hearing protector absent, and the A-weighted sound pressure level with the hearing protector fitted, with all other conditions identical. (definition in agreement with standard EN352-1).

The A-weighted levels under a hearing protector are determined in two steps:

(1) Calculation of the difference in dB of the 1/3 octave noise level and the estimated attenuation. This estimated attenuation is defined by the mean 1/3 octave attenuation minus one standard deviation (corresponding to ISO 1999), for all 1/3 octave frequencies,

(2) Calculation of the overall A-weighted sound level from the 33 1/3 octave sound pressure levels as determined under (1).

In this way the attenuation curve of a hearing protector can be specified by a single number.

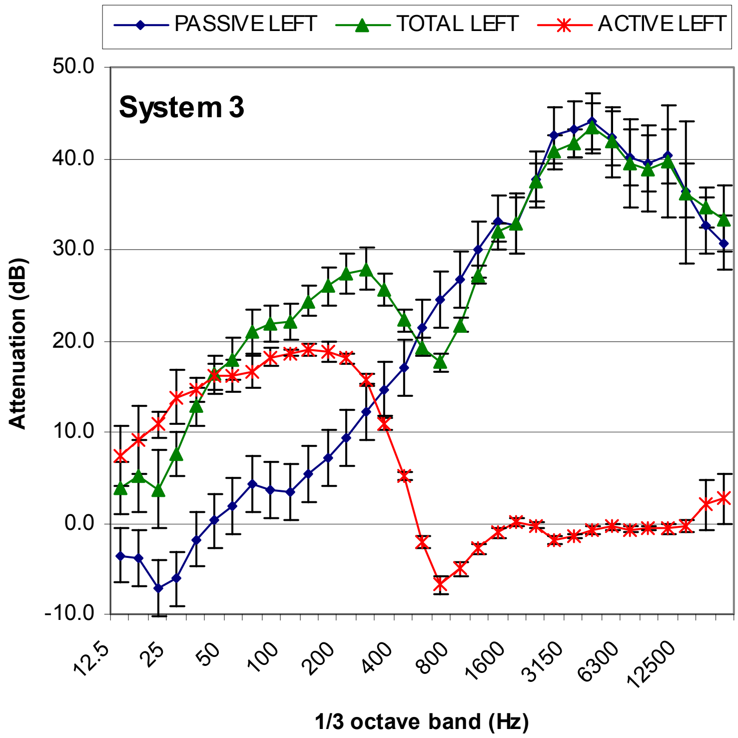


Figure 8: Example of the Total, Passive and ANR Attenuation of Headset 3 as a Function of Frequency. The curves present the mean attenuation obtained for 5 subjects measured on the left earcup. The vertical bars indicate the standard deviation.

Two ANOVA's were performed, based on these predicted A-weighted noise levels. One ANOVA based on the passive attenuation and one ANOVA based on the total attenuation (passive + active).

The following variables were used: *laboratories, left* and *right earcups, type of noise* and *headset*. The repeated measures for 5 subjects were used as replica.

Passive Attenuation

For the passive attenuation data the variables: laboratories, noise, and systems provided very significant differences ($p<0.01$). The left-right earcup variable provided a significant difference ($p=0.04$).

A Tukey test (detailed analysis within one variable) was applied for the variable laboratories. In Table 3 the significance of the difference between each pair of laboratories is given. The mean value of the A-weighted noise levels is also given for each lab.

Table 3: Significance of the Difference (p-values) of the Predicted Noise Levels under the Hearing Protector for the *Passive* Attenuation and Four Laboratories. The mean noise levels are also given. (p<0.01 very significant **, p<0.05 significant *, p>0.05 not significant ns.)

	DRDC	QinetiQ	TNO-HF	AFRL/HECB
DRDC		0.024 *	0.002 **	0.000 **
QinetiQ	0.024 *		0.550 ns	0.012 *
TNO-HF	0.002 **	0.550 ns		0.152 ns
AFRL/HECB	0.000 **	0.012 *	0.152 ns	
Mean Level (dBA)	93.3	91.6	90.8	89.7

The maximum difference of the mean noise level is between DRDC and AFRL/HECB and is 3.6 dBA. There is also a significant difference between the results of QinetiQ and AFRL/HECB. The results between DRDC and AFRL/HCEB are very significant and correspond with the maximum level difference of 3.6 dBA. Although the maximum difference is small, a difference of 3 dBA represents a factor of two in exposure time of a subject to a noise. The difference obtained above may be introduced by various factors such as the instruction of the procedure for placing the headset on the head of the subject (to minimize leakage), the particular group of subjects, and the measuring set-up.

The difference between the left and right earcup for the average of all other variables (*labs, systems, noises*) is only 0.6 dBA (left 91.6 and right 91.0 dBA).

Total Attenuation

For the total attenuation a similar analysis was performed as was conducted for the passive attenuation results. For this condition a significant difference between the laboratories (p=0.023) and the left-right condition (p=0.047) was found. The differences for the noises and systems were very significant. The results of the Tukey test for the variable *laboratories* are given in Table 4.

Table 4: Significance of the Difference (p-values) of the Predicted Noise Levels under the Hearing Protector for the *Total* Attenuation (Passive and Active) and Four Laboratories. The mean noise levels are also given. (p<0.01 very significant **, p<0.05 significant *, p>0.05 not significant ns.)

	DRDC	QinetiQ	TNO-HF	AFRL/HECB
DRDC		0.371 ns	0.126 ns	0.017 *
QinetiQ	0.371 ns		0.899 ns	0.329 ns
TNO-HF	0.126 ns	0.899 ns		0.715 ns
AFRL/HECB	0.017 *	0.329 ns	0.715 ns	
Mean Level (dBA)	86.6	85.3	84.8	84.0

All differences, except one (DRDC-AFRL/HECB) are not significant. The maximum difference between the mean predicted noise levels is 2.6 dBA. This value is smaller than the value obtained for the passive attenuation. As already shown in figure 8 the variance of the mean attenuation is smaller for the total attenuation than for passive only and smallest for the active attenuation.

Active Attenuation

Additional to the comparison of measuring methods for the passive and total attenuation, that is based on insertion loss for a specific noise spectrum, a comparison was made for the active attenuation data. In this analysis we reduced the data of the active attenuation response to three parameters. These were: the maximum active attenuation value, the minimum active attenuation value (this value is normally a negative number and is related to the stability), and the equivalent bandwidth expressed in a number of 1/3 octave-bands.

The maximum attenuation value was determined with the 1/3 octave-band attenuations over the frequency range from 25 Hz to 1 kHz, the minimum attenuation value in frequency range from 200 Hz to 1.6 kHz, and the equivalent bandwidth in the frequency range from 25 Hz to 1 kHz. The equivalent bandwidth refers to the number of 1/3 octave-bands at the maximum attenuation value with an equal, intensity based, transfer as the squared (intensity) surface under the actual attenuation curve.

We calculated the correlation coefficients, for a first order regression between the individual laboratory results. These were based on the three parameters for five systems and the active attenuation response derived with the MIRE method. The correlation coefficients are given in table 5. All coefficients are above 0.99, which indicates a close relation between the individual laboratory results. We did not find any systematic differences between the individual laboratory results based on the first order regression lines.

Table 5: Correlation Coefficients between Four Laboratories and for the Three Parameters that Describe the Active Attenuation. (Based on 5 systems and the mean active attenuation derived with the MIRE method for 5 subjects)

	DRDC	QinetiQ	TNO-HF	AFRL/HECB
DRDC		0.993	0.994	0.992
QinetiQ	0.993		0.998	0.994
TNO-HF	0.994	0.998		0.994
AFRL/HECB	0.992	0.994	0.994	

Analysis of Block Frequency Data

By close inspection of the data it can be seen that analysis by blocks of frequency data is feasible rather than by individual frequency bands.

A detailed inspection of the data shows that a dependent variable, fit, is one of the main sources of variance. In this particular analysis approach errors due to fit have been minimised by using the results of the two headsets that provided the most consistent fit results (Headsets 1and 5 which can be confirmed from variance data and the subjective perceptions from section 5.3). This allows analysis of data where the variance due to fit is minimised and a better comparison can be made solely between measurement techniques of the participating laboratories.

By inspection of the passive MIRE data, it can be seen that the passive attenuation characteristic falls into three main frequency bands and this is in line with the three theoretical acoustic mechanisms that control the production of the attenuation curves (see Figure 9). These are:

(1) the low frequency results, which are controlled by the mass, volume damping characteristics of the ear cup (up to 300-400 Hz);

(2) the mass and density characteristics of the earmuff , which provide the mass line, related to the attenuation from 400 to around 1 to 2 kHz);

(3) the high frequency end controlled by the damping material fitted in the earcup (above 2 kHz).

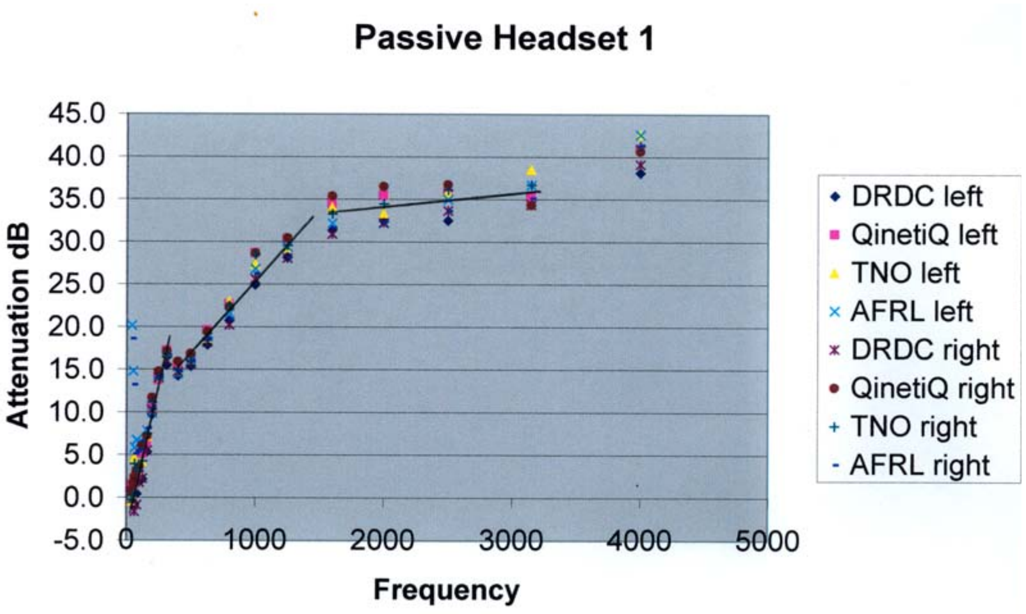


Figure 9: Passive Attenuation for Left and Right Earcup of System 1, Showing the Three Acoustic Mechanism.

Figure 10 shows the combined left and right cup data for Headset 1.The data here can be seen to be tightly bunched and well suited to a linear regression analysis. Similarly, the data for headset 5 shows a similar trend.

In the higher frequency band from 400 Hz to 1600 Hz, in the area where the mass law lines are in control, the data for headsets 1 and 5 show similar characteristics of tight bunching of results and are conducive to comparative analysis by linear regression. In a similar way to the passive data the approach can be applied to the total attenuation (passive + active) data.

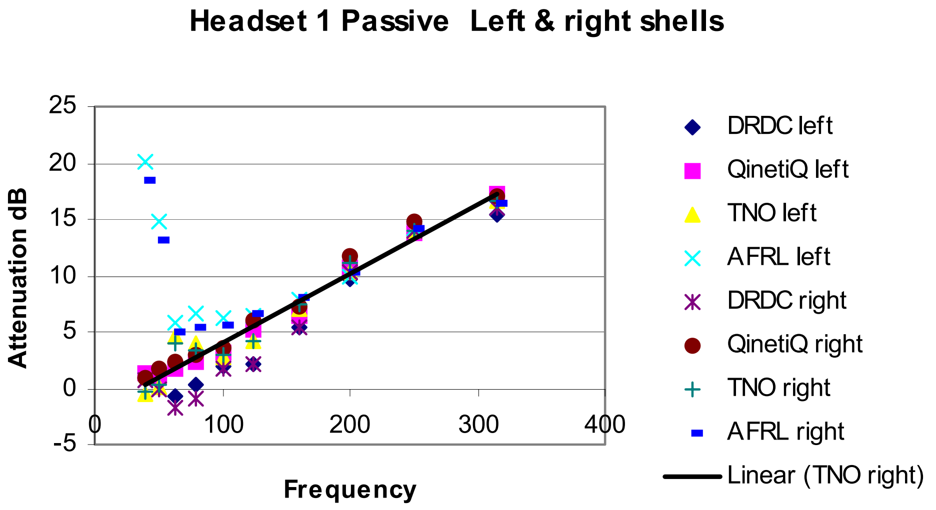


Figure 10: Passive (Left and Right) Cups of Headset 1 Showing the Potential for Linear Regression Analysis up to 315 Hz.

The full range of result falls again into three distinctive bands (at least up to 2 kHz) and this is shown clearly in figure 11 for headset 1.

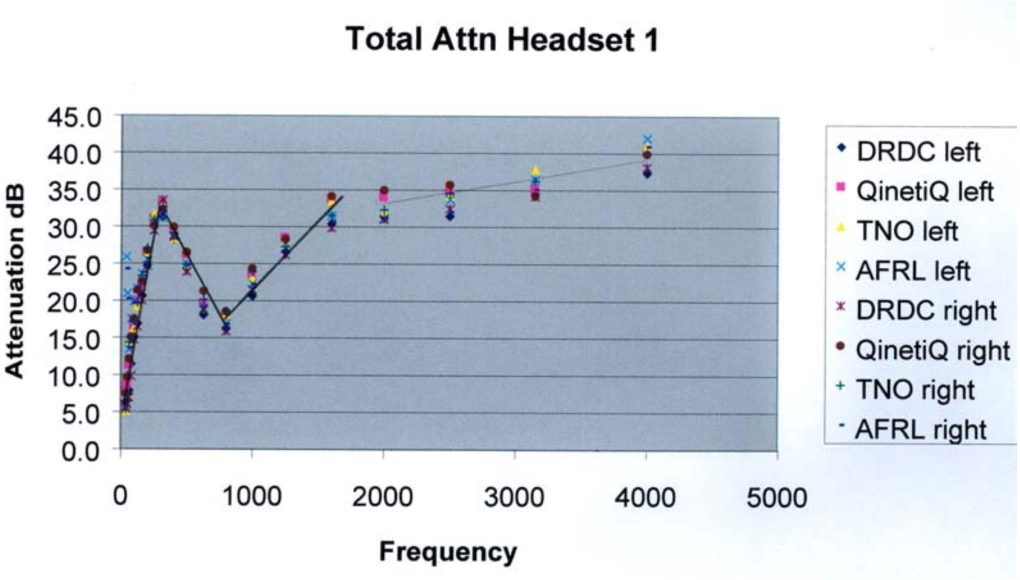


Figure 11: Total Attenuation Showing the Acoustic Mechanism Boundaries.

Thus, for this block frequency analysis, the frequency bands that were used to compare the data obtained from the laboratories involved were:

- Passive Attenuation: 40 – 315 Hz and 400 – 1600 Hz
- Total Attenuation: 40 – 315 Hz, 400 – 630 Hz and 800 – 1600 Hz

Each frequency band was analysed using a linear regression, from which the regression constants (slope and intercept) and the correlation coefficients could be statistically compared.

Total Attenuation

For the corresponding total (passive + active) attenuation this data is encompassed in the frequency bands 40-315 Hz, 400 – 630 Hz and 800-1600 Hz.

Figure 12 shows the data for headset 1 for the 40-315 Hz band and the corresponding regression lines. From these data sources the correlation coefficients are significant and that there are no significant differences between regression constants. Similar inspection of the data and the regression analyses for the two upper bands (400 – 630 and 800 – 1600 Hz) draw similar conclusions to the low frequency band.

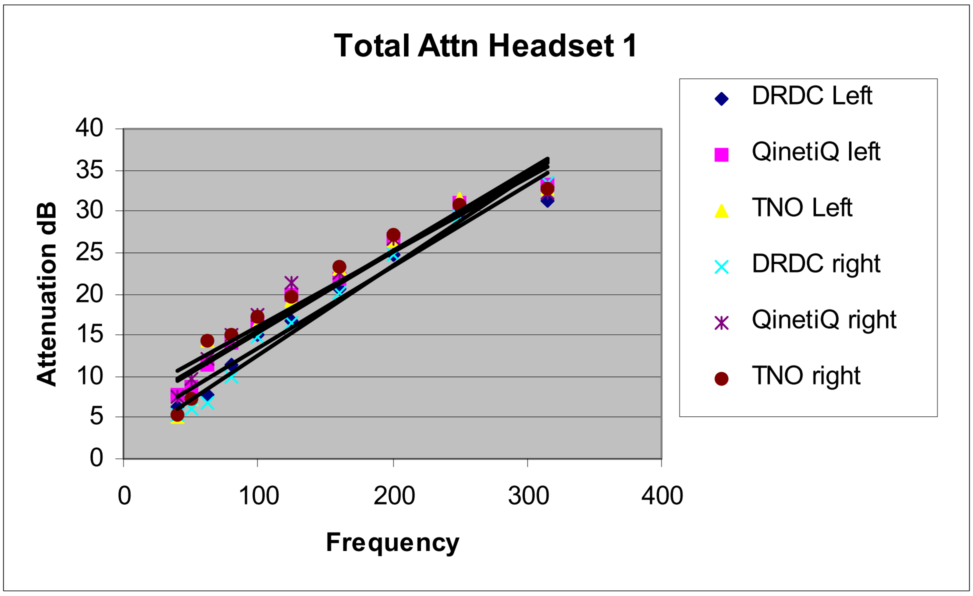


Figure 12: Linear Regression Lines for Total Attenuation; 40 to 315 Hz.

From this data and the subsequent analysis it can be seen that, in terms of block or combined data, the results between laboratories are not significantly different.

At this experimental level, the results clearly show that, as a method of analysis in those frequency bands that are important in the acoustic attenuation of headsets and helmets, it works well. The fit of the data to linear regression is significant and the equations of the regression line, both for slope and intercept, are not significantly different between laboratories.

5.1.2 MIRE-Method, Intra Subject Variation

Comparison of inter-subject variance with intra-subject variance provides information about the necessity to measure with many subjects and a few or no replica's or with few subjects and many replica's. For this reason some experimental designs include 5-10 replicas and a limited group of subjects.

In figure 11 an example is given for the mean attenuation curves obtained with one subject and repeated measures performed at different sessions and at different days. In figure 8 similar results are given, however these curves are for five subjects and no replications.

Smaller standard deviations are observed in figure 13 than in figure 8. This is not surprising as the fit on the same head may introduce a smaller variance than the fit of the same hearing protector on different heads.

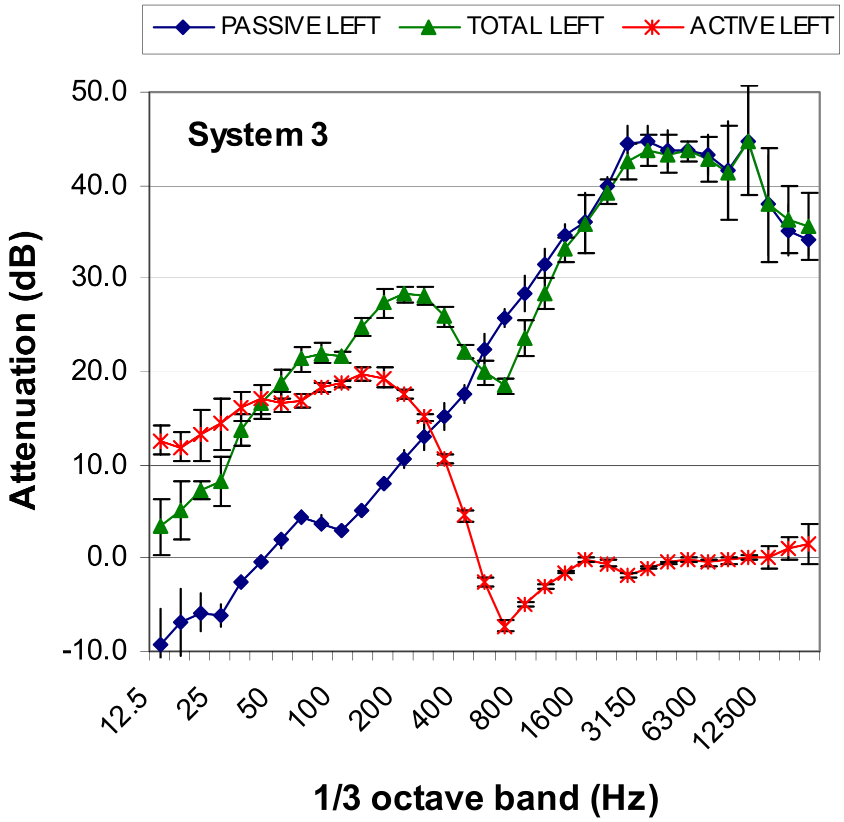


Figure 13: Example of the Total, Passive and Active Attenuation of Headset 3 as a Function of Frequency. The curves present the mean attenuation obtained for 5 replica's of the same subjects and the left earcup. The vertical bars indicate the standard deviation.

An ANOVA was performed based on the 5 systems, left/right ear, and 5 repeated measures. The inter-subject results were used as the error estimate. No significant differences were found between repeated measures for the same subject (p>0.140).

5.1.3 Artificial Ear, Flat Plate, and Artificial Head Measurements

As described in section 3.1, measurements were performed with artificial ears, flat plates and artificial heads. There were 7 different measuring set-ups using all five hearing protectors and both the left and right earcup. As for the MIRE-method the passive, total, and active attenuation was determined. However, repeated measurements were not conducted at all laboratories. The methods that were applied are shown in Table 6.

Table 6: Overview Objective Attenuation Measurements with Artificial Means (1-7) and with Subjects (8-11)

Label	Laboratory	Method	Remarks
1	DRDC	Flat plate	
2	DRDC	Artificial Head	
3	ISL	Artificial Head	With B&K ear simulator, pinna and circumaural flesh simulation (Figure 5)
4	QinetiQ	Artificial Ear	B&K ear simulator
5	TNO-HF	Artificial Head	With MIRE, (Figure 3)
6	TNO-HF	Artificial Head	With Artificial Ear, (Figure 2)
7	TNO-HF	Flat Plate	High Level Dome (Figure 6)
8	DRDC	MIRE	See 5.1.1 (Figure 1)
9	QinetiQ	MIRE	See 5.1.1 (Figure 1)
10	TNO-HF	MIRE	See 5.1.1 (Figure 1)
11	AFRL/HECB	MIRE	See 5.1.1 (Figure 1)

The results of the seven artificial head related measuring methods were compared with an ANOVA similar to the analysis described in section 5.1.1. The variables for this analysis were: 7 methods, 5 headsets, 6 noises, and left-right earcup. If there were no repeated measurements, the analysis was performed with the *left-right* variable as replica and an analysis with *left-right* as random factor. The A-weighted IL (insertion loss) was calculated for the six noises. These data were used for the ANOVA. The mean attenuation values were not corrected by subtracting one standard deviation as there were no repeated measurements. In Table 7 and 8 the significance, based on a Tukey test for the seven measuring methods, is given for both the passive attenuation and for the total attenuation. The ANOVA was based on the analysis with the *left-right* variable as random factor. The analysis with the *left-right* variable as replica provided similar results. The Tables also give the mean noise level for six noises, 5 systems and left-right earcup. The mean noise levels for the MIRE-method with subjects (described in section 5.1.1) are also given.

Table 7: Significance of the Difference of the Predicted Noise Levels under the Hearing Protector for the *Passive* Attenuation and Seven Measuring Methods. The mean noise levels are also given (p<0.01 very significant **, p<0.05 significant *, p>0.05 not significant ns).

Method	1	2	3	4	5	6	7	8	9	10	11
1		ns	ns	ns	**	**	ns				
2	ns		*	ns	**	**	ns				
3	ns	*					ns				
4	ns	ns	**				ns				
5	**	**	**	**			**				
6	**	**	ns	**	*		*				
7	ns	ns	ns	ns	**	*					
Mean Level dBA	87.3	87.3	89.8	86.4	94.1	91.2	88.6	92.9	91.4	90.7	89.4

Table 8: Significance of the Difference of the Predicted Noise Levels under the Hearing Protector for the *Total* Attenuation and Seven Measuring Methods. The mean noise levels are also given (p<0.01 very significant **, p<0.05 significant *, p>0.05 not significant ns).

