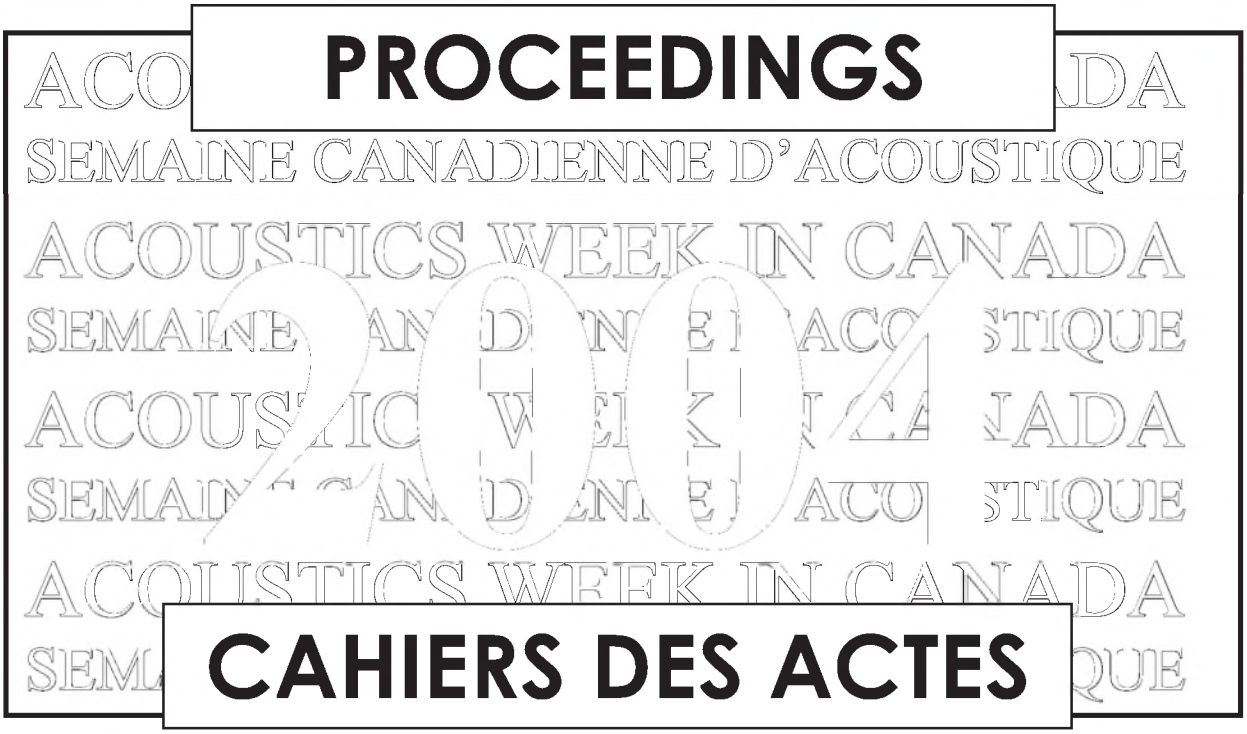
ISSN 0711-6659

canadian acoustics
acoustique canadienne

Journal of the Canadian Acoustical Association - Journal de l'Association Canadienne d'Acoustique

SEPTEMBER 2004
Volume 32 -- Number 3

SEPTEMBRE 2004
Volume 32 -- Numéro 3



canadian acoustics

acoustique canadienne

THE CANADIAN ACOUSTICAL ASSOCIATION
P.O. BOX 1351, STATION "F"
TORONTO, ONTARIO M4Y 2V9

L'ASSOCIATION CANADIENNE D'ACOUSTIQUE
C.P. 1351, SUCCURSALE "F"
TORONTO, ONTARIO M4Y 2V9

CANADIAN ACOUSTICS publishes refereed articles and news items on all aspects of acoustics and vibration. Articles reporting new research or applications, as well as review or tutorial papers and shorter technical notes are welcomed, in English or in French. Submissions should be sent directly to the Editor-in-Chief. Complete instructions to authors concerning the required camera-ready copy are presented at the end of this issue.

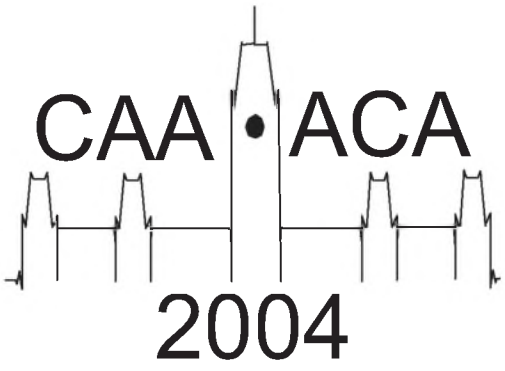
CANADIAN ACOUSTICS is published four times a year - in March, June, September and December. The deadline for submission of material is the first day of the month preceeding the issue month. Copyright on articles is held by the author(s), who should be contacted regarding reproduction. Annual subscription: $20 (student); $60 (individual, institution); $250 (sustaining - see back cover). Back issues (when available) may be obtained from the CAA Secretary - price $10 including postage. Advertisement prices: $500 (centre spread); $250 (full page); $150 (half page); $100 (quarter page). Contact the Associate Editor (advertising) to place advertisements. Canadian Publication Mail Product Sales Agreement No. 0557188.

ACOUSTIQUE CANADIENNE publie des articles arbitrés et des informations sur tous les domaines de l'acoustique et des vibrations. On invite les auteurs à soumettre des manuscrits, rédigés en français ou en anglais, concernant des travaux inédits, des états de question ou des notes techniques. Les soumissions doivent être envoyées au rédacteur en chef. Les instructions pour la présentation des textes sont exposées à la fin de cette publication.

ACOUSTIQUE CANADIENNE est publiée quatre fois par année - en mars, juin, septembre et décembre. La date de tombée pour la soumission de matériel est fixée au premier jour du mois précédant la publication d'un numéro donné. Les droits d'auteur d'un article appartiennent à (aux) auteur(s). Toute demande de reproduction doit leur être acheminée. Abonnement annuel: $20 (étudiant); $60 (individuel, société); $150 (soutien - voir la couverture arrière). D'anciens numéros (non-épuisés) peuvent être obtenus du Secrétaire de l'ACA - prix: $10 (affranchissement inclus). Prix d'annonces publicitaires: $500 (page double); $250 (page pleine); $150 (demi page); $100 (quart de page). Contacter le rédacteur associé (publicité) afin de placer des annonces. Société canadienne des postes - Envois de publications canadiennes - Numéro de convention 0557188.

Proceedings Issue

Canadian Acoustical Association
Annual Conference
Ottawa
6-8 October 2004



Conference Organising Committee

Convenor	John Bradley
Papers Chair	Brad Gover
Publicity & Translations	Christian Giguère
Exhibits	Hugh Williamson
Webmaster	Alf Warnock
Audio-Visual	Frances King

Conference Sponsors

Durisol Inc., Bruel & Kjaer , Eckel Industries of Canada Ltd., Jade Acoustics Inc., Logison (Environmental Acoustics Inc.)

Exhibitors

Norseman Allfoam, Durisol Inc., Dalimar Instruments Inc., Can-Cell Industries Inc., Acoutherm Insulation Ltd., Bruel & Kjaer, H. L. Blachford Ltd., Logison (Environmental Acoustics Inc.), Eckel Industries of Canada Ltd., Novel Dynamics Test Inc., Scantek Inc.

Noise Pollution

The SLARM™ Solution

The **SLARM™** developed in response to increased emphasis on hearing conservation and comfort in the community and workplace incorporates **ACOustAlert™** and **ACOustAlarm™** technology. Making the **SLARM™** a powerful and versatile sound monitoring/alarm system.

Typical Applications Include:

Community

- Amphitheaters
- Outdoor Events
- Nightclubs/Discos
- Churches
- Classrooms

Industrial

- Machine/Plant Noise
- Fault Detection
- Marshalling Yards
- Construction Sites
- Product Testing

FEATURES

- √ Wired and Wireless (opt)
- √ USB, Serial, and LAN(opt) Connectivity
- √ Remote Display s and Programming
- √ SPL, Leq, Thresholds, Alert and Alarm
- √ Filters (A,C,Z), Thresholds, Calibration
- √ Multiple Profiles (opt)
- √ 100 dB Display Range:
- √ 20-120 dBSPL and 40-140 dBSPL
- √ Real-time Clock/Calendar
- √ Internal Storage: 10+days @1/sec
- √ Remote Storage of 1/8 second events
- √ 7052S Type 1.5™ Titanium Measurement Mic

ACO Pacific

2604 Read Ave., Belmont, CA 94002 Tel: 650-595-8588 FAX: 650-591-2891

www.acopacific.com acopac@acopacific.com

ACOustics Begins With ACO™

EDITORIAL / EDITORIAL

Welcome to the proceedings of the 2004 CAA conference. I hope they will provide a lasting account of one of our best CAA conferences to date. There are a record number of abstracts and many special sessions on a range of acoustical topics. All Canadian acousticians should find something of interest at this conference. We will have a very full three days of presentations, a product exhibition, and other associated events.

From the programme, you will see there is a big emphasis on signal processing as it applies to various acoustical issues. The topic of environmental noise issues will be discussed in another major series of sessions. Of course, there are many other special sessions including those on: acoustical materials, hearing aids, hearing in the workplace, and underwater sound. In addition there are sessions on understanding speech, educational facilities, and instrumentation, as well as others on building acoustics and noise emissions. Many thanks must go to the session organisers who have been so successful in creating this impressive conference programme.

On behalf of the organising team, we hope our efforts will bring you to Ottawa to be a part of this acoustical event. We all look forward to seeing you. For all abstracts and last minute programme updates please check our website,
http://www.caa-aca.ca/ottawa-2004.html

John Bradley

Bienvenue au numéro spécial du congrès 2004 de l'ACA. J'espère qu'il reflétera bien ce qui devait être l'un de nos meilleurs congrès à ce jour. Nous avons reçu un nombre record de soumissions et il y aura plusieurs sessions spéciales sur différents domaines de la l'acoustique. Tous les acousticiens canadiens devraient être y trouver leur compte à ce congrès. Nous aurons trois jours bien chargés de présentations, une exposition technique et d'autres événements particuliers.

Au programme cette année, le traitement du signal est à l'honneur et ses applications dans différents domaines de l'acoustique. Différents aspects du bruit environnemental seront aussi abordés dans le cadre de plusieurs sessions. Il y aura plusieurs autres sessions spéciales comme par exemple: les matériaux acoustiques, les aides auditives, l'audition et le milieu de travail, et l'acoustique sous-marine. Il y aura aussi des sessions en perception de la parole, en acoustique des établissements éducationnels, sur l'instrumentation, en acoustique du bâtiment et sur l'émission sonore des sources bruyantes. Merci à tous les responsables de sessions pour leur excellent travail.

Au nom du comité organisateur, nous espérons que nos efforts vous amèneront à Ottawa pour participer à cet événement. Nous avons bien hâte de vous recevoir. Vous pouvez consulter notre site Internet (http://www.caa-aca.ca/ottawa-2004.html) pour connaître la dernière mise à jour de l'horaire des présentations.

John Bradley

WHAT'S NEW ??

Promotions
Deaths
New jobs
Moves
Retirements
Degrees awarded
Distinctions
Other news

Do you have any news that you would like to share with Canadian Acoustics readers? If so, send it to:

QUOI DE NEUF ?

Promotions
Décès
Offre d'emploi
Déménagements
Retraites
Obtention de diplômes
Distinctions
Autres nouvelles

Avez-vous des nouvelles que vous aimeriez partager

Steven Bilawchuk, aci Acoustical Consultants Inc., Edmonton, Alberta, Email: **stevenb@aciacoustical.com**

Table of Contents/Tables des matières

Guidelines for Environmental Noise

Hearing Aids

Hearing and the Workplace

Instrumentation and Measurements

Noise Emission and Noise Control

Outdoor Sound

Signal Processing I

Signal Processing II

Signal Processing III

Sound Transmission and Room Acoustics

Speech Understanding and Psychoacoustics

Underwater Sound

What's new at NRC-IRC?

Abstracts for Presentations without Summary Papers

ACOustics Begins With ACO™

ACOustical Interface™

Systems

PS9200KIT

SI7KIT

Simple Intensity™

New **7052SYS**

Includes:

4212 CCLD Preamp for ICP™ Applications

7052S Type 1.5™

2 Hz to>20 kHz

Titanium Diaphragm

WS1 Windscreen

Measurement Microphones

Type 1

1" 1/2" 1/4"

2Hz to 120 kHz

<10 dBA Noise

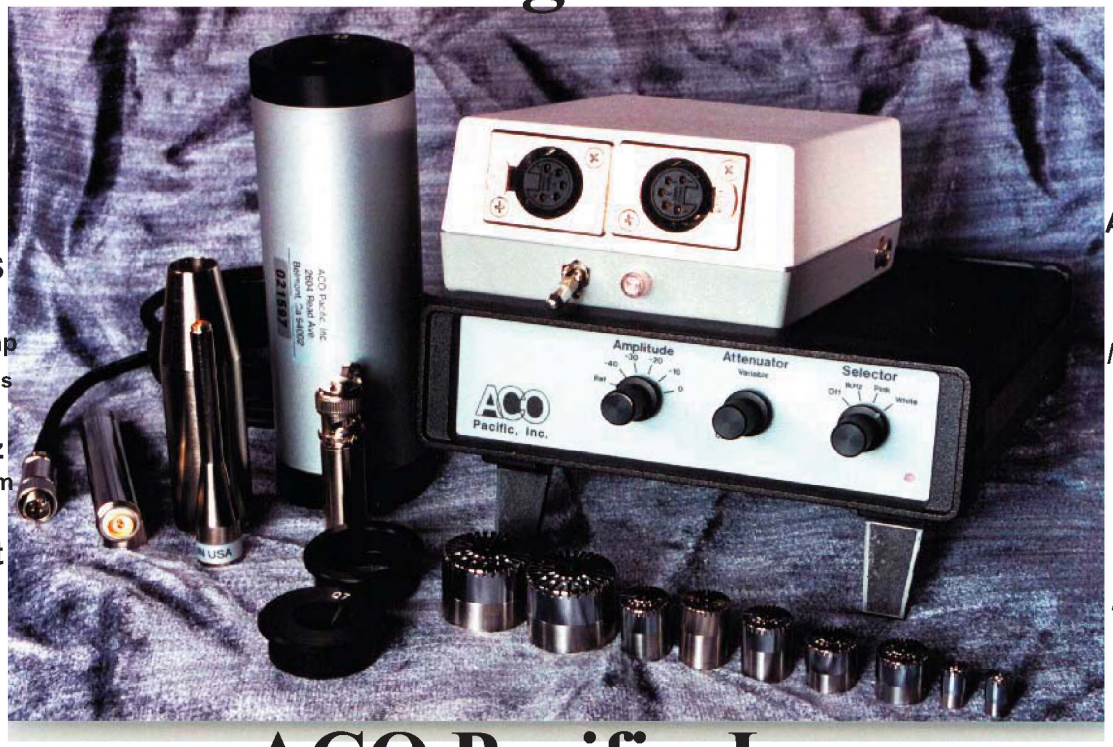
>175 dBSPL

Polarized and Electret

NEW **PSIEPE4**

***and* ICP1248**

ICP™(PCB) Adaptors for PS9200 and Phantom



Very Random™

Noise Generator

White, Pink, 1kHz

SPL Calibrator

***New* 511ES124**

124 dBSPL@1 kHz

ACOtron™Preamps

4022,4012,4016

4212 CCLD

for ICP™ Applications

NEW **RA and RAS**

Right Angle Preamps

DM2-22

Dummy Mic

WS1 and WS7 Windscreens

***NEW* -80T Family**

Hydrophobically Treated

NEW **SA6000 Family**

ACOustAlarm™ with ACOustAlert™

ACO Pacific, Inc.

2604 Read Ave., Belmont, California, 94002, USA

Tel: 650-595-8588 Fax: 650-591-2891

e-Mail: acopac@acopacific.com Web Site: www.acopacific.com

Wednesday, 06 October 2004

AM	**Plenary: "Outdoor Sound Propagation" Gilles Daigle** (Chair: Mike Stinson)		
	Outdoor Sound (Chair: Cameron Sherry)	**Signal Processing I** (Chair: Scott Norcross)	**Hearing and the Workplace** (Chair: Christian Giguère)
	Guidelines for Environmental Noise (Chair: Stephen Bly)		
Lunch			
PM	**Guidelines for Environmental Noise (cont'd)** (Chair: Stephen Bly)	**Signal Processing II** (Chair: David Havelock)	**Acoustics of Educational Facilities** (Chair: John Bradley)

Thursday, 07 October 2004

AM	**Plenary: "From Acoustics to Cognition: Some Surprising Connections" Bruce Schneider** (Chair: Kathy Pichora-Fuller)		
	Instrumentation and Measurements (Chair: George Wong)	**Signal Processing III** (Chair: Heping Ding)	**Acoustical Materials** (Chair: Raymond Panneton)
Lunch			
PM	**Sound Transmission and Room Acoustics** (Chair: Corjan Buma)	**Underwater Sound** (Chair: Nicole Collison)	**Speech Understanding and Psychoacoustics** (Chair: Kathy Pichora-Fuller)

Friday, 08 October 2004

AM	**What's new at NRC-IRC?** (Chair: Dave Quirt)	**Hearing Aids** (Chair: Vijay Parsa)	**Noise Emission and Noise Control** (Chair: Stephen Keith)
Lunch			
PM	**Laboratory Tours**		

OUTDOOR SOUND PROPAGATION

Gilles A. Daigle
National Research Council, Ottawa, ON K1A 0R6, gilles.daigle@nrc-cnrc.gc.ca

SUMMARY

The reality of sound propagation outdoors is more complicated than simple geometrical spreading above a flat hard ground. Most common grounds, such as grass covered ground and layers of snow, are acoustically soft. This implies a complex reflection coefficient leading to a measured spectrum that is strongly influenced by the type of ground surface between source and receiver. Grounds may not be flat, leading to shadow zones or alternatively multiple reflections at the ground. Gradients of wind and temperature refract sound either upwards (upwind or in a temperature lapse) or downwards (downwind or in a temperature inversion), also leading to shadow zones or multiple reflections, respectively. Atmospheric turbulence causes fluctuations and scatters sound into acoustical shadow zones. Many of these features mutually interact and accurate predictions of sound transmission from source to receiver must somehow account for all of these phenomena simultaneously. Thus for example, ISO 9613 Part 2 in wide use today, attempts to account for all the phenomena empirically. In recent years the application of numerical techniques has led to significant advances. This plenary will review the various phenomena. Emphasis will be put on field measurements and simple physical interpretations. In a few cases, the predictions of ISO 9613 Part 2 will be compared with physical or numerical models.

In recent years a number of review articles and book chapters have appeared in print and give a detailed summary of outdoor sound propagation. Thus no attempt is made here to write more material. For a detailed general review see Embleton and Daigle (1994) or Sutherland and Daigle (1997). For a tutorial on outdoor sound propagation see Embleton (1996). For a detailed treatise of computational aspects see Salomon (2001). Articles written for the non-specialist include Daigle (1992) and Daigle (2000). The paper by Daigle (1995) focuses on the noise control aspects of sound outdoors. Finally, for the practical engineering aspects of predicting sound propagation outdoors see Piercy and Daigle (1991).

REFERENCES

Anon., Acoustics – Attenuation of sound during propagation outdoors. Part 2: General method of calculation. ISO 9613-2, 1996.

Daigle, G.A. (2000). "Atmospheric acoustics," in the McGraw-Hill Encyclopedia of Science & Technology (McGraw-Hill, also on-line at www.AccessScience.com).

Daigle, G.A. (1995). "Acoustic of noise control outdoors," Proc. 15th ICA **II**, 49-56 (Trondheim).

Daigle, G.A. (1992). "Atmospheric Acoustics," in the Encyclopedia of Physical Science and Technology (Academic Press)

Embleton, T.F.W. (1996). "Tutorial on sound propagation outdoors," J. Acoust. Soc. Am. **100**, 31-48.

Piercy, J.E. and Daigle, G.A. (1991). "Sound propagation in the open air," *Chapter 3* in Handbook of Acoustical Measurements and Noise Control, edited by C.M. Harris (McGraw-Hill)

Embleton, T.F.W. and Daigle, G.A. (1994). "Atmospheric Propagation," *Chapter 12* in Aeroacoustics of Flight Vehicles: Theory and Practice, edited by H.H. Hubbard (Acoust. Soc. Am. Publications on Acoustics).

Salomon, E.M. (2001). *Computational Atmospheric Acoustics* (Kluwer Academic)

Sutherland, L.C. and Daigle, G.A. (1997). "Atmospheric Sound Propagation," *Chapter 32* in the Encyclopedia of Acoustics, edited by Malcolm J. Crocker (John Wiley & Sons).

From the manufacturer of Scamp® Sound Masking Systems comes the revolutionary LogiSon™ Acoustic Network: the first and only sound masking, paging and music system to provide full, digital control of individual speaker settings from a central location.

Acoustic Control

At Your Fingertips ™

The Network offers the highest level of component integration in the industry. Each hub contains a random masking sound generator, amplifier and even independent volume and equalizer controls for masking and paging, eliminating the need for most centralized audio equipment.

Timer and paging zones, volume and equalizer settings, and paging channel selection are configured through a centrally-located control panel or remotely with user-friendly LogiSon™ Acoustic Network Manager software.

This digital contol and accuracy, combined with small zone sizes and one-third octave equalization capabilities, allows you to custom tune the Network to suit the unique requirements of the space, increasing the masking's effectiveness.

And, as your needs change, the Network can be completely re-adjusted without rewiring or requiring access to the ceiling.

Though the LogiSon Network is typically installed above the suspended ceiling, it has been designed for visible applications and will compliment the investment you make in the professional appearance of your facility.

LogiSon™
ACOUSTIC NETWORK©

For more information, call 1 (866) LOGISON or visit www.logison.com to request an Information Package.

