

Technical Evaluation Report

Willard D. Larkin

Air Force Office of Scientific Research, AFOSR/NL
875 Randolph Street, Arlington, VA 22203
USA

Tel: 703-696-7793, Fax: 703-696-8449

Willard.Larkin@afosr.af.mil

1. INTRODUCTION

This Symposium was organized to consider the challenges and opportunities offered by new audio technologies, such as active noise reduction systems and new designs for spatial audio displays. The participants included experts in human factors engineering, acoustics, speech and hearing science, audiology, military and occupational health & safety, and end users of the various devices for hearing protection and audio information portrayal. Thirty-two technical papers were scheduled, representing research and development efforts in eight nations.

The Symposium convened in Amersfoort in the conference centre, “De Eenhorn,” from Monday, 11 April to Wednesday, 13 April 2005. Open to citizens from NATO member states and NATO Partnership for Peace (PfP) states, it was capably hosted by participants from the nearby TNO Human Factors Institute, Soesterberg. Dr. Bernd de Graaf (TNO) and Ms. Marjolein Klootwijk (TNO) coordinated the local arrangements. Dr. Adelbert Bronkhorst (TNO) and Dr. Douglas Brungart (US Air Force Research Laboratory) chaired the 9-member Program Committee, assisted by Ms. Danielle Pelat (RTA, Paris). The opening session featured a welcoming address by the Mayor of Amersfoort, Mrs. Albertine van Vliet-Kuipers.

Three keynote speakers, Dr. A. Dancer (France), Dr. R. Price (USA), and Cdr. F. Bigot (France) contributed additional context for the technical papers.

The Human Factors & Medicine (HFM) Panel coordinates an array of research and development efforts beyond the scope of this Symposium, including topics such as training, selection, health, safety, and survival of military personnel. It is one of seven Panels organized under the NATO Research & Technology Board (RTB). RTB is a component of the cooperative research and information exchange undertaken by NATO’s Research and Technology Organization (RTO).

2. THEME

The Symposium was organized to explore developments in three areas: noise control, speech communication, and audio displays. Several presentations reported on developments relevant to two, or all three of these inter-related topics.

The incidence of hearing loss from exposure to high levels of noise is a continuing concern, especially for military personnel whose normal work activity places them in proximity to jet aircraft engines, or exposes them routinely to impulse noise from explosives or artillery fire. Impaired hearing degrades human performance, but full protection against high noise levels may also carry a risk to human safety, because deep

attenuation also impairs the ability to localize nearby sound sources. On an aircraft carrier, for example, deck crew may face a cruel dilemma: Wear hearing protection to save the ears, or put the ears at risk to avoid bodily injury? Finding a solution to this problem has the highest possible urgency. It must be a solution that simultaneously protects human hearing and yet enables normal auditory awareness of the direction and proximity of physical danger.

Noise also impairs human safety by interfering with communication. In this case, “noise” may include a broad category of acoustic interference, such as multiple, overlapping voice communications that compete for the listener’s attention. Or it may include distortions of speech signals that stem from digital encoding or encryption. The Symposium was therefore concerned with the general question of how best to make speech intelligible, and with the closely related question of how best to measure intelligibility.

Helmets, cockpits, ship and submarine compartments have in common that they seal off the natural acoustic environment. Headphones then re-introduce electronically generated sound – speech or other audio signals – but now in mixed waveforms coinciding at the ear without distinction as to the direction or distance of these multiple sources. Recent scientific and technical developments have aimed at re-instating three-dimensional (3D) spatial perception, e.g., via head-related transfer functions (HRTFs). The Symposium addressed the technical challenges in this effort, including the assessment of 3D audio displays in the context of complex, high-workload operational tasks.

3. PURPOSE AND SCOPE

One of the Symposium’s goals was to identify the emerging technical possibilities for improving noise control, communication, and display. Another goal was to examine the problems that these new technologies may pose for adoption by end-users. This Technical Evaluation Report will describe the progress and continuing challenges brought to light in the Symposium papers and in the discussions provoked by these presentations.

The Symposium began with a session on noise, chaired by Dr. A. Bronkhorst (The Netherlands) and Dr. D. Brungart (United States). Dr. A. Dancer (France) presented a keynote address entitled, *Noise, The Limiting Factor for the Use of Modern Weapon Systems*. His address was an up-to-date review of the mechanisms of cochlear damage, the statistics of hearing disability in military occupations, the likelihood of future risks, and the need for improvements in equipment, measurement systems, and standards.

A later session on noise was chaired by Dr. Dancer and Dr. S. Abel (Canada). Thirteen technical papers were presented, overall, in the two sessions on noise.