Method	1	2	3	4	5	6	7	8	9	10	11
1		ns	**	ns	**	**	**				
2	ns		**	ns	**	**	**				
3	**	**		**	**	**	**				
4	ns	ns	**				**				
5	**	**	**	**			**				
6	**	**	**	**	**		**				
7	**	**	**	**	**	**					
Mean Level dBA	81.6	81.9	84.3	81.6	87.8	85.6	83.1	86.2	84.7	84.6	83.8

Active Attenuation

The active attenuation was characterised by three parameters (maximum, minimum and equivalent bandwidth) similar to the method described in section 5.1.1. The correlation coefficients for these parameters between the results of the seven methods based on artificial heads and for the MIRE-method (4 laboratories) are given in Table 9.

Table 9: Correlation Coefficients between 11 Measurements of the Active Attenuation (described by three parameters and for 5 systems)

Method		1	2	3	4	5	6	7	8	9	10	11
DRDC 1, FP	1		0.960	0.983	0.995	0.984	0.988	0.994	0.977	0.987	0.984	0.987
DRDC 2, MiArt	2	0.960		0.974	0.959	0.976	0.963	0.956	0.954	0.960	0.961	0.965
ISL, Art Head	3	0.983	0.974		0.988	0.994	0.987	0.983	0.986	0.987	0.987	0.992
QinetiQ, MiArt	4	0.995	0.959	0.988		0.987	0.989	0.993	0.984	0.990	0.987	0.992
TNO1, MiArt	5	0.984	0.976	0.994	0.987		0.993	0.981	0.987	0.990	0.990	0.992
TNO2, Head	6	0.988	0.963	0.987	0.989	0.993		0.987	0.985	0.993	0.990	0.991
TNO3, FP	7	0.994	0.956	0.983	0.993	0.981	0.987		0.971	0.982	0.980	0.986
DRDC, MIRE	8	0.977	0.954	0.986	0.984	0.987	0.985	0.971		0.993	0.994	0.992
QinetiQ, MIRE	9	0.987	0.960	0.987	0.990	0.990	0.993	0.982	0.993		0.998	0.994
TNO, MIRE	10	0.984	0.961	0.987	0.987	0.990	0.990	0.980	0.994	0.998		0.994
WPAFB, MIRE	11	0.987	0.965	0.992	0.992	0.992	0.991	0.986	0.992	0.994	0.994	
	mean	0.984	0.963	0.986	0.986	0.988	0.987	0.981	0.982	0.987	0.987	0.988

5.1.4 High Level Measurements

As stated in chapter 3 high noise levels may overload the electronic system of the ANR, and hence introduce distortion components rather than reducing the noise level. Therefore, measurements were performed with a stationary noise test signal at a level up to 126 dBA and with impulsive noise with a peak level up to 170 dB.

The results for stationary noise and the method described in chapter 3 (see Figure 6) are given for two ANR systems. System 1 (Figure 14) shows overload at a level of 126 dBA, the active attenuation reduces with 7-8 dB and distortion components are obtained in the frequency range between 1250 Hz and 3000 Hz. For system 5 (Figure 15) no overload is observed at a test signal level of 126 dBA and only small distortion components are obtained above 3000 Hz. For both systems similar results were obtained for the left and right earcups.

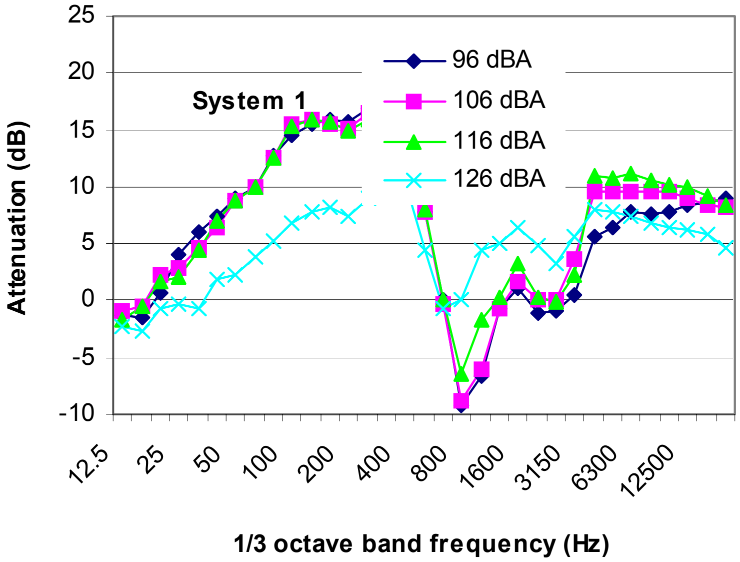


Figure 14: Active Attenuation for Frequencies between 12.5 Hz to 20 kHz (1/3 octave intervals) Derived with Test Signal Levels from 96 dBA up to 126 dBA. This graph represents the results for the left earcup of system 1.

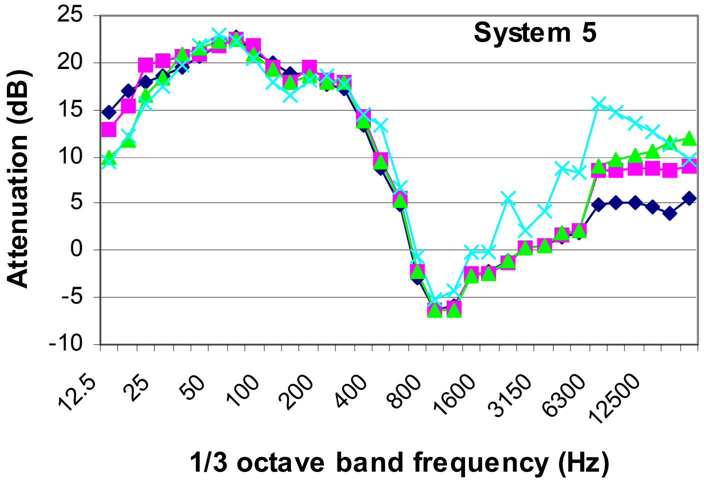


Figure 15: Active Attenuation for Frequencies between 12.5 Hz to 20 kHz (1/3 octave intervals) Derived with Test Signal Levels from 96 dBA up to 126 dBA. This graph represents the results for the right earcup of system 5.

The overload response of an electronic system is different for stationary and impulsive signals, an impulse may be less stressful than a continuous signal. However, in a military environment impulsive noise levels are much higher than stationary noise levels. Therefore, we also studied the overload response for impulsive noise up to 170 dB. The measuring method is described in chapter 3 and Annex D. In figures. 16 and 17 the active attenuation derived with impulsive stimuli with levels from 130 dB to 170 dB is given for the same systems as used for the stationary condition. The responses for the lower levels are similar with figures. 14 - 15. For system 1 a decrease of the active attenuation is observed from a 150 dB stimulus at levels from 160 dB the ANR effect has disappeared. System 5 can handle somewhat higher levels, at a 160dB stimulus a small ANR effect (5-6 dB) is obtained whilst at a stimulus level of 170dB no effect is obtained. The ranking of the two systems was the same for both noise conditions.

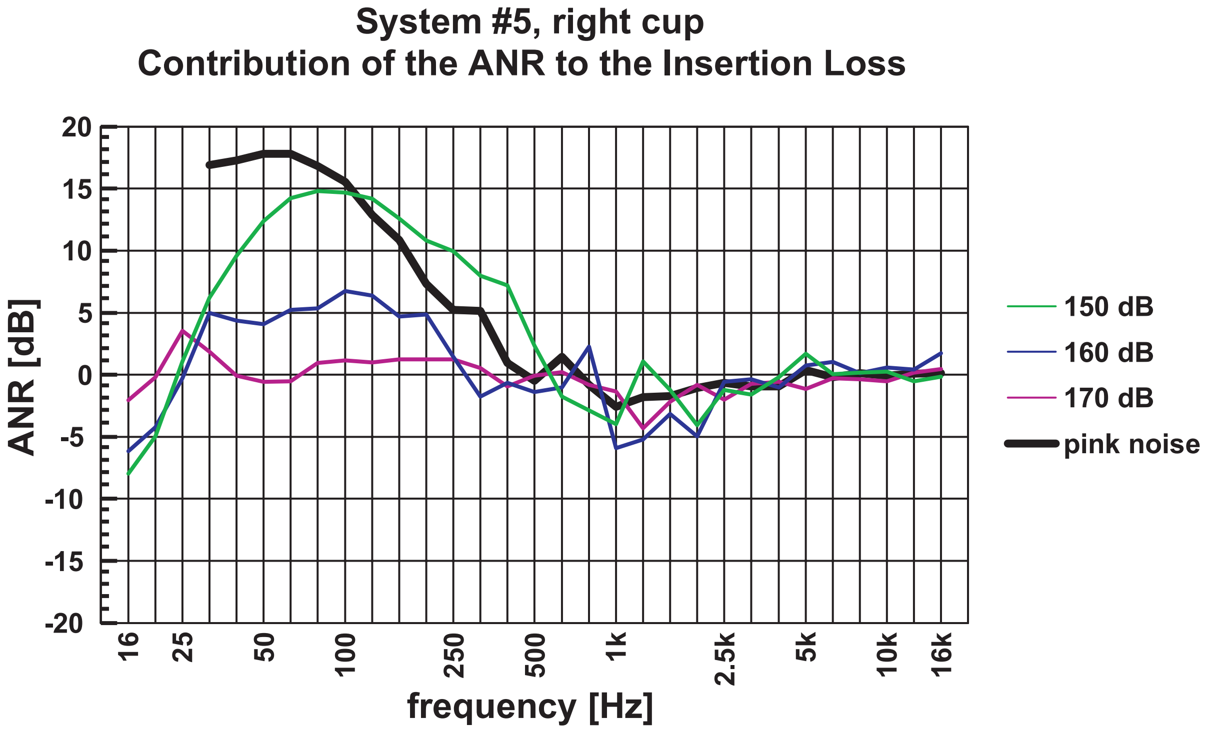


Figure 16: Active Attenuation for Frequencies between 16 Hz to 16 kHz (1/3 octave intervals) Derived with an Impulse from 130 dB up to 170 dB and for Pink Noise. This graph represents the results for the left earcup of system 1.

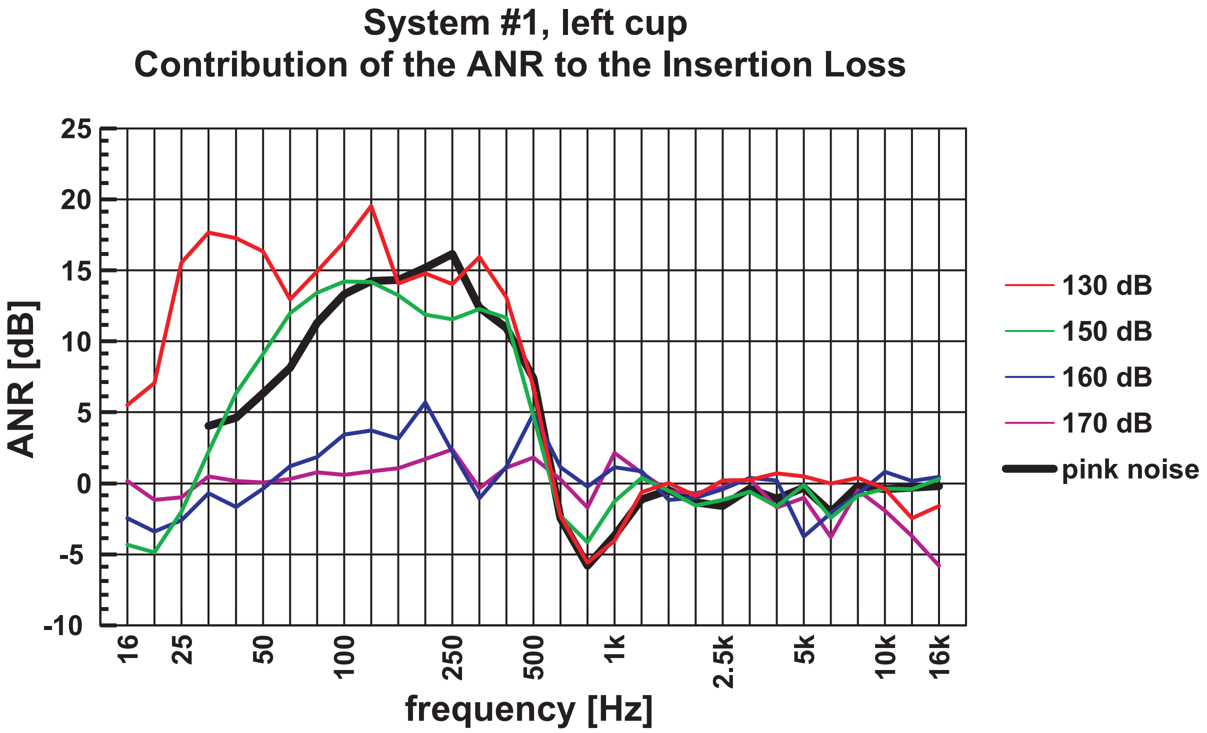


Figure 17: Active Attenuation for Frequencies between 16 Hz to 16 kHz (1/3 octave intervals) Derived with an Impulse from 130 dB up to 170 dB and for Pink Noise. This graph represents the results for the left earcup of system 5.

5.1.5 Self Noise

The dynamic range of the objective measurement of the attenuation of a hearing protector is defined by the maximum level of the test signal, the noise floor introduced by the measuring equipment, the active hearing protector, and the physiological noise from the subject.

The maximum test signal level (normally a wide band noise signal) should not exceed an 8 hour equivalent noise level (Leq8h) of 80-85 dBA at the earcanal of a subject in order to prevent a hearing loss.

With the assessment of active systems the noise floor is normally determined by the electronics of the system under test and sometimes by the physiological noise from the subject. A typical spectrum for the internal acoustical noise of an ANR system is given in figure 18. At lower frequencies (with smaller efficiency of the ANR) a higher internal noise level is observed. The mean self noise spectrum including the physiological noise from 5 subjects is given in figure 19. The vertical bars represent the standard deviation for the five observations. An increase of approximately 10 dB is obtained at very low frequencies.

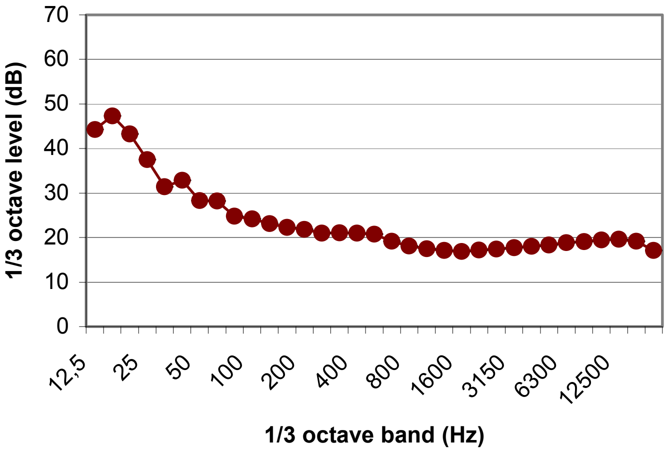


Figure 18: Self-Noise Spectrum Determined on an Artificial Head (no Physiological Noise).

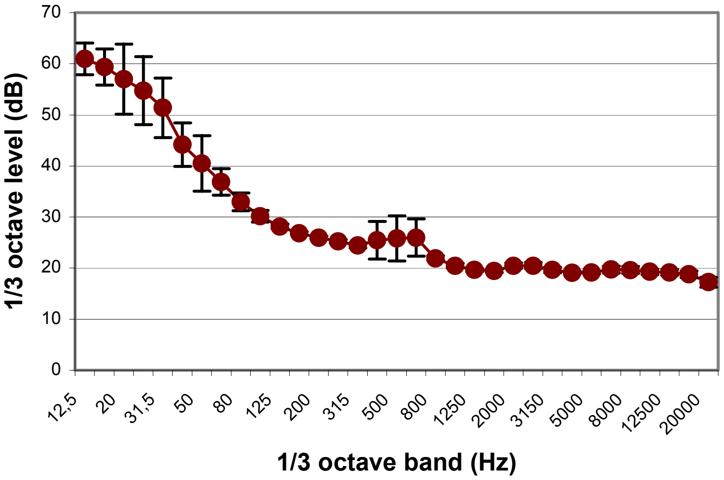


Figure 19: Self-Noise Spectrum Determined on the Head of a Subject, including Physiological Noise (Mean for 5 Subjects).

For all measurements performed in this study the self-noise spectra were determined and used for checking the validity of the calculation of the attenuation. This procedure was included in the spreadsheet described in section 5.1.1.

Only on a few occasions did a warning show that the calculated attenuation value in a certain frequency band was unreliable. Such a condition was due to a high attenuation value that caused a test signal level under the hearing protector of the same order as the self-noise level. It occurred at very low frequencies, where the self-noise spectrum level is highest and for system five which offers a relatively high active attenuation at these frequencies. The self-noise spectrum levels for the different laboratories were all very low.

5.1.6 Discussion

The primary goal of this study was to determine the validity of assessment methods for hearing protectors equipped with active noise reduction. In a practical situation this performance is defined by the *combination* of the passive and active attenuation. In this study various measuring methods for the attenuation that is provided by ANR headsets were validated. For this purpose the evaluation of identical parameters obtained at different laboratories were compared. For a representative sample of active headsets (5 systems) the passive, total and active attenuation was determined with different measuring methods. These measurements were all based on the objective methods using both subjects and artificial heads.

The results for the passive attenuation determined with the MIRE-method are reported in section 5.1.1. The relevant variable is the reproducibility between the laboratories. In a comparison between these laboratories it was found that for the six possible pair-wise comparisons three significant difference were found. The largest difference being 3.6A dB (significant) and the smallest 0.8 dBA (not significant). These levels express the mean difference of the A-weighted sound level for 5 systems and 6 representative noise spectra. Such a variance may be due to the uncontrolled parameters “leakage” and “morphology”. The use of different subjects and different procedures to fit the headset on the head will introduce such an effect. Further inspection of the passive data shows that the attenuation curve divides into three bands. This is in line with the three theoretical acoustic mechanisms that control the passive sound attenuation (see Figure 9).

For the comparison of the total attenuation (passive and active) no significant differences for five of the six pair-wise comparisons between the labs were found (see Table 4). This can be explained by the contribution of the active attenuation. The active attenuation transfers control of the A-weighted energy to the part of the spectrum that is less sensitive to leakage. This brings the data closer together. The maximum difference of the insertion loss between two labs is 2.6 dBA (just significant) and the minimum is 0.5 dBA (not significant).

A comparison of the active attenuation data only was expressed by the correlation coefficient between the labs these were all above 0.992.

The conclusion can be drawn that no major differences between the labs were found for the MIRE-method applied with subjects. However, a careful fit of the system on the head is required and it is obvious that for practical applications adequate fitting instructions will be required.

The comparison of methods based on artificial ear, flat plate and artificial head indicated that these methods correlate well with the MIRE-method. Correlation coefficients between these methods and the MIRE results are all above 0.95. However, the MIRE-method with 5 subjects provides a standard deviation that is related to the individual fit for subjects. This standard deviation is used to predict the insertion loss (see example in section 5.1.1). Single measurements on an artificial head do not provide an estimate of the spread. The use of an artificial head has the advantage that there is no risk of hearing loss for subjects. Systems can be tested to very high levels with stationary or impulsive noise.

In Table 3 and Table 4 the mean predicted noise levels for five systems and six noises were given. The passive attenuation ranges from 89.7 to 93.3 dBA while for the same systems with active noise reduction switched on the attenuation ranges from 84.0 to 86.6 dBA. This shows a gain of 5.7 to 6.7 dBA. The dependency of the “IL-A” on the noise type is not shown here. It is clear that for different types of noise, the “IL-A” can be considerably different for the same device.

5.2 SPEECH COMMUNICATION

5.2.1 MRT in Combination with Noise

The MRT according to the maximum likelihood procedure was performed at DRDC Toronto (see Annex C). The subjects were placed in a noise level of 90 dBA. The target MRT score was 60%.

The speech level at which this score was derived is determined. The mean speech level for 5 subjects and for the 5 systems considered in the test is given in Table 10. Hence, the lower the speech level the better the performance with respect to intelligibility of the system. The rank ordering of the performance according to this test is also given in Table 10.

Table 10: Speech Level in dBa for a MRT-Score of 60% at a Noise Level of 90 dBa. The test was performed for system 1 to 5 (a lower speech level indicates a better performance).

System	Speech Level in dB for (MRT= 60%)	Rank order
1	59.4	2
2	60.2	4
3	66.7	5
4	60.1	3
5	58.1	1

At AFRL/HECB the MRT score was determined for four systems as a function of the noise level. The test is described in Annex G. For the test, two speakers were used one male and one female. In Table 11 the mean MRT scores for 5 (well trained) listeners are given for four of the five systems that were included in the test (system 3 was not available). The female scores are in general slightly higher than the males scores for conditions with a higher noise level. This may have been due to the use of a single male and female speaker.

The rank order, similar to the DRDC results, is also given in Table 11. The rank order of the two methods is quite similar, only rank 2 and 3 are swapped. A systematic difference may have been introduced by using just one speaker for each gender at the full MRT test, however the two speakers were selected to be representative of the approximate mean of the population for their respective genders.

Table 11: Mean MRT Scores for One Male and One Female Speaker and 5 Listeners. There were four systems (1, 2, 4, 5) involved in this test. The rank order for male speech at a noise level of 90 dBA is given.

Noise Level	**90dB**	**95dB**	**100dB**	**105dB**	**Rank order**
System 1	MRT %	MRT %	MRT%	MRT%	
Male	96.2	95.2	89.7	83.0	3
Female	97.6	93.3	93.3	86.3	
Mean	96.9	94.2	91.5	84.6	
System 2					
Male	94.2	88.0	76.7	57.0	4
Female	92.3	89.2	85.8	71.0	
Mean	93.3	88.6	81.3	64.0	
System 4					
Male	97.1	92.6	87.8	76.2	2
Female	98.1	95.0	90.2	80.1	
Mean	97.6	93.8	89.0	78.2	
System 5					
Male	98.1	94.7	90.9	80.6	1
Female	97.4	97.4	94.2	87.3	
Mean	97.7	96.0	92.6	83.9	

5.2.2 STI$_r$ in Combination with Noise

At TNO-HF the STIr was measured as a function of a back-ground noise level in the high noise room shown in figure 3. The performance of the ANR systems may be subject dependent, therefore five subjects were used to determine the inter subject variability. For each subject the STIr was determined for the left and right earcup separately.

The noise used for this experiment had a spectrum similar to the long-term average speech spectrum, and was derived from the NATO-RTO noise-ROM [26]. The test signal level was adjusted to 70 dBA at a position close to the entrance of the earcanal. The transmitted test signal with the noise under the ANR earcup was measured with the MIRE method and analyzed by the STI analyzer.

In figures 20 and 21 the mean STIr for five subjects and the corresponding standard deviation was given as a function of the noise level (blue curve). In general an STIr value above 0.40 is required for a fair communication quality. This performance was obtained at noise levels around 90-100 dBA for the systems used in this test. It should be noted that a level of the test signal higher than 70 dBA resulted in a higher noise level for the same STI value (if the system was not overloaded).

Measurements with the artificial head were performed under the same conditions as with human subjects. The results (one measurement per condition) are also given in figures 20 and 21 (purple curves).

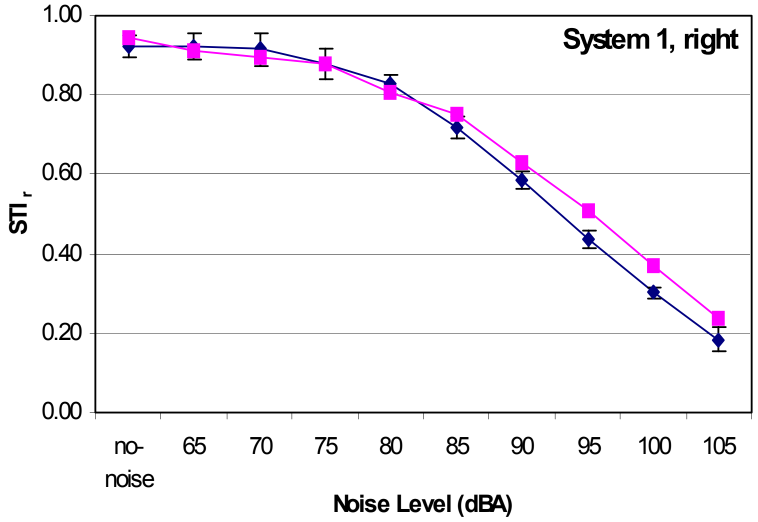


Figure 20: STI$_r$ as a Function of the Background Noise Level for System 1 (right earcup).

