Assessment and Standardization of Personal Hearing Protection including Active Noise Reduction

Dr. Herman J.M. Steeneken
TNO-Human Factors
The Netherlands

Email: herman@steeneken.com

Abstract

The performance of passive hearing protection is normally quantified by the sound attenuation or insertion loss (IL). The IL allows prediction of the noise level at the eardrum for a given ambient noise spectrum. A subjective test (with test-signal levels at the threshold of hearing) is normally used to measure the sound attenuation.

Active noise reduction requires a different assessment method. Due to self-noise and level dependency, assessment methods operating at threshold level cannot be used and have to be replaced by objective methods.

The effect of high noise levels and impulsive noises may introduce a non-linear behaviour of active systems. Therefore, the use of artificial heads is applied to avoid the risk of introduction of temporary or permanent hearing loss. Comparison of results from subjective and objective test methods will be discussed.

Prediction of the noise dose, representative for a certain noise condition, can be obtained by consideration of the environmental noise spectrum, the insertion loss of the hearing protector and an estimation of the variance of the insertion loss among individual users. Examples of such a prediction (by using a spreadsheet) will be given at the lecture.

Speech communication quality is an important issue for use at operational conditions. The noise level at the ear is one of the major variables that define the speech communication quality. Subjective and objective assessment methods for speech communication systems will be presented and discussed. Prediction of the speech intelligibility of a communication system (in a similar condition as presented for the noise dose) will be demonstrated by using an objective intelligibility measure.

Some performance measures for hearing protection, speech communication and criteria for speech quality are standardised by international bodies. *International* standards are provided by ISO, CEN, and IEC.

1. Introduction

An important characteristic of a hearing protection device is the Insertion Loss (IL), which quantifies the ability to attenuate environmental noise. In order to derive the IL the attenuation of the hearing protector as a function of frequency the spectrum of the ambient noise should be known.
For passive hearing protectors (such as earmuffs and earplugs) methods based on the detection of test-signals at threshold level are normally used.

Paper presented at the RTO HFM Lecture Series on "Personal Hearing Protection including Active Noise Reduction", held in Warsaw, Poland, 25-26 October 2004; Belgium, Brussels, 28-29 October 2004; Virginia Beach, VA, USA, 9-10 November 2004, and published in RTO-EN-HFM-111.

Assessment of headsets equipped with active noise reduction requires a different approach. ANR systems may introduce audible electronic noise and may show a non-linear behaviour at high noise levels, the use of test-signals at threshold level is therefore not appropriate.

As stated above the present standardized methods suitable for ANR devices (see: ISO 4869-1, and ISO 4869-2, EN 352-1 and associated standards EN 13819-1 and EN 13819-2) make use of the threshold of hearing of a subject. Therefore, a special Task Group was established by the NATO-RTO (2001) to develop and evaluate assessment methods for hearing protectors with ANR and specifically for military applications. The Task Group initiated a so-called Round Robin test series in which five laboratories assessed the same set of ANR headsets. This assessment included the passive and active attenuation and the quality of speech communication. Some of the results of this study are discussed in this lecture. The full results are published under the auspices of the RTO-HFM panel [22].

ANR based hearing protectors are equipped with an electro-acoustic system, which can also be used for intercom applications. Assessment should include the speech communication facility.

Hearing protection devices in general may give some discomfort to the user and some ANR based systems sometimes show instability that can result in a hazardous noise. This requires ergonomic assessment by users under representative conditions.

2. Insertion Loss and sound attenuation

A single number expresses the Insertion Loss provided by a (passive) hearing protector and represents the noise reduction in decibels. It is the difference between the mean sound level at the entrance of the ear canal with and without the hearing protector fitted on the head.

The official definition of insertion loss, according to standard EN 13819-2, is: “The mean algebraic difference in decibels between the one-third octave band sound pressure level, measured by a microphone of the acoustic test fixture in a specified sound field under specified conditions with the hearing protector absent, and the hearing protector on, with other conditions identical”.

This means that the IL is an *application* dependent measure. A different spectrum of the environmental noise will result in a different sound level for hearing protectors with the same IL value. For the assessment and comparison of hearing protectors we need a unique quantification of the performance. This is determined by the sound attenuation as a function of frequency. In general this sound attenuation will be *user dependent*. The main cause of this variance is the fit of the ear-cushions on the head of the user and for earplugs the fit of the plug in the ear canal. A poor system will show differences of the individual attenuation values up to 10 dB at low frequencies, for a good system this will be limited to 3-5 dB. Representative assessment of the sound attenuation should therefore be performed with many subjects. A typical number is 5-16 subjects (depending on the purpose of the assessment). The results are given by the mean sound attenuation in decibels and the corresponding variance expressed by the standard deviation. An example of the attenuation of a custom moulded earplug is given in figure 1. The mean sound attenuation and the standard deviation (the spread in individual results presented by the vertical bars) have been obtained with the REAT method (Real Ear ATtenuation) described in section 2.1.1.

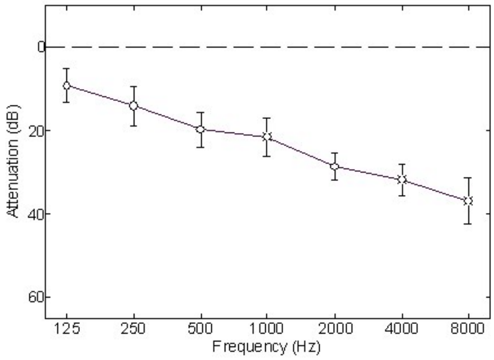


Fig. 1. Average sound attenuation and standard deviation in dB as a function of test frequency for a custom moulded earplug.