Dr. H. Steeneken (The Netherlands) and Dr. W. Tieleman (The Netherlands) chaired a session on communication, with seven technical papers. Dr. R. Price (United States) gave the keynote address, entitled *New Considerations in Promoting Speech Intelligibility, Situational Awareness, and Hearing Protection*. His address departed considerably from this topic, and dealt mainly with an algorithm the U.S. Army Research Laboratory has developed to assess auditory hazards.

Two sessions on auditory displays concluded the Symposium. The first, chaired by Dr. A. Leger (France) and Dr. D. Brungart (United States) began with a keynote address by Cdr. F. Bigot, an experienced fighter pilot who has served in Bosnia, Kosovo, and Iraq. His address, entitled *Constraints and Needs in the Auditory and Vocal Environments of Modern Air Combat*, reflected his expertise as a pilot as well as his experience in ergonomic design. He reminded the Symposium of the importance of voice and audio in the modern cockpit,

when tactical engagements may often take place at thirty kilometre separations between the combatants, far beyond visual range.

Mr. Richard McKinley (United States) and Dr. A. Bronkhorst (The Netherlands) chaired the final session on auditory displays. Altogether, there were 12 technical papers on this topic.

The Symposium's scope can be seen in a rough enumeration of the principal areas of technical interest. These areas range from basic scientific studies to surveys of usability in operational settings. Two papers discussed audio alert symbology. Three focused on active noise reduction (ANR) techniques, or methods of testing ANR. Four discussed bone conduction and/or contact transduction. Six papers discussed aspects of auditory perception and/or hearing mechanisms. Seven papers dealt with 3D audio technology and/or its application. Eight papers presented work on the evaluation and usability of hearing protection devices. There were also eight papers on speech intelligibility.

As HFM Symposium papers and abstracts are available in published form, this Technical Evaluation Report will not review or summarize them individually. Citations to other published papers are indicated by numerals in parentheses ().

4. EVALUATION

4.1 Background & Progress

In October, 1996, NATO sponsored a conference in Copenhagen on "Audio Effectiveness in Aviation" (1). Organized by the former AGARD Aerospace Medical Panel, the Copenhagen meeting provides one point of reference for the problems and progress discussed during the Amersfoort Symposium. Two other points of reference are the recently-published HFM studies, "Assessment Methods for Personal Active Noise Reduction Validated in an International Round Robin" (2), and "Reconsideration of the Effects of Impulse Noise" (3). Prior to these, AGARD sponsored a 1981 conference on "Aural Communications in Aviation," documented in AGARD-CP-311. Held in Soesterberg, The Netherlands, this conference had 27 papers on an agenda very similar to the Amersfoort Symposium.

Consistent with the mandate and instructions given by the HFM Panel, this Technical Evaluation Report adopts a yet broader perspective, to include a general consideration of any relevant scientific and technical developments.

Hearing protection for extreme noise environments has benefited from developments of new active noise reduction (ANR) techniques and from new designs for passive attenuation by earplugs, earmuffs, and helmets. The nonlinear earplug, developed at the French-German Research Institute of Saint-Louis, (ISL) France, is a noteworthy advance (4). The ISL plug has a venturi, or orifice, such that turbulence generated at high sound levels fractures the pressure wave associated with the acoustic impulse. Thus, the effective acoustic resistance of this device increases with the peak pressure of the incident sound. For example, with an input sound level of 170 dB, the attenuation is 25 dB greater, over a broad spectrum, than it is for an input level of 110 dB. Because the ISL earplug has a lower insertion loss for low and moderate intensity sound, it preserves the possibility for speech communication, as well as for identification and localization of environmental sounds, while affording greater protection against high-level impulse noise.

NATO's interest in noise intersects with that of the International Committee on the Biological Effects of Noise (ICBEN), including concerns with noise-induced hearing loss, effects on communication and human

performance, and the development of metrics, regulations, and standards. At five-year intervals, ICBEN convenes the International Congress on Noise as a Public Health Hazard. The most recent meeting was held in Rotterdam, in 2003. The first such meeting was held in Washington, D.C., in 1968, just two years after U.S. President Lyndon Johnson's office issued a special report on jet aircraft noise. In 1972, the U.S. Congress passed the Noise Control Act. This legislation directed the U.S. Environmental Protection Agency (EPA) to develop noise exposure criteria to safeguard public health. In 1981, President Ronald Reagan terminated the EPA's noise control program, and it has never been reinstated. Thus, it falls to international organizations, such as ICBEN and NATO, to provide leadership in dealing with the problems of noise. Several participants in the HFM Symposium were represented in the recent ICBEN Congress. It is expected that several will again participate in the next Congress, scheduled for 2008, which will mark the first return of this important meeting to the United States. The Chairperson for the 2008 Congress is Dr. Jerry V. Tobias, recently retired from the U.S. Navy Submarine Medical Research Laboratory.