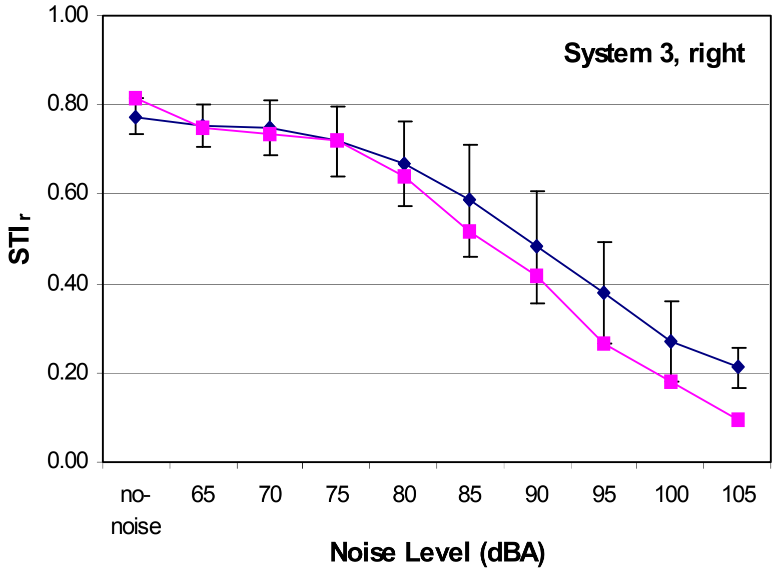


Figure 21: STI$_r$ as a Function of the Background Noise Level for System 3 (right earcup).

The results obtained with human subjects fit quite well with the artificial head results however, the results obtained with system 3 show a much larger standard deviation than those of system 1. This may be explained by leakage of system 3 and was in correspondence with the results obtained for the attenuation as presented in figure 8 (see Annex F).

A similar comparison with STI measurements for subjects and an artificial head were performed at DRDC. For these test three conditions were used: no noise, a noise level of 90 dBA and a noise level of 84 dBA. The comparison of the data points derived with subjects (MIRE) and the artificial head (Artf) are given in figure 22. The figure shows a good relation between both measuring methods.

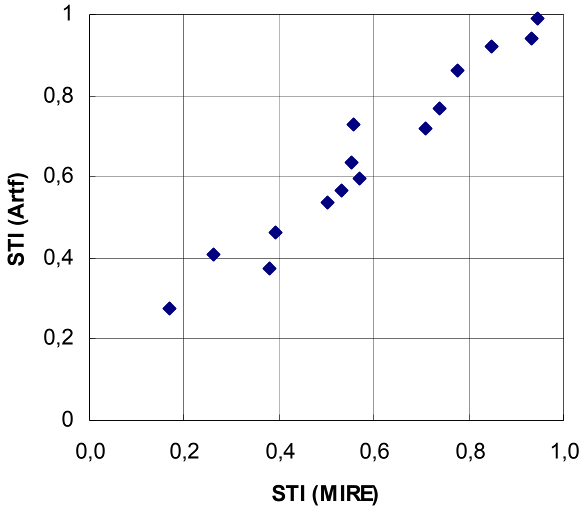


Figure 22: Comparison of STI Measurements Obtained with Subjects (MIRE) and with an Artificial Head (Artf).

5.2.3 Discussion

The mean scores for the five systems of the two MRT related experiments and the STI_r are summarized in Table 12. The rankorder of MRT and STI is equal, while for the maximum likelihood MRT only one confusion is obtained.

Table 12: Overview of Comparable Score of MRT and STI

System	HECB	DRDC	TNO	Rank order
Unit	MRT %	Speech level	STI_r	
1	60	59	0,60	2
2	39	60	0,40	5
3	not available	67	0,50	4
4	53	60	0,57	3
5	64	58	0,61	1

The MRT measurements performed at AFRL/HECB provided MRT scores as a function of the signal-to-noise ratio, similar conditions were obtained for the STI measurements at TNO. Therefore, the relation between STI and MRT could be plotted (see Figure 23, blue data points). This graph is similar to the standardised graphs (Figure E4) given in Annex E. The data points show a monotonic increasing relation. The curve shows a shift to lower STI values which was *not* typical of the original literature in the area but has been consistent with the performance of well trained, experience subjects. Therefore, we plotted the original (House et al., [9]) results in the same graph and an increase of 0.1 STI at a similar MRT was obtained. The difference between the House scores and the MRT scores from this study may be due to the method of measuring speech levels and/or the training/experience level of the subjects.

The saturation of the MRT versus STI is related to the limited response set of Rhyme test in general. This is also illustrated in figure E4. The redundant SRT (simple sentences) shows a small range between unintelligible and a high score, while nonsense words provide a more sensitive test but it is more difficult to accurately administer the test. Objective predictive measures of speech intelligibility (STI, SII) consider the effective signal-to-noise ratio and predict over a wide range.

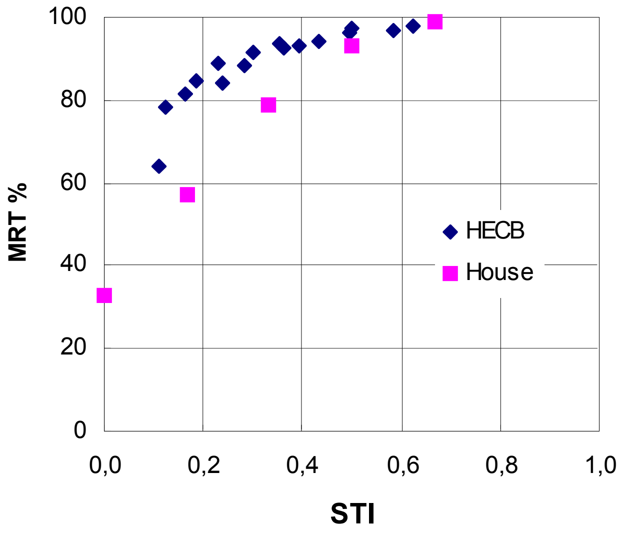


Figure 23: MRT Score as a Function of the STI$_r$ Value. The blue data points refer to the AFRL/HECB scores and the purple data point refer to House et al. [9].

5.3 SUBJECTIVE ASSESSMENT

Several subjective assessment sessions were held at 2 laboratories. The subjective assessment procedure was not standardized. In general subjects were asked to give their opinion on some specific aspects such as: fit of the headset, adjustment, the noise reduction, etc. The best we could do was to report on the individual experiments and to compare the subjective rankordering.

5.3.1 TNO-HF Results

Five subjects performed the subjective assessment. All subjects were familiar with ANR based personal hearing protection systems. The subjects had to judge four aspects of comfort and two on the ANR effect. Each aspect was qualified on a five point scale ranking: bad (1), poor (2), fair (3), good (4), excellent (5).

The subjects were asked to place the system on their head and to walk around for at least 15 minutes.

The questions raised were:

A Examine the fit of the seal of the earcup around your ear. Pay attention to leakage.

B Examine the fit of your ear inside the cup. Pay attention to pressure on your ears.

C Examine the pressure of the headband.

D Examine the fit of the system on your head. Pay attention to the stability of the fit.

E Examine the ANR effect when moving the head. Pay attention to side effects, noise in silence, stability (including warbling, rumble, motor boating, and oscillations).

F Examine the ANR effect without moving the head. Pay attention to side effects, noise in silence, stability (including warbling, rumble, motor boating, and oscillations).

Questions A-D were performed with a background noise of 100 dBA. Questions E-F were performed with and without the background noise.

The mean scores and the corresponding standard deviations for the 6 questions and the 5 systems are given in Table 13.

The mean scores vary between 3.0 and 4.0, corresponding with fair and good respectively. The judgments were performed within a relatively short time period, it may well be that using systems for a full working day would have given lower scores as the effect of weight and pressure on the head would have been increased.

Table 13: Mean Scores (μ) of the Subjective Assessment Judgments and Standard Deviation (s) for the 5 ANR Systems.

	System 1		System 2		System 3		System 4		System 5	
Question	**M**	**s**	**μ**	**S**	**M**	**s**	**μ**	**s**	**μ**	**s**
A	4.2	0.84	2.2	0.45	3.2	0.84	3.6	0.55	3.8	0.45
B	4.0	0.71	2.8	0.84	3.4	0.89	3.6	0.55	2.8	0.84
C	3.2	0.45	2.4	0.89	3.2	0.84	3.2	0.84	3.2	1.1
D	3.4	0.55	2.8	0.84	2.8	0.84	3.2	0.84	3.6	0.55
E	4.6	0.55	3.6	1.34	3.2	1.64	2.4	1.67	4.2	0.84
F	4.6	0.55	4.2	0.84	3.6	1.34	3.3	1.64	4.6	0.55
Mean A-F	4.0		3.0		3.2		3.2		3.7	
Rankorder	1		5		3/4		3/4		2	

5.3.2 Results of DRDC

Subjects were asked to evaluate the systems at the end of each session for MIRE testing and communication measurements. Two headsets were evaluated in each experimental session that took about two hours to complete.

The subjects scored three aspects on a 5-point scale (none-to-severe): howling, clicking and squealing. The scores (1-5) were converted to the TNO scaling (5-1).

In Table 14 the mean scores for the three aspects (A = howling, B = clicking, C = squealing) and the corresponding standard deviation are given. Also the mean of all three aspects and the corresponding rankorder is given.

Table 14: Mean Scores (μ) of the Subjective Assessment Judgments and Standard Deviation (s) for the 5 ANR Systems.

	System 1		System 2		System 3		System 4		System 5	
Question	**M**	**s**	**μ**	**S**	**μ**	**s**	**μ**	**s**	**μ**	**s**
A	4.8	0.45	4.8	0.45	4.4	0.89	4.2	1.8	4.8	0.45
B	4.0	0.71	4.0	1.0	4.6	0.55	4.2	0.4	4.6	0.55
C	4.8	0.45	4.8	0.45	4.8	0.45	3.6	1.95	4.8	0.45
Mean A-C	4.5		4.5		4.6		4.0		4.7	
Rankorder	3/4		3/4		2		5		1	

The query included a verbal report on the subjective performance. The general opinion on the five systems was:

System 1

The headset received positive comments referring to overall comfort. However, one subject found the headband slightly uncomfortable while another found the temperature became quite warm under the earcups. Two subjects noted an effective seal, another two subjects noticed a decrease in voice quality when the ANR was switched on. All subjects noticed the headset generated a click when the ANR was turned on (associated with the activation of the power supply).

System 2

Four of the five subjects commented on the discomfort they experienced from the headband of headset 2. One subject noticed a slight buzzing when the ANR was on. Another subject thought the passive attenuation of the headset was poor. Finally, one subject noticed vocal differences between the ANR on and off conditions as well as a lot of low frequency noise even when the ANR was turned on.

System 3

All subjects found the headset quite heavy and the headband particularly uncomfortable. Two subjects commented on difficulty in achieving a good seal. Comments also revealed that speech became muffled when the ANR was switched on.

System 4

Two subjects noticed that the earcups of the headset squeal, howl or “talk to each other” when the ANR is turned on and the headset is fully removed from the head. One of the subjects rated the headset severe for both howling and clicking for this reason. Most subjects found the headset quite comfortable; there was only one comment of slight discomfort on the top of the head from the headband.

System 5

Most subjects found headset 5 comfortable with the exception of pressure on the ears from small earcups. All subjects noticed that there was a large difference when the ANR was on.

5.3.3 Discussion

The subjective assessment was performed by subjects who are familiar with ANR devices. There was a wide variation is size, seals around the ear, weight and objective performance. However, the assessment results vary in a small range (for TNO good-fair-poor, for DRDC excellent-good). We could not find any correlation with the total sound attenuation. It means that this type of assessment may touch different aspects that may be relevant for a qualification for use over a longer period of time.

We suggest that a (standardized) test could be developed that includes comfort of use for a longer period of time, the presence of unwanted effects (instability, clicking, etc), and the ease of use in combination with other systems (gasmasks, helmets, spectacles, etc). The introduction of some reference systems is advised as it will allow comparison of results with other devices using the same baseline or reference.

Chapter 6 – CONCLUSIONS & RECOMMENDATIONS

6.1 ATTENUATION MEASURING METHODS

6.1.1 MIRE Measurements

From both the frequency based block data results and the A-weighted predicted overall levels, the differences of the results between the labs are small and generally statistically not significant. It is clear, from the complete range of data across headsets, that the most significant variable was the quality and consistency of fit of the headsets. Where the fit was good (e.g. Headset 1 and 5) and the acoustic leakage reduced to a minimum, the data across labs was all tightly bunched, for both the passive and total (passive + active) attenuation cases. Clearly in these cases the differences in measurements, using the same technique across the labs, are well within experimental error. However, in acoustic testing of this type (i.e. the testing of the attenuation characteristics of headsets and flight helmets), the inter-subject standard deviations are in the region of 1.1 to 1.5 dB at the frequencies that control the overall A-weighted levels. The differences in mean values between labs are between 1.5 and 2.5 dB at similar frequencies. Whilst the differences in mean values are well within experimental error, in real terms a difference of just 3 dBA is significant in terms of hearing damage risk exposure. Thus, while the differences in actual measuring processes are shown to be insignificant, the actual variability of the acoustic attenuation, caused by very small differences in acoustic leakage, can be significant in the real world.

By closer inspection of the data and the noise-at-the-ear results calculated from the range of military noise spectra, it can be clearly seen that a very few frequency bands control the overall level of the spectrum. For instance, if we take the Harrier cockpit noise case and the headset 5 attenuation right cup, the full spectrum level is 93.4 dBA. If we take away all the low frequency up to 400 Hz, the level with a spectrum from 500 Hz to 10 kHz is 93.3 dBA (i.e. 0.1 dB) down by removing 11 bands). If we just use the 500 to 1000 Hz bands (one octave) then the level drops just another 0.3 dBA to 93 dBA. If the 1000 Hz band is removed the overall level only drops to 92.8 dBA, which means that the overall levels are controlled by just three noise bands (500, 630, and 1000 Hz). Similarly for the Chinook cabin noise, the overall noise level is essentially controlled by the 800 Hz band (88.7 dBA against the full spectrum level of 89.2 dBA).

From this we can deduce that just 3 dB difference in attenuation measurement in the 800 Hz band would result in close to a 3 dBA difference in overall level and this can be seen illustrated in the two figures. In both cases the overall 'A-weighted levels are plotted against a single dominating band – 500 Hz (see Figure 24) for the passive case and 800 Hz band (see Figure 25) for the total case. Here we can see that in the passive case, the 500 Hz noise band, a difference in 1.5 dB in attenuation measurement will lead to a 1.9 dB difference in overall A-weighted energy and for the total attenuation a difference of 4.2 dB in attenuation results in a 4 dBA difference in overall A- weighted noise level. In both cases this is close to 1dB attenuation for 1 dB in overall level. This means that some noise spectra are dominated by a single band with respect to the A-weighting.

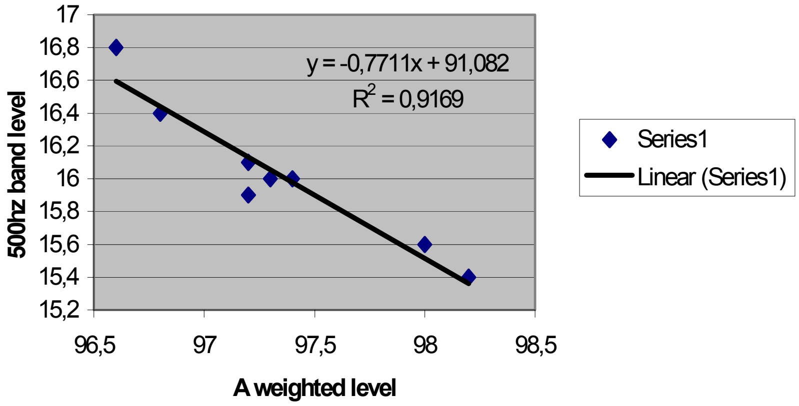


Figure 24: Level of the 1/3 Octave-Band 500 Hz as a Function of the A Weighted Level for the Condition with Passive Attenuation (Headset 1).

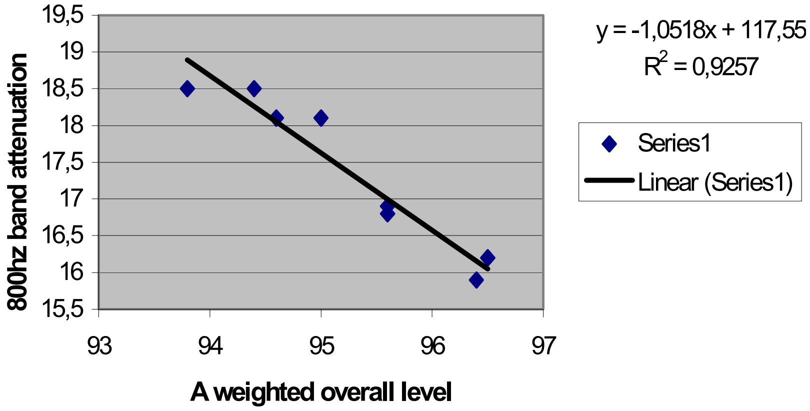


Figure 25: Level of the 1/3 Octave-Band 800 Hz as a Function of the A Weighted Level for the Condition with Total Attenuation (Headset 1, Passive + Active).

From this discussion it is reasonable to deduce that, as far as the primary object of this Round Robin was concerned, the differences between labs using the same measuring technique for MIRE is not significant. However, small differences in acoustic attenuation in critical acoustic bands, which are well within experimental error, can cause differences in overall A-weighted levels that are of some concern.

6.1.2 MIRE and Artificial Head Measurements

The use of a microphone in the concha of the ear of a subject is an accepted and standardized method to determine the insertion loss (reduction of the ambient noise level) of a hearing protector. This method is also valid for applications with active noise reduction (see Annex E). In this study we applied the MIRE-method both with subjects and with artificial means as an artificial ear, artificial head and a so-called flat-plate. It was found that these methods provide similar results. The MIRE-method with subjects has the advantage that inter-subject variance is included (head size and head shape, amount of hair in between the earcup and the head, etc). The use of minimal five subjects provides a reliable estimate of the variance. The mean attenuation together with the variance can be used to predict the sound level at the entrance of the earcanal. The artificial head measurements do not provide an estimate for the inter subject variance. The inter subject variance is also a measure for the “goodness of the fit” of a certain hearing protector. It was found that the user instructions for positioning of the hearing protector are very important. This is of course also the case when the system is used in a practical application. Personnel should get adequate users instructions.

Measurement with artificial means allow the use of very high noise levels as there is no risk for introducing temporary or permanent hearing loss. The advice is first to assess a system with an artificial means and investigate the maximum sound level for stable operation, followed by a MIRE-method assessment with subjects.

The majority of the data used for the comparison were based on the effect of the passive and total attenuation for six military noises and five different headsets.

As an example of "real life" levels, we compared the A-weighted initial noise level outside the hearing protector with the reduced noise level at the inside of the earcups. In Table 15 these initial levels are given together with the mean levels under the hearing protector (mean A-weighted value in decibel for five subject minus one standard deviation, five systems, and left and right earcup).

Table 15: Initial Sound Levels (in dBA) for the Six Military Noises (see Annex B) and the Reduced Levels (passive and total) under the Earcups. The levels under the earcups are mean values for 5 systems, left and right earcup, and based on the mean attenuation (determined with the MIRE-method for 5 subjects and minus one standard deviation.)

Noise type	1	2	3	4	5	6
Outside	83.8	122.8	117.5	113.2	102.5	110.9
Passive	69.8	101.6	96.6	102.9	85.4	91.5
Outside – Passive	14.0	21.2	20.9	10.3	17.1	19.4
Total	62.4	96.8	94.8	91.3	77.7	88.1
Outside –Total	21.4	26.0	22.7	21.9	24.8	22.8
ANR only	7.4	4.8	1.8	10.6	7.7	3.4

An example of the difference in sound attenuation among the five systems is given in Table 16. For the five systems the mean A-weighted sound level for the six noises and the left and right earcup are given. The optimal performance of a specific system is noise spectrum specific. Here only mean values are given as it is not the purpose of the study to compare these systems (we selected them to cover a wide range). However, the given values are representative and demonstrate the contribution of ANR.

Table 16: Mean Passive and Total Levels for Five Headsets. The mean value (in dBA) is based six noises, left and right earcup, and the MIRE-method (5 subjects minus one standard deviation).

Headset	1	2	3	4	5
Passive	91.9	91.8	91.9	93.0	88.1
Total	85.9	85.4	86.3	86.0	82.4
ANR only	6.0	6.4	5.6	7.0	5.7

6.1.3 High Level Measurements

It was demonstrated that active systems will show overload at a certain high level excitation. This effect is different for impulsive noise and stationary noise. Assessing a system at representative sound levels is therefore required. This procedure cannot be performed with subjects because of the risk of permanent hearing damage but has to be performed with an artificial head or flat plate.

6.2 SPEECH COMMUNICATION

Speech communication is an essential function of a military headset. Assessment of the performance of the communications, under adverse military noise conditions, is based in the intelligibility. Subjective measures and objective measures were compared in the Round Robin test. Both methods provided the same rankorder of the systems. However, the application of both methods is different. Subjective tests require many trained subjects as talker and listener. The selection of the test material (rhyme words, nonsense words, sentences) define the resolution and the effort required to perform the test. Objective measures require a specific test system, no subjects, and offer in general diagnostic results.

6.3 SUBJECTIVE ASSESSMENT

Assessment of the comfort of a system is required for applications when hearing protection is to be used for a long period of time. Also the combination with other systems (helmet, gasmask) have to be considered. A standardized test is not available.

6.4 RECOMMENDATIONS

Passive Hearing Protection

From these series of experiments it is clear that the fit of the headsets is a significant dependent variable, and work needs to be carried out to provide a technique for the objective assessment of quality of fit. This is particularly important for the next generations of military vehicles, where noise levels are liable to be higher and hearing damage risk will be increased.

The Use of Active Noise Reduction

Active noise reduction may increase the insertion loss by 5-10 dB depending on the noise spectrum. An increase of 6 dBA means that the maximum exposure time for the same noise dose increases with a factor of four. Such a gain is achieved for noises with strong components below 1000 Hz such as for the Hercules aircraft or Leopard tank.

Hearing Protection and Communication

There is no doubt that most military activities require adequate hearing protection. We considered in this study assessment methods for passive and active systems. However, in some situations a double protection (earplug and headset) may be required. Such a solution will reduce the speech communication if the earplug is not equipped with a telephone. A recent development resulted in a wireless communication earplug. Such a device requires a transmission part in the headset. The conclusion is that no generic optimal solution can be provided. Each application requires careful comparison of possible solutions. Standardised test techniques for assessing the attenuation and speech communication via an earplug or earplug/headset combination now needs to be addressed.

Subjective Assessment

Development and introduction of a standardized test for the subjective assessment of hearing protectors. Such a test should include the use of reference systems in order to allow comparison with other test results.

Chapter 7 – BIBLIOGRAPHY

International Standards

EN 352-1: 1993 Hearing Protectors – Safety Requirements and Testing – Part 1 Ear-Muffs.

EN 24869-1: 1990 Acoustics – Hearing Protectors – Part 1: Subjective Method for the Measurement of Sound Attenuation.

ISO 4869-2: 1994 Acoustics – Hearing Protectors – Part 2: Estimation of Effective A-Weighted Sound Pressure Levels when Hearing Protectors are Worn.

IEC 60268-16 third edition 2002-03 Part 16: Objective Rating of Speech Intelligibility by Speech Transmission Index.

ISO TR 4870 first edition 1991-12-15 Acoustics – The Construction and Calibration of Speech Intelligibility Tests.

ISO/CD 9921 Ergonomics – Assessment of Speech Communication (accepted as committee final draft 2002).