FROM ACOUSTICS TO COGNITION: SOME SUPRISING RESULTS

Bruce A. Schneider bschneid@utm.utoronto.ca
University of Toronto at Mississauga, 3359 Mississauga Rd., Mississauga, ON L5L 1C6

When the acoustics are poor, a listener's ability to navigate an auditory scene, communicate within it, or learn while immersed in it, is adversely affected (see, for example, Picard & Bradley's, 2001 analysis of classroom acoustics). Moreover, when poor acoustics are combined with virtually any kind of auditory problem (even those which would not normally merit clinical attention), all of these listening and learning difficulties are considerably exacerbated. For example, a number of studies have demonstrated that older adults with clinically normal hearing are considerably more disadvantaged than normal-hearing younger adults in adverse listening conditions (e.g., Schneider, Daneman, Murphy, & Kwong See, 2000). Indeed, hearing status in older adults is, arguably, the best predictor of their performance on a number of different cognitive tasks. For example, in the 1994 Berlin Aging study (Lindenberger & Baltes, 1994), the hierarchical model that provided the best account of age-related declines in cognitive functioning was one in which age effects on cognitive tasks were mediated, in large part, by age-related changes in auditory function. Because the proper functioning of higher-order cognitive processes can be highly dependent on the integrity of the information supplied by the sensory systems, it is not unreasonable to expect that cognitive functions dependent on sensory input might be adversely affected by poor acoustics. Hence, acousticians, audiologists, psychologists, and cognitive scientists need to understand how acoustics and cognitive functioning are related.

1.0 AN INTEGRATED MODEL OF INFORMATION PROCESSING

Both sensory and cognitive scientists study how humans detect, encode, process, store, and recall information. However, they concentrate on different aspects of this process. Sensory scientists typically study how information available in the pattern of energy falling on the sensory receptors is used to build up a representation of the external world. Cognitive scientists typically begin to study how information is processed *after* a perceptual representation has been achieved, and neither group, until recently, has been concerned with the nature of the interaction between sensory and cognitive processing. In other words, both groups tend to treat perception and cognition as separate modules (or boxes in a flow chart) with the perception module feeding the cognitive module. A more reasonable approach is to consider them as a unitary information-processing system, in which those processes we call sensory occur relatively early in the processing sequence, whereas those that are labeled cognitive are considered as elaborations of these early processes. Moreover, such an approach would have to explicitly recognize that, in addition to the upward (more central) flow of information, there is also a considerable amount of top-down control exerted over the upward flow of information.

2.0 HOW THE LISTENING ENVIRONMENT AFFECTS COGNITION

Evidence is accumulating that the acoustical environment determines not only how well we can hear but how well we can think. For example, Murphy, Craik, Li, and Schneider (2000) showed that the ability to memorize word associations is affected by background noise. These investigators assessed performance of young adults in a paired associates memory task in which listeners heard sets of five paired words, either in quiet, or in a moderate level of background noise (12 speaker babble). After each set, the first member of one of the paired-associates was presented to the listener, who was asked to supply the other word in the pair. When the paired associate tested was one of the last two presented (positions 4 & 5), performance was quite good and did not differ between noise and quiet conditions. Consistent with the memory literature, performance declined as the serial position of the word became more remote from the time of testing (positions 1, 2, & 3). However, the decline in performance was more severe for those listeners tested in a moderate level of noise than for those tested in quiet, indicating that background noise, even though it may have little or no effect on our ability to hear, can interfere with memory, and that the degree of interference depends on the serial position of the word to be recalled.

As a second example, we will consider the effect that source separation has on speech comprehension. It is well known that separating the position of a noise masker from that of the target stimulus improves our ability to detect, identify, and process the information coming from the target stimulus. In natural environments, much of this improvement comes from increases in signal-to-noise ratios that occur when the target and noise sources are physically separated. However, Freyman, Helfer, McCall & Clifton (1999) have recently shown that when the precedence effect is used to achieve a perceived separation between target and noise source (without substantially altering the signal-to-

noise ratio), perceived separation alone can significantly improve listeners' ability to recognize target words in a nonsense sentence. However, this release from masking with spatial separation only occurs when the masker consists of nonsense sentences spoken by other people, but not when the masker is speech-spectrum noise. Speech maskers, in addition to the masking effect that they produce along the cochlea, also interfere with speech recognition by eliciting activity in the semantic and linguistic (i.e., cognitive) systems. This elicited activity, if not cognitively inhibited or suppressed, competes with that elicited by the target, thereby adversely affecting target word recognition. Hence, a speech masker, because it produces cognitive interference in addition to peripheral masking, should reduce word recognition more than a noise masker, which provides the same degree of peripheral masking, but does not interfere on a cognitive level with the processing of the targeted speech. Shifting the perceived location of the masker away from that of the target helps to perceptually distinguish the target from the masker, thereby making it easier for the listener to cognitively suppress the competing activity elicited by a speech masker. Interestingly, Li, Daneman, Qi, & Schneider (in press) have shown that this release from informational masking is the same for young as it is for older adults in the early stages of presbycusis, indicating that this level of cognitive processing in older adults remains unaffected by aging.

We can also show that features of the acoustical environment can have a surprisingly large effect on the ability to encode and recall information from monologues and dialogues (Schneider et al., 2000), and can influence working memory (Pichora-Fuller, Daneman, & Schneider, 1995). In short, there may be very good reasons why people often say that "It is so noisy that I can't think straight."

3.0 THE ACOUSTIC ENVIRONMENT AND TOP-DOWN CONTROL.

A number of studies (e.g., Dai, Scharf, & Buus, 1991) have shown that listeners can "tune" their hearing to a particular frequency, and can "set" the degree of signal amplification in an auditory channel (e.g., Parker, Murphy, & Schneider, 2002). In other words, there is emerging evidence that auditory system is under top-down control, and that both the flow of processing and the emphasis given to certain kinds of processing may change according to context and task demands. For example in the absence of noise there may be little need to tune the auditory system to select the frequency regions that are required for source identification and for information extraction. However, as the listening situation becomes more difficult, we would expect increases in the degree of top-down control exerted over the upward flow of information. The imposition of a greater degree of top-down control would be expected to produce performance decrements in higher-order tasks because more processing resources would be allocated to controlling and improving lower level functions, to the detriment of higher-order cognitive processes such as working memory (e.g., Pichora-Fuller et al., 1995). Thus, within an integrated model of sensory and cognitive processing, the quality of the auditory environment and the status of a person's auditory system may have far-reaching consequences for cognitive performance. In particular, some portion of the age-related decline in cognitive functioning may be a direct consequence of age-related deterioration in hearing, and of the poor acoustic environments that older adults must function in.

Finally, the strong connections between acoustics and cognitive processing provides another argument as to why we need to be concerned about the auditory environment within which we function. We need environments that not only are "hearing friendly" but "thinking friendly". We need to ensure that students in learning situations need not "strain" to hear what is being said. For if they have to "work at" hearing, their ability to take in information, integrate it with past knowledge, and store it in memory for future use, is likely to be compromised, with the situation being even worse for those with any kind of hearing impairment.

REFERENCES

Dai, H., Scharf, B., & Buus, S. (1991). Effective attenuation of signals in noise under focused attention. Journal of the Acoustical Society of America, 89, 2837-2842.

Freyman, R. L., Helfer, K. S., McCall, D. D. & Clifton, R. K. (1999).The role of perceived spatial separation in the unmasking of speech. Journal of the Acoustical Society of America, 106, 3578-3588.

Li, L., Daneman, M., Qi, J.G., & Schneider, B. A. (in press). Journal of Experimental Psychology: Human Perception and Performance.

Lindenberger, U. & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A sensory connection. Psychology and Aging, 9, 339-355.

Murphy, D. R., Craik, F. I. M., Li, K. Z. H., & Schneider, B. A. (2000). Comparing the effects of aging and background noise on short-term memory performance. Psychology and Aging, 15, 323-334.

Parker, S., Murphy, D.R., & Schneider, B.A. (2002). Top-down gain control in the auditory system: Evidence from identification and discrimination experiments. Perception & Psychophysics, 64, 598-615.

Picard, M, & Bradley, J.S. (2001). Revisiting speech interference in classrooms. Audiology, 40, 221-244.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. Journal of the Acoustical Society of America, 97, 593-608.

Schneider, B. A., Daneman, M., Murphy, D. R., & Kwong-See, S. (2000). Listening to discourse in distracting settings: The effects of aging. Psychology and Aging, 15, 110- 125.

SOUND RADIATION FROM POROELASTIC MATERIALS

F.C. Sgard[1], N. Atalla[2]

[1]Laboratoire des Sciences de l'Habitat, DGCB URA CNRS 1652, Ecole Nationale des Travaux Publics de l'Etat, 69518 Vaulx-en-Velin CEDEX, France, Franck.Sgard@entpe.fr

[2]Groupe d'Acoustique de l'Université de Sherbrooke, Department of Mechanical Engineering, Univ. de Sherbrooke, Sherbrooke, QC,J1K2R1,Canada, Noureddine.atalla@Usherbrooke.ca

1. INTRODUCTION

Numerical approaches based on finite element discretizations of Biot's poroelasticity equations provide efficient tools to solve problems where the porous material is coupled to elastic plates and finite extent acoustic cavities. Sometimes, it may be relevant to evaluate the radiation of a poroelastic material inside an infinite fluid medium. Examples include (i) the evaluation of the diffuse field sound absorption coefficient of a porous material and/or the sound transmission loss of an elastic plate with an attached porous sheet, (ii) the assessment of the acoustic radiation damping of a porous material coupled to a vibrating structure. The latter is particularly important for the correct experimental characterization of the intrinsic damping of the material's frame. To date, the acoustic radiation of a porous medium into an unbounded fluid medium is usually neglected. The classical approach for modeling free field radiation of porous materials assumes the interstitial pressure at the radiation surface to be zero. This paper presents a numerical formulation for evaluating the sound radiation of baffled poroelastic media including fluid loading effects. The problem is solved using a mixed FEM-BEM approach where the fluid loading is accounted for using an admittance matrix solid phase-interstitial pressure coupling terms. Numerical results will be presented during the conference in order to illustrate the technique.

2. THEORY

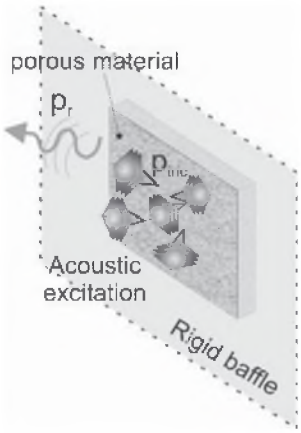


Fig.1: Configuration of the problem

Consider a rectangular porous material sample inserted into a rigid planar baffle excited acoustically. The porous material is coupled to a semi-infinite fluid on one of its face (excitation side) and has specific boundary conditions on the other faces. The modified weak integral form associated to the porous material is been given by [1]:

$$\int_{\Omega_p}\left[\underline{\underline{\tilde{\sigma}}}^s(\underline{u}):\underline{\underline{\varepsilon}}^s(\underline{\delta u})-\omega^2\tilde{\rho}\,\underline{u}.\underline{\delta u}\right]d\Omega+\int_{\Omega_p}\left[\frac{\phi^2}{\omega^2\tilde{\rho}_{22}}\underline{\nabla} p.\underline{\nabla}\delta p-\frac{\phi^2}{\tilde{R}}p\delta p\right]d\Omega$$
$$-\int_{\Omega_p}\frac{\phi^2\rho_0}{\tilde{\rho}_{22}}\delta(\underline{\nabla} p.\underline{u})d\Omega-\int_{\Omega_p}\phi\left(1+\frac{\tilde{Q}}{\tilde{R}}\right)\delta(p\underline{\nabla}.\underline{u})d\Omega \qquad (1)$$
$$-\int_{\partial\Omega_p}\phi[\underline{U}.\underline{n}-\underline{u}.\underline{n}]\delta p d\Gamma-\int_{\partial\Omega_p}\left[\underline{\underline{\sigma}}^t.\underline{n}\right].\delta\underline{u}d\Gamma=0\ \ \forall\ (\delta\underline{u},\ \delta p)$$

Ω_p and $\partial\Omega_p$ refer to the porous-elastic domain and its bounding surface. $\underline{u}$ and p are the solid phase displacement vector and the interstitial pressure in the porous-elastic medium, respectively. $\underline{U}$ is the fluid macroscopic displacement vector. $\delta\underline{u}$ and δp refer to their admissible variation, respectively. $\underline{n}$ denotes the unit normal vector external to the bounding surface $\partial\Omega_p$. ϕ stands for the porosity, $\tilde{\rho}_{22}$ is the modified Biot's density of the fluid phase accounting for viscous dissipation, $\tilde{\rho}$ is a modified density given by $\tilde{\rho}=\tilde{\rho}_{11}-\frac{\tilde{\rho}_{12}}{\tilde{\rho}_{22}}$ where $\tilde{\rho}_{11}$ is the modified Biot's density of the solid phase accounting for viscous dissipation. $\tilde{\rho}_{12}$ is the modified Biot's density which accounts for the interaction between the inertia forces of the solid and fluid phase together with viscous dissipation. $\underline{\underline{\tilde{\sigma}}}^s$ and $\underline{\underline{\varepsilon}}^s$ are the in-vacuo stress and strain tensors of the porous material. $\underline{\underline{\tilde{\sigma}}}^t$ is the total stress tensor of the material given by: $\underline{\underline{\tilde{\sigma}}}^s=\underline{\underline{\sigma}}^t+\phi\left[1+\frac{\tilde{Q}}{R}\right]p\underline{\underline{1}}$. Note that $\underline{\underline{\tilde{\sigma}}}^s$ accounts for structural damping in the skeleton through a complex Young's modulus $E(1+j\eta_s)$. $\tilde{Q}$ is an elastic coupling coefficient between the two phases, $\tilde{R}$ may be interpreted as the bulk modulus of the air occupying a fraction ϕ of the unit volume aggregate.

In this formulation, the porous media couples to the semi-infinite fluid medium through the following boundary terms:

$$-\int_{\partial\Omega_p}\left[\underline{\underline{\sigma}}^t.\underline{n}\right]\delta\underline{u}d\Gamma-\int_{\partial\Omega_p}\phi\left(U_n-u_n\right)\delta p d\Gamma \quad (2)$$

Since $\underline{\underline{\sigma}}^t.\underline{n}=-p\underline{n}$, at the excited surface, (2) becomes:

$$\int_{\partial\Omega_p}\delta\left(pu_n\right)d\Gamma-\int_{\partial\Omega_p}\left[\phi\left(U_n-u_n\right)+u_n\right]\delta p d\Gamma \quad (3)$$

In the semi-infinite domain, the acoustic pressure p_a is the sum of the blocked pressure p_b and the radiated pressure p_r. Applying the continuity of the normal displacement at the surface, (3) becomes:

$$\int_{\partial\Omega_p}\delta\left(pu_n\right)d\Gamma-\frac{1}{\rho_0\omega^2}\int_{\partial\Omega_p}\frac{\partial p_a}{\partial n}\delta p d\Gamma \quad (4)$$

Since $\partial p_a/\partial n=\partial p_r/\partial n$, the associated discrete form to the second term of (4) is:

$$\frac{-1}{\rho_0\omega^2}\int_{\partial\Omega_p}\frac{\partial p_a}{\partial n}\delta p d\Gamma=\frac{-1}{\rho_0\omega^2}\langle\delta p\rangle[C]\left\{\frac{\partial p_r}{\partial n}\right\} \quad (5)$$

where [C] is a classical coupling matrix given by $\int_{\partial\Omega_p}\langle N(M)\rangle\{N(M)\}d\Gamma(M)$ and $\{N(M)\}$ denotes the vector of the shape functions. The porous material being inserted into a rigid baffle, the acoustic pressure is related to the normal velocity using Rayleigh's integral:

$$p_r(x,y,z)=-\int_{\partial\Omega_p}\frac{\partial p_r(x,y,0)}{\partial n}G(x,y,0,x',y',0)d\Gamma \quad (6)$$

where $G\left(x,y,0,x',y',0\right)=\dfrac{e^{-jk_0R}}{2\pi R}$ is the baffled Green's function, $k_0=\omega/c_0$, is the acoustic wave number in the fluid, c_0, the associated speed of sound and R is defined by $R=\sqrt{\left(x-x'\right)^2+\left(y-y'\right)^2}$.

An associated integral form to (6) is given by:

$$\int_{\partial\Omega_p}p_r(x,y,z)\delta p d\Gamma=-\int_{\partial\Omega_p}\int_{\partial\Omega_p}\frac{\partial p_r(x,y,0)}{\partial n}G(x,y,0,x',y',0)\delta p d\Gamma \quad (7)$$

The associated discrete form is:

$$\langle\delta p\rangle[C]\langle p_r\rangle=-\langle\delta p\rangle[Z]\left\{\frac{\partial p_r}{\partial n}\right\} \quad (8)$$

with

$$[Z]=\int_{\partial\Omega_p}\int_{\partial\Omega_p}\langle N(M)\rangle G(M,M')\{N(M')\}d\Gamma(M)d\Gamma(M')$$

Since $\langle\delta p\rangle$ is arbitrary, one gets:

$$\left\{\frac{\partial p_r}{\partial n}\right\}=-[Z]^{-1}[C]\langle p_r\rangle \quad (9)$$

Substituting (9) into (5), and recalling that on the interface $p=p_a=p_r+p_b$, the discrete form of (3) reads finally:

$$\langle\delta u_n\rangle[C]\{p\}+\langle\delta p\rangle[C]^T\{u_n\}-\frac{1}{j\omega}\langle\delta p\rangle[A]\{p-p_b\} \quad (10)$$

where $[A]=\dfrac{1}{j\omega\rho_0}[C][Z]^{-1}[C]$ is an admittance matrix. The radiation of the porous medium into the semi infinite fluid amounts to an added admittance term onto the interface interstitial pressure degrees of freedom and to additional interface coupling terms between the solid phase and the interstitial pressure (first terms in (10)). The last term involving P_b is the excitation term.

Using classical notations [1], the discretized form of (1) combined with (10) leads to the following linear system:

$$\begin{bmatrix}-\omega^2\left[\tilde{M}\right]+[K] & -\left[\tilde{C}\right]+[C]\\ -\left[\tilde{C}\right]^T+[C]^T & \dfrac{\left[\tilde{H}\right]}{\omega^2}-\dfrac{\left[\tilde{A}\right]}{j\omega}-\left[\tilde{Q}\right]\end{bmatrix}\begin{Bmatrix}u\\ p\end{Bmatrix}=\begin{Bmatrix}0\\ F_f\end{Bmatrix} \quad (11)$$

with $\{F_f\}=-\dfrac{1}{j\omega}[A]\{p_b\}$.

This system is first solved in terms of the porous solid phase nodal displacements and interstitial nodal pressures. Next, the vibroacoustic indicators of interest can be calculated.

3. CONCLUSION

This paper presented an approach to predict the sound radiation of baffled poroelastic media including fluid loading effects. The problem has been solved using a mixed FEM-BEM approach where the fluid loading is accounted for using an admittance matrix and solid phase-interstitial pressure coupling terms. The method has been considered in the case of an acoustic excitation but the approach is general and can be used as soon as the porous material is inserted in a rigid baffle and radiates into a semi infinite fluid. Numerical examples will be presented during the oral presentation in order to illustrate the technique.

REFERENCES

[1] Sgard, F. C. Atalla, N. and Nicolas, J. (2000). A numerical model for the low-frequency diffuse field sound transmission loss of double-wall sound barriers with elastic porous lining," *Journal of the Acoustical Society of America* **108**(6), 2865-2872

[2] N. Atalla, M. A. Hamdi, and R. Panneton, "Enhanced weak integral formulation for the mixed (u,p) poroelastic equations," *Journal of the Acoustical Society of America* **109**(6), 3065-3068 (2001).

FROM MICROSTRUCTURE TO ACOUSTIC BEHAVIOUR OF POROUS MATERIALS

C. Perrot[1-2], R. Panneton[1], and X. Olny[2]
[1]GAUS, Mechanical engineering department, Université de Sherbrooke, Sherbrooke QC Canada J1K 2R1
[2]LASH, DGCB URA CNRS 1652, Ecole Nationale des Travaux Publics de l'Etat, 69518 Cedex, France

1. INTRODUCTION

Subject. To study how the microstructure (form and structure revealed using microscopy) of a foam can be used to determine its general sound absorption properties.

Background. (i) Foams can be seen as an arrangement of cells paving space [1], whose form and constitutive struts are determined by physical principles [2]. The two-dimensional ordered monodisperse foam is the celebrated honeycomb structure, the hexagonal structure. Kelvin's tetrakaidecahedron packed in a bcc structure is an acceptable equilibrum structure. (ii) The exact acoustic response of a microstructural system is restricted to the case of tubes of constant cross-section and slits. A real porous material is consequently seen as equivalent-fluid medium, of effective complex and frequency dependant density $\widetilde{\rho}$ and bulk modulus $\widetilde{K}$ (or equivalent set of dissipative functions) under the assumption of a rigid frame. Hence, a macroscopic Helmoltz equation provides a suitable paradigm for acoustic propagation and dissipation through porous media [3]. However, local geometrical variables (for example radius and thickness of tubes) do not appear explicitly in such a macroscopic description.

Purpose. To describe a novel procedure to characterize the sound absorption of reticulated foams (i.e. with open cells) from their microstructure.

Relevance. The enhancement of sound absorption properties of porous materials relies on the capacity to describe (i) microstructure and (ii) microstructure – acoustic energy interactions.

Outline. In Sec. II, methods to compute effective acoustic properties of reticulated foams are presented. Simulation results for effective acoustic properties of scaled cellular systems (Kelvin or honeycomb structure) are given in Sec. III and compared with direct experimental measurements.

2. METHOD

In this approach, one defines an elementary cell paving periodically space, experimentally identified or scaled by a simple geometrical model. We will develop here some of the essential notations relating to structure. Dissipation functions are computed by sophisticated numerical methods from the microstructure.

2.1 Input parameters, some necessary definitions

At microscale. Foam structure is experimentally identified by computed microtomography (μCT), including the average: (i) topology of a three-dimensional unit cell paving periodically space, (ii) shape of the struts of length ℓ and thickness t.

At macroscale. Porosity Φ and thermal characteristic length Λ' can be determined by independent and non-acoustical techniques. The porosity is actually the more common macroscopic parameter and is confidently measured. It is defined by the pore volume to bulk volume ratio (fluid volume fraction). The quantity Λ' is also known as the hydraulic radius and defined by twice the pore volume to pore surface ratio, or $2V_p/S$.

2.2 Cellular model, linking micro to macro scale

Once the morphology has been identified, a micro-macro relationship can also be established, leading to a cellular model of the foam. For example, in the case of (i) the Kelvin structure (a tetrakaidecahedron unit-cell or semi-regular polyhedron of fourteen sides paving efficiently space), (ii) having struts of regular triangle cross-section shape, the macroscopic indicators Φ and Λ' are written according to the microstructural indicators ℓ and t.

Therefore, these expressions for Φ and Λ' determine macroscopic surface and volume information for the Kelvin structure. For the purposes of our study, the solutions of this system are of primary interest as they describe the local dimensions of reticulated foams when microstructural information is not available.

2.3 Local computation of the dissipation functions

The result relating the porosity Φ and the thermal characteristic length Λ' to the sizes of the triangular struts arranged in a Kelvin structure, ℓ and t, allows for a simple cellular model to solve the viscous and thermal problems.