The 1996 Copenhagen conference inspired efforts to develop standards and metrics, especially with respect to noise exposure criteria and the evaluation of hearing protection devices. The more recent RTO-HFM study (2) reported on tests of active noise reduction (ANR) systems under conditions of use that represent adverse military deployments. The study adopted a "round-robin" testing plan, in which each participating laboratory examined several ANR headset systems in turn. To conduct these tests, the TNO Human Factors group organized groups working on military ANR in Canada, France, Germany, the Netherlands, the United Kingdom, and the United States. The study took two years to complete. Measurements among the different laboratories showed a high degree of agreement, with only small, generally non-significant differences in the test data. However, the performance of these ANR systems was found to be highly vulnerable to acoustic leakage around the headsets – i.e. to the success, or lack thereof, in fitting standard headsets comfortably to individual wearers. The problem of optimizing fits to individuals has received subsequent study, and was one topic of lively discussion in the Amersfoort Symposium.

Systems for active noise-reduction (ANR) at the ear have been developed to accommodate multiple functions, such as heart rate monitoring and communication. They also have been successfully combined with passive hearing protection. Laboratory tests of new hybrid control architectures, combining feed-forward and feedback algorithms, were reported at the Symposium (5, 6). These tests show that a hybrid system can provide improved attenuation of non-stationary noise, such as F-16 aircraft or helicopter cockpit noise, compared with the performance of either control component acting alone.

There has been significant progress toward the 50 dB attenuation goal for hearing protection in very high-noise environments, such as the 150 dBA sound fields near military jet engines. Much of this progress is very recent, and reflects success with custom-molded, deep-insertion earplugs. The process of molding the plug to individual ears has become simplified, efficient, and reliable. Anecdotal reports indicate that the plug is most comfortable only when correctly worn. Therefore it can be expected that wearers will have a comfort incentive to optimize the plug's performance. Current efforts to design custom-fitting ear cups and cranials (helmets) also show good prospects for improved hearing protection.

Despite these technical advances, noise-induced hearing impairment remains at the top of the list of military disabilities, in number of individuals affected. Yearly estimates for direct compensation for hearing loss among military personnel continue to rise at an alarming rate. In 2004, for example, the United States Veterans' Administration spent nearly \$634 million in compensation for hearing loss as a primary disability. Other countries reported similar substantial expenditures for hearing loss treatment and compensation. The urgency of this problem, obvious to all at the Symposium, calls for a vigorous and consistent research and development effort.

Another area of recent progress stems from scientific studies of binaural hearing, and the measurement of interaural differences in timing and level for acoustic waveforms reaching the two ears from various directions. These measurements provide the basis for head-related transfer functions (HRTFs), which have been exploited to enable 3D spatial perception in headphones. Although individual HRTFs differ from one person to the next, even generic versions prove useful. One important application is the use of spatial audio to localize targets and/or threats that either cannot be seen or would not be noticed if visual search were the only method of detecting them. Another important application is the imposed spatial separation of multiple audio signals or multiple speech sources, so that even if they overlap in time, they will occupy apparently different positions in the listener's acoustic space. One version of this spatial separation technology, reported in the Symposium by Dr. D. Brungart and Dr. B. Simpson, has been designed for AWAC personnel. It provides an optimal separation of speech channels, in virtual directions and distance, for up to seven simultaneous voices.

In a related Symposium paper, Dr. K. McAnally and his colleagues demonstrated the effectiveness of the spatial separation approach. They found that up to 5 communication channels, separated in azimuth, could be monitored in a simulated flying task. The 3D audio condition was better than non-spatialized conditions for speech detection and recognition, and it reduced the perceived workload.

The Symposium took place against a backdrop of greatly expanding interest and invention in the area of auditory displays. Much of this activity has exploited powerful new electronic capabilities to generate, organize, and manipulate waveforms, both for psychoacoustic study and for application in virtual audio displays. Work of this kind was under-represented at the Symposium, although its scope and variety easily could have overwhelmed a three-day program. Many ongoing developments in this area are relevant to human systems design in the military environment. A few examples will suffice to represent the scope of current efforts:

- Sonification of visual displays – e.g., to help diagnose complex turbulence patterns in the atmosphere or in ocean waves.
- Efficient methods to study perceptual classifications of sonic stimuli – e.g., multidimensional representations for very large “earcon” sets.
- Auditory navigation aids for virtual environments – e.g. to represent self-motion, platform motion, or target motion.
- Manipulation of non-traditional sonic parameters – e.g., to represent the size, shape, or composition of an auditory source.
- Control of virtual reverberation and acoustic shadowing – e.g., to assist distance and location judgments within a room or other closed areas.
- New stethoscope designs – e.g., with better impedance matching to the skin and ANR electronics for use in adverse military environments.
- Auditory alerts not based upon loudness – e.g., using temporal or spatial jitter to capture attention and better inform the listener.