With this attenuation curve and the noise spectrum, the IL can be calculated for a specific application. However, the result is based on the *mean* attenuation curve, hence only 50% of the population is covered by this prediction. A better coverage is obtained by reduction of the mean attenuation with one standard deviation value. This provides coverage of 84% of the population. This method is called the Assumed Protection Value (APV_{84}) [12].

Determination of the attenuation curve of passive systems can be accomplished with subjective and objective test methods. These will be discussed in section 2.1 and 2.2 respectively.

The following parameters are of interest for systems equipped with an ANR system:

1. Passive sound attenuation as a function of frequency,
2. Active sound attenuation as a function of frequency,
3. Variance among systems,
4. Variance among users,
5. Stability of the open system during placing on or removing from the head (donning and doffing of the headset),
6. Sensitivity to vibrations,
7. Maximum sound pressure level (dynamic range),
8. Overload response.

ANR systems are normally integrated in a standard passive hearing protector. This may reduce the passive attenuation. It is therefore of interest to compare the passive attenuation of the headset with and without the ANR system integrated.

A poor fit on the head of a user or significant volume reduction under the ear shell may have an affect on the acoustical properties hence, on the ANR performance.

Doffing of the headset may trigger oscillations of a feedback based ANR system.

Vibrations may introduce a (periodic) volume change under the ear shell and a corresponding sound-pressure variation. This effect may be introduced in the ANR system by the sensing microphone and may also introduce overload of the microphone circuit.

Overload of the system will introduce non-linear distortion. This will reduce the effectiveness of ANR [7].

2.1 Subjective Assessment

2.1.1 Real Ear ATtenuation (REAT)

The standardized REAT method is used for assessment of passive hearing protectors like earmuffs and earplugs. With REAT the sound attenuation is determined by measuring subjective hearing thresholds with and without a hearing protector. For this purpose a subject is placed in a diffuse sound field such as obtained in a room as shown in figure 7. A subject is positioned in this room (see figure 2) and exposed to periodic noise bursts from which he/she can control the level. The task is to set the level of the noise burst around the threshold of hearing. Thresholds are measured with a modified von Békésy procedure in which successive presentations are decreased in level by 2 dB as long as the subject indicates, by pressing a button, that the signals are above threshold. The button is to be released when the signal becomes inaudible, after which the level is increased in steps of 2 dB. When the signal is again above threshold, the button is to be pressed anew and the procedure repeats itself. The measurement continues until 10 reversals have occurred. The threshold is taken as the dB average of the last six reversals.

Preferred frequencies of the narrow band sound bursts are in octave steps at 125, 250, 500, 1000, 2000, 4000 and 8000 Hz. In figure 3 displays of the level changes for three completed trials at three different frequencies and one trial in progress are given. Presenting the frequencies to each subject in random order minimizes order effects for frequency and the influence of fatigue. The noise bursts have a bandwidth of one octave and duration of 250 ms; the interval between the noise bursts is also 250 ms.

Prior to the measurements, the subjects have to be informed on the test situation and procedures and they are instructed how to insert an earplug or placing an earmuff. To help them check the fit of the earplug or muff, a broadband noise with a level of about 70 dBA is presented in the test room. Before starting the measurement, the test leader also performs a visual check of the fit. ISO and CEN standardized this procedure.

A number of sixteen subjects is standardized for the measurement of the sound attenuation. They all should have a hearing threshold of at most 15 dB for frequencies of 2000 Hz and below, and at most 25 dB for frequencies above 2000 Hz.

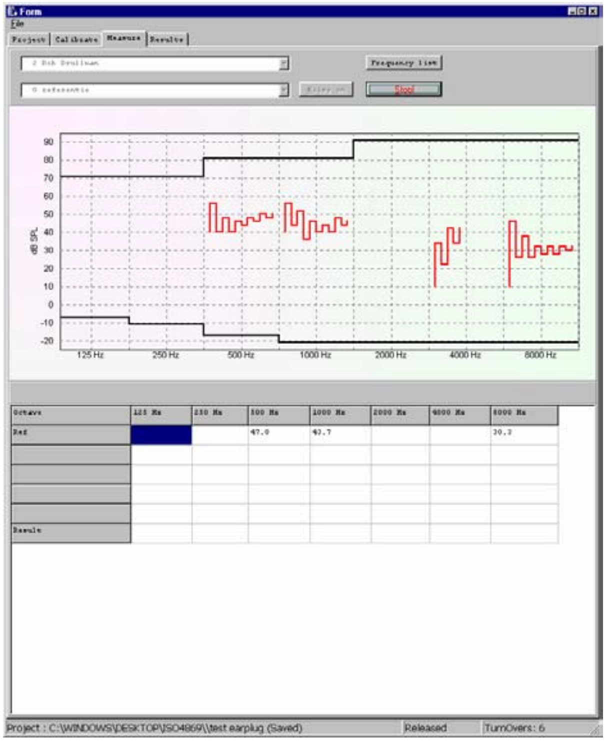
Fig. 2. A subject performing a REAT measurement.