Viscous problem. It has been shown that the Finite Element Method leads to a velocity field that is a solution of the Navier-Stokes equation [4]. Periodic boundary conditions are easily implemented under the Femlab® environment. The relevant physical properties are ensemble averages.

Thermal problem. The first passage sphere algorithm [5] was used to determine the frequency dependent bulk modulus of air saturating different arrangements of parallel solid cylinders [6]. We shall apply this algorithm to more complex three dimensional geometries, such as media comprised of Kelvin cells.

3. RESULTS

As a result, the dynamic bulk modulus can be computed in a periodic unit cell as shown in figure 1 and 2, as well as the macroscopic parameters (table 1).

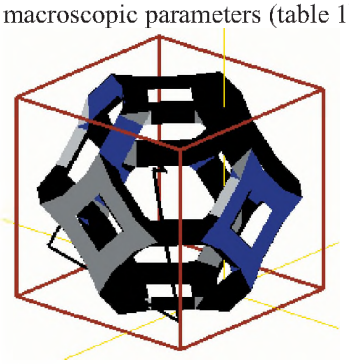


Fig. 1. The Kelvin structure paves periodically space. A "random walker" is seen. It serves to compute an essential thermal dissipation function, the dynamic bulk modulus.

In Tab. 1, the computed viscous macroscopic parameters are compared with measurements. The numerical resistivity σ is greater than the experimental one. In the hexagonal model, struts perpendicular to the flow direction increase resistivity.

In Fig. 2, the computed dynamic bulk modulus $\widetilde{K}$ is compared with measurements. Here, it is seen that there is a shift in amplitude for the polymeric foam. Therefore, the zero acoustic temperature boundary condition is not valid for this material. As shown by Tarnow in 1995, it is due to the fact that the ratio of the air/solid heat capacity coefficients is not small.

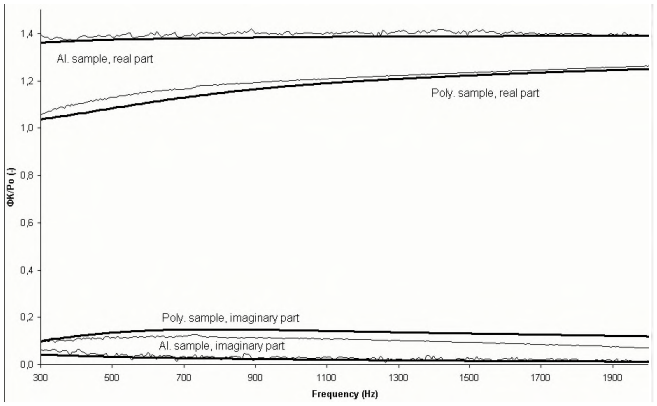


Fig. 2. Bulk modulus as a function of frequency for two real samples. Computed by Brownian motion simulations and compared to measurements for (i) identified (Duocel® al. foam 40 ppi) and (ii) scalded (polymeric foam) Kelvin structures. Top: real parts. Bottom: Imaginary parts.

Table 1. Macroscopic viscous properties

	porosity Φ (-)	thermal length Λ' (μm)	resistivity σ ($N.s.m^{-4}$)	tortuosity α_∞(-)	viscous length Λ (μm)
Measure	0.921 +/- 0.001	1926 +/- 431	177 +/- 21	1.07 +/- 0.01	988 +/- 57
Finite Element Method	0.911	1910	339	1.04	1047

4. CONCLUSION

We have proposed a method to determine the macroscopic parameters of absorbent materials from the knowledge of their cellular microstructure, either identified by μCT, or scaled by a macro-micro geometrical model. A good agreement is found between the microstructural approach and classical macroscopic measurements for the two samples studied. The principal contribution of the present work is that all the relevant quantities have been computed on cellular systems (Kelvin or honeycomb structure). In summary then, we have computed, for 3D cellular porous systems, the dynamic thermal permeability k', its static value k'_0, the static thermal tortuosity α'_0 , the transition frequency f_{tc} and the form factor M'. Study of the fluid velocity field for a hexagonal porous system (2D counterpart of the Kelvin structure) revealed significant quantitative agreement between macroscopic parameters derived and measured from real aluminium foam. This lends support to the idea that there is a deep connection between Kelvin structure and real reticulated foams: the scaled Kelvin structure is a representative configuration of reticulated foams.

REFERENCES

[1] Gibson, L. J. and Ashby, M. F. Cellular Solids (Structure and Properties) (2nd edition). Cambridge University Press (1997).

[2] Weaire, D. and Hutzler, S. The Physics of Foams. Oxford University Press (1999).

[3] Allard, J. F. Sound Propagation in Porous Media (Modelling Sound Absorption Materials). Elsevier, New Kork (1994).

[4] Craggs, A. and J.G. Hildebrandt. Effective densities and resistivities for acoustic propagation in narrow tubes. J.S.V. 92(3), 321-331(1984).

[5] Torquato, S. and Kim, I. C. Appl. Phys. Let. 55, 1847 (1989).

[6] Lafarge, D. Determination of the dynamic bulk modulus of gases saturating porous media by Brownian motion simulation. Poromechanics II, Auriault *et al* (eds.) (2002).

ACKNOWLEDGEMENTS

This work was supported by NSERC Canada and FQRNT Québec.

C. Perrot would like to thank R. Bouchard, F. Gauthier, E. Gautier, I. Jovet and M. Lefebvre for assistance; Alcan, CQRDA, and Région Rhône-Alpes for financial support.

A MISSING MASS METHOD TO MEASURE THE OPEN POROSITY OF POROUS SOLIDS

Emmanuelle Gros[1] and Raymond Panneton[2]
GAUS, Department of mechanical engineering, Université de Sherbrooke, Québec, Canada, J1K 2R1
[1]Emmanuelle.Gros@USherbrooke.ca , [2]Raymond.Panneton@USherbrooke.ca

1. INTRODUCTION

An air-saturated open cell porous material is a solid containing interconnected air cavities. In acoustics, one important physical property of such a material is indubitably its open porosity (ϕ). The open porosity is defined as the fraction of the interconnected air volume to the total volume of a porous aggregate. Sound propagation models for porous solids use the open porosity to relate the effective properties of the saturating air to the effective properties of the porous aggregate [1]. Consequently, from the measurements of the effective properties on a porous aggregate [2], this scaling is required to go back to the effective properties of the interstitial air, hence the importance of the open porosity.

In this paper, a non limitative method is proposed to evaluate the open porosity of air-saturated porous materials. The proposed method deduces the open porosity through the experimental determination of the volume of the solid phase of a porous aggregate. The method is based on the measurement of the apparent (in-air) and true (in-vacuum) masses of a porous aggregate, where an in-air "missing mass" is found and related to the volume of the solid phase through the Archimedes' principle. The originality of the method lies in its predictable accuracy, the simplicity of the experimentation, and in the fact that no isothermal process needs to be assumed.

2. THEORY

Let denote by V_s and V_f the volumes of the solid phase and interconnected air cavities, respectively, and by ρ_s and ρ_f the mass densities of the solid phase constituent and saturating air in a porous aggregate, respectively. While the mass density of the air may be determined from the ideal gas law, the mass density of the solid phase constituent as well as the solid phase and air-cavities volumes are usually unknown and not directly measurable. However, the total or bulk volume V_t of a porous aggregate and its vacuum mass M_1 are directly measurable with precision. The mass of solid per unit volume of porous aggregate is given by $\rho_1=(1-\phi)\rho_s$, where ϕ is the open porosity defined as $\phi = 1 - V_s / V_t$.

Since the mass density of the solid phase constituent is usually unknown, the mass density ρ_1 can be deduced from the measurable parameters following: $\rho_1 = M_1 / V_t$, where M_1 is measured with a balance under vacuum condition. Usually, a good estimate of M_1 may be obtained using a balance in air. Nevertheless, in cases where a high precision is required, the measurement in air is not appropriate since the measured apparent mass M_1' underestimates the true value of M_1. This underestimate is due to the Archimedes' principle [3] on the buoyancy. In air, the buoyancy is very small yet not negligible for the purpose of this work. Consequently, the difference M_1 - M_1' between the in-vacuum and in-air masses gives the "missing mass" error relative to any measurements performed on a balance in air. Following the Archimedes's principle, the "missing mass" of the porous aggregate measured in air is $\hat{M} = \rho_f V_s$, since the volume of air displaced by the solid phase of the aggregate is equal to V_s.

Following this observation, one can deduce a simple method for the measurement of the open porosity ϕ of the porous aggregate. This yields

$$\phi = 1 - \frac{\hat{M}}{\rho_f V_t} = 1 - \frac{M_1 - M_1'}{\rho_f V_t} \qquad (1)$$

The masses and the bulk volume in Eq. (1) can be measured with precision, and the density of dry air can be calculated from the ideal gas law equation.

3. PRECISION OF THE METHOD

Since the "missing mass" $\hat{M}$ is usually small compared to mass M_1', the accuracy of the balance, and the volume of the porous aggregate are important parameters to take into account in the measurement of the open porosity. An error analysis has been performed using a total differential method to predict the expected error on the open porosity due to the propagation of the errors relative to the measurement of the air density, bulk volume, and masses. The error predicted is valid only if the "missing mass" can be read by the balance. Figure 1 gives the minimum bulk volume per balance readability in function of the open porosity. Now, assuming $\hat{M}$ can be read by the balance, the absolute value of the expected error on the open porosity in function of the bulk volume of aggregate per balance readability is plotted in Figure 2.

4. EXPERIMENTAL TESTS

To validate the "missing mass" method and its precision, three experimental tests are studied.

4.1 Experimental set-up and measurement procedure

The experimental set-up includes a vacuum air pump, an analytical balance, and an airtight rigid chamber. A digital atmospheric station is used to measure ambient temperature, pressure and relative humidity.

The test procedure consists first in measuring the in-air mass M_1' of the test sample. Second, the empty chamber is pumped down to vacuum, and mass M_a is measured. Third, the test sample is placed in the chamber, the chamber is vacuumed, and mass M_b is measured. The true mass M_1 is deduced from the difference M_b - M_a. Finally, the open porosity and its expected error are obtained.

4.2 Air density test

The first test consists in measuring the "missing mass" of a solid sample of known solid phase volume V_s. Since the solid phase volume is known, the air density can be calculated, and compared to the air density predicted by the perfect gas law. The test samples are six cylindrical columns made from Delrin. Using the ambient conditions, the theoretical density of the humid ambient air is 1.146±0.001 kg/m^3. Following the test procedure described above, the "missing mass" $\hat{M}$ is deduced and is about 0.24 g. The statistics on the six measurements give an air density of 1.148±0.026 kg/m^3. This value compares well with the theoretical value calculated from the ideal gas law. To reduce the standard deviation, an optimal choice of volume and balance readability should be selected.

4.3 Low-porosity test

The second test consists of measuring the porosity of open-cell materials by mean of the "missing mass" method. In view of validating the method for low-porosity samples, eleven Delrin hollow cylinders of different bulk volumes have been machined. Their solid phase volume V_s and porosity ϕ are easily found. The results and the statistics are summarized in Figure 2 and Figure 3, where it is clearly observed that the precision on the measurements improves with the volume-to-readability ratio.

4.4 High-porosity test

In this third validation test, the "missing mass" method is applied on a high-porosity aluminum Duocel© foam. For the porosity tests, six samples of different bulk volumes have been used. Using the solid phase density ρ_s, the true mass M_1, and the bulk volume V_t, the theoretical open porosity of the sample can be found. The results and the statistics of the "missing mass" method are summarized in Figure 2 and Figure 3.

5. CONCLUSION

A simple method was proposed to measure the open porosity of open-cell porous solids. Contrary to existing methods, the "missing mass" method is not time consuming, and does not assume isothermal conditions for the test. Also, it only requires simple apparatuses: vacuum pump, small airtight container, analytical balance, thermometer, barometer, and humidity meter. The last three meters can be eliminated if the ambient air density is measured using the "missing mass" as per the first test.

REFERENCES

[1] Allard, J.-F. (1993). Propagation of Sound in Porous Media : Modeling Sound Absorbing Materials. Elsevier Applied Science.

[2] Lafarge, D. (1997). Dynamic compressibility of air in porous structures at audible frequencies. J. Ac. Soc. Am. 102, 1995-2006.

[3] Archimedes' principle. Physical law of buoyancy, discovered by the ancient Greek mathematician and inventor Archimedes, www.britannica.com/seo/a/archimedes-principle/.

ACKNOWLEDGEMENTS

The authors wish to thank N.S.E.R.C., F.Q.R.N.T., and REGAL for their financial supports.

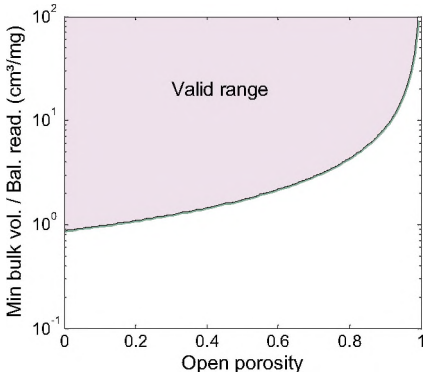


Fig.1 Minimum bulk volume of aggregate per balance readability in function of the open porosity.

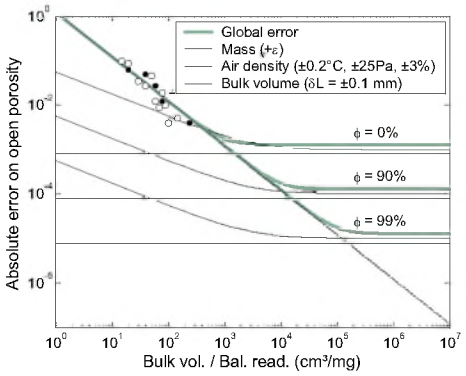


Fig. 2 Absolute error on the open porosity in function of the bulk volume per balance readability

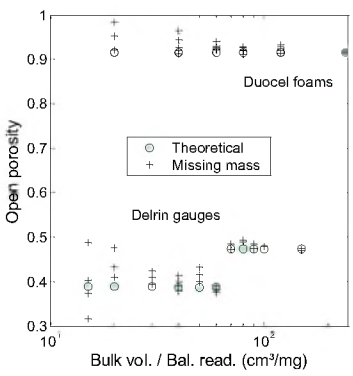


Fig. 3 Open porosity measured by the "missing mass" method compared to theoretical value

ACOUSTIC PROPERTIES OF LOOSE AND CONSOLIDATED GRANULATES

Kirill V. Horoshenkov[1], Siow N. Ting[1], Ian Rushforth[1] and Mark J. Swift[2]
[1]School of Engineering, Design and Technology, University of Bradford, Bradford, UK, BD7 1DP
[2]Armacell UK Limited, Mars St, Oldham, UK, OL9 6LY

1. INTRODUCTION

Loose granular materials can be regarded as an alternative to more common foams and fibrous acoustic absorbers. In their consolidated state, these materials often combine high structural strength with good acoustic performance and excellent vibration insulation properties Recently, it has been demonstrated that a particular class of acoustic absorbers can be developed from granulated elastomeric industrial waste (Swift, 2000; Miraftab *et al*, 2004). This type of process addresses directly the growing public concern for the environment and increased landfill taxation which is forcing manufacturers to look into alternative uses for their waste output. In this paper the acoustic performance of some porous materials made from loose and consolidated granulates have been investigated experimentally. It has been shown that a thin layer of recycled material can yield relatively high values of absorption coefficient. This is mainly attributed to the unique pore size distribution and fibre-grain composite structure resulting from the granulation and consolidation processes. A formulation has been proposed which yields samples with optimum acoustic absorption properties.

2. METHODOLOGY

For this study, loose mixes of granular and fibrous carpet tile waste were obtained from a typical recycling operation carried out on PVC-backed reject carpet tiles with nylon pile. The operation consisted of granulating the carpet tiles through a triple-blade, vertical rotation, granulator that had been fitted with a 2mm-aperture screen in order to control the particle size distribution ($< D > \cong 0.85$mm). The granulator output was then passed directly into a cyclone system that separated the material into fibrous (nylon pile) and granular (PVC tile backing) components. Consolidated samples were produced by adding 30% (by mass) PU binder to the loose aggregate of the selected particle size and leaving the sample in a mould to cure under a pressure of 850 Pa.

Figure 1 presents the measured normal incidence absorption coefficient for two 20mm hard-backed layers of loose mixes of purely granular component which differ by the particle dimension, *D*. Figure 1 also presents the measured absorption coefficient for a 20mm hard-backed layer of fibreglass to provide a direct reference to the expected acoustic performance. The results demonstrate that for the selected layer thickness the acoustic performance of the recycled carpet granulates is comparable or superior to that observed in the case of fibreglass in the low and medium frequency range. This performance can be associated with relatively high values of the porosity and flow resistivity which this particular type of materials normally exhibits. A summary of the non-acoustic parameters for these materials is provided in Table 1.

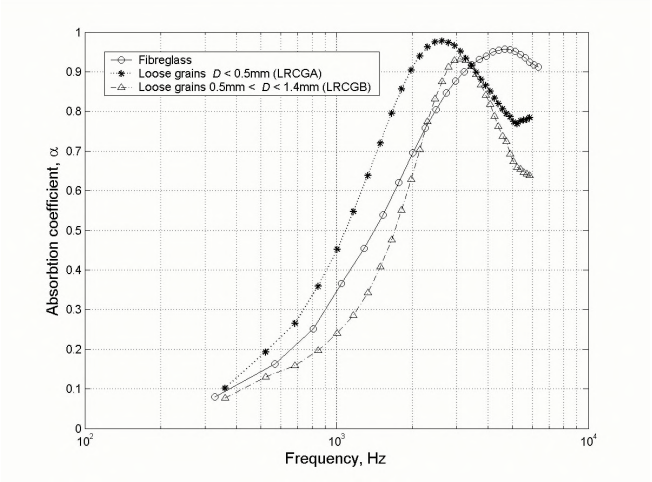


Figure 1. A comparison of the normal incidence absorption coefficient for a 20mm hard-backed layer of loose recycled granulates and commercial fibreglass.

Unfortunately, applications for loose materials are limited because it is more practical to use acoustic absorbers in the consolidated form. A major problem with the consolidation process in the case of granulates is the inevitable loss of porosity and the increase in the values of the real part of the characteristic impedance resulting in the reduced absorption performance. Figure 2 presents the measured real and imaginary parts of the normalized characteristic impedance and complex wavenumber for the loose and consolidated mixes which were prepared from the same type of aggregate (0.5-1.4mm base). The experimental data confirms that the consolidation process results in up to 50% increase in the real part of the characteristic impedance in the frequency range of 400-4000 Hz. Although the tortuosity (α_∞) is relatively unchanged by the compaction

process, the porosity value is strongly affected (see Table 1). In the case of material CRCGG, the porosity is reduced from 71% to 40% due to the binder and compaction pressure effects. The consequences are that despite some increase in the attenuation (see $\mathrm{Im}(k_b)$ in Figure 2) the absorption coefficient of the consolidated sample reduces by up to 20% in the higher frequency range.

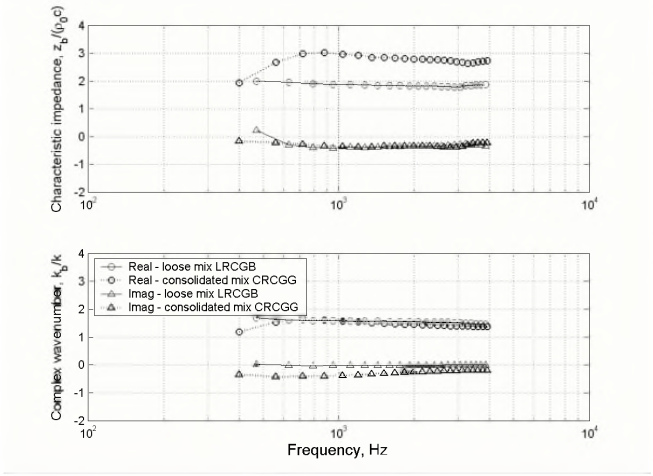


Figure 2. The effect of consolidation on the characteristic impedance and complex wavenumber of the 0.5 – 1.4mm material.

A noticeable improvement in the absorption performance can be attained if a 50% fibrous component is added to the purely granular mix (sample CRCGF). Adding the fibrous matter helps to recover the porosity to around 55% and reduce the mean pore size (r) from 230μm to 180μm. As a result, the value of the flow resistivity (R) increases from 18 kPa s m^{-2} in the case of the purely granular mix to 46 kPa s m^{-2} in the case of the grain-fibre composite mix. This combination of medium porosity and flow resistivity values yields a porous structure which is a more efficient acoustic absorber if manufactured in the form of 20-50mm layers. Similar values of porosity and flow resistivity are characteristic of Coustone or QuietStone which are granular acoustic absorbers manufactured by Sound Absorption UK Ltd. This phenomenon is related to the reduced sound speed and the increased attenuation in the consolidated grain/fibre mix (see Figure 3).

Table 1. Summary of material parameters (F – fibre, G – grains).

Material	R, kPa s m^{-2}	Ω	α_∞	D, mm	r, μm
LRCGA	41.8-48.7	0.75	1.39	$< 0.5^{(G)}$	N/A
LRCGB	10.2-14.1	0.71	1.26	$0.5 - 1.4^{(G)}$	N/A
CRCGG	18.0	0.40	1.31	$0.5 - 1.4^{(G)}$	230
CRCGF	45.9	0.55	1.22	$0.5 - 1.4^{(G+F)}$	180

Adding the fibrous component results in an increased imaginary part of the characteristic impedance which appears to have a noticeable positive effect on the measured values of the normal incidence absorption coefficient presented in Figure 4. This figure demonstrates an up to 25% improvement in the absorption coefficient in the low and high frequency regimes. This effect is closely predicted by the model proposed in (Horoshenkov and Swift, 2001).

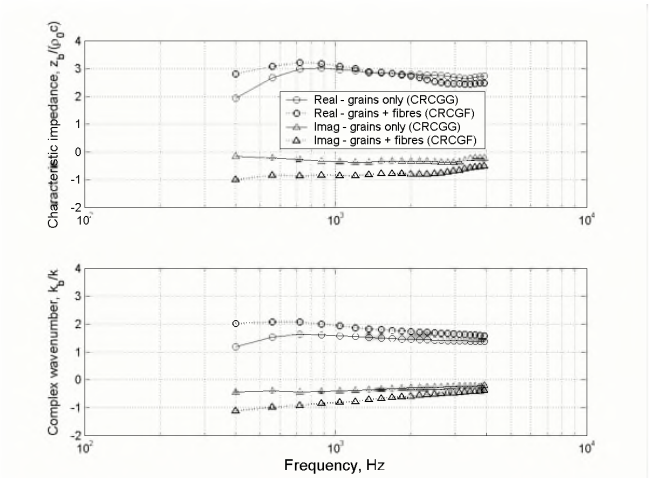


Figure 3. The measured characteristic impedance and wavenumber for two consolidated granular mixes.