Because of the growing commercial incentives in non-military applications, work in the area of auditory displays is likely to advance very rapidly. It will be increasingly important to monitor these innovations. International coordination in these areas is substantially aided by the annual International Conferences on Auditory Display (ICAD). The 11th ICAD meeting will take place in Ireland, July 6-9, 2005. The chairperson is Mikael Fernstrom, of the University of Limerick.

This rapid growth of new audio display techniques, although greatly needed, confronts a long-standing human factors problem: how to evaluate the consequences of any technical innovation for human performance. Several Symposium papers brought to light the extreme care – and often the extreme expense – required in such evaluations. Given the manifold complexity of many operational tasks, how should new auditory technology be used to optimize performance?

To take just one example, consider the control interface for an uninhabited air vehicle (UAV). Considering system costs as well as ergonomics, should haptic feedback or acoustic feedback be used to convey air turbulence? Would the accident rate decrease if the navigation map were provided with combined egocentric and exocentric directional cues? How would photographic time-on-target be affected? How would fuel use be affected? Should an auditory horizon be installed, in addition to the visual one? Aside from the obvious multiplicity of such questions, one of the great impediments to human factors optimization is that trade-offs among the various options are difficult to measure in advance of actually building and testing the entire system.

A traditional response to this evaluation problem has been to conduct empirical studies of various options, usually one option at a time, in laboratory set-ups that usually represent only a small fraction of the real-world task complexity. Over the past eight years, a new approach has been developed, combining computational modeling with cognitive task analysis (7). This approach develops a Synthetic Task Environment (STE), validated against expert performance in the real-world task, along with a validated computational model of human performance (representing all levels of training, from novice to expert). Design changes in the task interface can then be implemented quickly, and the consequences of these changes can be assessed, using all the metrics relevant to system performance. (One of the first STE's was constructed in the U.S. Air Force Research Laboratory for a UAV control task. This laboratory-based STE was not designed to train UAV operators, but it proved so successful that it was later adopted for that purpose, at one-tenth the cost of the originally planned training system.)

One key element in this approach is computational modeling that not only represents human performance within a complex system, but also opens quantitative pathways to define optimal performance. There always will be a role for the traditional laboratory-bench assessments of new ideas, including new types of audio displays, but in the expensive and often protracted development of complex military systems, there can be a much greater role for Synthetic Task Environments to evaluate these ideas in a rapid prototyping mode. To succeed in this approach requires a confluence of expertise, including cognitive science and optimization theory paired with traditional ergonomics.

4.2 Scientific & Technical Challenges

This section of the Technical Evaluation Report takes up a selected group of continuing challenges, all of which were discussed during the Symposium. Although much has been accomplished in each of these areas, much remains to be done. Solutions to these problems range from near-term to far-term in accessibility. They are considered, below, in a quasi-logical order, which should not be mistaken for an endorsement of relative priorities: All of these scientific and technical challenges are important.

4.2.1 Hearing Protection: Use and Non-Use

One consensus among many at the Symposium is that military personnel do not take full advantage of the hearing protection made available to them, and that the research community lacks a complete understanding of the phenomenon. For example, in a study of 301 U.S. Navy flight deck crewpersons, Ms. V. Bjorn and her

colleagues found that nearly half never wore earplugs & muffs that would have provided 30 dB protection. Dr. S. Abel, reporting on a Canadian study of 1,057 personnel, found that many avoided hearing protection because they found it incompatible with other gear. The cumulative effects of noise exposure in these personnel are substantial, as they are for the 252 pilots Spanish pilots described in another Symposium report by Dr. J. M. Lorente and colleagues.

Continuing field studies will be needed to define the problem, the elements of which will likely differ from one operational context to another. Continuing surveillance, educational efforts, and re-engineering will be needed to reverse the dismal statistics. NATO can have an important role in promoting international exchanges regarding successful and unsuccessful practices, stimulating reviews of policy and implementation, and coordinating field trials to improve compliance. The potential benefits to human welfare give this endeavor the highest urgency.

4.2.2 Impulse Noise

Understanding the effects of exposure to impulse noise remains an important, unsolved problem. Long-term effects of this exposure have not been quantified. Nor is there a reliable way to predict the long-term hazard. The heart of the matter is that dose-effect relationships have not been developed for relevant metrics of long-term audiological outcomes, such as a permanent threshold shift (PTS). Research effort in this direction should be a high priority.

