Applied Acoustics 74 (2013) 870–881

Contents lists available at SciVerse ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Assessment of military intercom headsets for maximum voice reproduction level in high noise conditions

S.M. Potirakis [a,*], N.-A. Tatlas [a], N. Zafeiropoulos [b], T. Ganchev [c], M. Rangoussi [a]

[a] Department of Electronics, Technological Education Institute (TEI) of Piraeus, 250 Thivon & P. Ralli, GR-12244 Aigaleo, Athens, Greece
[b] Acoustics, Audio and Video Engineering, Newton Building, University of Salford, Greater Manchester M5 4WT, United Kingdom
[c] Division of Electronics and Microelectronics, Faculty of Electronics, Technical University – Varna, 9010 Varna, Bulgaria

ARTICLE INFO

Article history:
Received 14 May 2012
Received in revised form 15 December 2012
Accepted 21 December 2012
Available online 14 February 2013

Keywords:
Electroacoustic assessment
Military intercom headsets
Headset testing
Reproduction level
Distortion
Communication quality
ANR
ASR
PESQ
STI
MUSHRA

ABSTRACT

Intercom headsets are mandatory communication apparatus in high noise environments (HNE). The headset selection in HNE, such as combat vehicles, is crucial for achieving the objectives of communication, as it serves the needs for both noise reduction and voice reproduction. Although military-grade intercom headsets are typically used under extreme environmental conditions, a standard performance evaluation method exists only for the earphone elements. In the present work we propose an integrated method for the assessment of the electroacoustic performance of HNE headsets in conditions of maximum reproduction level and high environmental noise, focusing on the voice communication quality. Objective methods, such as Automatic Speech Recognition (ASR), Perceptual Evaluation of Speech Quality (PESQ) and Speech Transmission Index (STI) are comparatively evaluated and their results are compared to subjective scores using Multiple Stimuli with Hidden Reference and Anchor (MUSHRA) in order to reveal the best fit metrics.

1. Introduction

Intercoms provide the means of speech communication in high noise environments (HNE), which is otherwise difficult or even impossible. The headset selection for such an application is crucial since it serves both noise attenuation (protection) and voice reproduction (communication) purposes. Only recently it has been acknowledged in telecommunications [1] that headset electroacoustic measurements should be performed on a Head and Torso Simulator (HATS). Actually, the existing standardized measurement methods regarding telecommunication-related sensitivity and frequency characteristics [2,3], are still used but a HATS is introduced as a measurement apparatus. However, in defense applications, where headsets are mainly used in HNE, a standard performance evaluation method does not exist. The single standard method used for sensitivity measurements refers to the earphone elements only [4], thus ignoring the impact of the acoustics of the earcup cavity, the absorbing materials, the head/face fitting quality and the headset positioning.

It is well known that the reproduced speech level should be at least 10 dB above noise level (without severe distortion) to achieve marginally acceptable intelligibility. Therefore, sensitivity, maximum reproduction levels under specific distortion limit and noise attenuation capability should be jointly assessed to evaluate a HNE intercom headset.

Inspired by the recently published work of Cui et al. [5] (followed by the relative ANSI S12.42 standard [6]), a systematic methodology for measurement and performance evaluation of HNE headsets has been proposed in [7,8]. This is based on the use of Acoustic Test Fixtures (ATFs), like HATS, and addresses both signal reproduction and noise reduction issues, while standardized methods are employed and measurements are maintained as close as possible to the existing telecommunication and military standards. Headset electroacoustic reproduction measurements, i.e., sensitivity, maximum sound pressure level (SPL) and harmonic distortion (HD), are proposed to be acquired as a function of frequency.

Nevertheless, beyond hearing protection issues, the most important evaluation parameter regarding the usefulness of a HNE intercom headset is the voice communication quality under

* Corresponding author. Tel./fax: +30 2105381550, mobile: +30 6947934056.
E-mail addresses: spoti@teipir.gr (S.M. Potirakis), ntatlas@teipir.gr (N.-A. Tatlas), N.Zafeiropoulos@edu.salford.ac.uk (N. Zafeiropoulos), tganchev@ieee.org (T. Ganchev), mariar@teipir.gr (M. Rangoussi).
URL: http://audio.teipir.gr/ (S.M. Potirakis).

0003-682X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.apacoust.2012.12.009

realistic conditions. This is not straightforwardly assessed by maximum SPL and HD measurements, especially for Active Noise Reduction (ANR) headsets. In this case, the environmental noise has a very important role on the intelligibility and the overall performance, as it potentially affects both reproduction frequency response and distortion. Moreover, the ANR functionality may also be influenced by maximum SPL reproduction conditions, leading to worst noise reduction or even higher distortion.