Furthermore, it has to be verified that for all subjects, the results of three consecutive threshold measurements, performed according to the standard procedure does not differ more than 6 dB at any test frequency.
The measuring procedure is controlled by a computer program, the measuring set-up includes a PC, noise generator, octave-band filter, gate for control of the sound burst, and a subject response unit. The experimenter has feedback on the actions of the subject and the corresponding sound level. After 10 level reversals, the mean level is calculated from the last 6 and displayed in the matrix. Then the program switches to the next, randomly selected, frequency.
The sound attenuation of the earmuff under test is to be taken as the decibel difference between the open- and occluded-ear threshold measurements. Over all 16 subjects, the average values and standard deviations are to be determined for each frequency (see figure 1). Subsequently, the Assumed Protection Value for 84% of the population (APV_{84}) is obtained by subtracting the standard deviation from the mean value.

2.1.2 Level matching for ANR-systems

For subjective measurement of the ANR-attenuation (not including the passive attenuation) two methods are available:

1. Subjective matching of the loudness of two sound levels, representative for the additional attenuation of the ANR system [13].
2. Determining the masking of a test tone as a function of frequency [21].



Octave	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	8000 Hz
Ref			47.0	43.7			30.2
Result							

Project : C:\WINDOWS\DESKTOP\ISO4869\\test earplug (Saved) | Released | TurnOvers: 6

Fig. 3. Display on the experimenter station during a REATmeasurement.

(1). For the subjective loudness matching a subject (with a separate ANR system for each ear) is placed in a diffuse sound field. The sound level alternates periodically between two levels (typically every second). An example of this level alternation is given in figure 4.

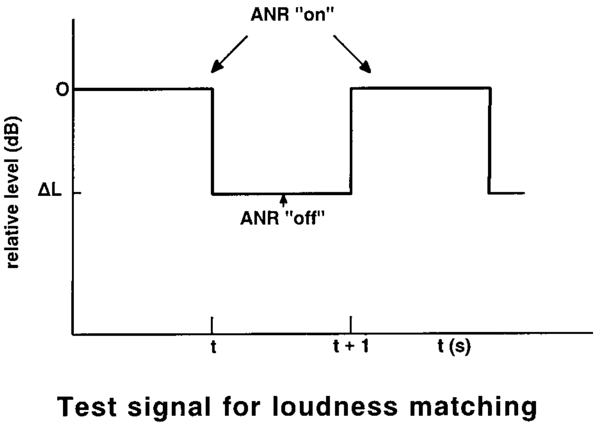


Figure 4. Test signal level as a function of time for the subjective measurement
of the suppression of an ANR system. The ANR system is switched
on and off simultaneously with the test signal envelope.

During the highest sound pressure level the ANR system is switched “on”, while during the lower sound level the ANR system is switched “off”. The subject only hears a small difference between the two sound levels, as the ANR system will attenuate the highest level only. The subject is asked to match both levels for equal loudness by adjusting the level difference, ΔL, between the two signals. The resulting difference in sound level outside the earmuff is equal to the (subjective) attenuation provided by the ANR system.

Since the subject perceives a continuous signal he/she may loose track with the alternation rhythm, therefore the on-off alternations are to be indicated by a visual display (light signal). A study showed that the reproducibility lies between 1-3dB. 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 results obtained with this method will be compared in section 2.3 with results of two objective methods described in section 2.2.

(2). The masking method as proposed by Zera et al. [21] is based on detection of a pure tone that is masked by noise. The detection threshold will shift when the ANR system is switched on/off. This is related to the attenuation provided by the ANR.

This method was not included in the round robin as the level calibration of the test tone is dependent on the ANR system and hence, no benefit with respect to the MIRE method (see 2.2.1) will be obtained.

2.2 Objective Assessment

2.2.1 MIRE-method

Two somewhat related methods might be used for measuring the sound attenuation of ANR-systems:

(1) Comparison of the sound pressure level measured under the earmuff with the ANR system switched on and off. The level difference between the two measurements gives the additional sound attenuation provided by the ANR system. The sound pressure level is obtained from the sense-microphone (loop-microphone) of the ANR-system.

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. It is normally not possible to tap the microphone output with a commercial system.

(2) Similar measurements as described under (1) by making use of an additional microphone, positioned close to the entrance of the ear canal (MIRE, Microphone In Real Ear). This method is also applicable for passive systems.

The active sound attenuation can be obtained by measuring the difference between the sound pressure level under the ear shell with the ANR system switched on and off.

The additional microphone is placed close to the entrance of the ear canal (Figure 5). This MIRE method will be considered as a international standard for measuring the acoustic attenuation of ANR based hearing protection devices (EN 352-5). The MIRE method allows a comparison of the levels at an occluded (ANR switched on and off) and unoccluded ear. This comparison provides the total, passive, and active attenuation of the ANR device.

The noise level and noise spectrum used for the assessment of the performance of the ANR headset should be identical to the noise level and spectrum of the environmental noise in which the device will be used.

ANR systems may have a level dependent performance therefore; it is advised to determine the attenuation as a function of the noise level.

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

A **B**

Figure 5. Microphone configuration used for the MIRE-method.

2.2.2. Microphone in artificial ear (MIArtE)

Exposure of a subject to a high noise level may introduce temporary or permanent hearing loss. Therefore, a pilot test can be performed in which the subject is replaced by an artificial ear or head (MIArtE, MIcrophone in Artificial Ear, Figure 6). An option is to use a standardized artificial ear with a test microphone inside the ear canal and with a representative cavity.

Figure 6. Artificial ear mounted in an artificial head (MIArtE).

The method can also be used with a similar microphone as is used with the MIRE-method but placed in the concha of an artificial ear.

A very simple method is to use a microphone mounted in a flat plate coupler; see Figure 8 [3, 6, 20].