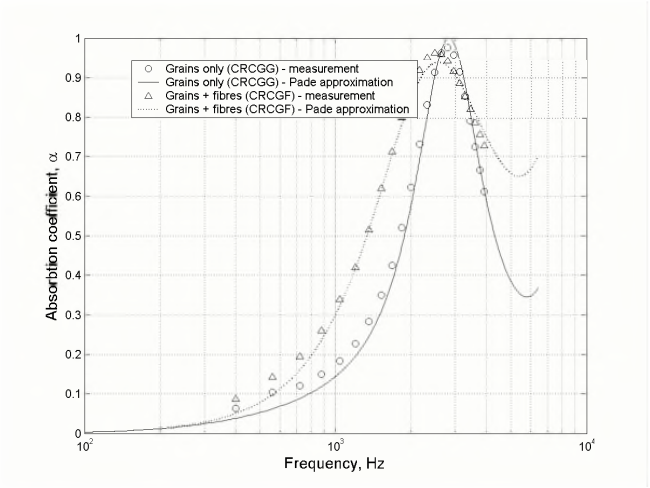


Figure 4. The normal incidence absorption coefficient of 20mm layers of consolidated granular mixes.

3. REFERENCES

Miraftab, M, et al. (2004). *Recycling Carpet Waste into Acoustic Underlay for Commercial Production*, Final Report for WRAP, UK http://www.wrap.org.uk/reports.asp

Swift, M.J. (2000). *The Physical Properties of Porous Recycled Materials.* PhD thesis, University of Bradford, UK

Horoshenkov K.V. and Swift, M. J. (2001). "Acoustic properties of granular materials with pore size distribution close to log-normal," J. Acoust. Soc. Am. Vol. 100(5), 2371-2378.

INFLUENCE OF MICRO-STRUCTURAL PROPERTIES ON THE ACOUSTIC PERFORMANCES OF NOVEL METALLIC FOAMS

Dominic Pilon[1,2], Raymond Panneton[1], Franck Sgard[2], Louis-Philippe Lefebvre[3]
[1]GAUS, Dept. of Mech. Eng., Université de Sherbrooke, Sherbrooke (QC) J1K 2R1, Canada
[2]LASH DGCB URA CNRS 1652, ENTPE, 2 Maurice Audin, 69518 Valux-en-Velin, France
[3]IMI – NRC, 75 Mortagne blvd., Boucherville (QC) J4B 6Y4, Canada

1. INTRODUCTION

Smart design of acoustic materials is fast becoming one of the main research interest of the acoustic community. This is a necessary step in order to develop new and more efficient sound absorbing materials for tomorrow's applications. Two main approaches can be used to achieve this goal: a microscopic approach and macroscopic one. The microscopic approach consist in developing an acoustic wave propagation model which is directly based on the microstructure of the given material. The macroscopic approach involves establishing links, through either analytical or empirical methods, between the micro-structural properties of porous materials and their macroscopic acoustic properties. Once these links have been established, they are substituted in the macroscopic models and the necessary tools for optimizing the microstructure of an acoustic material are then available. In both approaches, the ability to produce the porous materials with the desired micro-structural properties is also an important part of this equation.

The materials evaluated in this study are newly developed metallic foams having an interesting combination of properties, including good sound absorption qualities [1, 2]. A distinctive feature of the process is its ability to control some of the foam's final micro-structural properties such as the average pore D_p and window D_w diameters. With such a process, the concept of smart design of acoustical materials can be deemed attainable. The main objective of the work presented here is to establish links between the micro-structural properties of the metal foams and their acoustic properties and to then study the influence of the micro-structural properties on the acoustic performances.

Some models correlate wave propagation in porous media with some micro-structural parameters. Usually, they are only valid for simple pore geometries: cylindrical pores with constant circular, rectangular, triangular or slit section. More general models developed by Johnson et al. and by Champoux and Allard [3, 4] are valid for a wide range of porous materials with arbitrary pore geometry. They introduced two parameters, the viscous Λ and thermal Λ' characteristic lengths that are known to represent, from a macroscopic point of view, the radius of the small and large pores respectively [3, 4]. Generally speaking, these models are semi-phenomenological models and are not truly based on the microstructure of the studied materials. In 1999, Wang and Lu [5, 6] developed an analytical model for an acoustic material having an honeycomb structure. They optimized the microstructure of the material in order to maximize its acoustic sound absorption. In 2000, Lu et al [7] develop a theory based on the acoustic impedance of circular openings and of cylindrical cavities to model the acoustical behavior of an aluminum foam having this type of microstructure. They completed a parametric study to identify the influence of the micro-structural properties on the sound absorption of the foam and then identified an optimal pore size.

The approach of developing a micro-structural model for a new type of material is quite interesting, but can be difficult for complex microstructure geometries. Hence, as a first step, the approach that will be used here is based on linking the acoustic parameters of the Johnson-Champoux-Allard models, mainly the viscous Λ and thermal Λ' characteristic lengths, to the window and pore diameters of the metallic foams. This should lead to the development of a semi-phenomenological model which takes micro-structural properties as input parameters. Such a model will allow for the study of the influence of micro-structural properties on the acoustical performances of the metal foams.

2. MATERIAL CHARACTERIZATION

For the purpose of this research, five different metallic foams were produced. Various manufacturing parameters were used to obtain materials with different pore size distributions and densities. In the end, four disk shaped samples (diameter of 29 mm and thickness of 10 mm) were made for each metallic foam.

2.2 Micro-structural properties

The micro-structural analysis of the metallic foams was done measuring the pores and windows on micrographs taken with a SEM (model JEOL JSM-6100). The pore D_p and window D_w diameters were evaluated by image analysis on the digitalized SEM micrographs. The windows are defined here as the opening between two adjacent pores.

Each pore and window micrographs are taken at 20X and 50X magnification respectively. The large and small diagonals of the pores and windows were measured using an image analysis software. It is important to point out that windows smaller that 10 μm are not taken into account in the measurements, since they are too small to be measured on a 50X micrograph. The average pore and window diameter measurements are presented in Table 1.

Table 1. Comparison of acoustic and SEM measurements

	Window		Pore	
Metal foam	2Λ (μm)	D_w (μm)	$2\Lambda'$ (μm)	D_p (μm)
A	87	92	357	363
B	74	91	297	305
C	83	101	384	389
D	47	83	238	256
E	77	103	667	676

2.2 Macroscopic acoustic properties

The viscous Λ and thermal Λ' characteristic lengths of the 5 metallic foams are also presented in Table 1. All measurements described here were performed on all four samples of each metallic foam. The viscous and thermal lengths were measured using an acoustical [8] method. The mechanical properties of the metallic foams (E, ν, and η) were not measured here since the rigid frame hypothesis was used to model the acoustic behavior of the foams. The metallic nature of the frame justifies the use of this hypothesis.

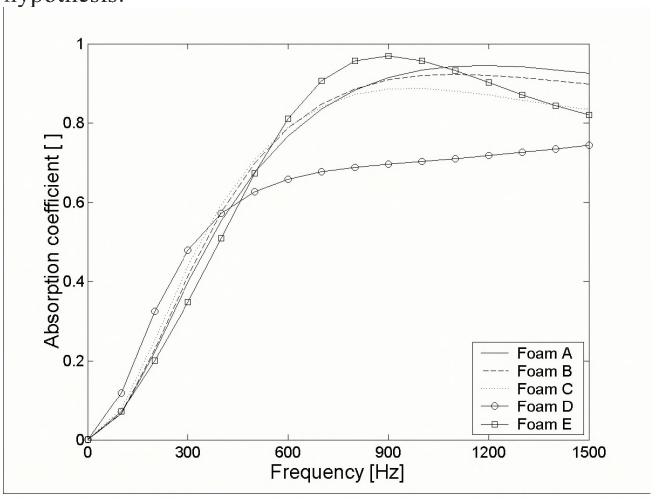


Figure 1. Absorption coefficient for the 5 metal foams with a thickness of 50 mm.

3. RESULTS

The agreement between the measured average pore diameter D_p and $2\Lambda'$ is good for all five metallic foams. This confirms that the thermal characteristic length is equivalent to the pore radius. As for the viscous characteristic length Λ, the agreement is initially not as convincing: Λ is smaller that the average window radius $D_w/2$. This result was nevertheless expected. Since windows smaller than 10 μm are not considered in the measurement of D_w, it is reasonable to assume that the actual average window diameter is smaller that the one measured.

Figure 1 presents the absorption coefficient for the 5 metallic foams with a thickness of 50 mm. As it can be seen, the metal foam with the smallest pore diameter, material D, is clearly the less effective from an acoustical point of view. As for the material with the largest pore diameter, material E, it presents the highest absorption peak of all 5 foams.

4. CONCLUSION

It has been shown that a direct link exists between the thermal characteristic length Λ' of the Johnson-Champoux-Allard model and the average pore diameter D_p ($\Lambda' \approx D_p/2$) and between the viscous characteristic length Λ and the average window diameter D_w ($\Lambda \approx D_w/2$) in the studied open-cell metallic foams. By introducing these relations in the Johnson-Champoux-Allard model, it becomes possible to asses the influence of the pore and window diameters on the acoustical performances of the foam and hence, identity the optimal microstructure. Considering that the metal foam production process used allows a good control of the final micro-structural properties, smart design of acoustical materials, namely optimization of the microstructure to obtain specific acoustical performances, can be deemed attainable with these new materials.

REFERENCES

[1] L.-P. Lefebvre and Y. Thomas, *Method of making open cell material*, U.S. Pat. No. 6660224, December 2003.

[2] L.-P. Lefebvre, M. Gauthier, M. N. Bureau, M. Le Roux, R. Panneton, and D. Pilon, "Properties of nickel foams having different pore size distributions and densities", in Proceedings of Metfoam 2003, Germany, 2003.

[3] D. L. Johnson, J. Koplik, and R. Dashen, "Theory of dynamic permeability and tortuosity in fluid-saturated porous media", J. Fluid Mech. **176**, 379-402, (1987).

[4] Y. Champoux and J.-F. Allard, "Dynamic tortuosity and bulk modulus in air-saturated porous media", J. Appl. Phys. **70**, 1975-1979, (1991).

[5] X. Wang and T. J. Lu, "Optimized acoustic properties of cellular solids", J. Acoust. Soc. Am. **106**, 756-765, (1999).

[6] T. J. Lu, "Ultralight porous metals: From fundamentals to applications", Acta Mechanica Sinica **18**, 457-479, (2002).

[7] T. J. Lu, F. Chen, and H. Deping, "Sound absorption of cellular metals with semiopen cells", J. Acoust. Soc. Am. 108, 1607-1709, (2000).

[8] X. Olny, R. Panneton, and J. T. Van, "Experimental determination of the acoustical parameters of rigid and limp materials using direct measurements and analytical solutions", in Proceedings of Forum Acusticum 2002, Spain, 2002.

ACKNOWLEDGEMENTS

The authors wish to thank NSERC Canada, IRSST Quebec, NRC Canada and the North-American office of the French University Association for their financial support.

SPEECH RECOGNITION BY GRADES 1, 3 AND 6 CHILDREN IN CLASSROOMS

J.S. Bradley and H. Sato[b]

Institute for Research in Construction, National Research Council, Montreal Rd. Ottawa, K1A 0R6

(b) current address: Institute for Human Science & Biomedical Engineering, Tsukuba, Japan

Introduction

This paper summarises the results of new acoustical measurements (for both occupied and unoccupied conditions) and speech recognition tests in 43 classrooms of grade 1, 3 and 6 students [1,2].

Speech and noise levels were measured during a regular teaching activity as well as during the speech tests. Room acoustics measurements were obtained from impulse response measurements for both occupied and unoccupied classrooms. The Word Intelligibility by Picture Identification (WIPI) test was used to measure speech recognition scores for varied signal-to-noise ratio (S/N).

Room Acoustics Measurements

Room acoustics quantities were obtained from impulse response measurements in the classrooms. A sine sweep signal (covering the 6 octave bands from 125Hz to 4kHz) was used to obtain the impulse responses and was reproduced by a small loudspeaker with directional properties similar to a human talker. The speaker was set 1.5 m above the floor at the front of the room, where the teacher might normally stand. Sound level meters with digital wireless transmitters were located 1.2 m above the floor at 4 locations in each classroom.

Oct. band center frequency, Hz	125	250	500	1k	2k	4k	A-weighted
Occupied							
Mean rev. time,s	0.58	0.51	0.45	0.40	0.38	0.39	0.41
S.D.	0.14	0.09	0.10	0.11	0.09	0.08	0.09
Mean C50, dB	5.34	6.39	7.98	9.75	11.12	11.46	10.49
S.D.	3.55	2.76	2.53	3.00	3.09	2.83	2.68
Unoccupied							
Mean rev. time,s	0.61	0.53	0.48	0.45	0.43	0.43	0.45
S.D.	0.15	0.10	0.11	0.12	0.12	0.11	0.11
Mean C50, dB	5.20	6.01	7.37	8.32	9.58	9.87	9.13
S.D.	3.71	2.59	2.36	2.90	3.07	2.70	2.63

Table 1: Mean reverberation times and early-to-late arriving energy ratios (C_{50}) measured in occupied and unoccupied classrooms.

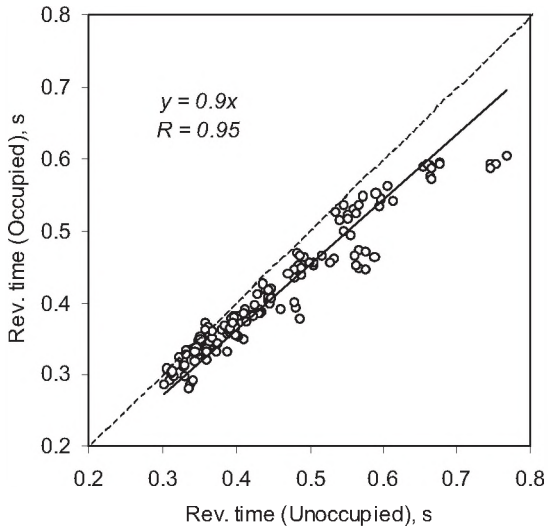


Figure 1. Relation between measured reverberation times in occupied and unoccupied classrooms.

For the unoccupied classrooms, mid-frequency reverberation times varied from 0.3 to 0.7 s with a mean of 0.45s. When the classrooms were occupied, reverberation times were decreased by approximately 10% as shown in Fig.1. Early decay times indicated similar results. Table 1 gives mean reverberation time and early-to-late arriving sound levels for both occupied and unoccupied conditions.

Measurement of Speech and Noise Levels

It is very important to know the levels of teachers' voices and classroom noises during actual teaching activity as well as levels during the speech tests. Distributions of recorded levels, at 200 ms intervals, were used to estimate separate speech and noise levels [3]. Two normal distributions were fitted to each histogram of A-weighted

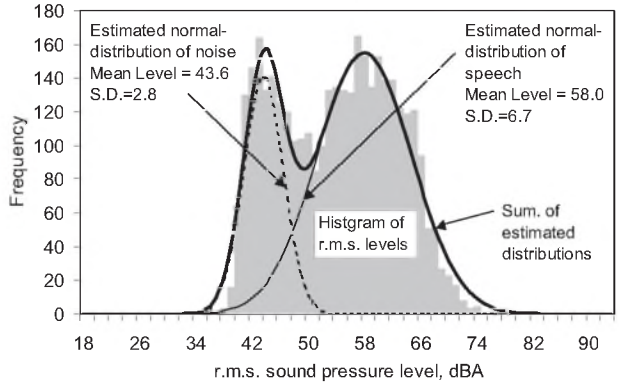


Figure 2. Example distribution of sound levels measured over 200ms intervals in an active class.

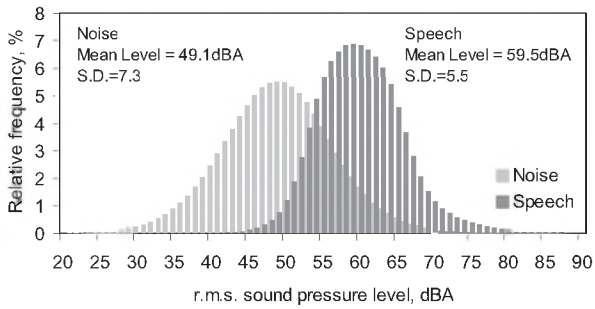


Figure 3 Relative frequency distributions of speech and noise levels from 28 classrooms.

levels as illustrated in Fig. 2. One distribution identified the noise and the other the teachers' voice levels.

Fig. 3 presents distributions of the average speech and noise levels. Mean speech and noise levels were 59.5dBA and 49.1dBA respectively. The corresponding free field speech level 1m from the talker was estimated to be 68.8dBA. The results in Fig. 3 indicate an average S/N ratio of about 10 dB. It was estimated that only 2% of the cases would satisfy a S/N ≥ 15dBA criteria.

Speech Recognition test Results

The WIPI test was used because it is easy to explain to listeners of a wide range of ages. It consists of simple test words familiar to 5 year olds, which were presented in the carrier phrase, "Please mark the _____ now." The students responded by placing a sticker on one of 6 pictures to indicate the correct word. The students sat at their desks in their regular classroom. The tests were carried out in 43 classrooms evenly distributed among grade 1, grade 3, and grade 6 students (6, 8, and 11 year olds). A total of 878 students were evaluated in 43 classrooms. Each grade 1 student was tested at 2 different S/N values and the other students at 3 different S/N values to give a total of about 2285 individual speech recognition tests.

The same sound source described above was used to reproduce the test sentences. Speech and noise levels were measured during each test using the statistical technique described above. These levels were used to determine S/N ratios for each test at each microphone position. There were on average 5 students near each microphone.

Figure 4 shows the mean speech intelligibility scores of each group of students associated with a particular measurement microphone position. They are plotted versus S/N separately for the grades 1, 3 and 6 students. An analysis of variance of the scores showed that there were highly significant main effects of student age and S/N as well as a significant interaction effect of these 2 independent variables. That is, although there is significant scatter in the results, there are highly significant effects related to the age of the listeners. The younger children clearly need higher S/N to obtain the same intelligibility scores as the older children in these tests. The large scatter at lower S/N values is probably indicative of how students react to more difficult listening conditions. At lower S/N, some students can still do quite well, but others more or less give up and get much lower scores.

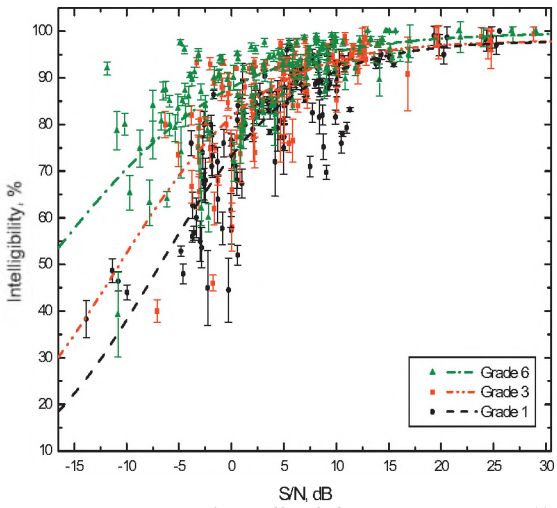


Figure 4. Mean speech intelligibility scores versus S/N by school grade.

The performance of the 3 age groups can be compared by considering the required S/N for a 95% intelligibility score as indicated by the mean trend lines. While grade 6 students could, on average, achieve 95% correct scores for a S/N of +8.5 dB, the grade 3 students required +12.5 dB S/N and the grade 1 students +15.5 dB S/N. In this case there is a 7 dB difference between the needs of grade 1 and grade 6 students. Of course many students score below this mean trend. For very high S/N cases (+25 to +30 dB), the grade 1 and 3 students scored ~98% correct and the grade 6 students ~99.5% correct, indicating that all students can do very well on the WIPI test in actual classrooms when there is minimal masking noise.

Conclusions

The results clearly show the importance of better conditions, with lower noise levels, for younger students. However, it will not be obvious to adult listeners that younger children cannot understand speech in moderately noisy conditions.

Acknowledgements

The authors are grateful for the financial support and collaboration of the Canadian Literacy and Language Research Network and for the help of Ms. Kimberlee Cuthbert in carrying out these experiments. They are also very appreciative of the help of the audio group at the Banff Centre for their help in editing the speech recordings and to the many teachers and students who participated.

References

1. Sato, H. and Bradley, J.S., "Evaluation of acoustical conditions for speech communication in active elementary school classrooms", Paper Tu4.B1.1, Proceedings of the 18th International Congress on Acoustics, Kyoto (2004).
2. Bradley, J.S, and Sato, H., "Speech Intelligibility Test Results for Grades 1, 3 and 6 Children in Real Classrooms", Paper Tu4.B1.2, Proceedings of the 18th International Congress on Acoustics, Kyoto (2004).
3. M.R. Hodgson, R. Rempel and S. Kennedy: Measurement and prediction of typical speech and background-noise levels in university classrooms during lectures, *J. Acoust. Soc. Am.* 105, 226-233 (1999).

INVESTIGATION OF THE OPTIMUM ACOUSTICAL CONDITIONS FOR SPEECH USING AURALIZATION

Wonyoung Yang and Murray Hodgson
School of Occupational and Environmental Hygiene, University of British Columbia, 2206 East Mall, BC, V6T 1Z3
wyang@interchange.ubc.ca

1. INTRODUCTION

Speech intelligibility is known to be mainly determined by the signal-to-noise level difference at a listener, and reverberation. It is directly related to signal-to-noise level difference and is inversely related to the reverberation time. However, their interaction results in a complicated situation in real rooms. Increased reverberation benefits speech intelligibility by increasing speech levels. Noise is also affected by the reverberation time, as are the speech levels. Thus, the spatial relationship between a listener and the sound sources – both speech and noise – affects the optimal reverberation for speech intelligibility in rooms [1].

In this project, an experimental approach to identify the optimal reverberation time in an idealized room, and validate theoretical prediction, considering babble noise sources inside the room, is presented using auralization. Realistic optimal reverberation times are found using speech-intelligibility tests with normal-hearing and hard-of-hearing subjects. The best metric predicting speech intelligibility is presented for both normal-hearing and hard-of-hearing groups.