[1] Anderson, B. Wayne, Garinther, G.R. (**1997**). "Effects of Active Noise Reduction in Armour Crew Headsets". *AGARD Conference Proceedings 596,* Audio Effectiveness in Aviation, Neuilly-sur-Seine: AGARD (pp. 20-1 - 20-6).

[2] Berger E.H. (**1986**), "Hearing Protection Devices", in Noise and Hearing Conservation Manual, edited by Berger, E.H., Ward, W.D., Morrill, J.C. and Royster, L.H., American Industrial Hygiene Association, 4th edition, 1986, (pp. 319-382).

[3] Buck, K. and Dancer, A. (**1998**) "Passive and Active Techniques for Hearing Protectors and Their Field of Application", Proceedings Nordic Noise Conference, Stockholm.

[4] Buck, K. and Parmentier, G., (**1996**). "Evaluation of Active Hearing Protectors with High Level Impulse Noise", Proceedings of Internoise, Liverpool, 989-992.

[5] Buck, K., Parmentier, G., (**1999**) "Artificial Heads for High Level Impulse Sound Measurement", Proceedings of International Congress on Sound and Vibration (ICSV).

[6] Crabtree, R.B. (**1997**) "Constraints in the Application of Personal Active Noise Reduction Systems". *AGARD Conference Proceedings 596,* Audio Effectiveness in Aviation, Neuilly-sur-Seine: AGARD (pp. 15-1 - 15-4).

[7] Crabtree, R.B., and Behar, A. (**2000**). "Measurement of Hearing Protection Insertion Loss at Ultrasonic Frequencies". Applied Acoustics 59, 287-299.

[8] Green, D.M. (**1993**). "Maximum Likelihood Method for Estimating Thresholds in a Yes-No Task", J. Acoust. Soc. Am., 93(4), pp. 2096-2105.

[9] House, A.S., Williams, C.E., Hecker, M.H.L., Kryter, K.D., (**1965**). "Articulation Testing Methods: Consonantal Differences with a Closed-Response Set." J. Acoust. Soc. Am. 37, 158-166.

[10] Houtgast, T., and Steeneken, H.J.M. (**1984**). A Multi-Language Evaluation of the RASTI Method for Estimating Speech Intelligibility in Auditoria. Acustica 54(1984), 185-199.

[11] James, Susan. H. (**1990**). “In-Flight Assessment of a Helmet Mounted Active Noise Reduction System in Sea Harrier FRS1”. KMT Harpur, January 1990.

[12] James, Susan H. and Rogers, I. (**1997**). “The Benefits of Improved Active Noise Attenuation to Communication Intelligibility”. DERA/AS/SID-884/CR97042/1.

[13] James, Susan H. (**2001**). “A Noise Survey of the Sea Harrier Cockpit and an Assessment of the Benefits an Active Noise Reduction Headset May Provide”. QINETIQ/FST/CSS/CR011650/1.0.

[14] Lueg, P. (**1936**). “Process of Silencing Sound Oscillations”, US Patent No. 2043416.

[15] McKinley, R.L. and Nixon, C.W. (**1993**) “Active Noise Reduction Headsets”. Proc. 6th Int. Conference on Noise as Public Health Problem”, Nice, pp. 83-86.

[16] Olson, H.F., and May, E.G. (**1953**). “Electronic Sound Absorber”, J. Acoust. Soc. Am., 25, pp. 1130-1136.

[17] Pan, G.J., Brammer, A.J., and Crabtree, R.B. (**1997**), “Adaptive Active Noise Reduction Headset for Helicopter Aircrew”, AGARD Conference Proceedings 596, Audio Effectiveness in Aviation, Neuilly-sur-Seine: AGARD (pp. 19, 1-6).

[18] Parmentier, G., Dancer, A., Buck, K., Kronenberger, G., Beck, C. (**2000**); “Artificial Head (ATF) for Evaluation of Hearing Protectors”, Acustica, **86**, 847-852.

[19] Parmentier, G., (**1993**) « Synthèse des resultants expérimentaux relatifs aux detonations d’explosifs sphériques », Rapport ISL, R-113/93.

[20] Pellieux, L., Sarafian, D., and Reynaud, G. (**1997**) “Evaluation de casques a reduction active de bruit: protection auditive et intelligibilité”. (Assessment of Active Noise Reduction Hearing Protectors: Noise Attenuation and Speech Intelligibility” *AGARD Conference Proceedings 596,* Audio Effectiveness in Aviation, Neuilly-sur-Seine: AGARD (pp. 16-1 - 1-19).

[21] Ryan, J.G., Shaw, E.A.G., Brammer, A.J., and Zhang, G. (**1993**), “Enclosure for Low-Frequency Assessment of Active Noise Reducing Circumaural Headsets and Hearing Protectors”, Canadian Acoustics 21, #3, pp. 19-20.

[22] Rood, G.M. (**1996**). “The Audio Environment in Aircraft”. AGARD-CP-596, October 1996.

[23] Rood, G.M. and James, S.H. (**2003**) “Evaluation of Assessment Methods of Active Noise Reduction Systems in an International Round Robin: Part1; UK National Report”. QinetiQ/FST/TR034699/1.0 April 2003.

[24] Smoorenburg, G.F., Bronkhorst, A.W., Soede, W., and Reus, A.J.C. de. (**1993**) “Assessment of Hearing Protector Performance in Impulse Noise. Feasibility Study Part II: Methods”. Report IZF 1993- C-32, Soesterberg, The Netherlands, TNO Human Factors Res. Inst.

[25] Shaw, E. (**1974**), “The External Ear”, in Handbook of Sensory Physiology, Vol. V/1, pp. 455-473, Springer, Berlin, Heidelberg, New York.

[26] Steeneken, H.J.M., Geurtsen, F.W.M. (**1988**). Description of the RSG-10 Noise Data-Base. Rapport TNO-TM 1988 3. TNO Technische Menskunde, Soesterberg.

[27] Steeneken, H.J.M. & Verhave, J.A. (**1997**). Personal Active Noise Reduction with Integrated Speech Communication Devices: Development and Assessment. *AGARD Conference Proceedings 596,* Audio Effectiveness in Aviation, Neuilly-sur-Seine: AGARD (pp. 18-1 - 18-8).

[28] Steeneken, H.J.M., and Houtgast, T. (**1980**). "A Physical Method for Measuring Speech-Transmission Quality", J. Acoust. Soc. Am., **67**, 318-326.

[29] Steeneken, H.J.M., (**1998**). "Personal Active Noise Reduction with Integrated Speech Communication Devices: Development and Assessment". Noise and Health 1998; 1:67-75.

[30] Steeneken, H.J.M., and Houtgast, T. (**1999**) "Mutual Dependence of the Octave-Band Weights in Predicting Speech Intelligibility". Speech Communication, 1999, Vol. 28, 109-123.

[31] Steeneken, H.J.M., Buck, K, Crabtree, B., James, S., McKinley, R. (**2001**). "Standardized Assessment of Personal Active Noise Reduction". International Military Noise Conference, Baltimore, Maryland, 24-26 April.

[32] Steeneken, H.J.M., Verhave, J.A. (**2001**). "Digitally Controlled Active Noise Reduction with Integrated Speech Communication". Proc CIOP Noise Conference, Kielce, Poland.

[33] Steeneken, H.J.M., and Houtgast, T. (**2002**). "Validation of the Revised STI_r Method". Speech Communication, 2002, Vol. 38.

[34] Steeneken, H.J.M., and Verhave, J.A. (**2003**). "Evaluation of Assessment Methods for Personal Active Noise Reduction in an International Round Robin". Report TNO-Human Factors TM-03-A015, TNO-Human Factors, Soesterberg, The Netherlands.

[35] Zera, J., Brammer, A.J., and Pan, G.J., (**1997**). "Comparison between Subjective and Objective Measures of Active Hearing Protector and Communication Headset Attenuation. J. Acoust. Soc. Am. **101** (6), 3486-3497.

[36] [Zera, J. (**in press**). "Speech Intelligibility Measured by Adaptive Maximum Likelihood Procedure" Speech Communication (in press, 2004).

Annex A – ANR SYSTEMS USED FOR THE TEST

Figure A1.

Figure A2.

Figure A3.

Figure A4.

Figure A5.

Annex B – NOISE SPECTRA USED FOR INSERTION LOSS CALCULATIONS

Table B1: Noise Spectra for the Six Military Noises used for the Calculation of the Insertion Losses (see 5.1.1). The 1/3 octave-levels are given in dB re 20 μPa, [11, 13, 26].

1/3 Oct	Hercules C130K	Sea Harrier Ext, on deck carrier	Sea Harrier (cockpit)	Leopard 2 70km/u insidel	Chinook cockpit	Chinook cabine
25	84	120	73.2	92.5	120.7	109.9
31	83	120	77.1	92.1	111.9	106.8
40	80	120	79.2	103.4	102.9	104.4
50	78	121	83.4	109.9	103.8	106.4
63	95	120	87.6	104.5	103.5	104.8
80	85	119	105.2	115.1	101.9	99.4
100	75	117	95.1	120.9	94.7	103.5
125	82	114	99.9	111.7	96.7	107.4
160	77	111	99.4	112.7	97.8	102.0
200	77	110	101.2	118.1	91.1	98.2
250	77	109	104.7	111.5	90.8	98.4
315	78	106	105.3	110.3	97.8	96.3
400	79	104	109	107.4	96.3	99.0
500	77	107	112	103.9	92.9	95.1
630	76	109	111	103.4	87.8	95.4
800	75	111	110	100.7	90.4	107.7
1k	74	112	109	97.4	84.5	94.7
1k25	72	111	106	98.5	92.6	95.0
1k6	69	111	105	96.7	92.2	99.6
2k	68	111	106	93.9	85.4	97.7
2k5	69	112	104	93.5	86.3	94.5
3k15	68	112	100	93.4	96.3	98.6
4k	66	112	98.8	91.9	86.3	98.9
5k	65	112	97.7	90.1	81.5	96.6
6k3	64	111	95.5	89.4	80.5	99.7
8k	64	110	94.4	87.3	75.3	93.4
10k	62	108	94.5	87.4	70.7	88.3
12k	61	0	95.1	84.4	72.1	87.6
16k	59	0	89.9	84.4	63.5	83.0
20k	56	0	65.5	79.7	57.7	71.5
LIN	**97.9**	**129.5**	**119.3**	**124.9**	**121.6**	**117.0**
AWT	**83.8**	**122.8**	**117.5**	**113.2**	**102.5**	**110.9**

Annex C – CONTRIBUTION OF DRDC, TORONTO, CANADA

Supplementary ANR Assessment Procedures

B.C. Crabtree, DRDC Toronto

C1.0 SCOPE

Details of some procedures used by DRDC Toronto that were supplementary to the common testing routines carried out by all laboratories are provided in this Appendix. One of these is a procedure to combine the traditional Modified Rhyme Test (MRT) for assessing speech discrimination with the Maximum Likelihood Procedure (section C2) for the determination of signal-to-noise ratio necessary for a given level of speech discrimination proficiency. The other is a procedure for determining the “Saturation Threshold” for ANR equipment (section C3), that is, the maximum level of noise at the ear that can be reliably accommodated by ANR at the onset of audible distortion. The data resulting from the application of these procedures are given elsewhere in this report.

The general procedure adopted by DRDC Toronto to simplify the process of carrying out the test battery on the Round Robin ANR headset sample is also described in this annex (see section C4 and [6, 7, 17]). Although the specific example of the Task Group Round Robin testing is described, the procedure is easily adapted to other headsets and situations as conditions warrant. The remaining routines run by DRDC Toronto were similar to those used by other laboratories, and are addressed in detail in other sections of this report.

C2.0 MAXIMUM LIKELIHOOD PROCEDURE VARIANT OF THE MODIFIED RHYME TEST

The Modified Rhyme Test for assessing speech intelligibility proposed by House et al. [9] and used for many years, notably by the US military for assessing the subjective efficiency of communications equipment. The MRT test material consists of 50 word sets containing six rhyming words each, for a total vocabulary of 300 words. In a modern implementation of the traditional paper-and-pencil test, one of the 50 word sets or “screens” selected at random is presented on a computer monitor as a 2 x 3 matrix, as one of these words selected at random is presented as an audio stimulus within a carrier phrase. The level of the stimulus is chosen to produce a desired signal-to-noise ratio (SNR) with respect to a continuous noise masker. The observer is provided with a response box having six buttons with the same matrix alignment as the words on the visual display. The observer’s task is to press the button that corresponds to the word that was thought to be heard, and the answer is logged by the host computer as a correct or incorrect response. This process is repeated without replacement until all 50 word sets have been presented, after which the percentage intelligibility score is computed. The stimulus level is adjusted and the test repeated, until all SNRs of interest have been run. The associated psychometric intelligibility function (word score vs. SNR) can then be plotted from the data.

The Modified Rhyme Test coupled with the adaptive Maximum Likelihood Procedure or MLP [8] is intended to establish the SNR at which a given percentage speech discrimination proficiency is achieved in a given noise background. This hybrid process converges rapidly to the associated SNR; thereby conserving time that would otherwise be spent in finding SNRs removed from the immediate area of interest.

In the application of the adaptive MLP method, it is assumed that the parameters of the MRT intelligibility function are already known. This function is fitted to the cumulative Gaussian distribution of MRT validation data [8]. The only variable is the function’s position along the abscissa (SNR scale), which is ultimately determined as the best fit to the data collected.

A “Target Intelligibility” (TI) of interest lying along this function is preselected before the experiment is run. A TI of 60 % was chosen by DRDC Toronto for this procedure, since according to Zera [36], this is where the standard deviation is a minimum (termed the sweet point) in a six-alternative forced-choice task. One could just as easily choose other TIs of particular interest without added complexity. Once the position of the function along the abscissa is established, the percentage intelligibility will be known over a range of SNRs with reasonable confidence.

When the MRT / MLP experiment is run, the carrier phrase is presented at a constant level during all trials, while the amplitude of the stimulus words is varied. The SNR of the carrier phrase with respect to the background noise is chosen for clear intelligibility of the phrase (typically the SNR is set within the zero to +5 dB range), since this forms the cue that the stimulus word is about to be presented. In the first trial, the level of the stimulus word is made the same as that of the carrier phrase (i.e., with high SNR). Based on whether the response is correct or incorrect (nearly always correct), the next trial is presented at a stimulus level having the maximum likelihood (greatest probability) of achieving the selected TI. In effect, the MLP algorithm computes a new position of the intelligibility function on the SNR axis.

In subsequent trials, the maximum likelihood calculation is based on a progressively larger history of responses accumulated within the run; hence subsequent predictions quickly converge to the stimulus level producing the desired TI. Over 50 such trials, this implies 30 correct responses and 20 incorrect responses, although typically, some dither is seen in the actual tally of responses over a number of runs. After 50 trials the MLP is terminated and the level difference between the carrier phrase and the final stimulus word, along with the associated percentage score, are taken as the results of the run. From knowledge of the background noise level at the ear and the level of the carrier phrase, the actual SNR may be calculated. These co-ordinates then define a point on the associated psychometric function, from which intelligibility proficiency at different SNRs may be predicted. Normally, several such sessions are conducted under identical conditions and the data are averaged to fine-tune the position of the psychometric function. In some instances, e.g., for ANR active and passive mode comparisons, the results are usually more meaningful if given in terms of absolute stimulus level, rather than SNR. This aids in demonstrating, for example, that when ANR is used, the level of presentation may be reduced with no associated decrement in speech discrimination proficiency.

Experience with this procedure suggests that there is no loss of generality if the order of the words appearing in each of the 50 screens is fixed. In fact retaining the order of the six words may assist for tasks in environments where scanning the set for the correct word may not be trivial. Furthermore, it seems unnecessary to exclude from the word sets and audio stimuli those already presented to the observer within the same run. Thus, presentation of the word screens and word stimuli are both purely random with replacement functions.

C3.0 MEASUREMENT OF ANR SATURATION THRESHOLD

The saturation threshold of an ANR system is taken as the sound pressure level at the ear above which the ANR control system contributes audible distortion while attempting to reduce the amplitude of the sound field. Several factors can contribute to perceived distortion. These include, but are not limited to, the excursion limits of the cancellation transducers; power supply limitations that cause signal clipping and inherent limiting in the control loop electronics.

The relationship between audible distortion produced by an ANR system and the system linearity discussed in other sections of this report was not rigorously investigated. It would seem safe to assume, however, that distortion of sufficient amplitude to degrade noise reduction performance would exceed the threshold at which distortion becomes audible. Thus the procedure to be described might be useful in predicting which systems are overly susceptible to non-linear behaviour in high noise environments.

The procedure involves the generation of a sinusoidal sound field within the earcup of an ANR headset at discrete frequencies over a range of 10 to 100 Hz. The fixture to accomplish this employs a KEMAR acoustical mannequin headform. The headform is modified from its usual configuration by removing one of the Zwislocki couplers and associated mounting plate from one of the artificial ear cavities and suspending a calibrated measurement microphone (Brüel & Kjær type 4134) in the opening so formed. This arrangement allows air to pass freely through the opening around the microphone. The hollow neck attaches the headform to the interior of a loudspeaker enclosure containing a 200-mm low-frequency driver, thereby providing an acoustical coupling path between the loudspeaker and cavity within the earcup of an ANR system placed on the headform.

When the loudspeaker is driven by a low-frequency pure tone, the earcup cavity of the ANR system can be excited to sound pressure levels exceeding 140 dB. The ANR system cannot distinguish between this type of inside-out excitation and that, which normally permeates the earcups. Thus the control loop attempts to establish an opposing noise field. Since the measurement microphone is located in proximity to the cancellation transducer, it is highly sensitive to the onset of distortion or overload. The resulting distortion (the creation of frequency components not related to the excitation frequency) is clearly audible over headphones arranged to monitor the output of the microphone, as the excitation level is raised and lowered in the vicinity of the threshold. Alternatively, the microphone signal may be monitored by a signal analyser in order to display the magnitude and spectrum of the distortion components. The saturation level and excitation frequency are noted before selecting a different excitation frequency and repeating the threshold determination. Conducting measurements at 10 Hz intervals seemed to provide sufficient resolution.

The procedure is repeated with the ANR system reversed on the fixture such that the remaining earcup performance is also assessed. The data are pooled across earcups before plotting as a function of frequency.

Although the above procedure provides a descriptor of ANR behavior within the earcup at very low frequencies, it clearly cannot define the magnitude of an *external* sound field, which will cause the threshold to be exceeded. This is a function of the passive mode attenuation provided by the device, the effectiveness of fitting and the characteristics of the disturbing noise field.

C4.0 COPING WITH MULTIPLE TESTS AND HEADSET VARIABLES

The complexity of the Round Robin test battery and the relatively large number of test samples agreed by the Task Group made it worth while to implement a computer-based configuration program that greatly simplified the tasks of headset and test randomisation, setting stimulus levels and response logging normally conducted by the experimenter. Among other attributes, the procedure was pivotal in determining the SNRs under which speech discrimination tests were performed.

What made the development particularly challenging was the number of variables that had to be properly accounted for. Each of the seven headsets tested by DRDC Toronto had a different attenuation characteristic, different input sensitivity and different frequency response; each of these variables depended in turn on whether active or passive mode was selected. As agreed at one of the first consortium meetings, “A” curve weighting was to be used as the “universal” measure of level, without regard to spectral differences. Nevertheless, A-weighted levels are still dependent on all of the variables listed above. Thus, the experimenter was faced with correctly assigning fourteen different signal levels in the MRT/MLP test (for example) during the course of each experimental session with human subjects – a truly formidable task to execute repetitively without error.

To simplify this process, a configuration/file program was created to prompt the experimenter and to preset the level corrections as the session proceeded. This application ran on the same computer that supported the MRT/MLP task and served as the host for the Tucker-Davis Technologies equipment suite. The program selected the headset to be tested and the mode of operation at random without replacement. The nominal signal levels at the ear for that headset were indexed and preset on a mechanical attenuator to compensate for gross differences in headset sensitivity. These settings were selected to produce levels slightly higher than the required A-weighted at-ear levels.

Corrections to the nominal levels resulting from pilot study measurements were programmed into the configuration file to compensate for the mode of operation and the resulting differences in speech noise background level at the ear. These corrections were used to automatically set the signal attenuator in the Tucker-Davis system under program control to match the speech noise background levels determined in the pilot study for that particular headset. In this way, the MRT carrier phrase was automatically set to zero dB SNR (clearly audible), regardless of the noise conditions or headset sample chosen.

In a similar fashion, corrections were established for all headsets for the presentation of the STITEL signal, except that in this instance, the level of the STI signal was set to 70 dBA at the ear, without regard for the actual speech noise level reaching the ear under each headset.

As the operator stepped through the program during each session, spectrum analyser and STITEL file numbers were logged and MRT level data were automatically stored for later retrieval. The program also prompted the operator to change experimental conditions at appropriate times, for example, to reduce the level of noise for open-ear measurements, or to change the noise spectrum from “mauve” noise (for MIRE measurements) to “speech” noise (for STITEL measurements) etc. These prompts had to be acknowledged via the keyboard before the program would advance. The integration of this program with the experimental apparatus proved invaluable as a means of organizing the experiment, setting parameters without error and producing a running log and cross-reference of all data entries.

C5.0 ACKNOWLEDGEMENTS

The author is most grateful for the dedicated assistance of April Straus in conducting the isolation measurements on the Kunov Head used in the evaluation of ANR attenuation and certain agreed physical measurements, for the professional efforts of Lisa Massel in organizing and conducting both objective and subjective phases of testing and for keeping excellent notes. Ms. Straus and Ms, Massel were at the time of testing undergraduate students at the University of Waterloo in Design Engineering and Kinesiology programs respectively. He also acknowledges the extensive contributions of Garry Dunn, a graduate of computer engineering at the University of Waterloo who developed the computer applications necessary for Round Robin testing, including complete recoding the MRT/MLP application. He also produced most of the tabular and graphical data DRDC Toronto contributed to this report and established the technical “pipeline” to the Netherlands. Mr. Dunn provides technical assistance to DRDC Toronto under a contractual arrangement. The MRT/MLP concept was originally proposed and implemented by Jan Zera, formerly of the National Research Council, Ottawa, Canada. The application was originally used in the joint development of the NRC/DRDC Toronto feed forward adaptive D-ANR technology demonstrator. The author is also indebted to Andrew Welker, who developed the acoustic test fixture used for assessing the saturation thresholds of the Round Robin headset sample. At that time, Mr. Welker was an undergraduate student in electrical engineering at the University of Waterloo.

Annex D – CONTRIBUTION OF INSTITUTE DE ST. LOUIS, FRANCE

Methods using Impulse Noise for the Evaluation of ANR Hearing Protectors

K. Buck, G. Parmentier, V. Zimpfer-Jost
French-German Research Institute (ISL), Group APC

D1.0 INTRODUCTION

Active Noise Reduction (ANR) hearing protectors and headsets may be useful for soldiers that are exposed to very high levels of continuous noise with the major spectral distribution in the low frequency range as we find in tanks, helicopters, propeller aircraft, etc. However, to date, no normalized procedures for the evaluation of these devices in the military acoustic environment exist. The ISL has joined this "Round Robin" test because of the need of widely accepted evaluation procedures for ANR hearing protection devices.

At the ISL all hearing protectors have been evaluated in continuous and impulse noise. The measurement procedures for continuous noise are very close to those used in the other participating laboratories. Therefore only the procedures for measuring the performance using impulse noise and some specific interpretations of stability are described in this Annex. For more specific information about all procedures used at the ISL a separate report is in preparation.

D2.0 GENERATION OF IMPULSE NOISE

To evaluate ANR headsets under impulse noise conditions, it is very important to be able to generate impulse noise in a way that is reproducible and comparable to munition noise encountered in the military environment. Under these conditions, the peak pressures range from about 140 dB to 190 dB and the A-Duration from 0.2 ms to 5 ms. The reproduction of these pressure time histories in the free sound field is almost impossible when using standard noise sources such as loudspeakers.

As the ISL possesses a shooting range that allows the use of explosive charges we have the opportunity to generate the impulses by means of explosive charges. This technique allows the generation of all needed peak levels and A-durations by modifying the distance from the Acoustic Test Fixture (ATF), the mass and/or the type of the explosive. Figure D1 shows a typical setup for the exposure of an ATF to an impulse noise generated by an explosive charge. In order to avoid reflections and to obtain a propagating shock wave, the explosive is placed close to the ground (concrete).