These three objective methods were compared in the round robin test (see section 2.4).

The objective attenuation is determined in 1/3 octave frequency bands by subtraction of the 1/3 octave spectra obtained with the test microphone, with and without the hearing protector placed on the head. The preferred frequency range is 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 ear cup 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 7.

Figure 7. 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 REAT experiments).

2.2.3 Testing at high stationary noise levels

For stationary noises a high continuous (stationary) noise level and a simple flat plate coupler in a small 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-simulator and the flat sides. The small volume allows easy generation of high noise levels, although the resulting sound field will not be diffuse. With the system shown in figure 8 levels up to 130 dBA can be achieved [20].

Figure 8. Dummy head with "flat-plate" microphone in a volume suitable for relatively high noise levels (130 dBA).

2.2.4 Testing at high impulsive noise levels

Impulsive noise will introduce much higher levels, for a short period of time. The Institute Saint Louis developed a specific test for this type of noise [1,2]. 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 is created using explosive charges.

In order to avoid overload of the microphones in an unoccluded condition the Transfer Function of the Open Ear (TFOE) has to be determined. The procedure can be performed at low levels. Once this function is known, the attenuation can be calculated:

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

The pressure-time history for the free field and at the ear underneath the ear cup is measured simultaneously using the set-up shown in figures 9 and 10. The attenuation is calculated using the above formula in 1/3 octave bands.

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 ear cup is measured with the ANR system switched on/off, the difference between these two measurements represents the pressure signal produced by the ANR system.

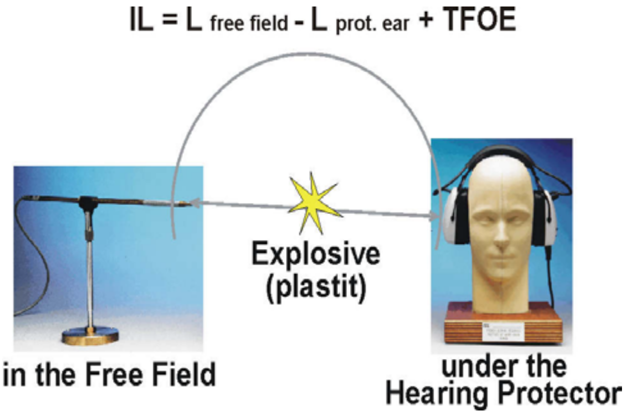


Figure 9. Measuring set-up for the attenuation measurement by using explosive charges.

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 determination of the limits of the electro-acoustic system and its behaviour in an overload condition.

Figure 10. Experimental set-up for attenuation measurements with impulsive noise at ISL.

2.3 Comparison of subjective and objective measuring results

Experimental results of subjective and objective attenuation measurements for an ANR system were compared. The subjective attenuation was measured according to the method described in 2.1.1 with four subjects and various signal levels. For one of the conditions the 1/3 octave-band signal level was 105dB SPL. The mean attenuation for this condition, as a function of frequency with one-octave steps, is given in figure 11.

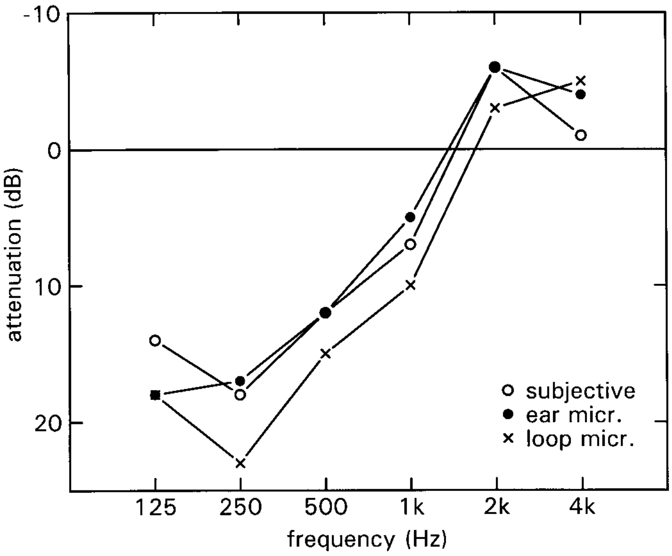


Figure 11. Mean sound attenuation measured with 4 subjects in one-octave intervals, for the subjective level matching and objective methods.

The objective attenuation (of the ANR system only) was measured with the loop microphone as well as with the MIRE-method (see 2.2.1.). 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 additional 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. The MIRE-method provides a good prediction of the subjective results.

2.4 Validity of subjective and objective measuring methods

The Task Group on “Assessment of personal active noise reduction” HFM-094 /TG-028 performed a Round Robin test. 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.

Various assessment methods were used, however it was suggested that all laboratories assess the attenuation of the systems for the passive, and the active conditions. The common experiments included measuring methods that use a subject in order to consider the fit of the earmuff to the human head. For some countries noise regulations do not allow the use of human subjects. In this case an artificial head or artificial ear had to be used. The use of high noise levels and tests with subjects may induce hearing damage. Therefore, participants were encouraged to also use artificial heads or artificial ears in order to compare results of these different methods. Five laboratories participated in the Round Robin test. These were: DRDC, Canada; ISL, France; TNO Human Factors, the Netherlands; Qinetiq, UK; and HECB, USA.

The tests included:

1. Passive, active, and total attenuation of five headsets,
2. The experimental designs include 5-10 replicas and a limited group of subjects,
3. Speech communication quality (intelligibility),
4. Human factors.