2. METHOD

2.1 Subjects

Subject groups for the study were normal-hearing and hard-of-hearing people with a mild to moderate sensori-neural hearing loss. The hearing-loss criterion for the hard-of-hearing subjects was less than 20 dB at 250 Hz and 500 Hz and greater than 30 dB at 1 kHz to 8 kHz. Twenty-four normal-hearing subjects (mean age=27) and ten hard-of-hearing subjects (mean age=58) completed the tests.

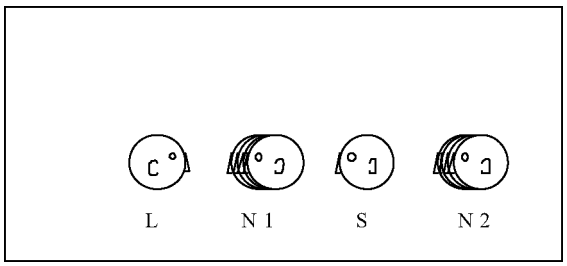


Fig. 1. Room elevation showing the speaker (S), listener (L) and noise-source (N) positions.

2.2 Test materials and simulated sound fields

The Modified Rhyme Test (MRT) was used as the speech-intelligibility test. The MRT lists were processed through the CATT-Acoustics room-acoustical prediction and auralization program with 4-talker babble noise. The volume of the virtual room used for the simulation was 385 m^3 (11 m x 7 m x 5 m). Fig. 1 shows the elevation of the virtual room, and the relative positions of the listener, the speaker, and the noise sources. In this study, the object was to model an idealized room with exponential sound decay. The effects of the different distributions of various surface materials on the walls, floor and ceiling were excluded by using the same absorption coefficients and diffusion coefficients for all octave bands and for all of the surfaces.

A total of 16 different sound fields were created, consisting of the combinations of 2 different speech- and noise-source output levels (SNS = 0 and +5 dB), 4 different reverberation times (RT = 0, 0.2, 0.4, and 0.8 s), and 2 different positions of the noise source (see Fig. 1). The speech-to-noise level difference at the listener (SNR) varied from -6 dB to 8.5 dB.

The completed auralization test materials were transferred to a compact disc for presentation using a CD player. Hearing screening tests were done prior to the main speech-intelligibility test to identify the hearing categories of subjects. The tests were processed through a Sony MDR V600 headphone in a soundproof room.

3. RESULTS

Fig. 2 shows the mean speech-intelligibility scores with 95 % confidence intervals. The speech-intelligibility scores were analyzed statistically by analysis of variance (ANOVA). Between normal-hearing and hard-of-hearing groups, there was a significant difference ($a = 0.05$) in the speech-intelligibility scores. However, there was no relationship between the mean difference and SNR.

When the distance of the listener from the noise is farther than the distance from the speaker (N2), there was an overall trend for speech intelligibility to decrease with increasing RT for both normal-hearing and hard-of-hearing groups (Figs. 2b and d); i.e. the optimum reverberation time was zero. At the listener position, the noise level decreased more than the speech level in this case. Therefore, reverberation had a detrimental effect on speech intelligibility. When the noise was positioned between

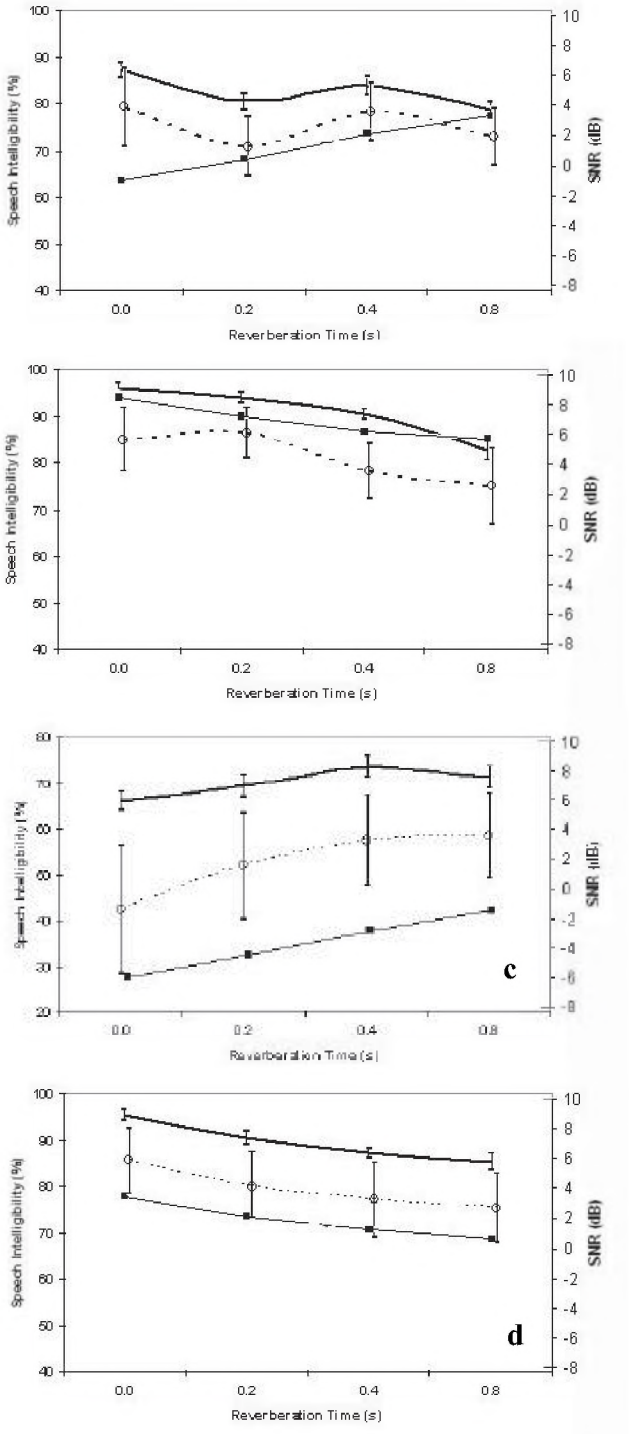


Fig. 2 Mean speech-intelligibility scores with 95% confidence intervals and signal-to-noise ratios (a: SNS=5 dB + N1; b: SNS=5 dB + N2; c: SNS=0 dB + N; d: SNS=0dB + N2. The lines have been offset to avoid overlapping. __: Normal-hearing; …: Hard-of-hearing; __: SNR).

Table 1. Strength of third-order polynomial regression

R^2	Normal hearing		Hard-of-hearing	
	Calculated	Measured	Calculated	Measured
U_{80}	84.8	83.3	91.3	91.3
U_{70}	86.5	84.3	92.9	91.5
U_{60}	87.6	85.1	93.0	91.5
U_{50}	88.4	85.9	92.2	91.2
U_{40}	88.2	86.2	89.6	89.9
U_{30}	86.2	86.7	83.4	87.4
SNR	-	72.7	-	84.9

the listener and the speaker (N1), there were significant differences among the RTs (a = 0.05). For normal-hearing with SNS = 0 dB (Fig. 2c), the speech intelligibility increased with increasing RT until 0.4 s then decreased (a = 0.05). For hard-of-hearing subjects with SNS = 0 dB, there were similar trends as for normal-hearing subjects. In this case, the speech level decreased more than the noise level at the listener's position. Therefore, reverberation increased the speech-intelligibility scores. When the SNS was 5 dB (Fig. 2a), unexpected results occurred. The peak occurred at RT = 0.4 s with SNS = 0 dB; however, at RT = 0 s, there was another peak.

For each sound-field configuration, useful-to-detrimental ratios were calculated from predicted impulse responses combined with speech and noise levels. As early-time limits of the useful and detrimental energies, 30, 40, 50, 60, 70 and 80 ms were used in the unweighted frequency spectrum. Third-order polynomial regression analyses were done with the mean MRT scores of each sound field configuration. Table 1 shows the strength of the relationship, R^2, between each measure and speech intelligibility for normal and hard-of-hearing people. Since the form of the fit and the number of data points was the same in every case, the success of each measure can be compared in terms of the corresponding R^2 values. In both calculation methods, there were similar trends. U_{50} values were best suited for predicting speech intelligibility for the normal-hearing group. However, for the hard-of-hearing group, the early-time limits were slightly higher than those for the normal-hearing group.

4. CONCLUSION

Optimal reverberation times depended on the signal-to-noise level difference and the spatial relationship between the listener and the sound sources for both normal and hard-of-hearing subjects, as predicted by theory. Hard-of-hearing people needed high signal-to-noise level difference rather than shorter reverberation time to achieve better speech intelligibility. They required longer early energy than normal hearing did.

REFERENCE

1 Hodgson and Nosal, "Effect of noise and occupancy on optimal reverberation times for speech intelligibility in classrooms", J. Acoust. Soc. Am., 111(2) 931-939.

Characteristics of the Noise, Reverberation Time and Speech-to-Noise Ratio Found in Day-Care Centers

Michel Picard
Ecole d'orthophonie et d'audiologie, Université de Montréal, Qc, Can., H3C 3J7. michel.picard@umontreal.ca

1. INTRODUCTION

Today's day-care centers are apparently extremely noisy environments. Several authors have been reporting high-level of exposure of children to noise with representative values ranging from about 65 to 92 dBA in groups of children aged between 18 and 60 months (Truchon-Gagnon & Hétu, 1988; Picard & Boudreau, 1999, McLaren & Dickinson, 2002). Given the paucity of data on this particular issue and causes of the problem (Golden, 2001), characteristics of the noise (exposure levels and spectrum), reverberation time and speech-to- noise ratios were studied in active day-care centers.

2. METHOD

Noise measurements were collected for 24 sites accommodating groups of 1.5 to 5 years-old children (N = 4-16 children per group). Ten minutes exposures to three types of children's activity noise were recorded on a minidisc recorder for further analysis in laboratory, namely: 1- unstructured activities of the children; 2- structured ones (e.g. story telling); and 3- lunch time. Noise levels were also collected for the unoccupied case before the children's arrival. Reverberation time (RT) at 1000 Hz in the unoccupied condition was also studied as a covariate and similarly for signal-to-noise ratio (SNR) when personnel is addressing children in the conduct of regular activities. RT was measured in accordance with ISO 3382 (1997) interrupted noise method. SNR was derived from samples of discourse when personnel is instructing children or conducting verbal activities with them in a manner to isolate meaningful discourse or instructions from the din. The principles behind this approach was to fit two normal distributions to statistical distributions of measured sound levels consisting of combination of sound sources. This method was proposed by Hodgson, Rempel & Kennedy (1999) and is currently being used by Sato & Bradley (2004) in the NRC classroom acoustics research program. All measurements were collected in the center of the room usually occupied by children with the microphone of a precision sound level meter (class 1) located 95 cm above the floor.

3. CHARACTERISTICS OF THE NOISE FOUND IN DAY-CARE CENTERS: EXPOSURE LEVELS AND SPECTRUM

Statistical distributions of A-weighted levels of exposure for the various sites and measurement conditions (in 10 min. integration interval) ranged from 66 to 94 dBA with representative values of 79.4 dBA for unstructured activities, 75.3 dBA for stuctured ones and 75.9 dBA for lunch time. By comparison, values in the unoccuppied condition varied from 29 to 62 dBA for the various sites with an average integrated level of 45.3 dBA. Third-octave band analysis of the noise found in day-care centers while activities are in progress revealed a long-term spectrum corresponding to Pearsons 'loud' (and possibly 'shout') speech uttered by children and females (Pearsons, Bennett & Fidell , 1977). The observed low frequency boost by 8 dB between 200 and 400 Hz may be the result of interacting room reverberation.

4. REVERBERATION TIME

Reverberation time at 1000 Hz obtained in ten of the 24 rooms taken as representative showed no correlation with noise exposure levels nor room volume (actually ranging from 99 to 319 cubic meters). More specifically, $RT_{1\ kHz}$ ranged from 0.3 to 0.56 s. in facilities with a suspended ceiling of sound-absorbing tile (n = 5), not so different from the five rooms without this feature ($RT_{1\ kHz}$ from 0.55 to 0.85). In both cases, $RT_{1\ kHz}$ was not far from values in the range of 0.4-0.6 s. recommended for classrooms and rooms used for speech communication (Bradley, 1986; ANSI, 2002).

5. SIGNAL-TO-NOISE RATIO

Derivation of SNR when personnel is interacting with children or giving instructions is currently indicating a wide range in values. Of the five sites analyzed to date, SNRs can be close to 0 dB as a result of some intrusive speech of children not participating in the on-going activity or not paying attention to it (or speech of nearby personnel). Conversely, SNRs in excess of 10 dB are not uncommon. On average, SNR of 6 dB was identified. This is far from recommended value of +15 dB for classrooms (Bradley, 1986). However, SNRs in excess of 10 dB indicate that unobstructed verbal communication is not beyond reach with appropriate controls of the communication situation in day-care centers. This is suggesting pedagogical style as an important ingredient of successful verbal communication beyond acoustical solutions.

6. CONCLUSION

Overall, current findings suggest that speech from

conversations in parallel is the principal component of the so-called 'noise' problem in day-care centers. Excessive reverberation does not seem to be a major contribution to speech interference. Given the high levels of sound signals in day-care centers, it is not clear that children could initiate or sustain any meaningful verbal communication activity of their own unless they use extreme vocal effort to speak loud enough to raise their voice above the din. This is far from an ideal listening situation - and generally speaking, communication environment, at a time when children are developing their language competence.

REFERENCES

ANSI (2002). Acoustical performance criteria, design requirements and guidelines forschools. ANSI 12.60. American National Standards Institute, Melville.

Bradley JS (1986). Speech intelligibility studies in classrooms. Journal of the Acoustical Society of America, 80, 846-854.

Golden MV (2001). An acoustical analysis of infant/toddler classrooms in child care centers. Unpublished Master thesis. Penn.State Univ., State College.

Hodgson MR, Rempel R & Kennedy S (1999). Measurement and prediction of typical speech and background-noise levels in university classrooms during lectures. Journal of the Acoustical Society of America, 105, 226-233 .

ISO (1997). Acoustics - Measurement of the reverberation time of rooms with reference to other acoustical parameters. 2nd Ed. ISO_3382. International Standards Organization, Geneva.

McLaren SJ & Dickinson PJ (2002). Noise in early childhood centres and its effects on staff and children. Paper 811. Internoise 2002.

Pearsons KS, Bennett RL & Fidell S (1977). Speech levels in various noise environments. EPA-600/1-77-025. Washington: Environmental Protection Agency.

Picard M & Boudreau C. (1999). Characteristics of the noise found in day-care centres. Paper presented at the 137th meeting of the Acoustical Society of America, Berlin.

Picard M, Bradley JS (2001). Revisiting speech interference in classrooms. Audiology, 40, 221-244.

Sato H & Bradley JS (2004). Evaluation of acoustical conditions for speech communication in active elementary school classrooms. XVIIIe International Congress of Acoustics, ICA-2004. Kyoto, Japan

Truchon-Gagnon C, Hétu R (1988). Noise in day-care centres for children. Noise Control Engineering Journal, 30, 57-64.

CLASSROOM ACOUSTICS AND ARCHITECTS

Terence Williams, PPMAIBC, PPFRAIC, RCA, Hon. FAIA
terence williams architect inc., 102-2957 Jutland Road, Victoria, BC V8T 5J9

As an architect, good acoustics, good sightlines, and a sense of intimacy have been the criteria used for classroom design and performing arts facilities. Although classrooms and lecture theatres tend to be smaller in size, volume and number of seats than performing art venues, we have tried to incorporate the above criteria to improved room acoustics in classrooms.

As architects with an appreciation for the benefits of an improved acoustical environment, we do not consider ourselves to be acoustical consultants, we have taken a pragmatic approach to acoustical design, in as much as that many of the installations are untested, either before or after design and construction, but we have tried to address acoustical concerns within the challenge of designing contemporary classrooms. We have used the criteria outline in Michael Barron's 1993 book entitled Auditorium Acoustics and Architectural Design for describing an acoustical environment as follows:

Clarity	is it muddy or clear?
Reverberance	is it dead or live?
Envelopment	is it expansive or constructed?
Intimacy	is it remote or intimate?
Loudness	is it loud or quiet?

The functional challenges of a classroom design are many.

There is a space requirement to seat a specified number of students usually from 25 to 250. The room may be flat floored or in larger classrooms and lecture theatres, sloped or tiered. The spaces are usually multi-purpose and provide for a single speaker without amplification to several speakers and a variety of interactive audio/visual requirements. In current classroom design it is not only the presenter that has access to power and data, but the students too. Writing benches are now being installed for laptop computers and requiring power and data connections. Projection booths provide audio/visual support in addition to the traditional chalkboard or whiteboard needs. Overhead projectors, digital data projection and video and film presentations are now the norm in classroom design. Structural, mechanical and electrical consultants provide necessary engineering input. Architectural, structural, mechanical and electrical needs must be met to satisfy current building codes to support life safety issues and indoor environmental considerations. Classrooms must be accessible to those in wheelchairs and if above 60 seats must make provision for those who are hard of hearing. This is usually in the form of infrared assistance systems. Classrooms are by building code definition "Places of Public Assembly".

In designing classrooms we have considered the following issues to be article in promoting and supporting an improved acoustical environment:

- Acoustic isolation of the classroom from the exterior or adjoining noisy spaces. E.g.: Specifying carpet in adjoining corridors to reduce impact noise.

- Reduction in the ambient noise level within the room. Within some auditoria we have measured NC15. Normal classroom use is substantially higher, NC25 has been our target. This has been achieved with quiet mechanical systems. The use of electronic ballasts in direct/indirect lighting fixtures and a mixture of reflective and absorptive surfaces.

- We have experimented with room shapes and modeling of the walls and ceilings.

- We have tired to consider reflections, echoes and reinforce intelligibility with early lateral reflections.

- We have incorporated the need for diffusion and absorption with various architectural finishes.

- We have experimented with ceiling reflectors and hard reflective surfaces on the floor and walls of the presentation area.

- Floors for the audience have generally been carpeted and rear walls absorptive surfaces.

- We have tried to consider reverberation times realizing that levels above 1.5 seconds will not be good for speech. But also realizing that too dry an environment will lack character and induce a surreal and a natural experience that lacks excitement. Teachers and instructors need to excite, motivate and inspire. It is suggested that a

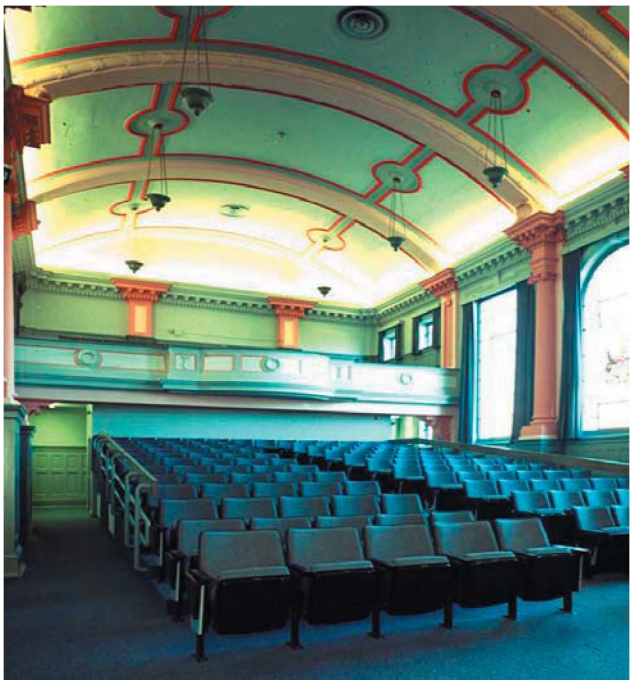
room that is too dry is even more difficult top excite students than a reverberant space.

- We have tried to deal with the two approaches to teaching and learning: 1) "the sage on the stage" 2) "the guide on the side". The Centre for Innovative Teaching at the University of Victoria illustrates these issues.

In addition to the Centre for Innovative Teaching many other examples of classroom, lecture theatre and music/drama education facilities will be illustrated.

References

Barron, Michael (1993). Auditorium Acoustics and Architectural Design.



Young Theatre, Camosun College

Lam Auditorium, University of Victoria

Centre for Innovative Teaching, University of Victoria

Victoria Conservatory of Music

Student Services Building Lecture Theatre, Malaspina University-College

CANADIAN ACOUSTICAL DESIGN STANDARDS FOR K-12 SCHOOLS

Darron Chin-Quee [1], Nadine Munro [2] and R.L. Scott Penton [3]
Rowan Williams Davies & Irwin Inc., 650 Woodlawn Road W., Guelph, Ont., N1K 1B8 [1]dcq@rwdi.com, [3]slp@rwdi.com
Simcoe Muskoka Catholic District School Board., 46 Alliance Blvd., Barrie, Ont., L4M 5K3, [2]nmunro@smcdsb.com

1. INTRODUCTION

The ANSI S12.60-2002 standard "Acoustical Performance Criteria, Design Requirements and Guidelines for Schools" has received much interest from education stakeholders since its release, garnering both support and opposition. The impetus for developing this standard stemmed from a realization that :

a) There were previously no U.S national standards and very few mandated state or local acoustical standards for school acoustics;

b) There are significant portions of the K-12 population whose education or learning potential may be compromised by poor classroom acoustics.

With respect to the latter, research has shown that young children and even those as old as about 15 years of age, do not comprehend speech as well as adults in moderately noisy and/or reverberant spaces. [1,2] Children with hearing loss - permanent (estimated at up to 15%) or temporary due to ear infections (up to 25 % of young children at any time), [2] those learning a second language (up to 45% in at least one major Canadian School Board) [3] and those with conditions such as attention deficit disorder (5% or more of all children) [2], require acoustics in classrooms favouring good speech intelligibility.

Given the experience in the U.S., should the ANSI standard for school acoustics or similar guideline/standard be adopted in Canada? Are the needs of Canadian school children adequately served by existing standards, guidelines and design practices? To answer these questions, available information from a number of provincial / territorial Education Ministries were reviewed. School board practices of some of the largest boards in Canada were also reviewed along with those of several architects who specialize in K-12 school design. Finally, the experience of educational audiologists was sought to assess if the status quo on school design is sufficient to meet the needs of those most at risk.

2. SCHOOL ACOUSTICS GOALS

Except for specialized spaces - music rooms, libraries and gyms, the desirable acoustical goal for most school space is good speech communication requiring:

i) low background noise

ii) low reverberation time

iii) sufficient sound isolation from both internal and external noise sources to prevent freedom from acoustical intrusions or distractions.