When a PTS is found, it is usually impossible to make a thorough retrospective analysis of the individual's exposure history. Impulse waveforms from small arms fire, large calibre weapons, blast noise, and the like, can vary in important details, such as duration, peak level, total energy, impulse repetition rate, and level of background noise. For example, as Dr. A. Dancer reported in his Symposium keynote address, a 120 mm mortar round can register a peak pressure of 185 dB, while a 0.50 calibre rifle registers about 175 dB at the shooter's ear. If damage risk criteria were to focus on peak pressure alone, then it is conceivable that current methods of hearing protection would not satisfy the new European Directive under consideration, which would limit residual peak pressure under earplugs to a maximum of 137 dB. On the other hand, it is likely not only that other waveform parameters matter, but also that these parameters enter a risk equation in a complex, nonlinear fashion, with an important dependence on long-term, cumulative exposures. Data on temporary threshold shifts (TTS) may be necessary, but they are not sufficient to understand this problem.

One Symposium paper, reported by Dr. J. Vos, noted that although a temporary threshold shift (TTS) can be measured under controlled (or at least well-monitored) exposure conditions, there are no data to support a fixed relationship between TTS and PTS at the individual subject level. Although a correlation may exist in aggregate epidemiological data, especially for continuous noise exposures, the problem is more complex for impulse noise. A relationship between TTS and PTS can be possible if the TTS definition were relaxed to allow a more prolonged interval between exposure to the sound and measurement of the consequent threshold shift. However, no scientific study can afford to prolong this interval without ensuring strict controls designed to actually prevent PTS.

An alternative can be to look for other psychoacoustic predictors of long-term hearing impairment. One approach discussed at the Symposium suggests that otoacoustic emissions can be useful. This research, presented by Dr. Agnes Job, asked if the time constant of tinnitus persistence after an acute acoustic exposure can be predicted from data on hearing thresholds and/or from the measurement of distortion-product otoacoustic emissions (DPOAEs). The data are encouraging. They distinguish clearly between groups of individuals whose tinnitus subsides before 72 hours, post-exposure, and individuals whose tinnitus persists

beyond 72 hours. That any such outcome difference exists is itself noteworthy, and relates to the general question of individual differences, to be considered next in this TER. Otoacoustic emissions are, of course, an incompletely understood phenomenon. Repeated measurements in the same ear do not always yield stable results. Their relationship to cochlear health and function is still an open topic for research, as they can be absent in some individuals with otherwise apparently normal hearing. Nevertheless, there is much to gain from a concerted effort to understand how the integrity of the cochlear mechanism could be assessed with this method.

A recent study (8) examined virtually all the available data on temporary threshold shifts (TTS) from impulse sounds. The analysis included results from the Blast Overpressure Project directed by the late Dan Johnson in Albuquerque, New Mexico, USA. Different results were found for short- and long-duration sounds. The data suggest strongly that frequency-dependent weighting is appropriate to reduce the contribution of low-frequency energy, but no single numerical rule emerged. The complexity of this problem demands an approach from first principles, rooted in auditory biophysics and neurophysiology. It also would be highly desirable to examine a broader array of audiological consequences, beyond the TTS and PTS, to discover, for example, how auditory masking, discrimination, and binaural processing might be affected, over the long term, by exposure to impulse noise.

4.2.3 Individual Differences

Some military recruits arrive with enough prior hearing damage, presumably from years of exposure to excessive noise, to disqualify them from many military assignments. Even those who can pass a basic audiological exam may differ greatly in hearing capabilities that can be important for human performance. The Symposium discussion repeatedly turned to this problem, as it was exposed in data from several research studies. The ability to localize acoustic sources, for example, is critical to many military tasks. To provide spatial audio information at the headphones, or inside a helmet, depends upon matching audio cues to this human ability. But tests of the ability to localize sound sources show large individual differences; apparently as variable as other differences in hearing sensitivity (e.g., pitch discrimination, absolute thresholds). It would be rather easy to develop efficient audiological tests for spatial hearing. A first priority should be to include these tests in standard audiological examinations. To understand the basis for the large differences among individuals, however, remains a major challenge for research.

A second aspect, also discussed at the Symposium, is the large variance between individuals in their apparent susceptibility to noise exposure. The word *apparent* is needed here because susceptibility is usually measured in a very limited way, such as a temporary threshold shift (TTS), and without complete data on relevant exposure histories. The notion of an “acquired resistance” to noise exposure has become popular in some quarters, based upon laboratory studies with animals. As with the now-discredited belief that people can adapt permanently to chronic sleep deprivation, this notion is likely to succumb, eventually, to more penetrating scientific inquiry. The important question is: Do some personnel require a more conservative approach to protect their hearing? Judging from the Symposium papers, the answer to this question is yes, but it is not yet possible to assess or to predict individual vulnerability with any degree of confidence. A much deeper understanding of long-term exposure is needed. To craft the next generation of damage-risk criteria, there must be a greater investment in longitudinal studies, coupled with psychoacoustic research to discover how the human auditory system adapts – or fails to adapt – to repeated sonic insults.