In this article a novel integrated method is proposed for the assessment of the electroacoustic performance of HNE headsets in conditions of maximum reproduction level and high environmental noise, focusing on the problem of assessing the communication quality of a HNE intercom headset via objective measurements. Specifically, an objective voice-based metric for the characterization of the voice communication quality in conditions of maximum SPL reproduction and high environmental noise is suggested. In brief, given a specific HD upper limit, the input level leading to maximum SPL is determined and the voice reproduction of different ANR military headsets is assessed by employing widely accepted objective evaluation methods for voice intelligibility and quality. Alternative objective evaluation methods are examined; correlation of the results to the corresponding subjective results is employed as eligibility criterion for the selection of the most suitable method for the problem at hand. It should be noted that the proposed method is mainly suitable for comparative study and quality control measurements.

2. Electroacoustic assessment of HNE headsets

2.1. Related work and motivation

The evaluation of the voice reproduced through HNE headsets has already been an active research topic for various users/environments of interest (aircraft pilots in flight, armored vehicle crew in operation, etc.) [9–12]. The quality of speech communication was found to be critical for the accomplishment of the mission at hand, as well as for the personnel safety and survival [13]. The need to take into account both noise attenuation capability and the electroacoustic properties, in respect to the received voice communication, for an intercom headset, has already been pointed out as early as in 1975 [14]. It has been found [15], that it is required to adequately raise the reproduced voice level above the noise finally reaching the ear in order to achieve acceptable intelligibility of speech.

A high voice level is necessary in order to obtain an acceptable intelligibility, even for a headset well attenuating the noise in the frequency range that is important for intelligibility. This necessity results from the psycho-acoustic masking of the high frequency components of the communication by the low frequency components of the noise (which are usually not enough attenuated) and by the poor quality of the transmission channel [16]. On the other hand, in extremely noisy environments, the voice reproduction levels may become sufficiently high to induce temporary hearing threshold shift in the user [13]; this variable has therefore to be restricted within the safety margin.

As a consequence, evaluation of the communication quality offered by a military headset is both a multi-parametric task and an important factor of its electroacoustic performance; this is especially true under the worst case scenario of maximum SPL reproduction in the presence of extremely high environmental noise. Related studies [9,17,18] rely on subjective and/or objective methods for the relative quality assessment of alternative solutions, the later being preferable from a practical point of view.

Communication quality primarily refers to the achieved *speech intelligibility*. The former term is adopted here, however, in order to stress the fact that the ultimate criterion is the end-user's opinion about the offered communication quality and not just intelligibility. Voice quality seems to be very important for end-users of military HNE headsets. In order to reduce distraction and facilitate the completion of their main task during a military operation, the cognitive load and the level of distraction due to the communication effort should be kept low; difficulty to identify the speaker and distortions such as strange-sounding or corrupted voice are therefore not desirable. Speaker recognition is, for example, extremely important during wired or radio communication of the (authorized) main battle tank (MBT) crew members to others outside the vehicle, whether on the same or between different military hierarchy levels. Therefore, the broader feature of communication quality and not just intelligibility is expected to be effectively offered by a HNE military headset.

2.2. The proposed approach

The proposed method for the assessment of the electroacoustic performance of HNE headsets in conditions of maximum reproduction level and high environmental noise is comprised of the following actions:

1. Given a specific HD upper limit, the maximum reproduced SPL is determined on ATF as a function of frequency, using a log-swept chirp signal, at office level noise conditions and in the presence of two different ambient noises (Pink noise and MBT – like noise), at three different levels (90, 110, and 120 $dB_{SPL(LIN)}$) (cf. Section 3.3.1).
2. A reference anechoic voice signal (of peak amplitude equal to the corresponding sinus leading to the maximum SPL for each case) is reproduced by the headset and recorded by the ATF in presence of the above noise environment cases (cf. Section 3.3.2).
3. In case of an ANR headset the above measurements are conducted both for activated (ANR-On) and deactivated (ANR-Off) ANR.
4. All measurements are performed for both sides of the headset, in order to investigate any differences in terms of amplitude and frequency content that can be introduced due to variability in the acoustics and the electroacoustics at each side.
5. Communication quality in conditions of maximum SPL reproduction in the presence of different noise types and levels is assessed using the previously recorded voice reproduction and a voice-based objective evaluation method (cf. Section 3.4).

Such a method would be equally valuable at the development, the production and the quality control stage of the headsets, which is of high importance for intercom developers and end-users.

3. Materials and methods

The following subsections outline the headsets, the measurement setup employed, as well as the measurement and the evaluation methods used in the experimental work.

3.1. Headsets

The ANR Combat Vehicle Crewman (CVC) headsets are considered state-of-the-art and recently employed for the intercommunication of the crew members of heavy armored vehicles by a number of modern armies. Three ANR military CVC headsets of different vendors and different design philosophy have been employed in this study. The first ANR headset, hereafter referred to as "HS1", uses two separate speakerphones, one for the voice reproduction and another for the anti-noise signal emission. On the other hand, the second ANR headset, hereafter referred to as