3. CANADIAN STANDARDS

Table 1 summarizes the findings on current Canadian practice /standards employed in school design There are no Federal standards or other national standards (e.g. CSA) pertaining to acoustics in schools,

3.1 Provincial / Territorial Guidelines

Capital planning guidelines and practices pertaining to school construction for seven provincial Ministries of Education (All Western provinces, Ontario, Quebec, New Brunswick) and one Territory (Nunavut) were reviewed. Combined, these constitute 95% of Canada's population. Alberta alone has sufficiently comprehensive acoustics guidelines for schools [4].

3.2 School Board / Architect Guidelines

School boards (owners) and architects (designers) determine the acoustics considerations constructed. School boards facilities procurement practices were reviewed for the five largest Canadian cities - Toronto, Montreal, Vancouver, Calgary, Ottawa.

Board guidelines or tenders are usually prepared as specifications. Acoustical considerations are mostly absent, hidden within other design requirements for specific spaces, indirectly considered , (e.g., "concrete block walls are preferred") or stated qualitatively (e.g., prevent disturbing mechanical noise). Performance

Table 1: Review Summary for Canadian Acoustical Education Facility Guidelines / Practice

Agency / Entity	% of Agencies Reviewed Having Guidelines (1) / Practice Which Address					
	Background Noise	Room Acoustics	Sound Isolation	Entire School	Specialty (2) Space Only	Constraints (3)
Provincial Ministries	14	50	33	14	43	29
School Boards	20	80	60	0	80	80
Architects	0	100	100	100	0	100

1) A guideline was considered to address an acoustical factor if any space requirement or performance objective had specific acoustic criteria (e.g., STC, RT, NC/RC) or inferred consideration of acoustical concerns (e.g., " sound proof and acoustic treatment"). Guidelines included facility building manuals, outline specifications for materials or specific spaces. Objective, performance based criteria were largely absent from the guidelines. Acoustics is typically addressed indirectly, (e.g., "classrooms shall have T -bar suspended ceilings")
2) Specialty spaces - gyms, music rooms, construction shops libraries and does not include classrooms and general instruction space..
3) Other consideration affecting acoustics: Constraints imposed on the acoustical design (e.g., costs, durability, maintenance)

specifications are rare , acoustical design objectives or rationale absent. Classroom acoustics is largely ignored with the focus on specialty space - music rooms, gyms.

School architects primarily rely on experience and prior project precedents. None were aware of specific school acoustics guidelines. All were cognizant for the need for good sound isolation and control of reverberation, most applying full height partitions and lay-in tile ceilings. Approaches vary substantially, especially for partition selection. None addressed background noise, deferring to the mechanical designer/contractor. Most primarily consider acoustical issues for specialized space - music rooms, auditoria, gymnasia, shop space.

4. AUDIOLOGISTS' VIEW

Educational audiologists serving school boards were interviewed to get the "front line" practitioners assessment of school facility performance. None were aware of acoustical design guidelines for schools in any of their Boards although some were aware of those issued by Alberta. Audiologists would embrace a comprehensive guideline, many having been active advocates within their boards. Key issues identified are:

- High background sound levels found in classrooms but especially those with wall mounted HVAC units. High sound levels in some classrooms have resulted in retrofit sound reinforcement to avoid teacher vocal strain and to address OHSA concerns.
- High background sound levels and reverberation affect student learning. In the worst schools, teacher fatigue and high teacher transfer rates have resulted.
- School acoustics in new schools are no better that those built 20 years ago

5. CONCLUSIONS

Few performance based guidelines exist in Canada which comprehensively address the key acoustical design considerations in schools. With no national and only one provincial standard specifically geared to school acoustics, most schools are being designed based on non-uniform and often qualitative approaches of the Boards and architects involved. The experience of educational audiologists and many in the acoustics design community suggests that acoustics is often not a major consideration in primary learning spaces. High background noise is a major concern.

A comprehensive national guideline and educational outreach program to inform school designers, board facility planners of the benefits of incorporating better acoustics into school designs is needed.

6. REFERENCES

1. Johnson, CE 2000. "Children's phoneme identification in reverberation and noise", *Journal of Speech Language and Hearing Research* 43, 144-157.
2. Nelson, Peggy, Feb2003. " Sound in the Classroom Why Children Need Quiet", *AHSRAE Journal*, February 2003
3. Sources: Toronto District School Board, www.tdsb.on.ca Commission Scolaire de Montreal, www.csdm.qc.ca
4. Alberta Infrastructure, "Standards and Guidelines for School Facilities", April 2001

DESIGN OF A CLASSROOM FOR DEAF / HEARING IMPAIRED STUDENT EDUCATORS

Darron Chin-Quee [1] and R.L. Scott Penton [2]
Rowan Williams Davies & Irwin Inc., 650 Woodlawn Road West, Guelph, Ontario, N1K 1B8
[1]dcq@rwdi.com [2]slp@rwdi.com

1. INTRODUCTION

Toronto's York University offers the only full-time Deaf Education Program in Ontario and the biggest one of four in Canada. Formerly known as the Teacher Preparation Program in the Education of Deaf and Hard of Hearing Students, this 10-month program teaches future teachers of the deaf and hard of hearing. The program graduates about 20 students per year.

To facilitate instruction of student teachers, many of whom are themselves deaf or hard of hearing, an existing seminar room was retrofitted as a special deaf education classroom in 1993. The classroom was equipped for multi-media presentation with assistive listening devices and designed with room acoustics consistent for high speech intelligibility, particularly in the context of the hearing impaired. Considerations incorporated as objectives into the design included: maximum 500 Hz RT60 values of 0.4 seconds; signal to noise ratios of 20 dB with normal vocal effort; reinforcing reflections within 20 milliseconds; and maximum background noise levels of NC-15 to NC-20.

Assessment of acoustical performance by measurements conducted in 1997 about three years after the facility opened, indicated the facility conformed to most design objectives [1]. Subjective feedback from users indicated a high degree of satisfaction, especially from the hard of hearing students. However, normal hearing students indicated concerns with the space and "have commented that they feel a bit cut off from the outside environment due to the quietness in the room." With 10 years of use, user satisfaction with the space was again recently evaluated. User feedback is discussed herein.

2. ROOM CHARACTERISTICS

Figure 1, illustrates the space layout and finishes. Table 1 summarizes measured acoustical parameters compared with the design criteria.

3. USER FEEDBACK

Hard of hearing students are very satisfied with the space as assistive listening devices and microphones work well in the low noise and low reverberation space.

Consistent with the earlier assessment, normal hearing students still feel isolated and the space is oppressingly quiet. Faculty have surmised that the sense of isolation felt by some students is the same phenomenon encountered in clinical audiology where patients dislike the audiometric booths due to the low reverberation and high sound isolation, characteristics atypical of spaces found in most buildings.

A new concern expressed by the normal hearing students and staff is insufficient signal (likely lack of strong early reflections). Staff have indicated the need to " raise my voice when I teach and that I am tired by the end of a lecture because of that; if I do not raise my voice, the normal hearing students do not hear well enough (similarly, the students need to raise their voices when they are making a comment to the class as otherwise, only the hard of hearing students can hear them with the hand held mics)".

Two other new concerns raised by users - dust buildup on acoustic finishes and poor air quality (stuffy room) , indirectly relate to acoustical considerations. Dust buildup is a by-product of normally porous acoustical finishes. Poor air quality potentially is related to low velocity ventilation systems (to reduce noise) and constrained by the original ductwork sized for higher air speeds. As a result of the stale air, the doors are often left open, degrading the sound isolation.

4. DISCUSSION / CONCLUSIONS

The subjective feedback of users is consistent with a bias in the design towards hearing impaired listeners. Given the mix of students (about 12% "hard of hearing", 25% deaf and the balance normal hearing) the design may have been too heavily weighted towards

Table 1: Summary of Design Criteria and Measured Performance

Parameter	**Proposed Design Criteria: Classroom for the Hearing Impaired**	**Typical Criteria: Classroom for Normal Hearing**	**Measured Value: York University Deaf Education Classroom**
Ambient Noise Level	PNC / NC 20 or less (30 dBA or less)	PNC / NC 30 - 35 (40-44 dBA)	NC 30 - 35 (41 dBA)
Reverberation Time (RT60)	0.4 seconds maximum @ 500 Hz	0.6 - 0.8 seconds@ 500 Hz	0.4 seconds @ 500 Hz
Minimum Signal to Noise Ratio (S/N), Normal Vocal Effort	20 dB	15 dB	16 dB- 21 dB
Arrival of Reinforcing Reflections	Within 20 msec of direct sound, Clarity ratio $C_{\text{arrival time +20ms}}$ of 10 dB or more	Within 35 msec of direct sound, Clarity ratio $C_{\text{arrival time +35ms}}$ of 10 dB or more	Clarity ratio $C_{\text{arrival time +20ms}}$ of 8 - 10 dB
Articulation Loss (% ALCons)	% ALCons < 3%	%ALCons <10%	%ALCons : 2.5 % -3.4%
Speech Transmission Index (STI)	STI > 0.75	STI > 0.55	STI: 0.73 - 0.78
Sound Isolation (STC or NIC)	STC or NIC: 55- 60 to traffic areas or adjacent classrooms	STC 50 (walls) to traffic areas, STC 25-30 (doors)	STC 30 to corridor NIC 34 to corridor (doors)

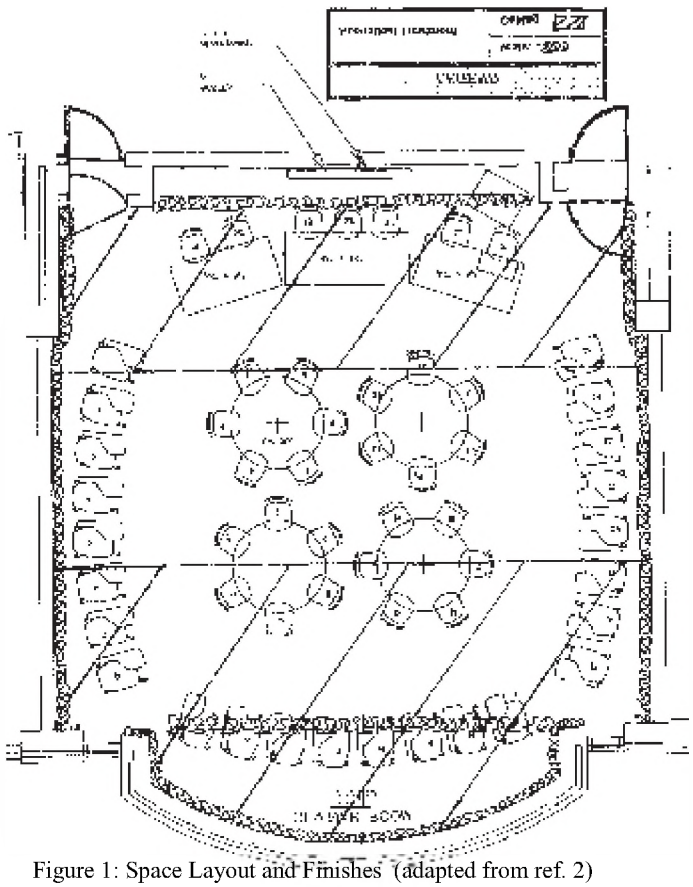


Figure 1: Space Layout and Finishes (adapted from ref. 2)

room acoustics for assistive listening. The major factor contributing to the perceived low signal levels are due to reduced strong early reflections (although, the measured clarity ratios are marginally within specification). While the plan view indicated in Figure 1 suggests central portions of the ceiling remain sound reflective, for architectural consistency these were made sound absorptive. User impressions are that the room functions extremely well with supplementary sound reinforcement.

Issues associated with ventilation and dust as related to acoustics / noise control, highlights the importance of a holistic design approach and coordination among all of the designers and users.

5. REFERENCES

1. Chin-Quee, D.A. "Acoustical Design Parameters for Hearing Impaired Classrooms: A Case Study", *Proceedings of Noise-Con 97,* pp 423 -426 June 1997.

2. D.A.Chin-Quee, "York Universtity Hearing Impaired Classroom Remedial Acoustical Treatment", *Valcoustics Canada Ltd. Report 92-093*, 1993.

6. ACKNOWLEDGEMENTS

We would like to thank York University staff for their help with this paper, especially: Deaf Education Program - Ms. Pam Millett and Dr. Neita Israelite; Facilites Planning - Mr. George Parker, for his help in the earlier 1997 review.

NATIONAL GUIDELINES FOR ENVIRONMENTAL ASSESSMENT: HEALTH IMPACTS OF NOISE

Stephen Bly[1], David Michaud[1], Stephen Keith[1], Carl Alleyne[2] and Diane McClymont-Peace[3]
[1]Health Canada, Consumer and Clinical Radiation Protection Bureau, 775 Brookfield Rd., AL6301B, Ottawa, ON, K1A 1C1
[2]Health Canada, Safe Environments Programme, 3155 Willingdon Green, Burnaby, BC, V5G 4P2
[3]Health Canada, Health Impacts Bureau, 2720 Riverside Drive, Ottawa, ON, K1A 0K9

1. INTRODUCTION

The *Canadian Environmental Assessment Act* (CEAA) requires certain projects with Federal Government involvement to undergo an environmental assessment (EA) before they are allowed to proceed. The intent is to identify, predict, evaluate and mitigate environmental and health impacts of a proposed project before it begins. Under CEAA, a project may not proceed if it causes significant adverse impacts, after mitigation is accounted for [1,2]. Impacts include health effects from project-related noise. Noise will likely continue to be an important issue in projects of major social, economic and military importance.

To cope with this issue, the efficiency and transparency of environmental noise impact assessment needs to be improved. The current situation is problematic because the number of assessments across Canada has increased considerably. Also, a wide variety of noise sources have been assessed, including: aircraft (civilian and military), rockets, rail, highway traffic, energy (wind turbines, gas pipeline compressor stations), construction and industrial. Furthermore, a wide variation has developed in the information and analyses used to assess the potential health impacts of project-related noise. An existing Federal-Provincial document, *National Guidelines for Environmental Noise Control* [3] was not designed to address these problems and elements of the document are out of date.

A new guideline document, *National Guidelines for Environmental Assessment: Health Impacts of Noise* is proposed to address the problems. This paper summarizes some options and considerations for development of the proposed guidance document.

2. HEALTH IMPACTS ASSESSMENT

The EA must consider direct and cumulative effects on human health, *i.e.*, physical, mental and social well-being [2]. The EA must also determine whether health effects are: (i) adverse, (ii) significant and (iii) likely. Significance is determined by: (i) severity, (ii) duration, (iii) frequency, and (iv) reversibility of the effects. The number of people affected may also need to be taken into account.

2.1 Characterizing the Noise Environment

To determine whether effects from project-related noise are adverse, the existing quality of the noise environment (baseline) must be compared to the quality of the noise environment after the project is in place [1]. This requires characterization of the noise environment through noise measurement and/or modeling.

The Guidelines will have to clarify requirements for measurement and/or modeling of noise levels, depending on noise source and type of noise. For example, impulsive noise, from shunting in rail yards and pile driving during construction, requires special consideration compared to more continuous noise, as from highway traffic. Ideally, the sound exposure level should be determined for each impulsive noise event and then appropriately adjusted [4].

Some guidance is also needed to ensure consistent prediction of noise levels from: (i) construction where multiple pieces of heavy machinery are in use simultaneously and (ii) noise sources at marine ports where propagation over water can be important.

The Guidelines will either need to reduce the number of noise metrics used to assess health impacts or provide guidance on conversion between metrics. Some Canadian guidelines require the A-weighted 24 hr time-average sound level, L_{eq24h}, whereas, others require a worst-case A-weighted 1 hr time-average sound level, L_{eq1h}. Annoyance assessments utilize the day-night sound level, DNL. However, civilian aircraft noise is mapped by Noise Exposure Forecast (NEF) contours in the airport vicinity [5].

Various national and international standards and commercial software/equipment are available for measuring/modeling environmental noise levels and these will need to be referenced/described in the Guidelines.

The proponent needs to determine representative baseline and project-related noise levels. The Guidelines will need to provide some guidance regarding the number of hours/days needed for monitoring and whether seasonal differences and differences between the weekend and weekday need to be taken into account. Guidance should also be provided for

determining noise levels so that all noise-sensitive sites are satisfactorily taken into account, including residences and noise-sensitive facilities such as schools or hospitals. Part of this guidance will be whether to use noise level contours or measurement/prediction at selected sites.

2.2 Standards and Guidelines

Ideally, existing Canadian guidelines and standards could be used to determine whether adverse health impacts are significant but, for noise, this is complicated by differences between various Canadian guidelines/standards. Important differences also exist between Canadian, international and U.S. standards and guidelines.

One important issue that needs to be addressed is whether, as the baseline noise level increases, health impacts become more severe for a fixed change in noise level. In some Canadian guidelines, the change in sound level is limited to a fixed value, regardless of baseline noise level. At the other extreme, some guidelines have sound level limits.

The "competition" between existing standards and guidelines also applies to noise sources such as: (i) gas compressor stations, where low frequency noise can be easily transmitted into residences, potentially causing vibration and rattling, and (ii) wind turbines.

In Canadian and U.S. guidelines, industrial, energy, highway, aircraft and rail noise sources are all treated differently. The Guidelines will need to indicate, with a rationale, which standard(s) and/or guideline(s) should be applied, depending on the noise source.

There are no guidelines for the important issue of determining the significance of noise impacts on cultural and ceremonial activities of First Nations people [2].

2.3 Dose-Response Relationships

Given the challenge of applying standards and guidelines to determine the significance of health impacts of noise, dose-response relationships need to be used to judge the severity and likelihood of effects. Various dose-response relationships for environmental noise have been established for speech interference, sleep disturbance and annoyance. Thresholds for associations with: (i) impaired reading comprehension in children and (ii) cardiovascular disease in adults have been suggested but causal relationships have not been demonstrated.

An international standard for environmental noise assessment [4] provides a dose-response relationship for the percentage highly annoyed as a function of rating level, for transportation and industrial noise sources in a typical community. The rating level arises from adjustments to the DNL to account for specific situations. The adjustments are: +10 dB in quiet rural settings, +12 dB for highly impulsive noise such as rail yard shunting, +3 to +6 dB for aircraft noise, 0 dB for road traffic and industrial noise and -3 to -6 dB for electric trains.

"Competing" dose-response relationships for construction noise have been used in EAs based on: (i) the ISO annoyance relationship and (ii) U.S. guidelines for a qualitative complaints response to noise [6]. The challenge is that effects are temporary but may be relatively severe, particularly at night from pile driving and backup alarms.

Critical analysis of available dose response relationships will be needed to decide on their use in the Guidelines.

2.4 Mitigation

To some extent, the Guidelines will also have to address mitigation measures. CEAA emphasizes community consultation and this is known to be a non-acoustical factor that can help mitigate annoyance with noise. With regard to acoustical factors, practices in other countries include controversial issues such as compensation of exposed residents for sound proofing within residences. These practices must be evaluated in a Canadian context.

3. CONCLUSIONS

National Guidelines are needed with criteria, methodologies and rationales for determining, as required by CEAA, whether project-related noise is likely to cause significant adverse health effects. A balanced working group must develop the document because of the complexity and broad range of issues. There must be wide ranging consultations with major Federal, Provincial, Territorial, industry and community stakeholders. The guidance document should also contain a checklist so that proponents would know the basic information needed for an environmental noise assessment, suggested criteria levels to be targeted and guidance for mitigation. This format would also serve to help make the environmental assessment more easily understood by all stakeholders.

REFERENCES

[1] Canadian Environmental Assessment Agency (2003). Reference Guide: Determining Whether A Project is Likely to Cause Significant Adverse Environmental Effects.
[2] Health Canada (1999). Canadian Handbook on Health Impact Assessment.
[3] Federal-Provincial Advisory Committee on Environmental and Occupational Health (1989). National Guidelines for Environmental Noise Control. Health Canada.
[4] ISO (2003). ISO 1996-1:2003. Description, measurement and assessment of environmental noise – Part 1: Basic quantities and assessment procedures.
[5] Transport Canada (1996) TP 1247E - Land Use in the Vicinity of Airports, Part IV – Aircraft Noise.
[6] U.S. EPA (1974) Information on Levels of Environmental Noise Requisite to Protect Public Health and Welfare with an Adequate Margin of Safety.

THE CURRENT STATE OF NOISE REQUIREMENTS FOR FEDERAL ENVIRONMENTAL ASSESSMENT PURPOSES

R. L. Scott Penton[1], Peter Vandelden[2], and David Chadder[3]
Rowan Williams Davies & Irwin Inc., 650 Woodlawn Road West, Guelph, Ontario, N1K 1B8
[1]slp@rwdi.com, [2]pv@rwdi.com, [3]dsc@rwdi.com

1. INTRODUCTION

The "National Guidelines for Environmental Noise Control" are intended to establish a common national basis for noise impact assessment. However, there are several provincial noise guidelines in place in Canada, which widely vary in their requirements. As a result, the type of information made available in Environmental Assessments varies widely from project to project.

Historically, noise is given very little emphasis in the federal environmental assessment process. Of the over 4000 projects which fell under CEAA review from January 2003 to August 2004, (ranging from joint panels to screening submissions), only eight projects listed noise as keyword in the Canadian Environmental Assessment Registry (CEAR) (www.ceaa-acee.gc.ca/050/index_e.cfm).

2. PROVINCIAL GUIDELINES

Each province and territory has different levels of noise requirements, ranging from none at all to requirements for highly detailed assessments. Table 1 summarizes the provincial and territorial noise guidelines for road transportation noise. Quebec guidelines are in flux, and will be discussed in detail in the paper "Noise impact assessment procedures in Quebec", presented as part of this session.

Table 1. Guidelines for Road Transportation Noise

Prov./Terr.	Comments	Descriptors
British Columbia	Ministry of Transportation, 55 dBA objective, mitigation varies depending on pre- and post-project noise levels, with 65 dBA threshold	L_{eq} (24 hr)
Ontario	MTO QST-A1, MOE/MTO Joint Protocol. 55 dBA objective, mitigation based on 5 dB change	L_{eq} (24 hr) L_{eq} (16 hr) L_{eq} (8 hr)

Of the provinces/territories reviewed, only BC and Ontario have set provincial noise guidelines for road transportation noise. Impacts within other provinces are either addressed on a case-by-case basis, are dealt with through municipal bylaws, or are not evaluated.

Table 2 presents provincial and territorial guidelines for "stationary" fixed facilities (e.g., industrial plants, etc.).