4.2.4 Ototoxic Exposures and Noise

It is known that exposure to high levels of noise generates reactive oxygen species (“free radicals”) in the cochlea, and that cochlear injury from noise can be exacerbated by exposures to solvents, jet fuel, antibiotics, and other chemical agents that contain reactive oxygen species (9, 10). The Symposium was not designed to include explicit consideration of this topic, but it relates to the question of hearing protection generally, as well as to the specific question of individual differences in susceptibility to noise. For individuals with hearing loss, an indeterminate prior history of exposures to various chemicals masks the effects of noise, and thereby complicates the audiologist’s task. Research on this question is beyond the scope of the Symposium, but not beyond the scope of the HFM panel. Modern research techniques, such as genomics and proteomics, could shed light on risk factors and biological mechanisms. Studies of this kind should be coupled to investigations of antioxidants, as some research has indicated they can be effective, at least in the short term, as a prophylactic against hair cell damage from noise or from chemical exposures (11, 12, 13).

4.2.5 Speech Intelligibility

Eight Symposium papers dealt directly with speech intelligibility. This topic represents a central challenge for audio effectiveness. One aspect of this challenge is to devise consistent metrics to evaluate the intelligibility of speech as it is received via electronic communication channels of one kind or another, and under circumstances in which there may be interference from ambient noise, channel noise, distortion, and overlapping non-speech signals. Algorithms for this purpose are under development in the signals intelligence community, as well as in the human factors community, as both groups attempt to keep pace with the rapid expansion in applications of electronic signal processing to speech transmission, speech separation, and speech synthesis. The era of single, dedicated transmission lines for voice communication is ending. In many cases, the same channel is used for voice, data, and video. There is an increasing use of vocoders (often in tandem, with entirely different types of voice coding at the sending and receiving stations). Signals are encrypted. Information flows in packet networks with multiple possibilities for distortion from peak-clipping, jitter, delays, packet loss, etc.

Against this backdrop, methods of intelligibility analysis that once served the need must be re-invented or replaced. The Symposium included excellent papers on this theme. In one, Dr. S. van Wijngaarden showed how to adapt the Speech Transmission Index (STI) sensibly to allow assessments with vocoders. In another, Dr. J. Beerends demonstrated that PESQ, an algorithm originally developed at TNO, Soesterberg, to assess speech quality rather than intelligibility, could be adapted to the latter task with low-bitrate encoders. That PESQ can work well in this domain deserves to be better known in the speech research community.

Efforts to improve the metrics for speech intelligibility will be continually challenged as new technologies for speech signal transmission continue to expand. This challenge equally applies to machine algorithms for speech detection, speaker identification, and speech restoration. But another, more difficult aspect of the speech intelligibility problem pleads for research attention, and offers a high potential payoff: How to achieve genuine enhancement of transmitted speech?

The term “speech enhancement,” as it is currently employed, often refers only to methods that attempt to separate speech signals from their noisy backgrounds. These methods can be categorized, roughly, either as “blind separation” techniques based upon a purely statistical analysis of the signal, such as Independent Component Analysis (ICA), or as “model-based” techniques that exploit aspects of human hearing and/or human speech production mechanisms. Both methods can improve the audibility of the speech signal. Neither approach has yet produced speech enhancement in the stronger sense of the term, in which a speech signal would be remarkably more intelligible than the original, noise-corrupted signal. If the original speech signal

can be heard at all, the human ear is a very good instrument for making sense of it. Machine techniques sometimes render speech less intelligible, because although the speech may be easier to hear, it is often truncated in ways that put a greater burden on the listener's ability to compensate for important, but missing articulatory, prosodic, and psycholinguistic features (14). This effect is the opposite of what is needed to ensure reliable voice communication. The problem is compounded, of course, when the communication is among people from diverse linguistic cultures, such as the NATO member states.

The goal of genuine speech enhancement sets a difficult course for research, but it should be embraced. It aims beyond present concerns with separating speech from noise and/or "high-fidelity" recovery of the original speech source. It is now possible to study new methods of speech modification that could sacrifice fidelity to a limited degree, yet provide a large gain in the listener's ability to understand or to identify a speaker. For example, Hideki Kawahara (15) has shown how it is possible to simulate altered vocal tract resonances and vocal tract sizes to produce exceptional or unusual voice characteristics. Roy Patterson (16) has begun to exploit this technique to study how these altered voices are perceived. Such techniques expand the palette of audible speech in ways that could provide a human factors benefit. Further benefit can be expected from research into the basic mechanisms of speech perception, especially the so-called "cognitive" operations of the human auditory system. Such mechanisms are at present almost entirely absent from speech-processing algorithms. The human factors goal should be to define the optimal characteristics of the speech signal, for intelligibility purposes, and then work to restore these features when they are missing or degraded. "Supranormal intelligibility," rather than recovery of the original speech signal, should be the goal.