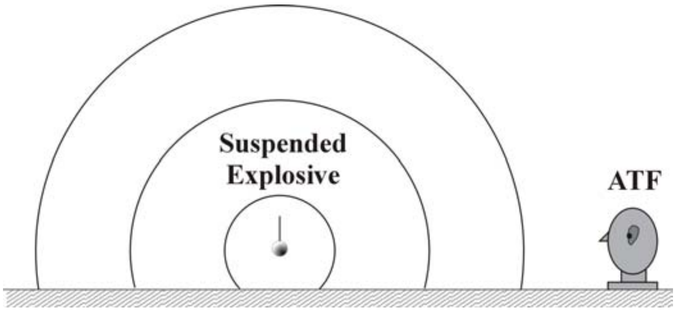


Figure D1: Schematic Setup for Impulse Noise Exposure.

These shock waves are very reproducible and the pressure time history is very similar for different peak pressures. This is shown in the figure. D2 where the generated waves for two peak pressures (150 and 160 dB) are superimposed. The amplitudes of the waves are normalized with respect to the peak pressure.

The measurements are made with four different peak pressures 130 dB, 150 dB, 160 dB and 170 dB. As all systems showed complete overload at the peak level of 170 dB, no measurements for higher levels were made.

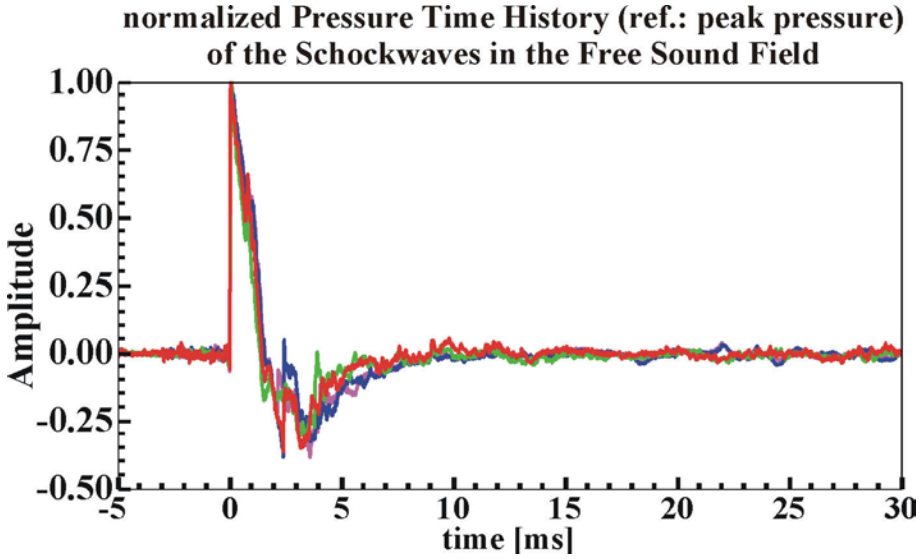


Figure D2: Normalized Pressure Time Histories for Blast Waves Created with Explosive Charges (A-duration ~1.5 ms). The reference is the peak pressure. red, green - L_{peak} = 150 dB, blue, cyan - L_{peak} = 160 dB.

D3.0 SET-UP AND INSTRUMENTATION

For the determination of the parameters with impulse noise, the pressure time history in the free field and at the microphone of the artificial head have to be measured with identical impulses [4]. To do this, a method allowing the simultaneously measurement of the impulse in the free field and underneath the hearing protector (microphone of the ATF) has been used.

D3.1 Set-Up

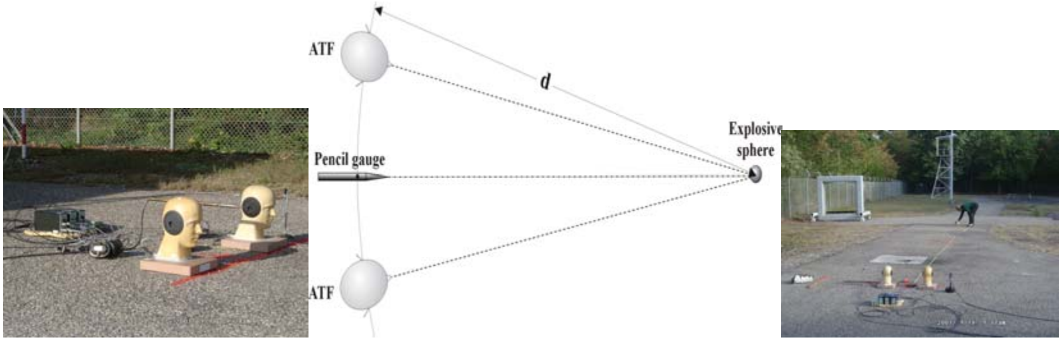


Figure D3: Physical Set-Up of the Experience at the Shooting Range.

Figure D3 shows the placement of the different parts of the measurement equipment. All measurement elements (ATF, Pencil gauge) are positioned at the same distance (d) from the explosive sphere. As for the configuration shown an undisturbed wave front can be assumed, the pressure time history at all measurement positions (ATF, pencil gauge) may be considered to be identical. Proceeding this way, both sides of two hearing protectors as well as the free field pressure time history can be measured simultaneously.

When changing the peak pressure, the artificial heads are repositioned to the modified distance.

D3.2 Instrumentation

Due to the limited attenuation that is provided by the commercially available artificial.

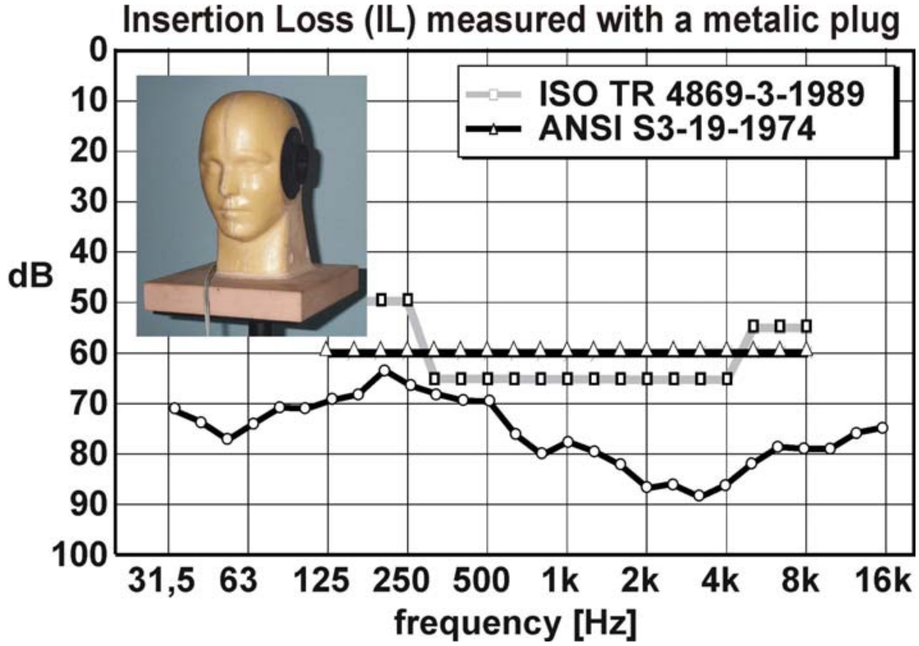


Figure D4: Insertion Loss of the ATF Used for the Measurements and Picture of the Artificial Head.

Heads, the ATF used for the evaluation of hearing protectors has been designed and made at the ISL (Figure D4). It has been especially adapted to measurements with very high acoustic levels like weapon noise or other noise found in the military environment. The IL of the ATF (measured when the earcanal is blocked by a metallic earplug) is better than 60 dB (Figure D4) for all frequencies.

In order to reproduce the impedance of the ear the artificial head is equipped with an ear simulator (B&K 4157), as well as with an earcanal, a pinna and flesh simulation in the circum-aural area (HeadAcoustics®). The TFOE (**T**ransfer **F**unction of the **O**pen **E**ar) measured with this ATF is very close to those measured on human ears by Shaw [25]. The comparison between the TFOE of human ears and the TFOE of the ATF is shown in figure D5.

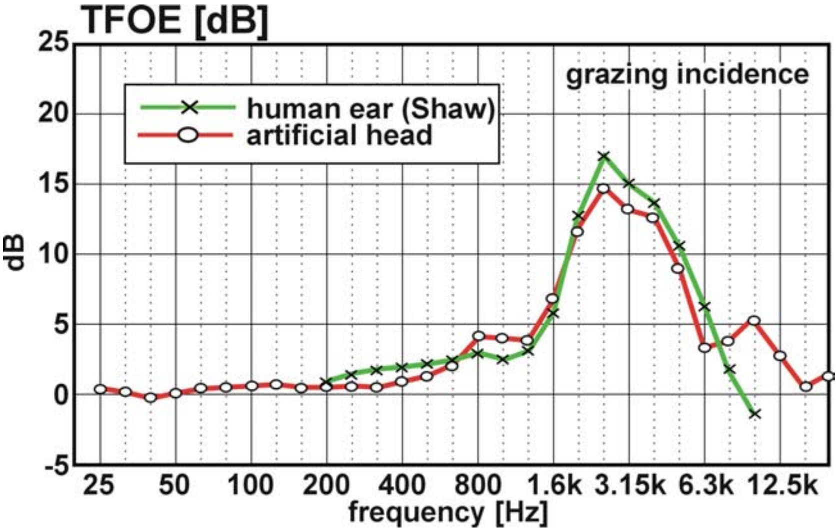


Figure D5: Transfer Function of the Open Ear of the Artificial Head (red line) Compared to the Values Published by Shaw [25] (green line).

The use of the ATF gives, in general, reproducible results that are comparable with results that are acquired using other methods (REAT, MIRE). As no human subjects are exposed, this method can be used for all levels admissible to the device (up to 185 dB at the unprotected “ear”). It is suitable for earmuffs as well for earplugs. However it may overestimate the IL for some types of earplugs. Another disadvantage of such a device is that it does not take into account the distribution of anthropometric variations as do REAT and MIRE.

D3.3 Data Acquisition

Figure D6 shows the instrumentation used for the data recording on the shooting range. If two artificial heads are used simultaneously, two digital recorders are used for the measurement. In order to synchronize all signals, the free field pressure is recorded on both recorders. A sampling frequency of 100 kHz is used. The measurements are in conformance with ITOP 4-2-822.

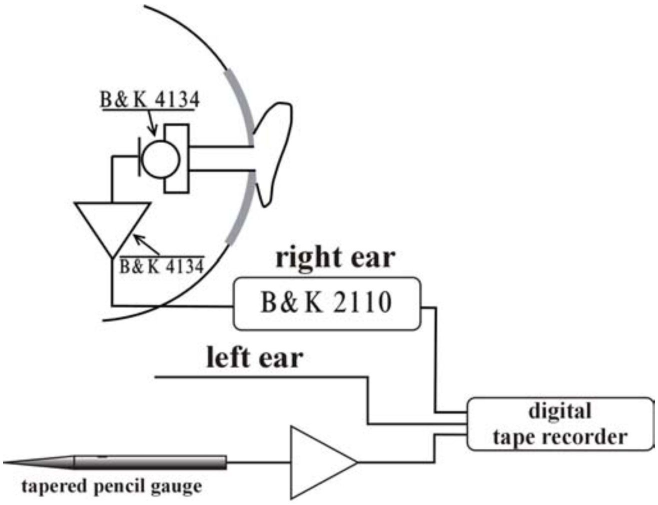


Figure D6: Schema of the Instrumentation, Used with the Artificial Head, Left and Right Ear of the Artificial Head have Identical Instrumentation. The tapered pencil gauge may be replaced by a blunt cylinder gauge for lower peak pressure levels.

D4.0 DATA ANALYSIS

In order to analyze the recorded data, the taped data have been re-sampled and stored as files on a workstation. Two types of data files are then available for further analysis.

- The pressure time history of all recorded events;
- The 1/3 octave-spectra of all recorded events. These spectra were typically calculated using the FFT data of 16,000 samples of the time signal. The procedure allowed for the calculation of the 1/3 octave-bands from 16 Hz to 16 kHz.

All results were derived from these two sets of data.

D4.1 Calculating the Insertion Loss (IL) from Single Shots

As described in section D2 the insertion loss is the difference between the level at the eardrum for the unprotected ear ($L_u(f)$) and the level at the eardrum when the ear is protected ($L_p(f)$).

$$IL(f) = L_u(f) - L_p(f) \tag{1}$$

Using this method, two measurements are needed for the determination of the IL.

However, if the IL is to be determined with impulse noise, a method using only a single measurement (single explosion) gives more reproducible results. If the TFOE (**T**ransfer **F**unction of the **O**pen **E**ar) representing the spectral difference between a signal measured at a reference point and the signal measured with the microphone of the ATF are known, the following method may be used (Figure D7).

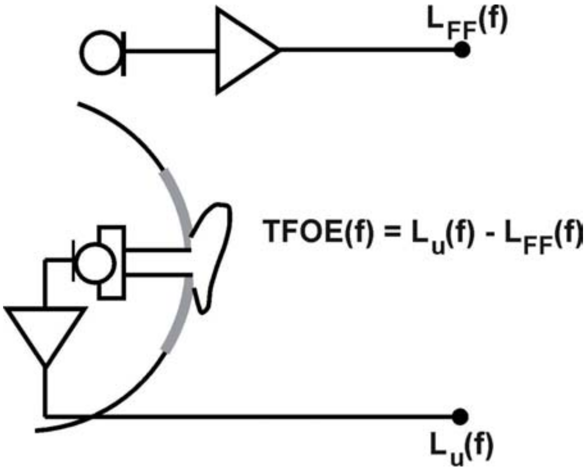


Figure D7: Measurement Scheme for the Determination of the TFOE.

As the TFOE(f) is known and independent of the hearing protector, the following equations may be applied:

$$TFOE(f) = L_u(f) - L_{FF}(f) \tag{2}$$

$$L_u(f) = TFOE(f) + L_{FF}(f) \tag{3}$$

(3) in (1) $$IL(f) = L_{FF}(f) - L_p(f) + TFOE(F) \tag{4}$$

Using this formulation (4) for the evaluation of the IL, once the TFOE is determined, a single two channel measurement of the free field level and the level underneath the hearing protection (at the "eardrum") allows the calculation of the insertion loss of a hearing protector.

The TFOE is determined prior to the IL measurements. It is the difference between the spectrum of the signal measured in the free field and the spectrum measured at the "eardrum" of the artificial head. The TFOE of the ISL artificial head at grazing incidence (head looking towards the explosive, configuration used for the measurements) is shown in figure D5.

Determination of the attenuation added by the ANR system.

D4.2 Linear Behavior

At sound pressure levels for which the device does not show any overload or other non-linearity's, the measurements may be made at one single level. The IL is measured for the hearing protection with the ANR system switched on and off.

$$IL_{ANR} = IL_{ANR\,on} - IL_{ANR\,off}$$

The difference of these two measurements is the contribution of the ANR system. Figure D8 shows the typical data for the measurement of the ANR contribution.

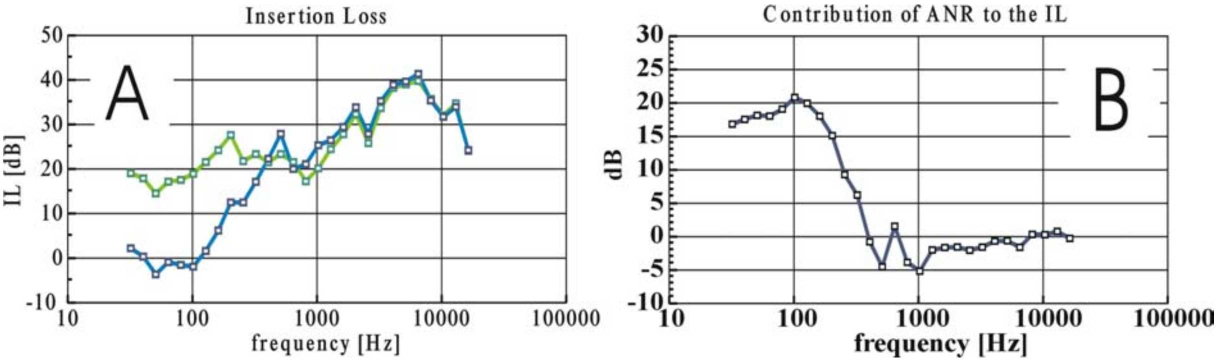


Figure D8: Determination of the Contribution of the ANR to the IL.
A) Insertion Loss, B) contribution of ANR blue: ANR off , green: ANR on.

This figure also shows that the contribution of the ANR to the IL may be divided in two regions:

- A region where the curve shows a positive attenuation. This part typically is located at low frequencies and shows the effectiveness of the system. For most of the ANR earmuffs this region goes up to about 400 Hz to a maximum of 600 Hz.
- A region where negative attenuation (amplification) is found. This part of the curve reflects the stability of the feedback system. The amplification in this region is inversely proportional to the distance of the closed loop transfer function to the point of instability ((-1,0) in the Nyquist diagram). This means that the greater the amplification, the more susceptible the system is to instability.

D4.3 Non Linear Behavior

As the active parts (electronics, loudspeaker …) of an ANR hearing protector can only generate noise up to a certain level, the contribution of the ANR will become non linear for very high sound pressure levels. In a military environment this typically happens when the devices are submitted to weapon noise.

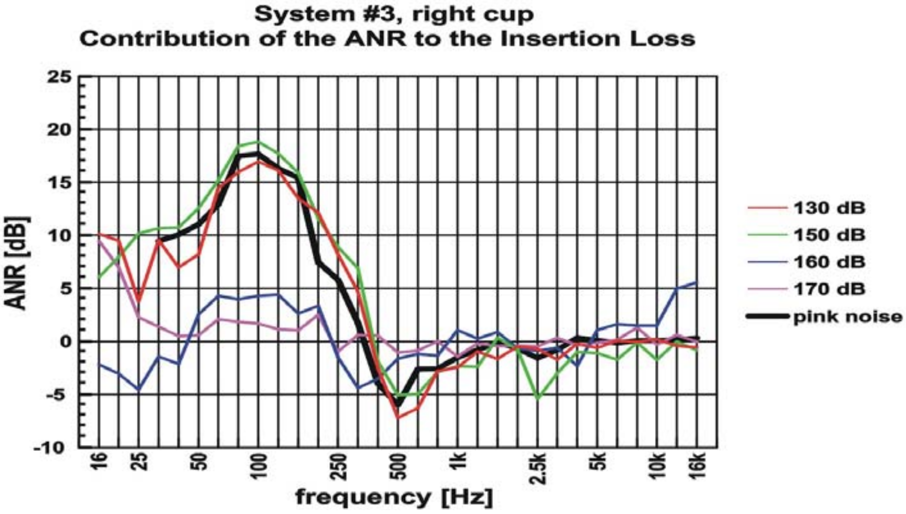


Figure D9: Contribution of the ANR System to the Insertion Loss for Impulse Noise with Different Peak Pressure Levels. The A-duration is 1.5 ms for all levels. Average of three measurements.

The determination of the starting point of the non-linear effects may be established with impulse noise measurements. Figure D9 shows the contribution of the ANR system to the insertion loss for four different peak pressure levels of an impulse noise with an A-duration of 1.5 ms. For the two lower levels (130 dB (red) and 150 dB (green)) the contribution of the ANR is equivalent to that measured with continuous pink noise (black). For the 160 dB peak pressure level (blue) the effect of the ANR system is reduced to about 4 dB. For higher peak levels, e.g. 170 dB (mauve), the ANR is no longer effective and the hearing protector acts like a passive device.

D4.4 Overload of the ANR System

The effects described above may be explained by an overload of the electronic and/or electro-acoustic part of the feedback loop. This overload will be observed, when the system is exposed to noise levels above its operating range. Situations like these may occur when a soldier is exposed to weapon noise. For systems designed to be used in military noise, the overload behaviour for continuous and impulse noise should be evaluated. Figure D10 shows a way for evaluating the overload behaviour using the pressure time histories underneath the hearing protector when exposed to impulse noise. In this figure the measured time pressure histories induced by impulse noise of different peak pressure levels (150 dB, 160 dB and 170 dB) are displayed for two conditions: ANR off (black line) and ANR on (blue line). The third pressure time history (red line) is the calculated difference between the ANR on and the ANR off condition. This curve is equivalent to the pressure input into the cavity by the loudspeaker of the ANR system, when it is switched on.

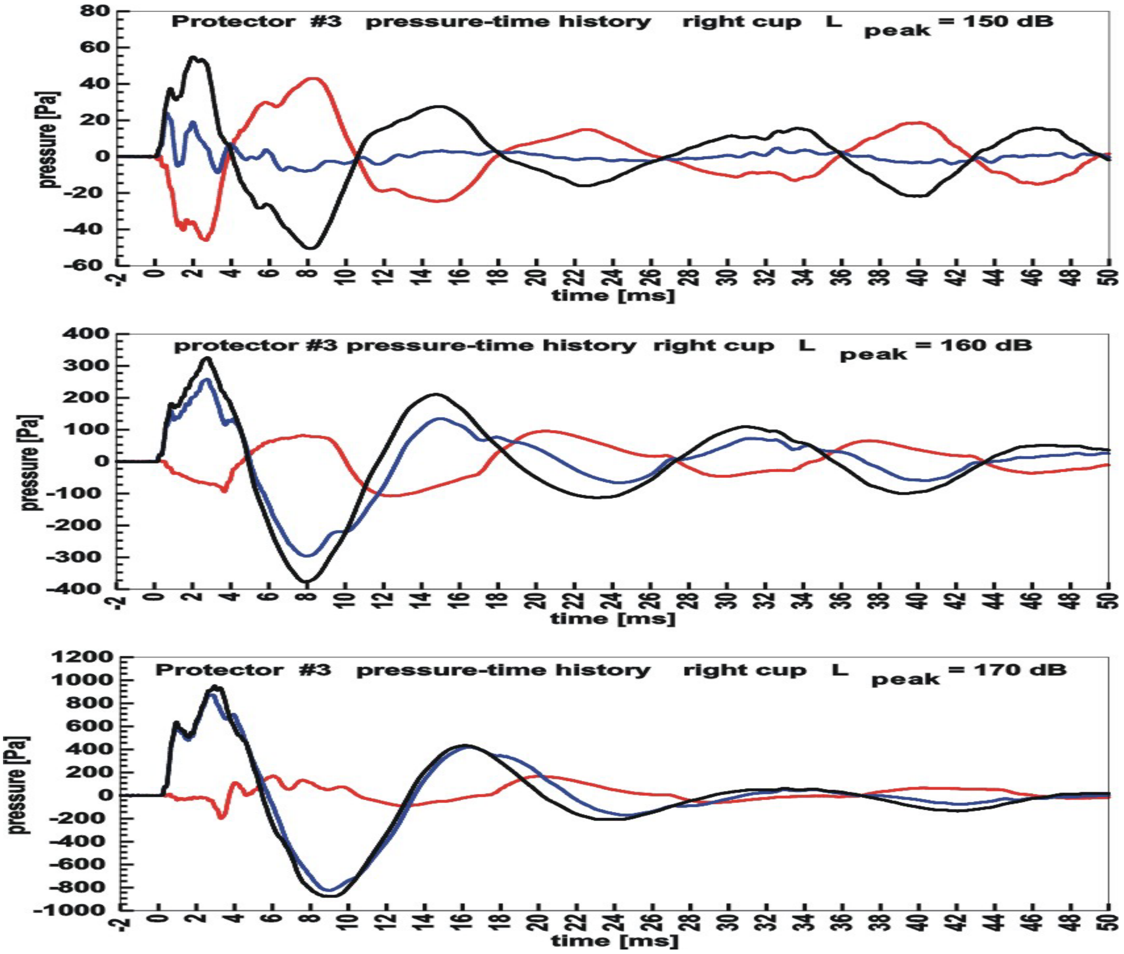


Figure D10: Time Pressure Histories Underneath the Hearing Protector with ANR On and Off for Different Peak Pressure Levels of the Impulse Noise in the Free Sound Field.

Black – Pressure time history with ANR switched off ($P_{off}(t)$)

Blue – Pressure time history with ANR switched on ($P_{on}(t)$)

Red – Calculated pressure time history of the “anti-noise”

$$P_{ANR}(t) = P_{on}(t) - P_{off}(t)$$

It can clearly be observed that for the 150 dB condition this “anti-noise” is close to the pressure signal without ANR, but in opposite phase. For the 160 and the 170 dB condition, a saturation of the “anti-noise” can be observed. This leads to a reduced ANR efficiency in the case of 160 dB and makes almost no contribution at the peak level of 170 dB. However, we can observe that the calculated maximum pressure delivered by the loudspeaker (broken lines) is close for both cases. It is about 134 dB for 160 dB peak pressure level and 137 dB for the 170 dB condition.