Table 2. Noise Guidelines for Fixed Facilities

Prov./Terr.	Comments	Descriptors
Alberta	Energy sector only, under Alta EUB ID-99-08 guidelines.	L_{eq} (15 hr) L_{eq} (9 hr)
Manitoba	"Guidelines for Sound Pollution" set by Manitoba Conservation. Maximum desirable and acceptable levels are set out.	L_{dn} L_{eq} (1hr)
Nova Scotia	"Guidelines for Environmental Noise Measurement and Assessment"	L_{eq} (1 hr)
Northwest Territories	Draft energy sector guidelines, adopting Alta. EUB approaches	L_{eq} (15 hr) L_{eq} (9 hr)
Ontario	"NPC" series of noise guidelines, enforced by Ministry of the Environment	L_{eq} (1 hr)

Prince Edward Island and New Brunswick list noise as an air contaminant, but do not require noise impacts to be assessed prior to construction. In the presence of complaints, these jurisdictions generally refer to the Ontario noise guidelines for guidance.

Alberta regulates noise only from energy-related facilities (e.g., oil and gas plants, pipelines, power plants, etc.) through the Energy Utilities Board (EUB). Other industrial land uses are not regulated. A draft Air Quality Code of Practice for the Northwest Territories (NWT) adopts the EUB approach. The EUB guidelines are unique, in that they consider cumulative noise impacts from all energy related facilities in an area, rather than impacts on a facility-by-facility basis.

Table 3 presents the various fixed facility limits

Table 3. Fixed Facility Noise Guideline Limits

Prov./Terr.	Comment
Alberta and NWT	• Cumulative impacts of all energy-related facilities, receptor based • Permissible Sound Level is ambient level + 5 dB + other adjustments as applicable • In remote areas, suggested limit of 40 dBA at 1.5 km distance, where no noise sensitive receptors exist
Manitoba	• Facility under review only, receptor based • Maximum acceptable level in residential area of L_{dn} of 60 dBA, L_{eq} (1 hr) of 60 dBA daytime, 50 dBA night-time • Desirable levels are 5 dB lower than maximum acceptable.
Nova Scotia	• Facility under review only, receptor based • 65 dBA daytime, 60 dBA evening, 55 dBA night-time
Ontario	• Facility under review only, receptor based • NPC-205 Urban and Semi-Rural S must not exceed greater of ambient or 50 dBA daytime, 47/45 evening, 45 night-time • NPC-232 Rural S must not exceed greater of ambient or 45 dBA daytime, 40 dBA evening and night-time

Remote and/or rural limits are generally consistent (~40 dBA during the night, 45 dBA during the day). The guidelines are all receptor based, rather than property line or distance from activity. However, the Alberta EUB guidelines recognize that even in the absence of permanent residences/receptors, uncontrolled noise mitigation should not occur. A suggested limit of 40 dBA at a 1.5 km distance is provided.

3. ISSUES WITH EA REVIEW PROCESS

A review of EAs conducted in the Northwest Territories are illustrative of some of the variation in noise assessments, and the resulting difficulties in determining environmental impacts.

There have been four recent EA studies conducted in the NWT, including:

- the Mackenzie Gas Project (pipeline)
- the Devon Beaufort Sea exploratory drilling (oil)
- the Ekati Diamond Project (mining)
- the Diavik Diamond Project (mining)

The first two listed, Devon Beaufort and the Mackenzie Gas Project, have conducted extensive environmental noise impact assessments as part of their comprehensive studies. These assessments included predictions of off-site noise levels from facility operations, construction, and infrastructure (e.g., roadways, aircraft). In comparison, detailed noise assessments were not conducted for either the Ekati mine and Diavik mine projects. Instead, only vague motherhood statements concerning potential noise impacts were provided by the proponents.

A similar diamond mine project in northern Ontario, the DeBeers Victor Mine, is currently undergoing a comprehensive study review. The DeBeers mine site is in a similar remote rural environment as the Ekati and Diavik sites. There are similar concerns with respect to the potential for adverse environmental noise impacts on wildlife and traditional activities of native peoples. However, a detailed noise impact assessment report was required for this facility, using Ontario NPC-232 guidelines.

4. INFORMATION REQUIREMENTS

When required, environmental noise impact assessments generally examine the following activities and provide the following information:

Construction, Decommissioning

- Identification of numbers and types of equipment, duration, etc. Identification of construction code of practice to reduce potential noise impacts.

Operations

- Identification of receptors of concern, existing ambient noise levels, predicted levels at receptors, noise contours, required noise mitigation measures.

5. DISCUSSION

As seen in the above examples, the variations in provincial / territorial noise guidelines means that projects with similar potential to cause adverse effects are dealt with in widely different manners depending on the location and type of facility. Of particular concern are noise impacts within remote rural areas, particularly in the far north. Land uses in the north are significantly different than in the south, and strict receptor-based approach may not be adequate to address potential impacts. We advocate the adoption of a common review process and national standard for assessing noise impacts in remote areas, similar to the Alberta EUB approach of:

- 45 dBA at 1.5 km for daytime L_{eq} (15 hr), and
- 40 dBA at 1.5 km for night-time L_{eq} (9 hr).

The guideline would apply to the facility under review only, and would be applicable to all fixed industrial activities, not just energy sector work, We feel that this would adequately limit the adverse "noise footprint" of facilities, while not requiring undue amounts of noise mitigation by project proponents, and would help to "level the playing field" in the EA review process.

ENVIRONMENT NOISE IMPACT ASSESSMENT IN QUEBEC

Allard, Jean-Luc
SNC-Lavalin Environment Inc., 2271 Fernand-Lafoantaine Blvd., Longueuil,Quebec, Canada J4G 2R7
Jeanluc.Allard@snclavalin.com

In Québec, there are currently no regulations on noise control. This, however, does not impede the occurrence of legal sanctions against excessive nois, which is considered an environmental contaminant. The Loi sur la qualité de l'environnement does prohibit any kind of environmental pollution.

Moreover, authorization certificates have to be issued by the Ministre de l'Environnement du Québec. To obtain such certificates, a project deemed to have a noise impact must submit a noise impact assessment proving their acceptability according to the applicable noise criteria.

In Quebec, more specifically, there are regulations and by-laws that set noise limitations, i.e.:

- Regulation respecting pits and quarries;
- Regulation respecting hot-mix asphalt plants;
- An Instruction Note applicable to stationary sources;
- An Instruction applicable to all road projects developed by the ministère des Transports du Québec);
- Criteria and methodologies applicable to construction activities;
- Over forty municipal by-laws on noise nuisances or prescribing general or specific noise limits for noise sources such as heat pumps.

Generally, the Instruction Note du ministère de l'Environnement and the By-law by of the ministère des Transports are the most commonly used.

The Note d'instruction 98-01 on community noise issued in 1998 by the ministère de l'Environnement du Québec sets up the maximum noise level from stationary sources according to the land use, as :

« The maximum noise level allowed from stationary sources shall be lower at any time and at any noise receiving point, than the highest of the following levels:

1. *Maximum permitted noise levels according to the zoning category*

Zoning	***Night (dBA)***	***Day (dBA)***
I	*40*	*45*
II	*45*	*50*
III	*50*	*55*
IV	*70*	*70*

ZONING CATEGORIES

Sensitive zones

I. Land intended for single-family detached or semi-detached dwellings, schools, hospitals and other educational, health or extended care facilities. Land used as existing dwelling in an agricultural zone.

II. Land intended for multiple dwelling units, mobile parks, institutions or camping sites.

III. Land intended for commercial purposes or recreational parks. However, the night noise level established shall only apply within the limits of residential sites property. In any other cases, the maximum day noise level shall apply overnight.

Non-sensitive zones

IV. Land zoned for industrial or agricultural use. However, for the land used as dwelling in an industrial zone and built up in compliance with the municipal by-laws in force at the construction time, the criteria are 50 dBA at night and 55 dBA during the day.

2. *Noise level equal to the ambient noise level measured at the same point during the shut-down of the company operations." (Our translation. Original in French)@*

In the early 1980s, the Service de l'environnement du ministère des Transports established a noise impact assessment grid for road noise. Their directive is used in case of complaints, modifications or implementation of a new road corridor in order to assess the noise impact projected according to the current noise environment. Considering the widespread and intensity of projected impacts, mitigation measures are put into place so that any residual impacts to which residents are exposed are

acceptable. Generally, level Leq_{24h} =55 dBA and below are acceptable. A corrective approach should be considered when the external noise level is equal or above Leq_{24h} 65 dBA.

The Direction des politiques du secteur industriel is currently considering the revision of the Instruction Note 98-01 and the noise generated by construction sites for any project submitted to environmental impact assessment procedure.

Among the orientations under current consideration are the following:

- The Selection of the sensitive receptor(s) that are the most exposed to noise;
- Noise measurement conditions;
- The Methodology of ambient noise measurement for the sector;
- Objectives of noise levels on construction sites in residential areas:
  - $Leq_{12\ h}$ = ambient noise or at 55 dBA minimum between 7 a.m. and 7 p.m.
  - Leq_{h} = ambient noise or at 45 dBA minimum between 7 p.m. and 7 a.m.;
- Surveillance and monitoring program including any justification of noise excess occurred;
- Noise level objectives for ambient noise:
  - Increased protection for specific sensitive noise-saturated environments;
  - LDEN index use to represent any nuisances on a 24 hour basis (Day-Evening-Night);
- Noise assessment method: ambient noise and noise from an identified source considering any corrective actions (impact noise, tone component or low frequency, etc.).

Laws on environmental noise have barely changed over the past 20 years even if the needs have evolved and if the population seems more demanding than ever.

Revisions to applicable standards are already on their way in a number of countries and we can only wish that they will consider technology progress on measurement instruments and the improved noise control techniques while keeping regulation applicability easy.

Technical knowledge in acoustics has evolved over the past 20 years, environmental, economic and social integration will also be a challenge.

HOW URBAN HUM BECAME THE NOISE LIMIT IN ONTARIO

Tim Kelsall[1]
[1]Hatch, 2800 Speakman Drive, Mississauga, ON L5K 2R7 tkelsall@hatch.ca

Since 1978 MOE has used the equivalent sound level of urban hum, that amalgam of distant traffic and other sounds which forms the background noise in most urban and suburban areas, as the primary limit for assessing the noise impact from industry and other sound sources. This means that the limit to be used at a specific location is based on the sound actually received at that location, rather than an arbitrary number. This has made sense to industry, residents and consultants, who have used this limit for over 25 years. It has also been used in a Federal-Provincial guideline and many practical limits are found to be based indirectly on urban hum. In this paper, the origins of this limit are explored and some of the issues which were examined at that time and later are discussed. How urban hum behaves, how widespread it is and how it is used as a limit are also reviewed.

1. INTRODUCTION (12PT FONT)

In 1978, the Ontario Ministry of the Environment published their Model Municipal Noise Control Bylaw[1]. This document included both a qualitative (no measurement required) and a quantitative noise bylaw and included NPC documents giving the procedures required for measurement and assessment under the quantitative bylaw. With a few updates to the procedures, this publication has formed the guideline used throughout the province for controlling noise. These guidelines have also been adopted and used in large part within the Federal Provincial Guidelines on Noise[2]published by the Federal Government and provinces.

The heart of the guidelines is NPC 205, which is an update of the original NPC105. This document states:

"the sound level limit expressed in terms of the One Hour Equivalent Sound Level (L_{eq}) is the background One Hour Equivalent Sound Level (L_{eq}) typically caused by road traffic".

This is subject to a minimum limit where there is little traffic noise as follows:

"No restrictions apply to a stationary source resulting in a One Hour Equivalent Sound Level (Leq) or a Logarithmic Mean Impulse Sound Level (LLM) lower than the minimum values for that time period specified in Table 205-1.

TABLE 205-1
Minimum Values of One Hour Leq or LLM by Time of Day
One Hour Leq (dBA) or LLM (dBAI)

Time of Day	Class 1 Area	Class 2 Area
0700 - 1900	50	50
1900 - 2300	47	45
2300 - 0700	45	45

This table is the main difference between NPC 205 and the original NPC105, where the limits were 50, 45, 40 for Day, Evening and Night in all urban areas.

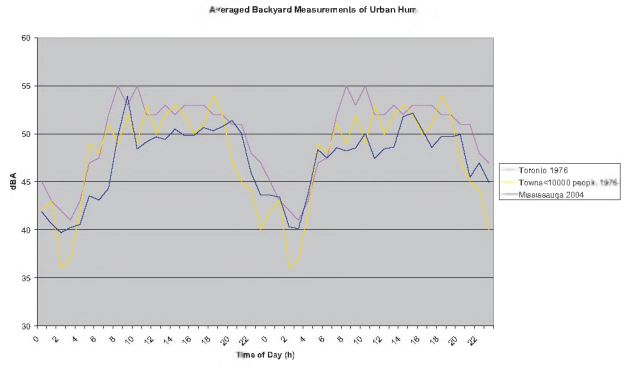
In any case, subject to lower limits, the bylaw restricts noise from industrial, commercial and some residential sources to produce an equivalent sound level no higher than the background equivalent sound level due to road traffic, i.e. the background in the absence of the noise source under investigation.

The original draft of the bylaw had not used this approach. Instead, it had set sound level limits by time of day and area. This was a common approach to regulating noise at the time and is still widely used. The idea was that the municipality would define these areas. The difficulty was that there was no established method for doing so and in any case it would end up being an ongoing job as the municipality grew and changed.

A study was carried out to examine how sound levels varies across communities.[3,4] It quickly came to the conclusion that almost all urban communities of any size were dominated almost exclusively by road traffic noise. This could be predicted successfully from traffic volumes, mix and speed on nearby roads. This has now advanced sufficiently that the EU uses the principle to map entire cities. It also found that away from major roads, in backyards and other sheltered areas, sound levels tended to minimum values which were consistent across communities. These sound levels, which were typical of urban areas, were given the name "urban hum".

Figure 1 shows the results of averaged community noise measurements from 1976 carried out at that time in areas dominated by urban hum. They show the typical near plateau during the day and drop of 10 dB at night found in all urban areas. It also includes some more recent measurements in a community in Mississauga which are

very similar and indicate that urban hum is a stable limit, i.e. similar values are found in most urban areas and the measurements are reproducible. The author has measured similar values in the Caribbean and in a town in South Africa. Even in areas dominated by specific roads, the same 10 dB variation with time of day was observed. The values are simply elevated.



It was clear in 1976 that the sound level limits or L_{DN} used in most legislation was simply a crude attempt to codify the variation seen in traffic noise in urban areas.[5] Almost all community noise regulations included a 10 dB drop at night, in some form or other.

Given that it was necessary to measure the offending source anyway, the extra time required to measure traffic noise at the same location or nearby for a similar length of time. For sources which could not be turned off the option of using a traffic noise prediction or typical values for urban hum, whichever was higher, was available.

The one change between NPC105 and NPC 205 is that the former used a lower limit based on a linear regression to measured urban hum values while the latter uses more arbitrary figures. NPC 205 also requires identifying an area as Class 1, 2. (There is a Class 3 for rural areas). This usually does not cause an issue because they both have the same lower limits at night and during the day and are only minimally apart in the evening.

In the late 70's (and even now) the concensus of many studies of noise descriptors measured against community reaction to noise showed L_{eq} to be as good or even a better description of people's reaction to noise than other descriptors. Even when a specific class of noise is measured using a descriptor designed for it, e.g. NEF for aircraft noise, L_{eq} proves nearly as good a descriptor. By 1975 MOE had determined that 1h L_{eq} was the best way to assess noise sources and it made eminent sense to compare it directly to the 1h L_{eq} of traffic noise.

This procedure, with some minor changes has been in place since 1978. Generally it has proved satisfactory to the public, to consultants and to regulators. The only difficulty is that while it is generally easy to prove an excess above the limit by showing an increase in L_{eq} when the source is operating, it is more difficult to prove that a source is in compliance by comparing two measurements and a variety of different approaches have been used over the years to get around this difficulty. However, the criterion has proven to be sufficiently successful to still be in use today. While some members of the public would consider any audible noise to be offensive, most understand a limit where industry muut not be louder than other noise sources in the area, when measured the same way.

REFERENCES

1. Model Municipal Noise Control Bylaw, Ontario Ministry of the Environment, 1978

2. National Guidelines for Environmental Noise Control, ISBN: 0-662617014-8, 1989

3. T. Kelsall and O. Friedman, "Development of a Prediction of the Equivalent Sound Level Due to Traffic Noise in Residential Areas", presented at Acoustical Society of America Meeting, December, 1977.

4. T. Kelsall, "Measurements & Prediction of Sound Levels in Quiet Urban Areas - Urban Hum", Acoustics & Noise Control in Canada, June, 1979.

5. T. Kelsall, "Use of Transportation Noise as a Standard for the Assessment of Other Noise Sources in Ontario", presented at Canadian Acoustical Association, 1977.

APPLICATION OF ENVIRONMENTAL NOISE GUIDELINES – TWO CASE STUDIES

Ramani Ramakrishnan
Associate Professor, Dept. of Architectural Science, Ryerson University
350 Victoria St., Toronto, Ontario, Canada, M5B 2K3, rramakri@ryerson.ca

1. INTRODUCTION

Many provincial agencies in Canada provide a set of procedures to use when preparing noise impact assessments for proposed and/or existing industrial operations. The application of these guidelines is more of an art than science in many instances. Ontario will be used as a test case for the application of the guidelines contained in the Ministry of the Environment (MOE) publications [1, 2, 3, 4]. Two case studies will be used to highlight the difficulties of interpreting and applying the guidelines to obtain Certificate of Approval (CoA) for industrial operations. The details of these case studies will be presented in this paper.

2. REGULATORY GUIDELINES

Noise, a pollutant under provincial law, must satisfy emission limits as per the guidelines of the Environmental Protection Act of the Province of Ontario. The noise limits are more stringent if located in a rural setting [3]. For plants in an urban setting, the ambient noise is as per Reference 4. The noise limit to be satisfied by the plants is 45 dBA in an urban setting and 40 dBA if located in a rural setting. These levels are expressed as one-hour energy averaged sound level, L_{eq}, in dBA.

2.1 Regulatory Procedures

The Province of Ontario requires all new industries to obtain a Certificate of Approval (CoA) for their operations. It also requires existing industries to obtain or upgrade a CoA to keep their operations current. The industry must show that it is in compliance with the limits of environmental pollutants to obtain the CoA. The noise compliance procedures are outlined in References 1 thru' 4. A brief outline of the procedures is presented below.

- Establish the existing ambient sound levels at the property boundary or at the receptor locations. The ambient sound is the combined noise level of the road traffic and any other plant that is not under investigation;
- Establish the noise limits to be satisfied by the plant under study;
- Evaluate the noise levels of the plant noise sources, either through predictive analysis from the sound power of the sources if the plant is a new development or from measurements (both near field and far field) if the plant is existing and operating during the study;
- Include penalties for source character [1, 2];
- Evaluate the noise impact by comparing the receptor noise levels to the ambient sound levels;
- Design suitable noise control measures if the there is excess;
- Prepare a noise assessment report in a format acceptable to the Ministry of the Environment.

The application of the above procedures is highlighted through the following two case studies.

3. CASE STUDY 1

The first case study deals with an existing plant that requires a comprehensive CoA. Noise impact assessment is part of the approval process. The general lay out of the manufacturing plant is shown in Figure 1. The plant is adjacent to a major freeway with substantial truck traffic. The plant is surrounded by other industries on the other three sides. However, a residential dwelling is located across a local road from the plant and the home is about 70 m from the plant. Most of the in-plant noise levels are well shielded by the building envelope itself. The only possible noise sources are roof top exhausts from plant equipment. The traffic noise dominates the noise environment in this area. As per the guidelines, the planes of the 2nd storey windows at Locations 1 , 2 and 3 during the night time are the receptor locations, where the noise impact must be evaluated.

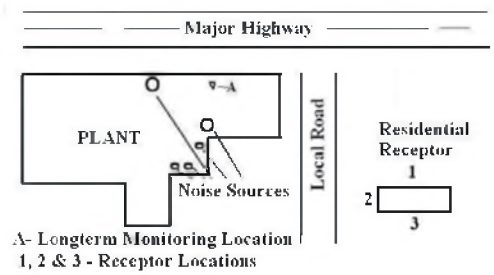


Figure 1. General Layout of the Plant and the Receptor.

The receptor Side 1 is completely dominated by road traffic. The shielded side (Side 3) not only shields the road traffic, but also shields the dominant rooftop exhausts. The receptor Side 2 is also dominated by road traffic and the plant noise is just audible. How does one determine the plant noise levels in such a complex situation when the highway noise is the dominant source and how does one apply the regulatory procedures? The following observations aid in the conclusions and procedures used for this noise assessment:

- Since the plant is located in a city, the urban regulations apply [4];
- Night time limit apply since the plant operates 24 hrs;
- Worst-case scenario conditions are at Location 3, where the traffic noise is well-shielded;
- Reference 2 allows the measurement of 20-minute Leq as a representative sample; MOE also requires measurements over a 48-hr period.

A simplistic method, not necessarily wrong, would have been to obtain a 20-minute Leq at Location 3 when the plant is shut down for a few hours during the earling Sunday morning to represent the ambient sound limit. The noise levels from plant sources is then measured within the plant property, shielded from the highway noise, adjust for distance correction to Location 3. The impact is then assessed. However, the assessment is not representative of realistic conditions. Instead, the following was used: A longterm monitor was set-up on the roof of the plant at Location A; One-hour Leq levels in dBA were measured over one week; The plant was shut down for four of the seven days; the traffic count along the highway was obtained and the noise levels at Locations A and 1 were predicted using methods provided by MOE; the average of the monitored levels agreed well with the predicted results; Five of the roof-top exhausts, as shown in Figure 1, were identified as the most dominat sources; the noise levels from these five sources were not audible at all at Location 3; And hence Location 2 (plane of the second storey window) was chosen for the assessment; Near-field measurements of these five sources were used to predict the noise levels at Locatoion 2. The results at Location 2 are : Plant noise – 56.6 dBA; Highway Noise – 64.7 dBA. It is seen that even if one applied a 5 dB penalty, the plant noise is within the guideline limits.

4. CASE STUDY 2

The second case study also deals with an existing plant that requires a comprehensive CoA. The general lay out of the manufacturing plant is shown in Figure 2. The plant is adjacent to a major freeway with substantial truck traffic. The plant is surrounded by other industries on two of the other three sides. There are a number of single family residences located across a local road from the plant. These houses are located 400 m from the plant. Most of the in-plant noise levels are well shielded by the building envelope itself.

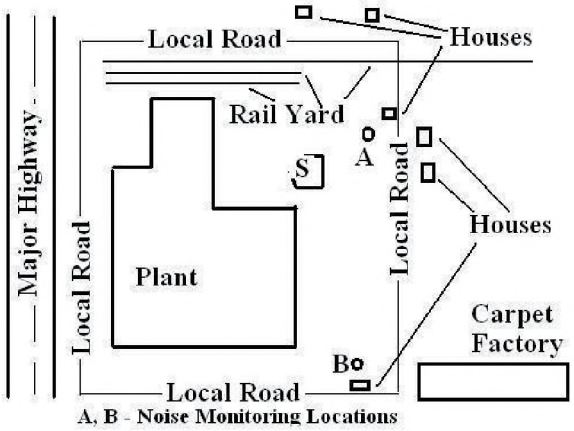


Figure 2. Layout of the Plant and the receptor locations.