4.2.6 Contact Transducers and Bone Conduction

A contact microphone, placed directly on the skin of the face or throat, has the possibility to eliminate much of the airborne background noise that could otherwise mask a speech signal. Such microphones could be extremely useful if the speaker is inside a noisy vehicle or helicopter. Of course, the microphone's placement can be problematic, because the biophysical machinery for speech production is widely distributed, from the lips to the tongue, to the back of the mouth, to the larynx: No single contact position will capture the full acoustic signal. As Dr. B. Acker-Mills reported in the Symposium, such microphones, acting as low-pass filters, may well pick up vowels, but not fricatives, resulting in poor overall speech intelligibility compared with noise-canceling air microphones. The idea is still rather new, and deserves further study. It may be possible to combine air- and contact microphones to good advantage. The restoration of missing fricatives through signal processing is another approach, though an ambitious one, in line with the speech-enhancement goals outlined above.

Contact transducers that bypass the ear canal also have potential to reduce audible noise in comparison with standard earphones. In the Symposium, Dr. L. Pellieux showed encouraging data for speech intelligibility using a vibrator in direct contact with the listener's head. As expected, the mechanical interface and the precise placement and geometry of contact on the skin surface can be critical; many such issues remain to be investigated.

In very high noise environments, adequate protection of the cochlea may require consideration of sound transmission paths that by-pass the ear canal. Helmets add some protection against non-airborne conduction, especially at frequencies above 1 kHz. However, little is known to answer very basic questions about bone conducted sound: How does acoustic energy reach the cochlea via multiple pathways through the head and torso? How dispersed and how focused is this energy? What is its masking potential? What is its contribution to long-term hearing impairment? Can it be suppressed with active noise reduction (ANR) techniques? Two Symposium papers reported new investigations of these primary biophysical and psychoacoustical problems.

Dr. A. Dietz summarized an extensive project that compared mechanical impedance calculations with measurements using a head simulator and with psychoacoustic data. Dr. W. O'Brien reported progress in the development of a model that tracks the propagation of acoustic waveforms through the human head. An ongoing French-American collaboration (A. Dancer, H. von Gierke, and R. McKinley; not reported at the Symposium) is pursuing psychoacoustic measurements related to bone-conducted loudness and masking effects. None of this research yet translates into revolutionary new techniques to control acoustic input to the cochlea, but it is an entirely new effort, it is being carried out with great care and expertise, and it deserves high priority for continued support.

4.2.7 Restoration of Ambient Spatial Perception

To the extent that hearing protectors and helmets succeed in blocking unwanted noise, they also reduce perceptual access to the ambient acoustic environment. It can be important to restore this access, especially if the environment is complex and dangerous. The human factors challenge is to devise equipment that makes headgear as "transparent" as possible to the nearby, low-level sounds of interest, yet affords maximum protection against high-level sonic intrusions. In one Symposium paper, Dr. A. Bronkhorst reported some progress toward this dual requirement. Concentrating on the need for spatial awareness of the local acoustic environment, he and Mr. J. Verhave tested a multi-microphone system that gave improved localization judgments, compared with a conventional earmuff lacking in sound pass-through. The idea is similar to that under intensive investigation for use in hearing aids, and for similar reasons: The human auditory system depends critically on spatial perception to disambiguate the acoustic world. Without this ability, human performance is disabled.

There is ample room for new research to solve this problem. One puzzle is to understand how a microphone system, such as the one reported above, should match interaural time and intensity cues to the perceptual requirements of the user. If microphones are spaced more widely than the user's two ears, the interaural timing will be inevitably mismatched to the user. Yet, preliminary data seem to indicate that people can adapt to non-optimal HRTFs. A general study of this possibility could be very informative.

5. RECOMMENDATIONS

This Symposium was well organized to consider a wide range of important technical challenges. By limiting the sessions to inter-related topics on noise, communication, and audio displays, the Symposium maintained good coherency. This plan can be recommended for future NATO meetings, even if some of the additional questions discussed in this TER are taken up.

In general, future meetings could benefit if there were a greater representation from the scientific community. This would foster stronger awareness of, and collaborations with, the community of technology users. At this Symposium, many papers demonstrated excellent connections with current scientific developments related to the technical problems under review, but some did not.

Quite aside from the value of bringing scientific perspectives to meetings such as this, it is clear that many of the more urgent technical problems require a greater investment in basic research aimed at understanding the human auditory system. As already mentioned, the need for deeper insight into speech perception is key to developing more effective speech processing technology. To cite just one further example, the hearing-protection problem should become a motivating beacon for basic scientific efforts aimed at understanding the peripheral auditory system. Most current models represent the auditory periphery as a group of feed-through

circuits, without taking into account the important role of efferent feedback. Neural efferents are widely believed to function as a neurological gain control system at the hair cell level, and they may also have a role in auditory selective attention, detection, and masking. A deeper understanding of this system could inform efforts to develop damage risk criteria and new protective technologies.