A more detailed analysis of the “anti-noise” signal will also show the type of clipping (soft or hard) if there is a recovery time after saturation of the system, or if saturation may induce instabilities.

D5.0 CONCLUSIONS

In the military environment hearing protectors are often subjected to impulse (weapon) noise. It therefore seems important to analyze the behavior of ANR systems under this condition. The evaluation of ANR hearing protectors with impulse noise, as done at the ISL, gives interesting and important information about the behavior of the devices in high level noise environments.

The measurements using impulse noise give good information about:

- the linearity of the ANR device;
- the overload behavior of the ANR device;
- the stability and recovery after overload of the ANR device.

Together with an evaluation in continuous noise, these measurements are needed if an ANR hearing protector is to be used in a military environment.

Annex E – CONTRIBUTION OF TNO HUMAN FACTORS, SOESTERBERG, THE NETHERLANDS

Objective Assessment of Attenuation and Speech Communication Quality of Headsets

Herman J.M. Steeneken and Jan A. Verhave
TNO Human Factors, The Netherlands

E1.0 INTRODUCTION

In 1936 Lueg [14] proposed to reduce the level of a primary sound by the addition of a secondary sound, the secondary sound being an inverted image of the primary sound. However, due to the lack of qualified electro-acoustical transducers his research was not successful. Twenty years later Olson and May [16] performed a renewed study but again, the transducer quality limited the success of the system.

Following an AGARD conference held in Soesterberg in 1981 TNO began development of their own ANR system and their first analogue system was produced in 1983. After many years of further research the present, state-of-the-art digital system features dynamic personal adaptation and advanced signal processing for the control of speech communications [1, 2, 5, 6, 18, 20, 24, 32, 34].

Successful development of such a system requires diagnostic feedback from the assessment methods. For that reason TNO developed methods for flexible determination of the sound attenuation, stability and speech quality.

A common method for determining the sound attenuation of hearing protection devices is the MIRE-method (microphone in real ear) where the sound level at the entrance of the earcanal is determined. In order to understand how this measurement is related to the sound level at the eardrum and with the subjective level sensation, a series of tests were performed and are described in summary in this Annex.

Speech communication over ANR headsets is another0relevant feature that has to be assessed and tuned for optimal performance. The speech quality can be determined subjectively with talkers and listeners but also with objective methods. The use of such an objective method, developed by TNO [28, 30, 33, and IEC60268-16] is also summarised in this Annex.

E2.0 ASSESSMENT METHODS

The performance of a hearing protector equipped with an ANR system and with an integrated audio facility depends on a number of technical properties. In order to predict the amount of protection and safety the device provides, the following parameters are of interest:

- passive sound attenuation as a function of frequency,
- active sound attenuation as a function of frequency,
- variance among systems,
- variance among users,
- stability of the open system during placing on or removing from the head,
- sensitivity for vibrations,
- maximum sound pressure level (dynamic range),

- overload response,
- speech intelligibility of the integrated communication system.

Some of these parameters and their effect on system performance were included in the Round Robin test and are described in detail in the main report. This annex will focus on the sound attenuation and speech intelligibility aspects.

E2.1 Sound Attenuation

With the introduction of active hearing protectors, which may introduce some system noise at the users ear, the assessment of the sound attenuation according to the standard measuring methods (ISO4869-1) is no longer valid. The ISO method is based on the threshold of perception and, thus, limited to low sound levels. The noise introduced by the ANR systems will interfere with the measurements. Also the sound attenuation of ANR systems is level dependent due to possible overload. Hence, measurements should be performed at various representative levels.

Four alternatives for measuring the sound attenuation may be used:

(1) Comparison of the sound pressure levels measured under the earmuff with the ANR system switched on and off. The level difference between the two measurements gives the sound attenuation of the ANR system. The measurements are performed by making use of the loop-microphone included in the ANR-system;

(2) Similar measurements as described under (1) by making use of an additional microphone, positioned close to the entrance of the earcanal;

(3) By subjective matching of the loudness of two sound levels, representative for the additional attenuation of the ANR system;

(4) By determining the masking of a test tone as a function of frequency [35].

E2.1.1 Objective Measurements

The active sound attenuation can be obtained by measuring the difference between the sound pressure levels under the earcup with the ANR system switched on and off. The measuring microphone may be the loop-microphone (1) or an additional microphone placed near the entrance of the earcanal (2).

The loop-microphone is part of the ANR system and positioned close to the loudspeaker or telephone cartridge in order to minimize the time delay in the feedback loop.

The additional microphone is placed close to the entrance of the earcanal (Figure E1). This method is called MIRE (Microphone In Real Ear) and is being considered as a new international standard for measuring the acoustic attenuation of hearing protection devices (Technical Committee CEN/TC 159). The MIRE method allows a comparison of the levels at an occluded (ANR switched on) and unoccluded ear. This comparison provides the total passive plus active attenuation of the ANR device.

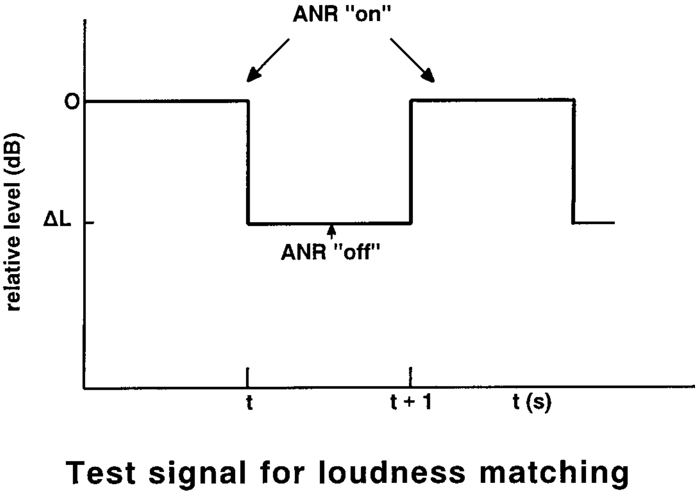
Figure E1: Measuring Microphone Near the Entrance of the Hearing Canal.

Ideally, the noise level and noise spectrum used for the assessment of the performance of the ANR headset would be identical to the noise level and spectrum of the true environmental noise the device will be used in. As ANR systems may have a level dependent performance it is advised that the attenuation is determined as a function of the noise level.

The attenuation may be determined for both the left and right earcups as a function of the frequency in 1/3 octave-bands using a spectrum analyser.

E2.1.2 Subjective Measurements

For the subjective assessment of the attenuation of active hearing protectors a subject (with an ANR system for each ear) is placed in a diffuse sound field which alternates periodically between two levels (typically every second). An example of this level alternation is given in figure E2.



Test signal for loudness matching

Figure E2: Relative Test Signal Level as a Function of Time for the Subjective Measurements of the Suppression of an ANR System. The ANR system is switched on and off simultaneously with the test signal envelope.

During the highest sound pressure level the ANR system was switched “on”, while during the low sound level the ANR system was switched “off”. The subject only hears a small difference between the two sound levels as the ANR system attenuates only the highest level. The subject was asked to match both levels for equal loudness by adjusting the level difference, L, between the two signals. The resulting difference in sound level outside the earmuff is equal to the subjective attenuation provided by the ANR system. The adjustment was made by changing the sound level during the “ANR-off” interval. Since the subject adjusts a continuous signal, the on/off rhythm was indicated with a light signal. The study showed that the accuracy lies between 1-3dB. However, it should be noted that the subject provides a response based on two ear listening.

This type of measurement has to be performed in a specific room with a diffuse sound field. The test signals are 1/3 octave-bands of noise and measurements are performed in one-octave steps. The absolute signal level can be adjusted to any level that is high enough not to interfere with the system noise. However, as the noise reduction of ANR systems may be level dependent, the measurements should be performed systematically as a function of the level.

The masking method as proposed by Zera et al [35] is based on detection of a pure tone that is masked by noise. The detection threshold shift when the ANR is switched on-off is related to the attenuation provided by the ANR.

This test method was not compared as the level calibration of the test tone is dependent on the ANR system and hence no benefit with respect to the MIRE method is obtained.

E2.1.4 Comparison of Subjective and Objective Measuring Results

Subjective and objective attenuation measurements were compared. The subjective attenuation was measured with four subjects and various signal levels. For one of the conditions the 1/3 octave-band signal level was 110dB SPL. The mean attenuation for these conditions, as a function of frequency with one-octave steps, is given in figure E3.

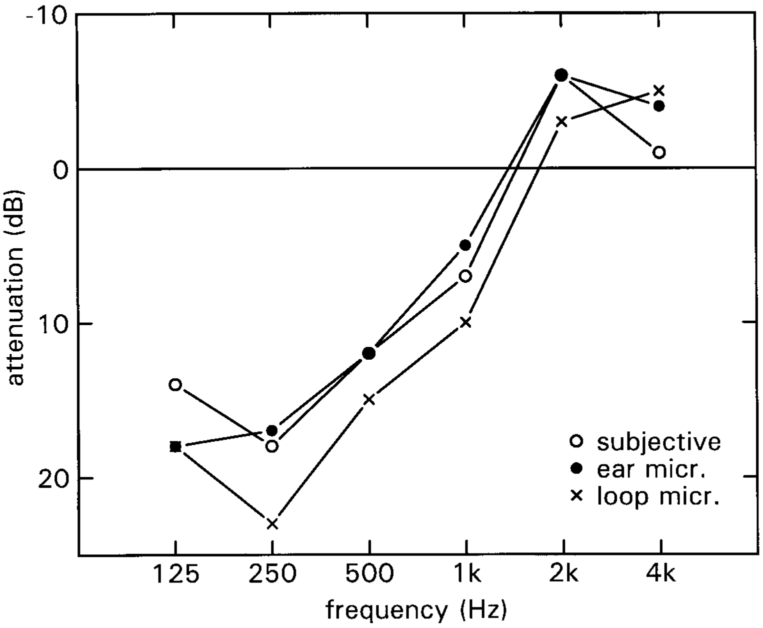


Figure E3: Mean Sound Attenuation Measured with 4 Subjects in One Octave Intervals for the Subjective and Objective Methods.

The objective attenuation (of the ANR system only) was measured with the loop microphone as well as with a special (MIRE) electret microphone positioned close to the entrance of the earcanal. For the objective measurement a pink noise (level 105 dB SPL) was used. The results indicate that the attenuation values obtained with the subjective level matching method and those obtained with the ear-microphone (MIRE) are in close agreement. The attenuation values obtained with the loop microphone are somewhat higher (2-5 dB). Obviously, the sound field under the earmuff is not homogeneous and is minimal at the sensing position of the loop microphone.

E2.2 Speech Transmission Quality

The speech quality over an ANR headset depends on the method used for the injection of the speech signal. Some ANR systems make use of a method that compensates for the suppression by the ANR system, other methods inject the speech signal directly at the loop microphone input. Sometimes designs make use of a correction amplifier.

As the speech transmission quality is defined by these design criteria of the ANR system and the speech injection method, it is important to assess the speech intelligibility in a representative condition.

Such an assessment may be performed with subjective measures (by making use of speakers and listeners) or by objective methods (by making use of a measuring device). During this study an objective method (the Speech Transmission Index, STI) is used. This method is standardized by IEC 60268-16 3rd edition 2002.

The STI-method assumes that the intelligibility of a transmitted speech signal is related to the preservation of the original spectral differences between speech sounds. These spectral differences may be reduced by band-pass limiting, masking noise, temporal distortion (echoes, reverberation, and automatic gain control), and non-linear distortion (system overload, quantisation noise). The reduction of these spectral differences can be quantified by the effective signal-to-noise ratio obtained for a number of frequency bands. Also, human related hearing aspects such as masking, the reception threshold, hearing disorders, and non-native speakers and listeners may reduce the effective signal-to-noise ratio. The method is based on the calculation of the effective signal-to-noise ratio in seven relevant frequency bands (octave-bands, centre frequencies ranging from 125 Hz to 8 kHz). Weighted contributions of the quantified information transfer function in seven octave bands results in a single index, the STI_r.

The relation between STI and three subjective measures is given in figure E4. Also a qualification of the transmission quality is given. These data were obtained from an international Round Robin held in 1984 [10].

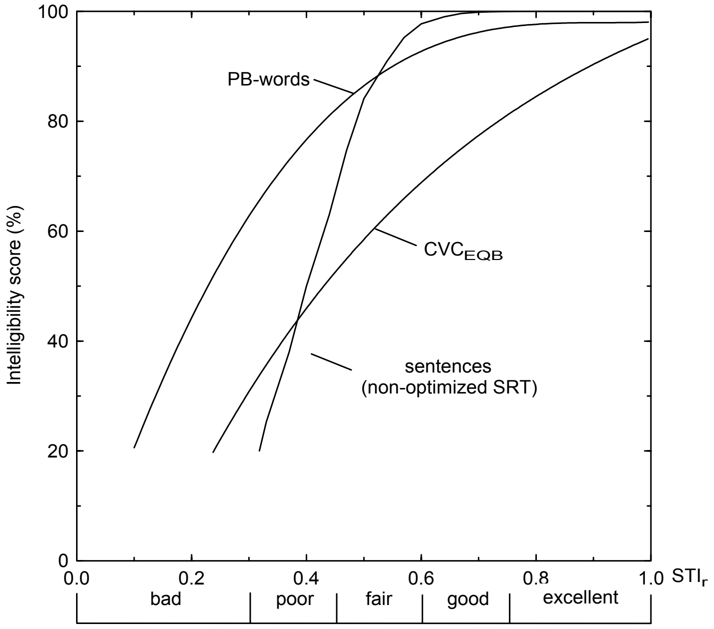


Figure E4: Qualification of the STI$_r$ and Relation with Various Subjective Intelligibility Measures for MALE Speech [10, 33].

The STI is obtained by applying a specific speech-like test signal at the audio input and by analysis of this transmitted test signal through the same measuring microphone as used with the MIRE attenuation measurements.

The STI for a specific communication system with ANR as a function of the noise level is given in figure E5. The STI is given for two conditions: ANR switched on and off.

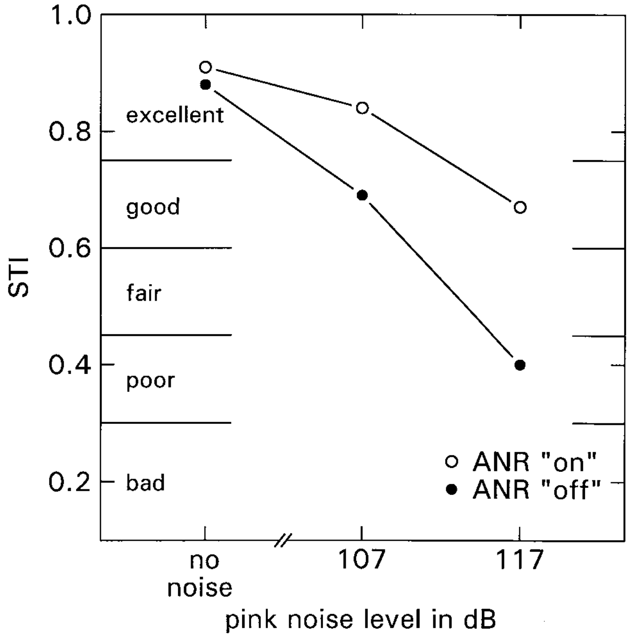


Figure E5: STI at Three Noise Levels for an ANR System Switched On and Off.

Hence, the effect of the ANR on the STI-value can be obtained by comparing the two conditions. Additional to the STI-value also a qualification (based on STI) is given. The improvement of the speech transmission quality is obvious. It is shown that for a constant speech intelligibility (STI=0.7) a 10-dB higher noise level can be applied. The *effective* gain in this situation and for this type of noise is 10 dB.

E3.0 SUMMARY

E3.1 Attenuation Assessment

The quantity that defines hearing protection is the noise dose at the eardrums of a person. It was found in earlier studies that, in case of using a hearing protector, this noise dose can accurately be predicted by reduction of the environmental noise level with the attenuation of the hearing protector (Insertion Loss). The normally used resolution in the frequency domain is with 1/3 octave steps.

The MIRE method, as used in the Round Robin test described in this report, provides results that are highly correlated with the subjective determination of the attenuation

E3.2 Speech Communication Assessment

The objective STI-method was applied for prediction of the speech communication quality. The advantage is that the same measuring set-up as used for the MIRE measurements can be used. STI provides diagnostic information on the speech transmission.

A technical report [34] provides the description and results of the TNO-HF experiments for the Round Robin project.

Annex F – CONTRIBUTION OF QINETIQ, FARNBOROUGH, UK

G.M. Rood, and S. H. James

F1.0 INTRODUCTION

Increasing noise levels in military vehicles (land, sea and air), particularly low frequency noise in some vehicles, can give problems in hearing protection, communications and detection of speech and non-speech signals that conventional passive hearing protection cannot resolve. The use of Active Noise Reduction (ANR) techniques can provide significant increases in attenuation and provide the added protection needed to resolve many of the noise-induced problems.

The use of ANR is now becoming more prevalent in military service and a number of European and US Government research establishments are working with such systems and measuring the active performance in a number of ways. Hence, some standardised methods of measurement between the various nations is becoming increasingly important. In 1999 an international collaborative task group (TG028) was formed under the NATO RTO/RTA HFM panel, to conduct a Round Robin series of measurements on a number of military ANR headsets to allow a comparison of the different assessment methods used by each nation, based on measured acoustic performance results.

For the series of measurements made across the international laboratories, the primary aim was to compare results using essentially the same acoustic measurement techniques and to ascertain and explain any major differences in (the same) headset performance. Thus the primary criterion was to compare measurement techniques and not the absolute and relative performance of the range of headsets. The ultimate aim of the collaboration being to provide a standardised assessment method, such that an international benchmark test can be achieved to allow internationally reproducible tests and performance results.

This report is a summary of the testing procedures and facilities used for the UK tests and the analysis carried out at DERA Farnborough (now QinetiQ). QinetiQ technical Report [23] provides more comprehensive technical content, results, data and figures.

F2.0 EXPERIMENTAL SET UP FOR UK TRIALS

F2.1 Facilities, Instrumentation and Calibration

The experimental test facilities equipment and calibration procedures are detailed fully in reference 1.

F2.2 Headsets Available

Six headsets were made available from the research establishments involved in the testing routine and were provided from TNO (Soesterberg) Holland (headset 6), QinetiQ (Farnborough) UK (headset 2), two from AFRL (Dayton) USA (headsets 1 and 5, remark: headset 5 was not used for the Round Robin), DRDC (Toronto) Canada (headset 4) and ISL Franco-German Institute (St. Louis) France (headset 3). To ensure that any complex fitting procedures did not complicate the results the headsets were all of the simple circumaural type and were used in a simple headband.

F2.3 Noise Field(s)

The noise field for this series of international tests was agreed as essentially a pink noise spectrum (flat in 1/3 octaves) rising in level at 3dB/octave (approximately 1dB/ 1/3-octave) above 1 kHz to accommodate the higher levels of headset or helmet attenuation at the higher frequencies. The spectrum had an overall Sound Pressure Level (SPL) of some 100 dB.

F2.4 Measuring Equipment

The measuring equipment comprised a B&K half inch microphone feeding a B&K Type 2133 Real Time Analyser to measure room noise fields, and two Knowles BT1759 miniature electret microphones to measure the noise field at the subjects' ears in the unoccluded and occluded state. To measure the noise floors of the measuring equipment a B&K Artificial Ear, suitably encased in a steel cylinder to prevent spurious flanking paths at these levels, was used with a heavy brass occluding cap that was sealed with a grease seal supported with a ‘plasticene’ outer seal. Additionally, the B&K low noise microphone was available to confirm the absolute low-level SPL of the room.

Measurements of the noise floor of the sensing microphone system showed an acceptable dynamic range for all acoustic measurements.

F3.0 ASSESSMENT METHODS

F3.1 Attenuation Measurements on Human Subjects: Microphone in Real Ear (MIRE)

The attenuation of the six headsets was measured using the MIRE technique involving six subjects with three replications of fit. All of the 6 headsets were tested in a single session, but each replication was in a different session, at least a day apart. The miniature microphone was fitted at the entrance to the concha and the ear remained fully unoccluded (i.e. no earplug was used). Unoccluded and occluded measurements were made on both ears simultaneously, with the unoccluded measurements made at the beginning and end of each session. Measurement of the levels was made using a 32-second integration time constant. Checks were made to ensure that this 32-second time period was sufficient to provide a fully stable noise measurement, by comparing to a spectrum measured using a 64-second time constant.

At the end of the third session for each headset the subject was asked to break the seal, (with the ANR system switched on but the noise field switched off) to see if the headset showed any signs of instability. The seal leakage was increased until either some instability, in terms of acoustic feedback occurred, or no instability was found.

This overall procedure was repeated with each of the remaining five headsets.

F3.2 Objective Attenuation Measurements: Artificial Ear

All measurements were made on a B&K artificial ear enclosed in a steel cylinder to minimise flanking paths.

The headsets were fitted on the artificial ear using their own headband and centralised over the measuring microphone. No damping or absorbing material was used under the earcup on the ‘flat plate coupler’. A single unoccluded measurement and occluded measurements on both left and right earcups were made with the ANR on and off. From this data the objective passive, passive plus active and active attenuation figures were obtained for comparison with the MIRE results.

Frequency Response

The frequency responses of the headsets were measured in passive and active mode on the B&K 2012 Audio Analyser using an artificial ear with a B&K 4134 half-inch microphone. Each headset was tested in eight conditions (2 earcups, 2 input voltages and with the ANR off and on). Comparisons were made between the responses at the two voltages (1v and 0.1v).

In theory, it is preferable for each headset to have a flat frequency response for both passive and active modes and to present no change of sensitivity of the telephone (dB-re1V-re1Pa). Under these conditions the communication response to speech and non-speech signals would remain unchanged in both the passive and active modes. However, in these series of international tests, the main parameter is the differences, if any, in response curves and/or sensitivity, measured by the same, or different, analysis tools.

High Level Linearity

The linearity of the active attenuation with increasing noise level was tested using the artificial ear system. The external noise levels produced in octave bands centred on 63, 125 Hz and 250 Hz were incrementally raised in 10dB steps from 80dB to 115dB. These frequencies were chosen, as generally they are where active attenuation performance is required.

Self Generated Noise Levels

To show the electronic noise generated by the electronics of the ANR system, each headset was placed on the artificial ear (in the quiet), and the ambient noise measured with the ANR system on and off on the right shell. The external noise levels in the room were low enough not to interfere with any acoustic measurements.

Comparison of MIRE and Artificial Ear Results

Comparisons were made of the active attenuation of the mean values of the MIRE tests and the single run of the artificial ear test of the left and right shells at each test frequency for each headset.

F4.0 EXPERIMENTAL RESULTS & DISCUSSION

F4.1 Attenuation

Microphone in Real Ear (MIRE)

The data analysed from the basic acoustic attenuation measurements were entered into a standardised analysis spreadsheet provided from TNO. This spreadsheet calculated the mean and standard deviation figures and plotted them in graphical form. Using this method all countries followed the same format and all results were directly comparable.

To calculate the acoustic attenuation, where the attenuation is measured on a real human using a miniature microphone, the unoccluded spectrum (i.e. measured at the subjects' ear with no protector in place) was subtracted from the relevant occluded spectrum (i.e. measurement at the ear with the protector in place and in either passive or active mode). This is an Insertion Loss method of measurement and requires no calibration corrections on the miniature measuring microphone.

Initially the passive attenuation of the system was calculated followed by the attenuation of the system in active mode (i.e. the passive + active attenuation). The active attenuation was calculated by subtracting these two sets of results. The attenuation was measured over the frequency range 12 Hz to 10 kHz.

The results of each headset are shown in figure F1, the plots displaying the passive, active and passive-plus-active acoustic attenuation and the associated standard deviations.