The methods that were used for the attenuation were: MIRE, MIArtE, High noise, and Impulsive noise. For speech communication the subjective MRT and the objective STI were used. Two laboratories performed a human factors test. The final report of the Round Robin is published by the RTO [22], in this overview we will give some of the results.

2.4.1 Attenuation measurements

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 12 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 sensitivity 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.

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 replicas. In figure 13 an example is given for the mean attenuation curves obtained with one subject and repeated measures performed at different sessions and at different days. The intra standard deviations are much smaller than the inter (subject) standard deviations.

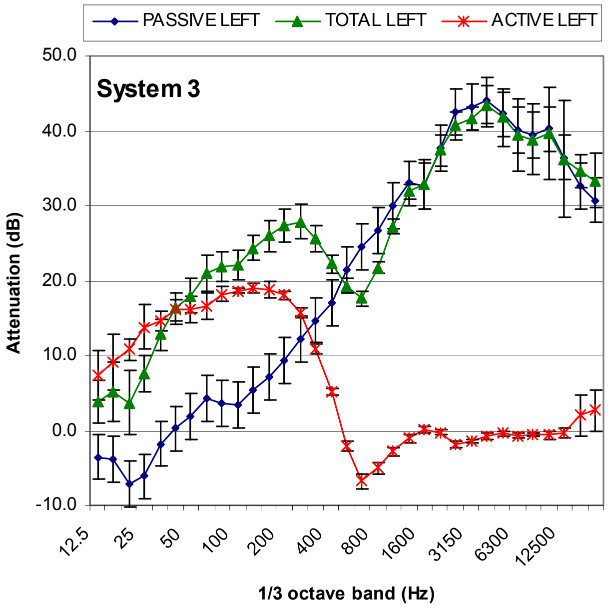


Figure 12. Example of the total, passive and ANR attenuation of headset 3 as a function of frequency. The curves present the mean attenuation obtained for 5 subjects measured on the left earcup. The vertical bars indicate the standard deviation.

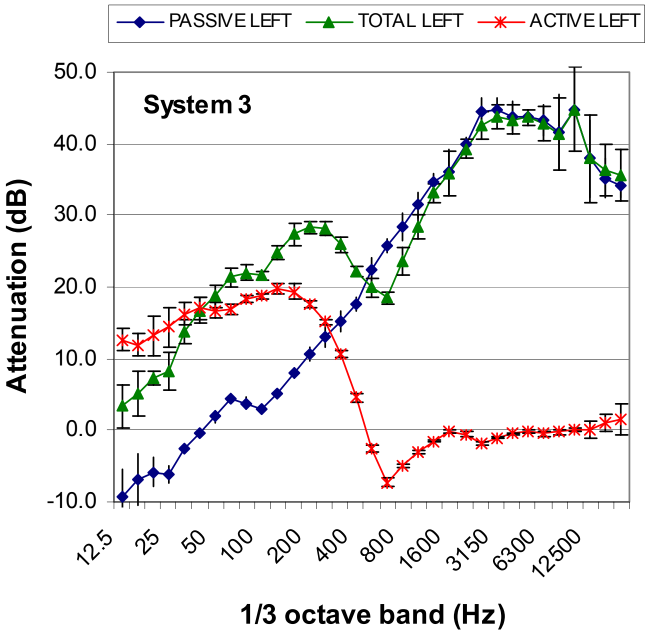


Figure 13. Example of the total, passive and active attenuation of headset 3 as a function of frequency. The curves present the mean attenuation obtained for 5 replica's of the same subject and the left earcup. The vertical bars indicate the standard deviation.

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 signal at levels up to 126 dBA and with impulsive noise with a peak levels up to 170 dB. The example given in figure 14 shows the reduction of the active attenuation for a stationary noise at a level of 126 dBA.

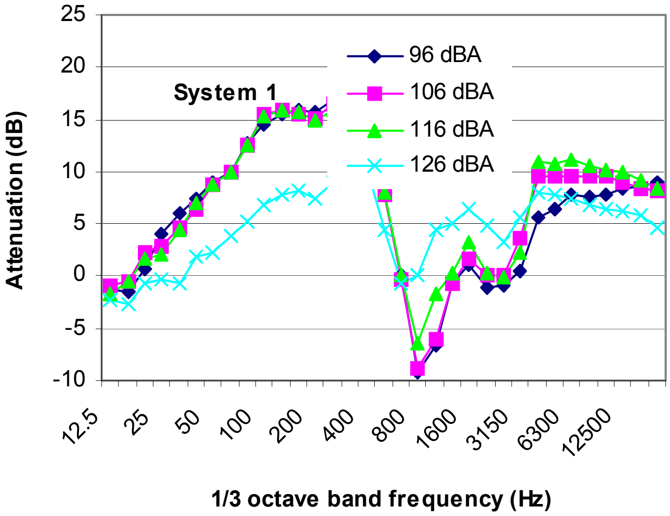


Figure 14. Active attenuation for frequencies between 12.5 Hz to 20 kHz (1/3 octave intervals) derived with test signal levels from 96 dBA up to 126 dBA. This graph represents the results for the left earcup of system 1 of the Round Robin test.

3. Communication quality

3.1 Subjective intelligibility measure-ment

Subjective intelligibility tests can be largely categorised by the speech items tested and by the response procedure used. The smallest items tested are at the segmental level, e.g., phonemes. Other test items are CVC combinations (Consonant-Vowel-Consonant), nonsense words, meaningful words, and sentences.