Section 4.2, above, describes a group of seven primary challenges for future research and technical development:

1. Hearing Protection: Use and Non-Use
2. Impulse Noise
3. Individual Differences
4. Ototoxic Exposures and Noise
5. Speech Intelligibility
6. Contact Transducers and Bone Conduction
7. Restoration of Ambient Spatial Perception

Aligned with these seven are many usability issues and technical questions that also deserve attention, but are included above by implication: Among them are the effects of whole-body vibration, the compatibility of hearing equipment with night vision equipment or goggles, the use of machine algorithms to separate one speech channel from another, the robustness of HRTF algorithms to noise and reverberation, and the ever-present problem of how new technologies contribute to, or reduce task workload.

The scope and importance of all these challenges for human effectiveness demands a broad and sustained effort. International coordination is vital. Although many non-governmental organizations take an interest, NATO has a leadership role in fostering the cross-disciplinary efforts that are the engines of technical progress.

6. REFERENCES

- [1] Audio Effectiveness in Aviation. AGARD Aerospace Medical Panel Symposium, Copenhagen, October, 1996. AGARD/AMP, 7 Rue Ancelle, 92200, Neuilly-sur-Seine, France.
- [2] RTO-TR-HFM-094. Assessment Methods for Personal Active Noise Reduction Validated in an International Round Robin. August, 2004. Research and Technology Organisation (NATO), BP 25, F-92201 Neuilly-sur-Seine Cedex, France. ISBN 92-837-1121-1.
- [3] RTO-TR-HFM-017. Reconsideration of the Effects of Impulse Noise. April, 2003. Research and Technology Organisation (NATO), BP 25, F-92201 Neuilly-sur-Seine Cedex, France. ISBN 92-837-1105-X.
- [4] Dancer, A., Hamery, P. (1998). A new nonlinear earplug for use in high-level noise environments. *Journal of the Acoustical Society of America*, 103, p.2878.
- [5] Cartes, D., Ray, L. R., Collier, R. D. (2002). Experimental evaluation of leaky least-mean-square algorithms for active noise reduction in communication headsets. *Journal of the Acoustical Society of America*, 111, 1758-1772.

- [6] Streeter, A., Ray, L. R., Collier, R. D. (2004). Hybrid feedforward-feedback active noise control. Proceedings of the June 2004 American Control Conference, Boston, U.S.A., 2876-2881.
- [7] Gray, W. D. (2002). Simulated task environments: The role of high-fidelity simulations, scaled worlds, synthetic environments, and microworlds in basic and applied cognitive research. *Cognitive Science Quarterly*, 2(2), 205-227.
- [8] Smoorenburg, G. F. (2003). Damage risk with impulse-noise exposure. Proceedings of the 8th International Congress on Noise as a Public Health Problem. ICBEN 2003, Schiedam, The Netherlands. Pp. 1-4. ISBN 90-807990-1-7.
- [9] Ohinata, Y., Miller, J. M., Altschuler, R. A., Schacht, J. (2000). Intense noise induces formation of vasoactive lipid peroxidation products in the cochlea. *Brain Research*, 878, 163-173.
- [10] Rao, D. B, Moore, D. R., Reinke, L. A., Fechter, L. D. (2001). Free radical generation in the cochlea during combined exposure to noise and carbon monoxide: an electrophysiological and an EPR study. *Hearing Research*, 161, 113-122.
- [11] Yamasoba, T., Nuttall, A. L. Harris, C., Raphael, Y., Miller, J. M. (1998). Role of glutathione in protection against noise-induced hearing loss. *Brain Research*, 784, 82-90.
- [12] Schacht, J. (1999). Antioxidant therapy attenuates aminoglycoside-induced ototoxicity. *Annals of the New York Academy of Science*, 884, 125-130.
- [13] Ohinata, Y., Miller, J. M., Schacht, J. (2003). Protection from noise-induced lipid peroxidation and hair cell loss in the cochlea. *Brain Research*, 966, 265-273.
- [14] Divenyi, P. (Ed.), (2005). *Speech Separation by Humans and Machines*. Dordrecht, The Netherlands: Kluwer Academic Publishers. ISBN: 1-4020-8001-8.
- [15] Kawahara, H., Irino, T. (2005). Underlying principles of a high-quality speech manipulation system STRAIGHT, and its application to speech segregation. In P. Divenyi (Ed.) *Speech Separation by Humans and Machines*. Dordrecht, The Netherlands: Kluwer Academic Publishers. ISBN: 1-4020-8001-8.
- [16] Smith, D.R., Patterson, R. D., Turner, R., Kawahara, H., Irino, T. (2005). The processing and perception of size information in speech sounds. *Journal of the Acoustical Society of America*, 117, 305-318.