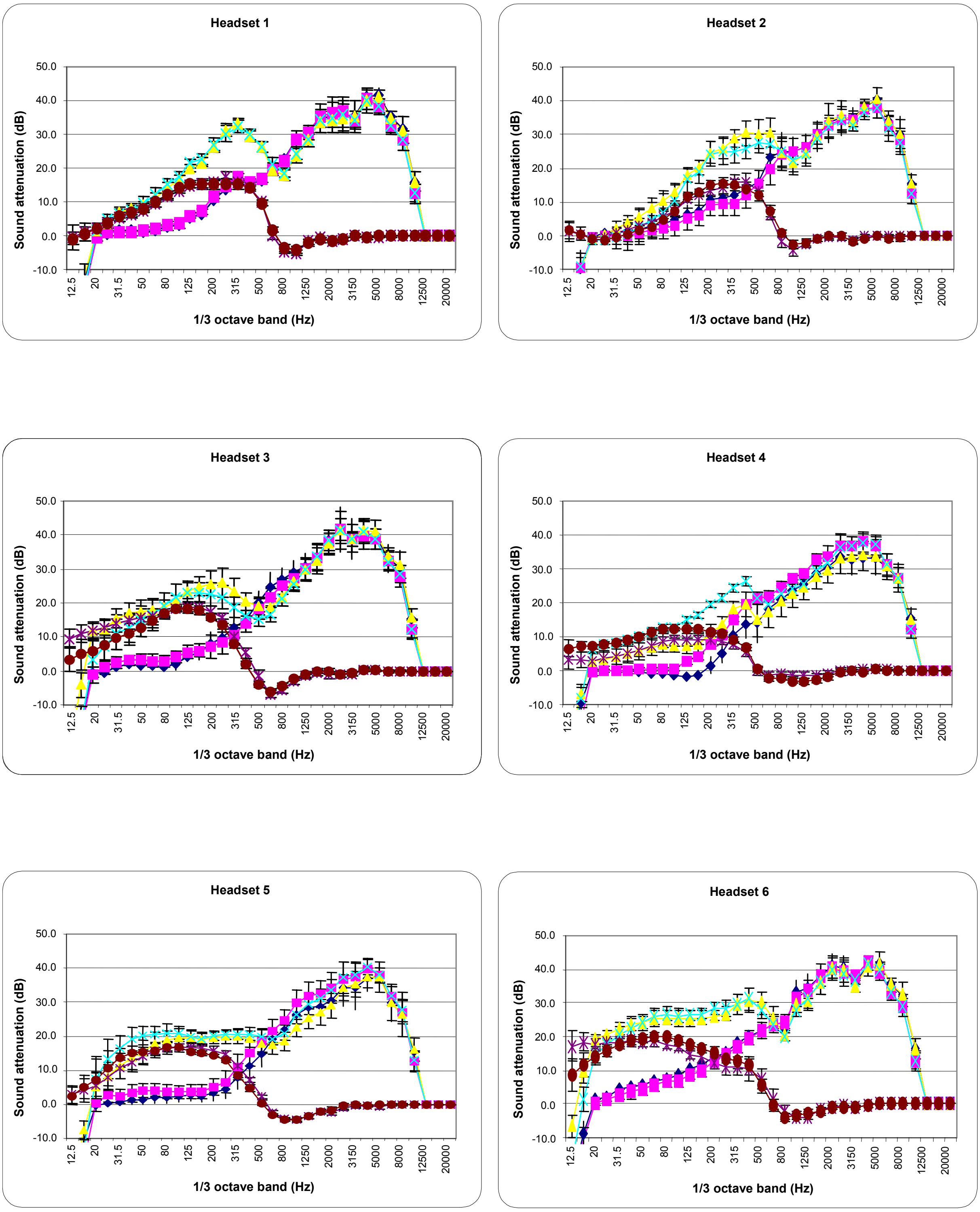


Figure F1: Passive, Passive + Active (Total) and Active Attenuation Curves for Left & Right Earcups of All Headsets Tested (headset 5 was not used for the round robin).

The standard deviations shown in figure F1 are a combination of the inter and intra subject variance (6 subjects with 3 fits each) and give some indication of the quality of fit to the head. The intra subject variance (i.e. the variance due to the fits on the same subject) is generally smaller than the inter subject variance (that due to the difference between subjects). Also, by using the miniature microphone technique and measuring the attenuation at each ear separately it is possible to get an indication of the quality of fit of each ear. However, whilst this is important when the absolute performance of the protector is concerned, in these experiments it is of lesser importance and the interest lies in the comparison of measurement results based on nominally the same procedure across different laboratories.

F4.2 Passive Attenuation

Looking at the passive attenuation data across the six headsets the main source of differences in measurements is the variance of fit. In some headsets the variance is small indicating, as noted above, a good and consistent fit.

Figure F2 shows the data for headset 1 (H/S 1), the 36 data points resulting from measurements on 6 subjects, 3 replications and left and right earcups. It is clear that below 1 kHz, where fit is important, the standard deviations are in the order of 1dB, which, in experimental measurements of this type, is good. At the higher frequencies (>1 kHz) the standards deviations increase to around 3dB, which is more typical for this type of headset measurement.

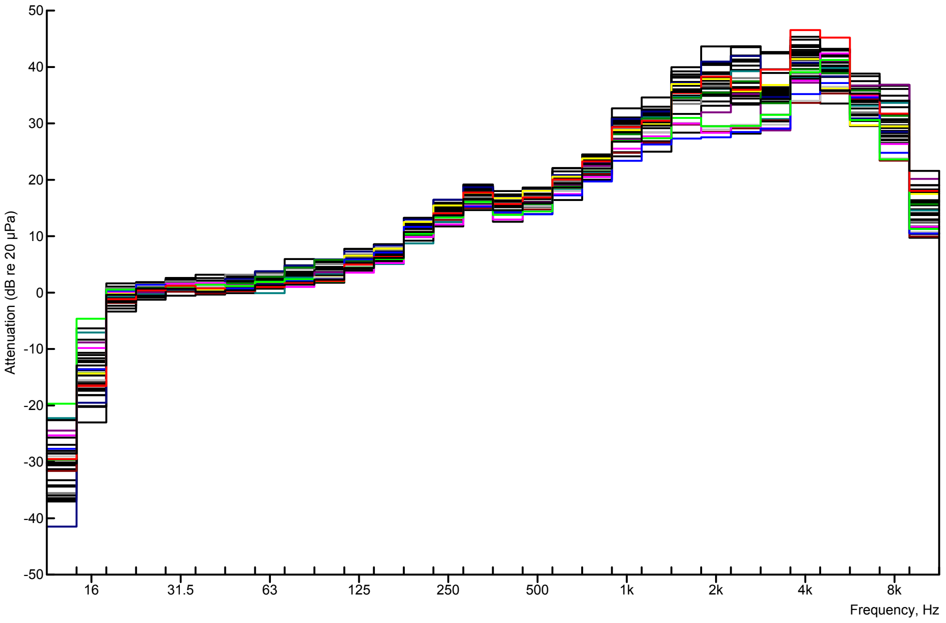


Figure F2: Headset 1, Passive Attenuation, 36 Data Points.

Where the fit is less consistent, for example in H/S 2 (see Figure F3), the standard deviations are more consistent with that found in measuring flight helmet attenuation. Here the earcup fit is not a simple and repeatable headband, but relies on the helmet fitting and tensioning devices to provide the quality of fit.

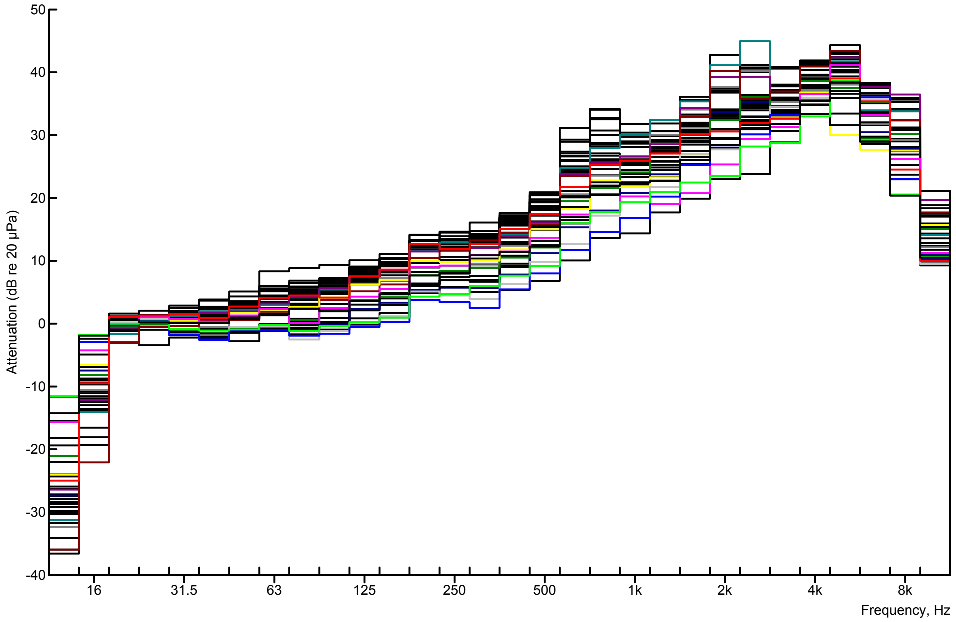


Figure F3: Headset 2, Passive Attenuation, 36 Data Points.

During the measurement process using miniature microphones to measure the SPLs at the ear directly, the attenuation figures are measured for the left and right earcup independently, and sometimes these attenuation figures, although nominally the same for identical earcups, can be different. This can be due to differences in earcup construction but in general it is more likely to be due to differences in quality of fit. If the attenuation figures are not statistically different for the left and right earcups, the data are combined to provide an overall attenuation figure for the acoustic attenuation of the headset or helmet.

The data from H/S 4 shows a case where this summation may be difficult. Differences in mean attenuation between shells are up to 6dB and this is statistically significantly different. A plot of all the files for the left and right earcups shows clearly that the left shell has a considerably wider spread of data than that of the right shell, and this is clearly evident in the variance figures. In this particular case, the fit of the left shell was being impaired by the design of headband.

Whilst these problems of fit found in H/S 4 are important where absolute figures of acoustic attenuation are being evaluated, in these series of tests under TG028, it is less important as the primary aim of the trials is to evaluate the reliability and similarity of measurement techniques across competent laboratories. In this case it will be important to see if the other laboratories replicated these differences.

F4.3 Passive plus Active Attenuation

The attenuation figures for this phase of testing mirror, to a large extent, the trend of data from the passive measurements. The results from H/S 1 (all 36 data files) display highly consistent data (as with the passive data). This closeness of data confirms the reliability and repeatability of the measurement and experimental technique, where the aim is to reduce the error variance to the minimum possible and to confirm that the variance measured is the true variance due to parameters in the design and construction of the measured device and not to experimental design and measurement techniques. This is clearly shown in this set of data, where the measurement technique is clearly reliable and repeatable.

The trend of the passive data is repeated for H/S 2 where the data is more widely spread, but well within experimental acceptability. Similarly for H/S 4, the measurement data for the left and right shells is different and in line with the passive results where the mean values across subjects for each replication (runs 1 to 3) for the left and right shells. As a further indication of the consistency of the left and right earcup measurements the runs for H/S 6 are a good example and this can be implied form the variance bars shown in figure F1. The left and right data is all closely bunched, and is typical for all the headsets other than H/S 4. Some headsets have a slightly greater variance of the mean, but all are within statistically acceptable limits.

F4.4 Active Attenuation

As the active attenuation is calculated by subtracting the ‘passive + active’ attenuation from the ‘passive’ attenuation, the results of the active attenuation are fully reflected by the trends of the previous attenuation data. The discrepancy between the left and right shells of H/S 4 is maintained, with the left shell having a peak active attenuation of some 3dB less than the right shell. However, it is interesting to note that the shape of the active attenuation characteristic of the left shell is virtually identical to that of the right shell. This provides further confirmation of poor fit and greater acoustic leakage affecting the acoustic impedance and the effective acoustic volume of the left earcup.

F4.5 Artificial Ear

The results of the single attenuation measurement made for the left and right earcups show differences that are small and generally within 2 to 3dB of each other, which is experimentally acceptable for a single run. It would be anticipated that, if a number of measurements were made on each headset earcup, then the mean values of the left and right shells would not be statistically significantly different. The exception to this is on H/S 4, where the differences between left and right shells are generally 2 to 3dB apart and reflect the MIRE results for each shell.

F4.6 Frequency Response

The responses were analysed in terms of consistency between the left and right earcups, the consistency between the voltage inputs of 1 volt and 0.1 volt (20dB) and in the change in effective sensitivity of the earcup telephone.

For all headsets, apart from H/S 4, the left and right earcup data was consistent to within 2dB. This is not unexpected, as the earcups are only different in terms of production tolerances. H/S 4 had differences of 14.5dB between the left and right shells (the right shell having the higher sensitivity), but the expected differences in the voltages inputs were highly consistent at 20dB. This input voltage consistency was also apparent on the other 5 headsets.

Where the main differences were found was in the changes in effective sensitivity when the ANR systems were active, compared to the passive state. Headsets 1, 2 and 5 had changes of around 9dB (at 1 kHz) whilst the other 3 sets were less than 3dB.

F4.7 High Noise Linearity

From the data plotted, it is clear that, apart from H/S 2 in the 63 Hz band noise, all headsets are effective in providing the same active attenuation performance up to external noise levels of 115 dB SPL in the tested frequency bands. The reason for H/S 2 dropping off marginally in performance is the design of the active circuitry, which has been optimised for the higher frequency bands found in fast jet aircraft, and this is confirmed by the performance in the 125 and 250 Hz noise bands.

F4.8 Self Generated Noise

All 5 systems produced some electronically generated noise when switched to the active mode. This is not surprising as the telephone becomes part of the feedback system and, in many cases it is useful for the operator to have some feedback that the system is operating when switched on. However, it is obviously important that the levels are not high enough to interfere with the prime purpose of the telephone, communications, or add to any hearing damage risk. As the ANR systems are intended to run in high noise environments, then low levels of self generated noise are acceptable, unless, in some special cases, the levels of the signals to be detected have a similar spectral content to that of the self generated system noise that may cause masking.

All of the systems produced noise peaking around the 500 to 800Hz 1/3-octave bands, at levels of around 30dB(A), H/S 4 was a bit higher at 38dB(A), but in all cases these levels were considered acceptable. In addition, for all headsets there was some low frequency noise (generally below 100Hz) and in all cases, apart from H/S 5, the use of the active mode reduced the overall levels by 2 or 3dB. Again this low frequency noise is not at levels high enough to cause concern.

F4.9 Comparison of MIRE & Artificial Ear Results

The comparison of the attenuation measured by both methods shows that the results of the two are remarkably similar. There is a trend towards the differences increasing as the frequency rises up to where the active attenuation ceases to be effective (around 1 kHz), but, apart from headsets 1and 2, where the differences at 630 Hz and 800 Hz are around 4dB, the differences are between 1 and 2.5dB. This might indicate that the use of an Artificial Ear or Artificial Head could provide a simple and quick measurement of active systems of this type, and it will be instructive to compare the results of the other laboratories.

However, this is the case for the measurement of the active component of the overall attenuation and some work is needed to correlate the measurements of the passive plus active attenuation with MIRE and Artificial Ear, which is the data used in calculations of the effective acoustic attenuation, and to complete more replications of the objective (i.e. Artificial Ear) measurements.

F4.10 Rejection of Outliers

There are well recognised statistical techniques for the rejection of data points or files, if, with some common sense and expertise and experience of testing, the data is considered incorrect due, perhaps, to values relating to other parameters than those directly involved in the experimental design. For instance this could be due to experimenter errors in, say, setting the measurement parameters incorrectly in the computer or inadvertently allowing the measuring microphone(s) to distort in over or underload. Whilst it is sometimes poor experimental control which causes these outliers, it must also be incumbent on the experimenter to use only data that is considered to be reliable, and can be shown statistically to be reliable, and to refrain from contaminating the results with known bad data.

DERA/QinetiQ normally uses Chauvenet’s Criterion as a statistical procedure for the rejection of these outliers. In this particular case such rejections are not specifically necessary (although they were performed) as the tests are solely to look at measurement procedures. However, differences in mean values and their associated standard deviations may be important in the comparisons of such measurement techniques, and these types of rejection procedures were used to clarify the data. To provide a comparison, the results were also calculated using the full set of data, potential outliers included. From a comparison of these two data sets it is clear that the differences are relatively small, generally less that 2dB, and this would be the expected outcome of the mean values and associated standard deviations when measuring 36 points with only 1 or 2 outliers. Thus if the differences are small, it would be possible to use either set of data, but for scientific reliability (or integrity) it is considered better to use the reduced set of fully reliable data.

F5.0 CONCLUSIONS

A number of methods for the assessment of the acoustic attenuation of the ANR systems were compared. In these series of national tests the consistency of the test method is generally shown in the variance of the results that is displayed when the fit of the headset is reliable and repeatable. This has been shown for those headsets that fitted well and the error variance is small compared to the real variances caused by the parameters of the headsets, such as repeatability of fit.

A comparison is made between the MIRE method and that method using a flat plate coupler and ‘artificial ear’, and this shows a potential promise for an inexpensive method of measurement for the determination of the active component of the ANR systems. However, further work is needed to indicate whether it can be reliably used for determining the overall (passive + active) attenuation that is the figure used in the determination of the hearing protection ability.

F6.0 RECOMMENDATIONS

The series of tests have highlighted the spread of variance caused by the variability of fit. In some headsets the fit is consistent and good and the variability of the acoustic performance between subjects small. In others it is less effective.

In terms of the assessment of hearing damage risk from the acoustic attenuation performance, the smaller the variance of attenuation, the easier it is to meet the requirements of the Health & Safety regulations in terms of the hearing risk population. In practical terms this would mean that a greater number of the noise exposed population could be accommodated within the guidelines for a given acoustic attenuation and that high levels of acoustic attenuation, needed to meet the statistically broader population, could be ameliorated. This could mean that, with a consistently fitting and acoustically efficient erase for a flying helmet, headset or carrier flight-deck helmet, the need for such high levels of acoustic attenuation to meet the population acoustic protection levels could be reduced, thus making it less complex to reach.

From this argument it is recommended that research is funded and undertaken to find new or existing earseals for headsets and helmets that will provide a consistent across-subject fit combined with acoustic efficiency.

As a potential example the calculations show that for a flight-deck helmet to be used for JSF operations, the acoustic attenuation required to meet the UK services Health and Safety requirements could be met with a less stringent acoustic attenuation requirement.

Annex G – CONTRIBUTION OF AFRL/HECB, DAYTON, USA

Speech Intelligibility versus Speech-to-Noise Ratio and Attenuation Performance with Active Noise Reduction Headsets

R.L. McKinley, M.A. Ericson, B.D. Simpson, D.S. Brungart
AFRL/HECB, WPAFB, OH

K.L. Youngblood
General Dynamics, Dayton, OH, USA

G1.0 INTRODUCTION

Active Noise Reduction (ANR) headsets have been a goal ever since the description of the concept by a German scientist, Paul Lueg, in 1936. The predecessor laboratory to the Air Force Research Laboratory reported a demonstration and performance measurement of an ANR headset in 1957. That system, designed by Willard Meeker, had performance characteristics which are remarkably similar to performance of modern ANR system. However, the 1957 Meeker system used a large tube amplifier, shown in Figure G1, and limited bandwidth in the cancellation loop. Also, the cancellation microphone was large by current standards. The demonstration of this system eventually led to the design of modern ANR systems in the USA which began in 1978. ANR systems are now being applied in several military systems including airborne command and control, high performance jet fighter aircraft, tracked ground vehicles, and jet engine ground maintenance. Characterization of the performance of ANR systems is important for the safety, health, and hearing of the users of ANR. This effort was the result of AFRL participation in HFM TG-028.

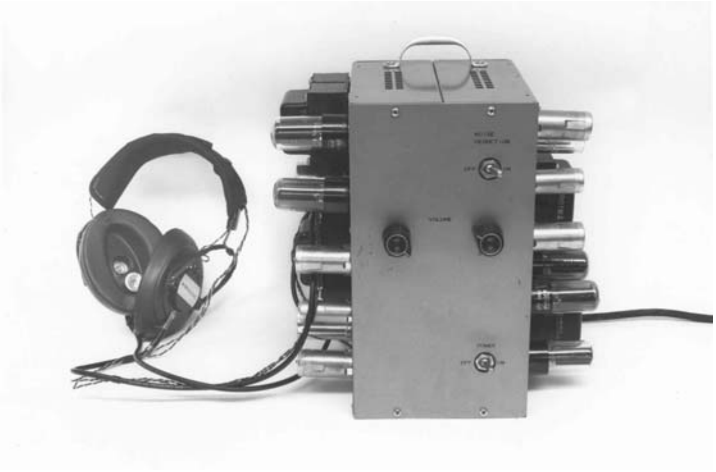


Figure G1: Active Noise Reduction Headset Wright Aeronautical Development Center Circa 1957.

G2.0 BACKGROUND

The speech communication and attenuation performance of five ANR headsets was measured using the Modified Rhyme Test (MRT) as described in ANSI-S3.2 (reference 3) and insertion loss/attenuation as described in ANSI S12.42 (reference 2). Speech to noise ratios (SNR) in the various applicable frequency bands are the primary parameter affecting speech intelligibility. While passive, active, and total attenuation, in frequency dependent dB, were the measures of attenuation.

The speech intelligibility objective was to measure the speech intelligibility versus SNR function to compare with the STI speech intelligibility predictions for the five different ANR headsets in the TG-028 study. The approach was to fix speech presentation levels at 75 dB SPL and to use five different ambient noise levels (90 dB, 95 dB, 100 dB, 105 dB, and 110 dB) to achieve a range of SNRs at the listeners' ears.

The attenuation objective was to measure the total, active, and passive attenuation for 1/3 octave-band frequencies from 50 Hz to 12.5 kHz in a 105 dB pink noise for the TG-028 Round Robin study. The approach was to generate the ambient noise in a medium sized reverberation chamber using a high intensity sound source, shown in figure G2, and measure the insertion loss on a group of human subjects using the standardized Miniature Microphone in Real Ear (MIRE) technique described in ANSI 12.42 (reference 2).

Figure G2: Air Force Research Laboratory High Intensity (142 dB) Reverberation Room.

G3.0 METHODS

G3.1 Speech Intelligibility

The experiment was a within subjects design using five SNRs, five ANR headsets, ten listeners, and two talkers. One male and one female talker were used. A total of 50 trials were conducted. Five listeners participated per session. The MRT speech stimuli were presented via a PC sound card. For all conditions the ANR system was "ON". A research intercommunication system was used with a 10 kHz bandwidth. The Air Force Research Laboratory Voice Communication Research System (VOCRES) facility, figure G3, was used to generate the ambient sound field, to generate the MRT stimuli, and collect the subject responses to the stimuli.

Figure G3: Air Force Research Laboratory.

VOCRES Facility

G3.2 Subjects

The five male and five female listeners had hearing threshold levels no worse than 15 dB HTL at 250, 500, 1000, 2000, 3150, 4000, 6300, and 8000 Hz. One male and one female talker were selected from the AFRL/HECB digital MRT database. The talkers in this database had no strong regional accents and did not have any speech pathologies.

G3.3 Procedure

The long term RMS speech spectrums were measured for MRT speech stimuli spoken by each talker. A fifty word list was analysed using a B&K 2131 spectrum analyser set for 1/3 octave-band analysis and using a 128 second linear averaging time. The MRT stimuli with carrier phase were presented at 5.1 second intervals. Ten MRT phases were measured for speech and non-speech time. The average level correction was computed using the following equation:

$$meanlevelcorrection = 10 * \log(meantotaltime / meanspeechtime)$$

If the average speech time was 50% of the average total stimulus time, the correction should have been 3 dB and would be added to the overall level of the speech. This was to correct for averaging the non-speech time in the long-term RMS level and spectrum of the speech. This procedure was accomplished using one 50 word MRT list for each talker.

The speech presentation level and spectrum for each headset was computed using the same procedure with the addition of the ANR headset being coupled to a B&K artificial ear with a 1 inch pressure microphone and a flat plate coupler. The same 50 word MRT list used to calibrate the talkers was used to calibrate the headsets. The long-term RMS 1/3 octave-band spectrum and overall sound pressure level was recorded for each combination of talker and ANR headset. Gain factors were computed for each combination of talker and ANR headset to give a 75 dB SPL speech presentation. These gain factors were used during the speech intelligibility data collection.