Besides intelligibility scores the speech quality can be determined by questionnaires or scaling methods, using one or more subjective scales such as: overall impression, naturalness, noisiness, clarity, etc. Speech quality assessment is used for communications with a high intelligibility, for which most tests based on intelligibility scores cannot be applied because of ceiling effects.

The overview given below describes representative tests from this segmental level up to sentence level, as well as tests giving a general impression of transmission or speech quality.

3.1.1. Tests at phoneme and word level

A frequently used test for determining phoneme scores is the rhyme test. A rhyme test is a forced-choice test in which a listener, after each word that is presented, has to select his response from a small group of visually presented alternatives. In general, the alternatives only differ with respect to the phoneme at one particular position in the test word. A frequently used rhyme test is the Modified Rhyme Test (MRT, testing consonants and vowels). The MRT is based on six alternatives [5].

A general approach is obtained with a test with an open response, such as with monosyllabic word tests (Fletcher, 1929). Open response tests make use of short nonsense or meaningful words of the CVC type. Sometimes VCV words, CV words, VC words, CCVC words, or CVCC words are used. This may depend on features of the particular language or the wish to evaluate specific clusters such as consonant clusters or diphone clusters. With nonsense words and an open response, the listener can respond with any combination of phonemes corresponding to the type of word as defined beforehand. This procedure requires extensive training of the listeners. Listeners give their response on a keyboard that allows automatic scoring of the responses. In figure 15 a panel of 4 listeners performs a CVC listening task in an anechoic room.

The test results can be presented as phoneme scores and word scores but also confusions between the initial consonants, vowels, and final consonants can be derived.

The confusion matrices obtained with open response tests provide useful (diagnostic) information for improving the performance of a system. Multidimensional scaling techniques may help to visualize the relations between the stimuli.

With word tests it is recommended to use embedded words in a carrier phrase. Such a carrier phrase (which is neglected in many studies) will cause representative echoes and reverberation in conditions with a distortion in the time domain. Also automatic gain control (AGC) settling will be established by the carrier phrase. An important aspect of using a carrier phrase is also that it stabilizes the vocal effort of the speaker during the pronunciation and that it reduces the vocal stress on the test words. Finally it can function as a cue to the listener that the next test word is going to be presented.

Figure 15. Listening panel performing a CVC listening task.

3.1.2. Tests at sentence level

Sentence intelligibility is sometimes measured by asking the subjects to *estimate* the percentage of words correctly heard on a 0-100% scale. This scoring method tends to give a wide spread among listeners. Sentence intelligibility saturates to 100% at poor signal-to-noise ratios (SNR 0 dB), the effective range is small.

3.1.3. Quality rating

Quality rating is a generic method, used to evaluate the user's acceptance of a transmission channel or speech output system. For quality ratings, normal test sentences or a free conversation are used to obtain the listener's impression. The listener is asked to rate his impression on a subjective scale such as the five-point scale: bad, poor, fair, good, and excellent. Different types of scales are used, including: intelligibility, quality, acceptability, natural-ness, etc. Quality rating or the so-called Mean Opinion Score (MOS) gives a wide variation among listener scores. The MOS does not give an absolute measure since the scales used by the listeners are not calibrated. Therefore the MOS can be used only for rank-ordering conditions. For a more absolute evaluation, the use of reference conditions is required as an anchor.

3.2 Relation between various measures

Fig. 16 gives, for five intelligibility measures, the score as a function of the signal-to-noise ratio of speech masked by noise. This gives an indication of the effective range of each test. The relation between intelligibility scores and the signal-to-noise ratio is valid only for noise with a frequency spectrum similar to the long-term speech spectrum, which makes the signal-to-noise ratio the same for each frequency band. This is for instance the case with voice-babble. A signal-to-noise ratio of 0 dB means that speech and noise have an equal spectral density.

As can be seen from the graph, the CVC-nonsense words discriminate over a wide range, while meaningful test words[1] have a slightly smaller. The digits and the alphabet (not shown) give saturation at a signal-to-noise ratio of 5 dB. This is due to: (a) the limited number of test words and (b): the fact that recognition of these words is controlled mainly by the vowels rather than by the consonants. Vowels have an average level approximately 5 dB above the average level of consonants, and are therefore more resistant to noise. On the other hand non-linear distortions, such as clipping, will have a greater impact on vowels than on consonants. Therefore the use of the digits and the alphabet, for which recognition is based mainly on vowels, may lead to misleading results.

3.3 Objective prediction of intelligibility

There are various methods to predict speech intelligibility, either by direct measurement or by making use of the physical properties of a channel under test. A standardized objective method to predict speech intelligibility (either by measure-ment or by calculation) is provided by the STI (Speech Transmission Index [14], and by a revised version STI_r [16, 19].

---

[1] Meaningful test words are normally phonetically balanced (PB), hence the frequency distribution of the phonemes is representative for the language used.

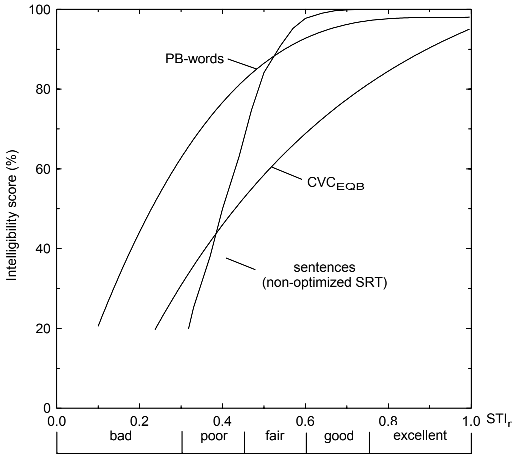


Figure 16. Qualification and relation between some intelligibility scores.

The method is standardized by IEC (IEC 60268-16, 3rd edition, 2002).
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 method.
For application with an ANR system the STI is measured as a function of the environmental noise level at a condition with and without the ANR system switched on.
The STI for a specific communication system with ANR as a function of the noise level is given in figure 17. Hence, the effect of the ANR on the STI-value can be obtained by comparing two conditions. In addition to the STI-value a qualification (based on STI) is also given. The improvement of the speech transmission quality for this example is obvious. It is shown that a 10 dB higher noise level can be applied at a constant speech intelligibility of STI=0.7. Hence, the *effective* improvement in this situation and for this type of noise is 10 dB.

3.4 Round Robin Speech communication performance

Objective measurements are not laborious and therefore allow studying the effect of a variable as a function of its level. In figure 18 the results of such a study are given for the STI as a function of the environmental noise level. The results show that this system provides a fair intelligibility for noise levels above 95 dBA (STI > 0.45).

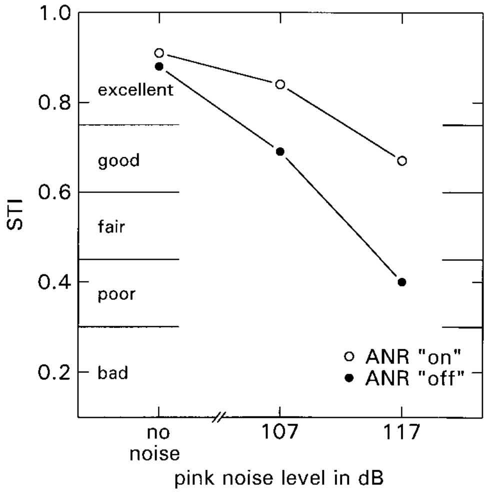


Figure 17 STI at three noise levels for an ANR system switched on and off.

The black curve presents the mean value and standard deviation for measurements with five subjects; the red curve presents the results for a measurement with an artificial head. These two methods give similar results for this headset. However, figure 19 gives similar results for a different headset. This graphs shows bigger standard deviation for the subject related results. This corresponds with the results of the attenuation measurements.

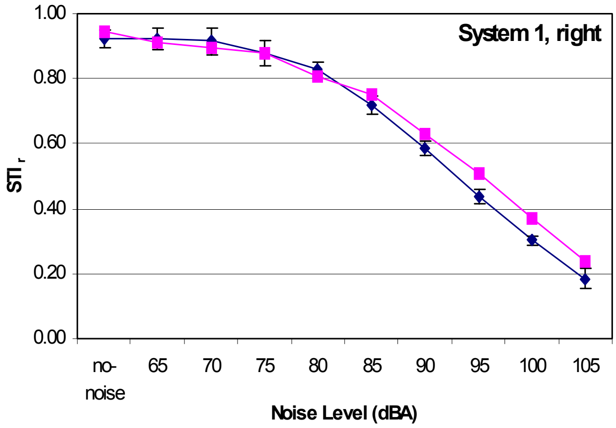


Figure 18. STI_r as a function of the background noise level for system 1 (right earcup).

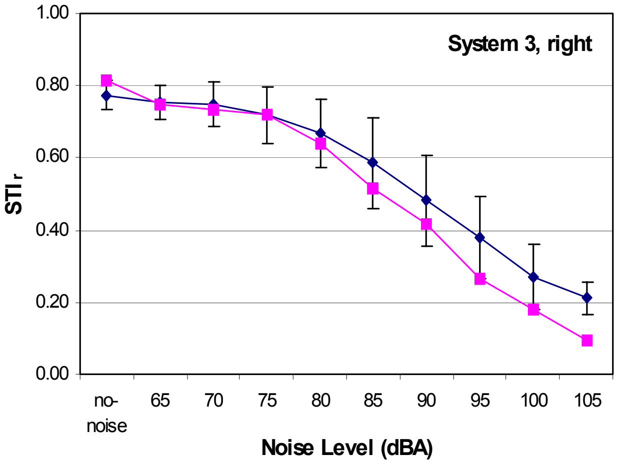


Figure 19. STI_r as a function of the background noise level for system 3 (right earcup).

MRT measurements were performed at AFRL/HECB. The MRT scores could be compared with STI scores that were obtained at TNO. The relation between STI and MRT is plotted in figure 20 (blue data points). The data points show a monotonic increasing relation. The curve shows a shift to lower STI values but this is not in correspondence with the original MRT literature. Therefore, we plotted the original (House et al., [5]) results in the same graph that resulted in an increase of 0.1 STI at a similar MRT value. The difference between the House scores and the MRT scores from the Round Robin study may be due to the method of measuring speech levels and/or the training/experience level of the subjects at AFRL/HECB.

The saturation of the MRT versus STI is related to the limited response set of Rhyme test in general.