The noise levels under the earcups of the ANR headset were measured for each of the five ANR headsets and each of the five noise levels. The ambient noise spectrum was similar to the long-term speech noise [26]. The long-term RMS 1/3 octave-band spectrum of the ambient noise was measured using the B&K 2131 and a 128 second linear average and recorded. The long-term RMS of the noise under the headset was measured by placing the ANR headset with a one pound weight on the B&K artificial ear with 1" pressure microphone and flat plate coupler. The B&K 2131 was set to conduct a 1/3 octave-band analysis using a 128 second linear averaging time. The overall level and 1/3 octave-band spectra from 63 Hz to 12.5 kHz were recorded for each combination of ANR headset and noise level.

Five listeners participated in the experiment in each session. During a five trial session, each listener listened to a 50 word MRT list (trial) one time with each headset. The presentation order of talkers and SNRs (ambient noise levels) was randomised. Number correct for each MRT trial was collected for each listener. After five trials, the subjects were given a 6 minute rest break. Following the rest break, an additional five trials were measured followed by another 6 minute rest break. Following the rest break, another five trials were collected followed by the last 6 minute rest break and the final five trials for the day. This collection block was conducted over approximately two hours. The second set of five listeners followed the same process. Approximately 5 hours of participation distributed over 3 days were required for each subject.

G3.4 Attenuation

The attenuation method used a repeated measures design where insertion losses/attenuations measured 3 times on each of the 10 subjects. This method gave both the inter and intra subject differences. The insertion losses/attenuations were simultaneously measured on both the left and right ears by using a dual channel real-time analyser, B&K 2133. The sound field was generated by a high power (12 kWatt) sound system installed in a medium sized (approximately 2,000 cubic feet) reverberation chamber. The sound system was capable of generating a maximum overall sound pressure level of 142 dB spanning a frequency range from 16 Hz to 12.5 kHz. Sound field data was collected using a miniature microphone mounted on a foam earplug for open, closed, and closed active conditions and the insertion loss-attenuations were calculated for each condition.

G4.0 ANALYSIS

G4.1 Speech Intelligibility

The data was computed for per cent correct with an adjustment for guessing. The algebraic reduction of the correction for guessing is in the following equation:

Percent correct (adjusted for guessing) = (2.4*number correct)-20

A mean percent correct (adjusted for guessing) and standard deviation was computed for each combination of headset and SNR collapsed over both talkers and the ten listeners. A second set of means and standard deviations was computed for each combination of headset and SNR and talker collapsed over the ten listeners.

G4.2 Insertion Loss/Attenuation

Individual insertion losses/attenuations were calculated for each subject, each device, left and right ear, and each fit. Means and standard deviations relative to frequency were computed collapsing across ears, fits, and subjects for each device.

G5.0 DATA

G5.1 Speech Intelligibility

The mean speech intelligibility data for each of the two talkers and the average of the two talkers are shown below in Figures G4-G6.

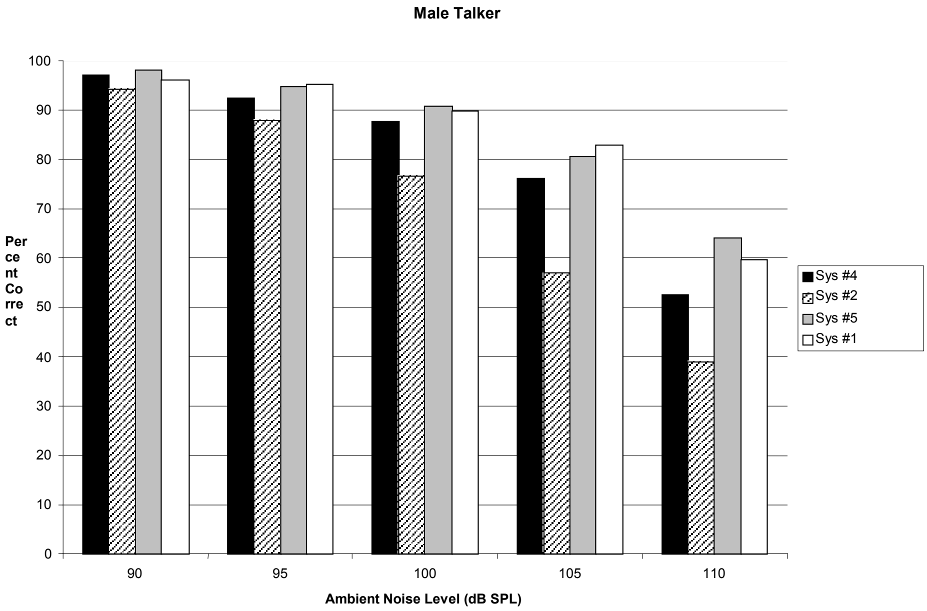


Figure G4: MRT Intelligibility vs Noise Level for a Typical Male Talker.

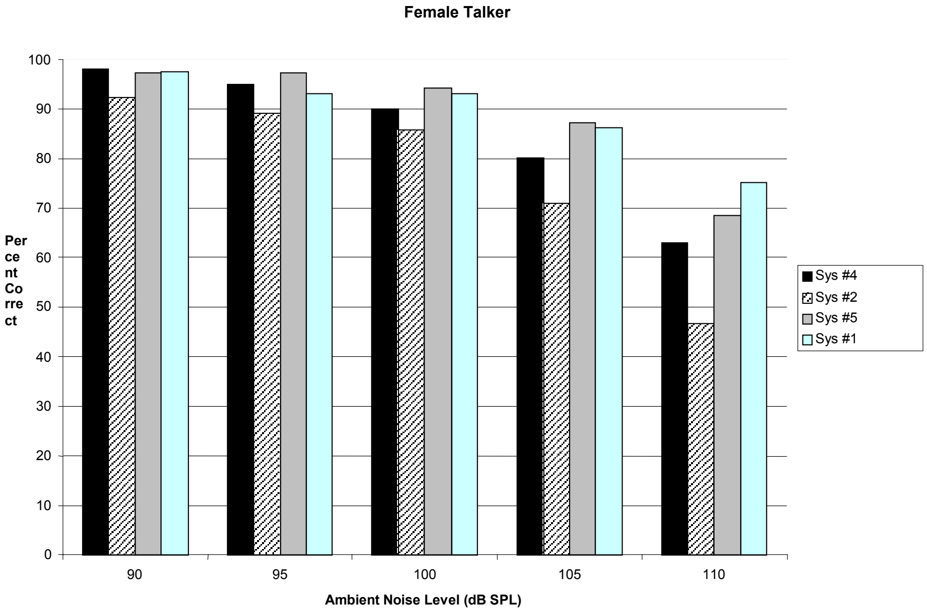


Figure G5: MRT Intelligibility vs Noise Level for a Typical Female Talker.

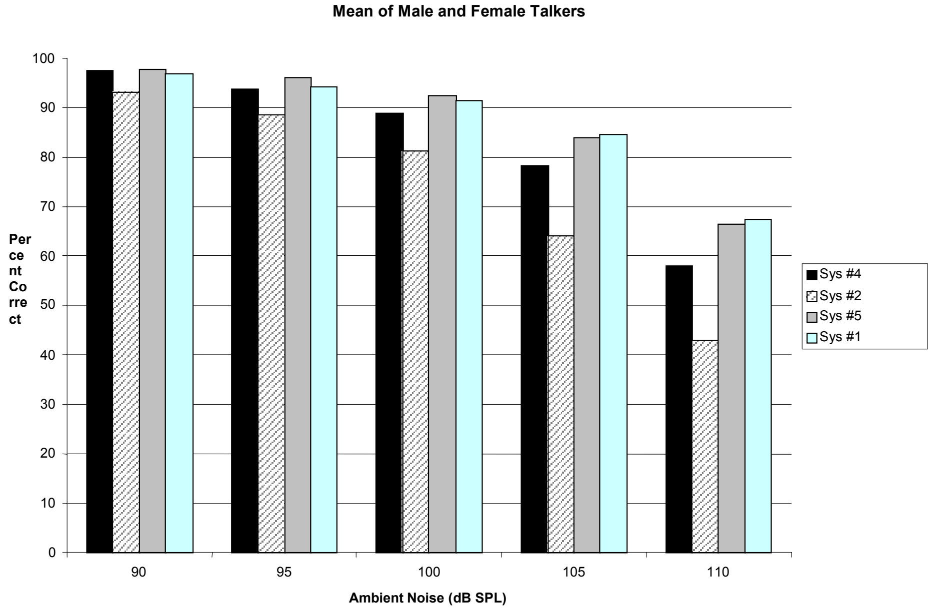
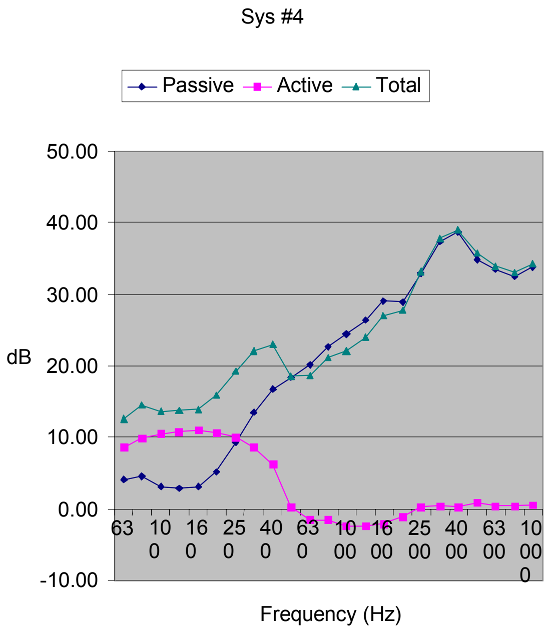


Figure G6: MRT Intelligibility vs Noise Level for Two Typical Talkers.

G5.2 Attenuation



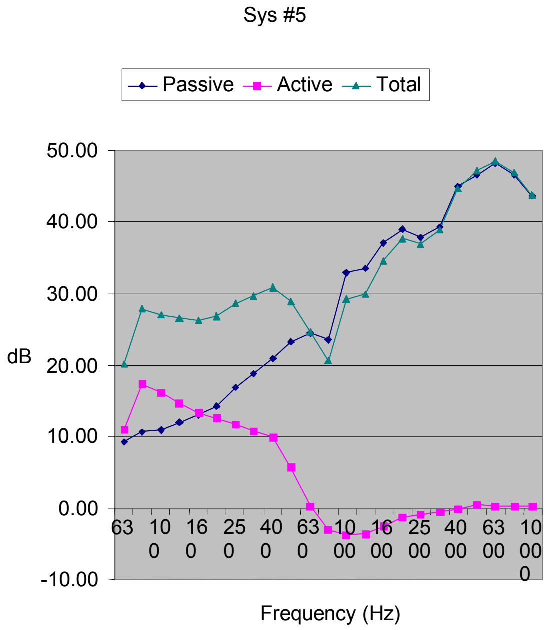


Figure G7: Attenuation vs Frequency for System #4.

Figure G8: Attenuation vs Frequency for System #5.

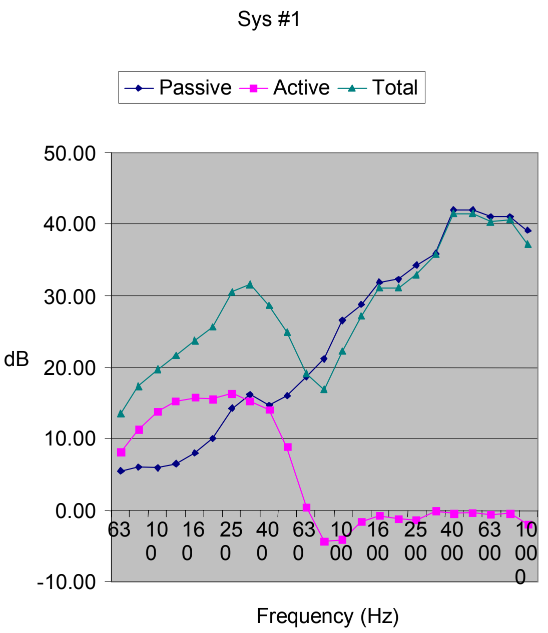


Figure G9: Attenuation vs Frequency for System #1.

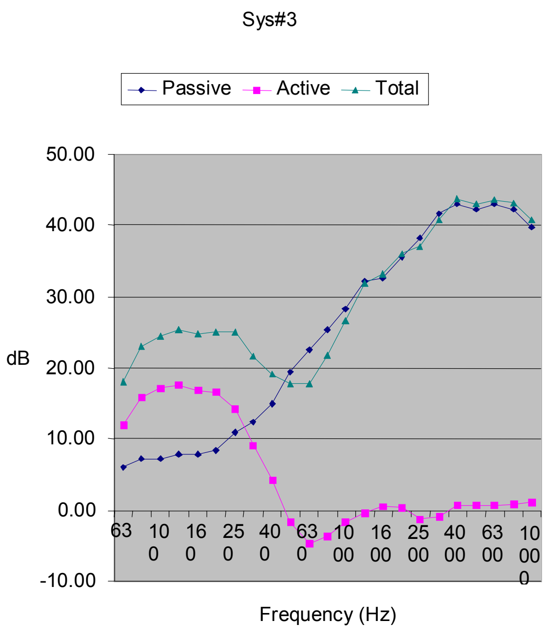


Figure G10: Attenuation vs Frequency for Telemit ANR.

G6.0 SUMMARY

The speech intelligibility and insertion loss/attenuation was measured on the group of ANR headsets used in the NATO HFM TG-028 Round Robin study using ANSI standard methodologies using human subjects. The performance of the systems were characterized using 5 subjects for speech intelligibility and 10 subjects for insertion loss/attenuation. Overall, ANR systems vary in performance, but the better systems can reduce the level of noise at the ear thereby reducing hearing damage risk.

G7.0 ACKNOWLEDGEMENTS

The authors would like to acknowledge the significant contributions of Alex Kordick, Paul Schley, Vernie Fisher, and Mike Ward in the collection and analysis of the speech intelligibility and insertion loss/attenuation data.

G8.0 USA STANDARDS

[1] ANSI S12.6-1997 (R2002) American National Standard Methods for Measuring the Real-Ear Attenuation of Hearing Protectors.

[2] ANSI S12.42-1995 (R1999) American National Standard Microphone-in-Real-Ear and Acoustic Test Fixture Methods for the Measurement of Insertion Loss of Circumaural Hearing Protection Devices.

[3] ANSI S3.2-1989 (R1999) American National Standard Method for Measuring the Intelligibility of Speech Over Communication Systems.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	**2. Originator's References** RTO-TR-HFM-094 AC/323(HFM-094)TP/53	**3. Further Reference** ISBN 92-837-1121-1	**4. Security Classification of Document** UNCLASSIFIED/ UNLIMITED

5. Originator

Research and Technology Organisation
North Atlantic Treaty Organisation
BP 25, F-92201 Neuilly-sur-Seine Cedex, France

6. Title

Assessment Methods for Personal Active Noise Reduction Validated in an International Round Robin

7. Presented at/Sponsored by

The RTO Human Factors and Medicine Panel (HFM-094/TG-028) as a result of a project on "Assessment of Personal Active Noise Reduction".

8. Author(s)/Editor(s) Multiple	**9. Date** August 2004
10. Author's/Editor's Address Multiple	**11. Pages** 106

12. Distribution Statement

There are no restrictions on the distribution of this document. Information about the availability of this and other RTO unclassified publications is given on the back cover.

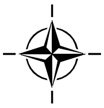
13. Keywords/Descriptors

Acoustic measurement
Active Noise Reduction (ANR)
Aircraft noise
Attenuation
Comparison
Defence programmes
Ear protectors
Helicopters
Helmets
Human factors engineering
Military aircraft
Noise (sound)
Noise reduction
Protective equipment
Research management

14. Abstract

Hearing protection with standard passive earmuffs is for many military applications insufficient. Low frequency noises (diesel engines, helicopter rotor) require increased attenuation. Active Noise Reduction (ANR) is specifically suited for reduction at low frequencies. Assessment of ANR for military applications is an important topic as the adverse military environment may cause unexpected reduced performance (e.g. very low rotor-blade noise may overload a for civil applications designed system).

This study contains assessment methods for ANR in military applications from five laboratories.



NORTH ATLANTIC TREATY ORGANISATION

BP 25
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@rta.nato.int

RESEARCH AND TECHNOLOGY ORGANISATION

DIFFUSION DES PUBLICATIONS
RTO NON CLASSIFIEES

Les publications de l'AGARD et de la RTO peuvent parfois être obtenues auprès des centres nationaux de distribution indiqués ci-dessous. Si vous souhaitez recevoir toutes les publications de la RTO, ou simplement celles qui concernent certains Panels, vous pouvez demander d'être inclus soit à titre personnel, soit au nom de votre organisation, sur la liste d'envoi.

Les publications de la RTO et de l'AGARD sont également en vente auprès des agences de vente indiquées ci-dessous.

Les demandes de documents RTO ou AGARD doivent comporter la dénomination « RTO » ou « AGARD » selon le cas, suivi du numéro de série. Des informations analogues, telles que le titre est la date de publication sont souhaitables.

Si vous souhaitez recevoir une notification électronique de la disponibilité des rapports de la RTO au fur et à mesure de leur publication, vous pouvez consulter notre site Web (www.rta.nato.int) et vous abonner à ce service.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Streitkräfteamt / Abteilung III
Fachinformationszentrum der
Bundeswehr (FIZBw)
Friedrich-Ebert-Allee 34, D-53113 Bonn

BELGIQUE
Etat-Major de la Défense
Département d'Etat-Major Stratégie
ACOS-STRAT – Coord. RTO
Quartier Reine Elisabeth
Rue d'Evère, B-1140 Bruxelles

CANADA
DSIGRD2
Bibliothécaire des ressources du savoir
R et D pour la défense Canada
Ministère de la Défense nationale
305, rue Rideau, 9^e étage
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Defence Research Establishment
Ryvangs Allé 1, P.O. Box 2715
DK-2100 Copenhagen Ø

ESPAGNE
SDG TECEN / DGAM
C/ Arturo Soria 289
Madrid 28033

ETATS-UNIS
NASA Center for AeroSpace
Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076-1320

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc
BP 72, 92322 Châtillon Cedex

GRECE (Correspondant)
Defence Industry & Research
General Directorate, Research Directorate
Fakinos Base Camp, S.T.G. 1020
Holargos, Athens

HONGRIE
Department for Scientific Analysis
Institute of Military Technology
Ministry of Defence
H-1525 Budapest P O Box 26

ISLANDE
Director of Aviation
c/o Flugrad
Reykjavik

ITALIE
Centro di Documentazione
Tecnico-Scientifica della Difesa
Via XX Settembre 123
00187 Roma

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25, NO-2007 Kjeller

PAYS-BAS
Royal Netherlands Military
Academy Library
P.O. Box 90.002
4800 PA Breda

POLOGNE
Armament Policy Department
218 Niepodleglosci Av.
00-911 Warsaw

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide
P-2720 Amadora

REPUBLIQUE TCHEQUE
DIC Czech Republic – NATO RTO
LOM PRAHA s. p.
o.z. VTÚL a PVO
Mladoboleslavská 944, PO BOX 16
197 21 Praha 97

ROYAUME-UNI
Dstl Knowledge Services
Information Centre, Building 247
Dstl Porton Down
Salisbury
Wiltshire SP4 0JQ

TURQUIE
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi Başkanlığı
06650 Bakanliklar – Ankara

AGENCES DE VENTE

NASA Center for AeroSpace
Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076-1320
ETATS-UNIS

The British Library Document
Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
ROYAUME-UNI

Canada Institute for Scientific and
Technical Information (CISTI)
National Research Council
Acquisitions, Montreal Road, Building M-55
Ottawa K1A 0S2, CANADA

Les demandes de documents RTO ou AGARD doivent comporter la dénomination « RTO » ou « AGARD » selon le cas, suivie du numéro de série (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Des références bibliographiques complètes ainsi que des résumés des publications RTO et AGARD figurent dans les journaux suivants :

Scientific and Technical Aerospace Reports (STAR)
STAR peut être consulté en ligne au localisateur de
ressources uniformes (URL) suivant:
http://www.sti.nasa.gov/Pubs/star/Star.html
STAR est édité par CASI dans le cadre du programme
NASA d'information scientifique et technique (STI)
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
ETATS-UNIS

Government Reports Announcements & Index (GRA&I)
publié par le National Technical Information Service
Springfield
Virginia 2216
ETATS-UNIS
(accessible également en mode interactif dans la base de
données bibliographiques en ligne du NTIS, et sur CD-ROM)

NORTH ATLANTIC TREATY ORGANISATION

RESEARCH AND TECHNOLOGY ORGANISATION

BP 25
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@rta.nato.int

DISTRIBUTION OF UNCLASSIFIED RTO PUBLICATIONS

AGARD & RTO publications are sometimes available from the National Distribution Centres listed below. If you wish to receive all RTO reports, or just those relating to one or more specific RTO Panels, they may be willing to include you (or your Organisation) in their distribution.

RTO and AGARD reports may also be purchased from the Sales Agencies listed below.

Requests for RTO or AGARD documents should include the word 'RTO' or 'AGARD', as appropriate, followed by the serial number. Collateral information such as title and publication date is desirable.

If you wish to receive electronic notification of RTO reports as they are published, please visit our website (www.rta.nato.int) from where you can register for this service.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Etat-Major de la Défense
Département d'Etat-Major Stratégie
ACOS-STRAT – Coord. RTO
Quartier Reine Elisabeth
Rue d'Evère
B-1140 Bruxelles

CANADA
DRDKIM2
Knowledge Resources Librarian
Defence R&D Canada
Department of National Defence
305 Rideau Street
9th Floor
Ottawa, Ontario K1A 0K2

CZECH REPUBLIC
DIC Czech Republic – NATO RTO
LOM PRAHA s. p.
o.z. VTÚL a PVO
Mladoboleslavská 944, PO BOX 16
197 21 Praha 97

DENMARK
Danish Defence Research
Establishment
Ryvangs Allé 1
P.O. Box 2715
DK-2100 Copenhagen Ø

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc
BP 72
92322 Châtillon Cedex

GERMANY
Streitkräfteamt / Abteilung III
Fachinformationszentrum der
Bundeswehr (FIZBw)
Friedrich-Ebert-Allee 34
D-53113 Bonn

GREECE (Point of Contact)
Defence Industry & Research
General Directorate, Research Directorate
Fakinos Base Camp, S.T.G. 1020
Holargos, Athens

HUNGARY
Department for Scientific Analysis
Institute of Military Technology
Ministry of Defence
H-1525 Budapest P O Box 26

ICELAND
Director of Aviation
c/o Flugrad, Reykjavik

ITALY
Centro di Documentazione
Tecnico-Scientifica della Difesa
Via XX Settembre 123
00187 Roma

LUXEMBOURG
See Belgium

NETHERLANDS
Royal Netherlands Military
Academy Library
P.O. Box 90.002
4800 PA Breda

NORWAY
Norwegian Defence Research
Establishment
Attn: Biblioteket
P.O. Box 25, NO-2007 Kjeller

POLAND
Armament Policy Department
218 Niepodleglosci Av.
00-911 Warsaw

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide, P-2720 Amadora

SPAIN
SDG TECEN / DGAM
C/ Arturo Soria 289
Madrid 28033

TURKEY
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi Başkanlığı
06650 Bakanliklar – Ankara

UNITED KINGDOM
Dstl Knowledge Services
Information Centre, Building 247
Dstl Porton Down
Salisbury, Wiltshire SP4 0JQ

UNITED STATES
NASA Center for AeroSpace
Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076-1320

SALES AGENCIES

NASA Center for AeroSpace Information (CASI)
Parkway Center
7121 Standard Drive
Hanover, MD 21076-1320
UNITED STATES

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
UNITED KINGDOM

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council
Acquisitions
Montreal Road, Building M-55
Ottawa K1A 0S2, CANADA

Requests for RTO or AGARD documents should include the word 'RTO' or 'AGARD', as appropriate, followed by the serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Full bibliographical references and abstracts of RTO and AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
STAR is available on-line at the following uniform
resource locator:
http://www.sti.nasa.gov/Pubs/star/Star.html
STAR is published by CASI for the NASA Scientific
and Technical Information (STI) Program
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
UNITED STATES

Government Reports Announcements & Index (GRA&I)
published by the National Technical Information Service
Springfield
Virginia 2216
UNITED STATES
(also available online in the NTIS Bibliographic
Database or on CD-ROM)

ISBN 92-837-1121-1