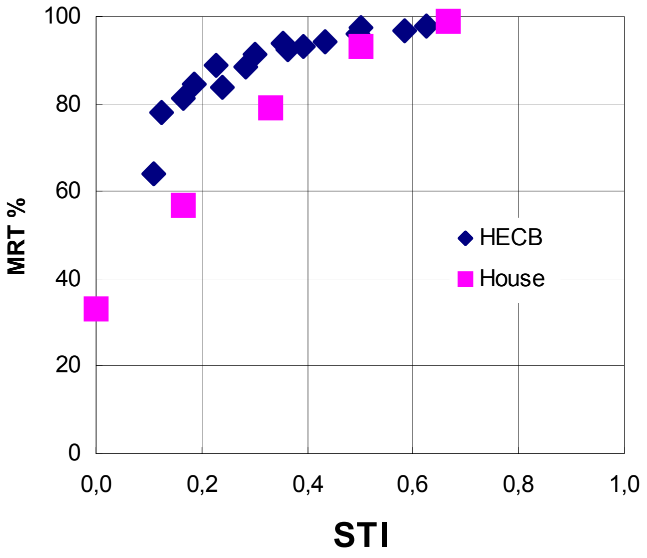


Figure 23. MRT score as a function of the STI$_r$ value. The blue data points refer to the AFRL/HECB scores and the purple data point refer to House et al. [5].

4. Ergonomics of hearing protection

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. For such a test subjects are 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.

Items to be tested are:

Stability:

Verification of instability or oscillations is 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 ear canal does not affect the acoustical conditions, which may influence the results. However, it is rare that closure of the ear canal 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.

5. Standards

An overview of some international standards is given below. National standards are not considered here because these mainly follow the international standards.

International Standards

EN 352-1: 2002 Hearing protectors – Safety requirements and testing – Part 1 Earmuffs.
EN 352-5: 2002 Hearing protectors – Part 5 Active noise reduction earmuffs –.

Associated standards: EN 13819-1: 2002 Acoustics - Physical test methods, and EN 13819-2: 2002 Acoustics - Acoustic test methods.
EN 24869-1: 1990 Acoustics – Hearing protectors – Part 1: Subjective method for the measurement of sound attenuation.

ISO 4869-1: 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.

IEC 60849 edition 1998, "Sound systems for emergency purposes" (This standard will be replaced by ISO).

ISO TR 4870, first edition 1991-12-15. Acoustics: The construction and calibration of speech intelligibility tests.

ISO 9921, "Ergonomics -Assessment of speech communication" (First edition, 2003-10-15).

ISO 11904-1:2002, Acoustics - Determination of sound immissions from sound sources placed close to the ears - Part 1: Technique using microphones in real ears (MIRE-technique)

6. Summary

The performance of passive hearing protection is normally determined by subjective tests in which the threshold of hearing for a number of subjects is obtained with and without a hearing protector. The difference between the two threshold levels quantifies the insertion loss of the hearing protector. The insertion loss is determined at a number of frequencies.
Active noise reduction systems require a different assessment method. Due to self-noise and noise-level dependency, methods operating at threshold level cannot be used. The effect of high noise levels and impulsive noise may introduce a non-linear behaviour of active systems. Therefore the use of artificial heads (to avoid the risk of introducing hearing loss of subjects) is applied. Comparison of results from subjective and objective test methods is a good agreement.
Speech communication quality is an important issue for a user in operational conditions. The noise level at the ear is one of the major variables that define the speech communication quality. Subjective and objective assessment of speech communication systems is a method to predict performance under realistic conditions.
Some of the performance measures are standardised by international bodies.

7. References

[1] Buck K. and Parmentier G., (**1996**). "Evaluation of active hearing protectors with high level impulse noise". Proceedings of Internoise, Liverpool, 989-992.

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

[3] Crabtree, R.B., and Behar, A. (**2000**). "Measurement of hearing protection insertion loss at ultrasonic frequencies". Applied Acoustics, 59, 287-299.

[4] Green, D.M. (**1993**). "Maximum likelihood method for estimating thresholds in a yes-no task", J. Acoust. Soc. Am., 93(4), pp.2096-2105.

[5] 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.

[6] James, Susan. H. (**1990**). “In-flight assessment of a helmet mounted active noise reduction system in Sea Harrier FRS1”. KMT Harpur. January 1990.

[7] 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.

[8] Lueg, P. (**1936**). “Process of silencing sound oscillations”. US Patent No. 2043416.

[9] McKinley, R.L. and Nixon, C.W. (**1993**) “Active noise reduction headsets”. Proc. 6th Int. Conference on Noise as Public Health problem, Nice, pp83-86.

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

[11] 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.

[12] 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-C32, TNO Human Factors, Soesterberg, the Netherlands.

[13] 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).

[14] Steeneken, H.J.M., and Houtgast, T. (**1980**). “A physical method for measuring speech-transmission quality”, J. Acoust. Soc. Am., **67**, 318-326.

[15] Steeneken, H.J.M., (**1998**). "Personal Active Noise Reduction with integrated Speech Communication devices: development and assessment". Noise and Health 1998; 1:67-75.

[16] 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.

[17] 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.

[18] Steeneken, H.J.M., Verhave, J.A. (**2001**). “Digitally controlled Active Noise Reduction with integrated Speech Communication”. Proc CIOP noise conference, Kielce Poland.

[19] Steeneken, H.J.M., and Houtgast, T. (**2002**). “Validation of the revised STI_r method”. Speech Communication, 2002, vol. 38.

[20] 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-TM-03-A015, TNO-Human Factors, Soesterberg, the Netherlands.

[21] 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.

[22] “Assessment methods for Personal Active Noise Reduction validated in an International Round Robin”, (Méthodes d'évaluation des protecteurs auditifs à atténuation active du bruit). Report TR-HFM-094, 2004, NATO-RTO, Neuilly sur Seine.