The above assessment is simple. A brief outline of the applied procedures and results are:

The houses are more than 1000 m from the highway and would be considered a rural setting [3]. However, because there are medium to large plants surround the main plant, urban setting guidelines are applicable. Long-term monitoring results were obtained from Locations A and B. The night time limits are 50 dBA for house near Location B and 45 dBA for houses near Location A. The plant noise was barely audible at these houses except for two sources: the rail yard activities and metal dumping noise for Source S. (Figure 2). The scrap handling is just audible at night time and the assessment showed that with a tonal penalty, the scrap handling exceeded the guideline by 6 dB. In addition, depending on the load, the locomotive noise, whistling within the yard and revving-up within the yard, exceeded the guideline by more than 10 dB, even though these operations lasted less than a few minutes in each hour. The plant is currently designing an enclosed scrap handling facility. The yard activities are being monitored so as to design an activity plan so that the locomotive sounds would be within the guideline limits.

REFERENCES

1. Model Municipal Noise Control By-Law, Ministry of the Environment, Ontario, Canada, 1978.
2. NPC – 233 "Information to be Submitted for Approval of Stationary Sources of Sound," Ministry of the Environment, Ontario, Canada, October 1995.
3. NPC – 232 "Sound Level Limits for Stationary Sources in Class 3 Areas (Rural)," Ministry of the Environment, Ontario, Canada, October 1995.
4. NPC – 205 "Sound Level Limits for Stationary Sources in Class 1 & 2 Areas (Urban)," Ministry of the Environment, Ontario, Canada, October 1995.

DEVELOPING NOISE CONTROL LEGISLATION (SILENCING THE CRITICS)

Anita Lewis and David DeGagne
Alberta Energy & Utilities Board, 640 – 5Avenue SW., Calgary, Alberta, Canada, T2P 3G4

1. INTRODUCTION

Environmental noise has been defined as unwanted sound that is annoying, distracting, or physically harmful. The World Health Organization believes that environmental noise can have serious affects on people such as; interfering with daily activities at school, work, home, and during leisure activities and even affect health defined as *"a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity"* [1]. Today, environmental noise remains an important issue in Alberta. With increasing rural population and the growth of industrial development, it is essential to have effective industrial noise control regulations in place to minimize the impacts on the environment.

2. BACKGROUND

As an emerging issue crossing all industry sectors, in 1974 the Environment Council of Alberta (the Council) took on the task of examining noise legislation that may be appropriate for Alberta and made recommendations to the Minister of Environment. The Council appointed a committee of experts, many from the ranks of the Canadian Acoustical Association, to assess the issue. Working independently, the expert committee produced a two-volume report in 1979 titled *"Noise in the Human Environment"*. The committee was able to identify a number of areas to address and made suggestions to the Council on how they should deal with these concerns. A key intent of the expert committee's report was to spark public reaction. Although the Council did not necessarily support all the views of the committee, it did believe that the points raised were important and should be considered by the people of Alberta. Ultimately, the Council held a series of public hearings throughout the province to obtain reaction to the subject of noise pollution. Special attention was given to noise sources and problems, the effects of exposure to noise on health and the human environment, and technological and other practices that may be adopted to control noise levels and resolve problems.

The comprehensive findings included 43 recommendations and contained information on the impacts of noise (auditory effects and economic effects), noise sources (transportation, work-related, domestic, recreational), and creation of quiet charter (right to quiet, health, education and research programs, and engineering controls). These findings were presented to the government of Alberta. Three of the recommendations related to the energy industry and suggestions for revisions to the EUB Noise Control Directive (Directive) were made. The Directive at that time was a one-page document that required energy industry facilities meet 65 dBA daytime or 55 dBA nighttime sound pressure levels 15 metres from nearby residences. The recommendations were:

1. That noise levels be measured at the property line of an energy development,
2. That permissible sound levels be lowered for permanent facilities, and
3. That, in rural areas, a noise standard of 5 dBA Leq (24) above ambient noise levels be adopted.

3. REVISING REQUIREMENTS

Based on the recommendations in the Council's report, the EUB began to re-evaluate its Noise Control Directive (Directive) taking into consideration that measuring and controlling environmental noise was going to take a much more rigorous approach than the earlier versions. The EUB believed that to be truly effective the new regulatory requirements would have to contain several important elements including:

- Statutory mandate and authority: The EUB, as a regulatory body, needed to have in place the legislative authority to create noise guidelines. This was determined to exist under the Oil and Gas Conservation Regulations, Hydro and Electric Energy Act, and the Pipeline Regulations.
- Clearly defined goals: The goals of the new guidelines were to provide clarity and certainty to affected parties including consequences for non-compliance, to have easily understood technical components, provide a consistent cost effective approach, contain a clear process for its implementation and identification of remedial action(s), and use of best practical technologies to minimize the impacts of industrial noise.
- Acceptance by industry: To assure acceptance of the guidelines the involvement of stakeholders in regulatory development would be necessary. Also there needed to be consistency with similar requirements in other jurisdictions. Finally a mandatory review process would be necessary to incorporate any improvements in technology and new understandings of industrial noise.

- Technical integrity: The Directive must be consistent with accepted acoustical standards, have specifications for instrumentation, site selection for instrumentation, specific acceptable sound levels and appropriate methodologies for the analysis of monitoring results.
- Enforceable: The Directive would have clearly defined compliance targets, appropriate consequences for non-compliance and reasonable expectation and timelines for noise mitigation programs as agreed to by all parties.

To ensure these elements would be addressed appropriately, a diverse task force was established comprising of academics, acoustical engineering consultants, the energy industry, provincial government, members of the public and EUB representatives to develop a comprehensive noise control guide. The objective was to develop an effective guide for industrial facilities taking into consideration the human psychological response to environmental noise and the many technical challenges that were faced by the energy industry.

After many months of dedicated work, the task force presented its recommendations to the EUB. The EUB, in turn, adopted the task force's recommendations and published its first comprehensive noise control regulation, *Interim Directive (ID) 88-1* and accompanying *User Guide 38*. The new policy provided a consistent and fair process to ensure noise impacts were considered in the design of a facility. It also attempted to take a balanced viewpoint by considering the interests of both the nearby residences and the licensee of the facility.

4. FUTURE OF EUB NOISE CONTROL DIRECTIVE

Since the release of *ID 88-1*, a great deal has been learned about industrial noise, its complexities and management resulting in several revisions. Coincidentally, as required by the periodic mandatory review process, the current version of the Noise Control Directive (*ID 99-08*) is undergoing an update by a multi-stakeholder review committee. The following are some of the key areas that are being considered for the next version of the Directive and Guide:

- Mandatory use of complaint investigation forms – In order to determine the nature of the noise concern, the use of the complaint investigation form is required to capture representative conditions when noise from an energy facility is a nuisance. The investigation form may convey important information such as the characteristic of the noise and weather conditions that may be important during a survey.
- Recognition of low frequency noise (LFN) – It has been determined that low frequency noise may exist in certain complaint situations where the comprehensive sound level is satisfactory but the concern is a dominant or resonant low frequency resulting in a high degree of annoyance. The new Directive outlines how the presence of LFN is to be determined and what corresponding adjustments must be made to the comprehensive sound level to determine compliance with EUB requirements.
- Consideration for wind turbines – Wind turbines posed an interesting challenge with regard to the potential for noise from the turbine and blades. Existing requirements for noise modeling and noise surveys used for energy facilities were inadequate for wind turbines. In hopes of better regulating the noise emitted by wind turbines, the review committee is developing modeling parameters and noise survey guidelines to be incorporated into the Directive.
- Standardized criteria for modeling – Differences occur in predicted noise levels depending on which noise propagation algorithm is used in modeling. As a way of providing more consistency for modeling results, the Directive will include a list of parameters that the model must incorporate and input conditions that must be used in determining predicted noise levels at the receptors.
- Process for consecutive monitored nights – To ensure representative conditions have occurred, multiple nights of monitoring may be a solution where there is uncertainty regarding what representative conditions might be prior to monitoring or where the licensee and residences have agreed prior to the survey. In cases where 2 or 3 nights are deemed to be representative of noise complaint conditions, the worst-case condition is used to determine compliance with EUB requirements.

5. CONCLUSION

As the Noise Control Directive evolves, it continues to serve industry, the public and the EUB as a useful tool to manage environmental noise. The involvement of a multi-stakeholder committee in the development of the Directive and a balanced viewpoint that considers both industry and residents is the basis for the effectiveness and acceptance of the Directive as a fair regulatory process. A logical next step in controlling industrial noise is to make it universal throughout the province. This would require meeting the elements noted above by the appropriate authority.

REFERENCES

[1] Environment Council of Alberta (1982). Public Hearings on Noise in Alberta.

[2] DeGagne, D. C. (1999). The Evolution of Environmental Noise Legislation for Alberta's Energy Industry Over Three Decades. Journal of the Canadian Acoustical Association, vol27, 76-77.

[3] Noise Control Directive ID 99-08 (1999), Alberta Energy and Utilities Board

ON ENVIRONMENTAL NOISE IMPACT ASSESSMENTS: AN ALBERTA PERSPECTIVE

Corjan Buma, M.Sc., P.Eng.
aci Acoustical Consultants Inc., #107, 9920 – 63 Avenue, Edmonton, AB, T6E 0G9
corjanb @ aciacoustical.com

1. INTRODUCTION

"Is it realistic, let alone workable, to expend (possibly-great) effort in compiling a set of guidelines on environmental noise impact assessments, applicable on a national scale, in Canada ? The population density ranges from one person per hundreds of kilometres to thousands per one square kilometre, and the range on human experience and expectations is no less divergent." The challenge is acknowledged, but consider the following …

2. THE RECEPTOR

100+ years' experience with quantifying human response to auditory stimuli has yielded some very consistent and specific results. For example, (a) a 40-dBA, non-tonal ambient sound in a residential yard is acceptable to most people (b) a 30-dBA broadband ambient but with prominent 60-Hz tone would likely be deemed "annoying" by most (c) 50-STC is statistically acceptable to most people as suite-to-suite airborne sound isolation in a multi-unit building and (d) it's "noise" when I haven't been invited to the party. The Fletcher-Munson curves are an early example and all modern academic research an on-going affirmation of the high degree of uniformity in human response to noise. Acknowledging that there are always exceptions for any number of reasons, there is high consistency in how humans respond to noise.

This underlies virtually all noise legislation/guidelines/ safety-codes worldwide to date. The EU has drafted its Noise-Control Directive and mandated that all Europe be noise-mapped by 2007. The WHO has extensive documentation casting noise as a health effect. Even the National Building Code of Canada requires a **rated** STC-50 between adjacent suites in a multi-unit building and a **rated** STC-55 between suite and higher-noise area. Similarly, CMHC has set an L_{eq}-24hr of 55 dBA as its criterion in outdoor residential amenity space.

3. IN THE CROSS-HAIRS: "ALBERTA"

Considering the experience with environmental noise in Alberta provides good examples of chaos and success.

In Alberta the energy sector is the only industry where specific legislation has been compiled on a provincial basis governing environmental noise. Other over-riding provincial legislation permits regional jurisdictions to draft noise legislation as they see fit and even defers to such legislation, especially if this more stringent. As to this type of community noise legislation: many towns and municipalities have gotten as specific as "do-not-annoy-your-neighbor" legislation, while the legislation of some urban areas actually requires meeting certain sound level metrics. Over against that, a late-90's re-write of one City's "Noise Bylaw" was declared by a Court-of-Law as being in-sufficiently specific, while in another, a meeting in early-2004 between Consultants and City-Administrative staff disclosed a profound resistance to updating the Noise Bylaw from its current 1970's formulation. (In the latter example, it was indicated that the current formulation is strongly preferred because it allows the City-representative conducting a measurement the discretion to declare a result in compliance or in violation of the Noise Bylaw.)

One example involving urban traffic noise legislation: if the Administration of an urbanized area decides to institute "trigger" legislation, typically, the trigger sound level is set so high (e.g. 24-hour L_{eq} sound level of 65 dBA) that action need not be taken "in our lifetime" (this particular example being for introducing noise-attenuation devices along existing roadways adjacent to existing communities).

As a few industry noise examples: a sawmill, because it is not involved in generating or processing "energy", is not governed by the provincial (energy-) noise directive. However, when that sawmill decides to add a co-generation unit, the new co-gen unit is required to comply with the energy-noise directive. The energy-noise directive was early on, by collective decision, chosen to be a "receptor-based" directive. Thus, on the assumption of a residence at some distance from an energy-facility, the resultant sound level at the residence due to the facility may not exceed a context-dependent specific value (typical default of 40 dBA, night-time L_{eq}). Further, in the absence of residences the facility owner must design for 40 dBA at 1.5 km. However, if an agricultural representative (i.e. farmer) decides, because of good weather, to continue harvesting operations within 200 meters of a village boundary such that the resultant L_{eq}-Night within the first row of homes is in the speech-interference range (60+ dBA), there is no provincial

noise legislation governing such activity (whether the regional do-not-annoy legislation applies across a boundary would necessitate involving legal advice). As one final example, in an area with multiple types of heavy industry, compliance with noise legislation is required of a refinery but not of an equal-or-larger chemical plant to which some not-yet-processed energy by-products are pipe-lined.

The intent thus far in this section has been to underscore a few very real examples of existing discrepancies in noise legislation. Interestingly, Alberta's energy industry has quite cooperatively participated in establishing and promoting this legislation, in spite of there being an "un-level playing field". True, there are enough instances where resolution over noise between industry and land-owner has been difficult to achieve. Sometimes resolution has been achieved only by adjudication by Provincial representatives or in the form of a land-owner choosing to move away in frustration. However, in the majority of cases the energy industry has introduced engineered noise control to achieve compliance. Experience has shown that of all the "tweaks" introduced in updating the Directive since 1988, the Permissible Sound Level grid has not been adjusted (suggesting this grid assesses typical human response quite adequately). Further, Consultants routinely apply the methodology of the energy-based noise Directive as a design tool in all manner of non-energy application. Those administering the energy-noise Directive report being frequently contacted as to whether it can be enforced in other areas. Alberta's energy noise directive is a good example of a legislative noise management tool that has achieved a high degree of success in balancing the needs of one industry with residents' quality-of-life.

4. TOWARDS NATIONAL ENIA-GUIDELINES

Given that we human receptors generally respond to noise in the same way, and given the example of a successful legislative noise management tool, it is considered that there is merit in establishing a national base-line by means of environmental noise impact assessment guidelines. Some deemed advantages are: (a) nation-wide consistency (b) more-level playing field (all industries treated alike) (c) function as a target in setting basic sound levels (similar to aspects of National Building Code) and methodology to achieve them (d) assist towards a better-informed and more-equally informed public (e) provide a tool for other levels of government to set their own context-specific criteria (f) …

In developing such a set of Guidelines it is suggested to consider, among others, at least the following items (no prioritization implied). (1) Whether environmental noise is (implied to be) an issue of quality-of-life vs. health; while these are related the latter would require a more conservative approach. (2) How much latitude to allow for differences of context (provincial vs. regional; rural vs. urban; pristine vs. more-developed) and receptor sensitivity (do the Guidelines assume only "typical human hearing"? do they allow for differences in human hearing ? do they reference noise impacts on wildlife and/or domestic animals?). (3) Is pre- or post-commissioning noise monitoring suggested, recommended or required ? (4) What tolerance on instrumentation (defer only to commonly-accepted standards ? frequency of re-certification of equipment ? re-certification to what degree ? on-site calibration ?) ? (5) If on-site noise monitoring is conducted, what are the limits on atmospheric conditions to obtain valid data ? are weather-data a required part of a noise monitoring ? is a different approach required in winter vs. in summer ? (6) Operation vs. construction phases of a project. (7) Fixed-asset vs. associated-traffic noise levels permitted. (8) To what degree do the Guidelines constitute "legislation" ? (9) What guidance is there to a proponent on absence-of-complaint developments vs. a specified complaint situation ? (10) To whom does a noise-affected person complain ? (11) Is there a grading of type-of-assessment: do some applications require only a broadband analysis, some a logged-spectral analysis and some a spectral analysis with audio-recording ? (12) How do the Guidelines account for low frequency noise ("LFN"), both as to the reduced human sensitivity for and reduced tolerance of LFN ? (13) Why do we want/need a set of specifically-Canadian Guidelines: why not just adopt a WHO, ISO or American Guideline ? (14) When is computer-modeling suggested, recommended or required ? which software is acceptable ? (15) Are there certain circumstances that require "grand-fathering" ? (16) Do the Guidelines reference associated non-noise social issues ? (experience repeatedly shows that noise is often used as an easy front for other issues). (17) Establish clearly what noise metrics are used and why. etc.

REFERENCES

Alberta Energy and Utilities Board (1999). "ID99-8 Noise Control Directive" and "Guide G38 User Guide" (available at www . eub . gov . ca)

ACKNOWLEDGEMENTS

The author offers a general acknowledgement to all those who retained his services for environmental-noise and other studies over the years which provided the experience-base for the discussion above and specific acknowledgement to those on the EUB Committee currently reviewing the Noise Control Directive for sharing their insights and experience.

ASSESSMENT AND MITIGATION OF COMMUNITY NOISE IMPACTS FROM MAJOR HIGHWAY PROJECTS – A PROPONENT'S APPROACH

C.W. Wakefield, M.A. Sc., P.Eng[1], M.J. Kent, P.Geo.[2],
[1]Wakefield Acoustics Ltd., #301-2250 Oak Bay Avenue Victoria, B.C. V8R 1G5, nonoise@shaw.ca
[2]B.C. Ministry of Transportation, 4B-940 Blanshard Street, Victoria, B.C. V8W 3E6

1. THE B.C. MoT NOISE POLICY

In 1989 the B.C. Ministry of Transportation (MoT) first adopted a policy to address traffic noise impacts at residences and schools associated with projects involving new or substantially-upgraded Provincial highways. Strictly speaking, the policy applies only to freeway and expressway projects. However, it has been used on some MoT projects that do not fully qualify as controlled access highways. In addition, the policy has been used as a guideline on projects sponsored by the BC Transportation Financing Authority (BCTFA) and the Greater Vancouver Transportation Authority (TransLink).

The MoT policy for noise impacts at residences contains a mitigation eligibility scale which permits larger project related noise increases where pre-project (or baseline) noise levels are lower. In its initial form, this policy included a "hard" upper limit of $L_{eq}(24)$ 65 dBA so that if, in the design year, the overall community noise level due to highway traffic and all other ongoing sources equaled or exceeded 65 dBA, mitigation would be considered. In 1993, the policy (Ref. 1) was revised and the 65 dBA limit was replaced with a maximum 3 dBA increase criteria. This change was made principally because the MoT policy was intended to focus on project-related noise impacts rather than function as a "retrofit policy"- that is, one which attempts to correct for historically high traffic noise levels. Further, it had been recognized that such a hard limit could require mitigation even for minor highway improvements (such as slight widenings or realignments, or traffic diversions) that would result in only very small noise impacts. This approach was consistent with that in use in Australia at the time and that which was adopted in 1995 by the U.S Department of Transportation (Ref. 2) in regards to transit noise.

Figure 2.1 provides a graphical representation of the revised MoT noise policy. The horizontal axis shows the pre-project noise environment expressed in terms of the $L_{eq}(24)$ and usually established through measurement. The vertical axis shows the predicted post-project noise environment (10-years after project completion), established either through a "baseline adjustment" approach or modeling using either manual or computer methods depending on the project scale, complexity and design/assessment stage.

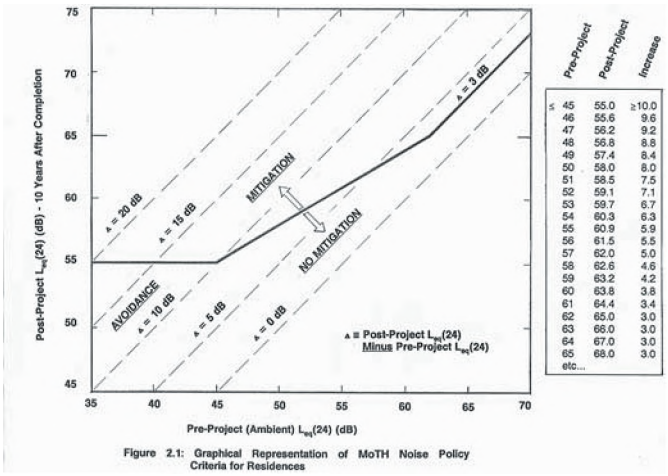


Figure 2.1: Graphical Representation of MoTH Noise Policy Criteria for Residences

2. MoT POLICY RATIONALE

Figure 2.1 indicates that mitigation will not be considered where the combined post-project noise environment does not reach 55 dBA. This minimum mitigation threshold was included for two reasons. Firstly, agencies such as the Canada Mortgage and Housing Corporation (CMHC) and the U.S. Environmental Protection Agency (EPA) consider an average noise environment of 55 dBA or less to be fully compatible with residential land use. Secondly, it is generally not possible to effectively mitigate highway traffic noise at these levels since they tend only to be reached at quite large distances (100 m or more) from highways where traditional roadside noise barriers do not perform well. Where pre-project noise levels are very low (less than 45 dBA), as, for example, where a new highway is planned in a rural area (such as the Inland Island Highway developed on Vancouver Island in the 1990's), large increases in average noise levels are therefore possible. However, to the extent possible, these impacts are "avoided" through route selection and the use of natural land forms and forested areas as noise buffers.

At post-project levels between 55 and 65 dBA, the allowable growth in noise due to a project decreases steadily from 10 dBA (where pre-project levels is 45 dBA) to 3 dBA where pre-project levels are 62 dBA or more. This sliding scale was adopt to reflect the philosophy that it would be inequitable for the Ministry, in proving necessary highway improvements, to expose those living in areas with high pre-

project noise levels to noise increases of the same magnitude as those living in quieter areas. This approach is consistent with many studies which have shown that the effects of intrusive noise on speech intelligibility, sleep quality and annoyance do not grow steadily with the level of the noise - rather they begin to increase more and more rapidly as noise levels exceed a threshold of about 60 dBA.

3. APPLICATION OF MoT POLICY

It is acknowledged that both the measurement of baseline community noise levels and the prediction of future levels are not exact procedures. As a result, the mitigation criteria of Figure 2.1 are not interpreted rigidly. Rather, mitigation is considered, and carried out where it can be shown to be effective (i.e., achieve at least a 5 dBA average noise reduction), economically feasible (benchmark cost per household of $15,000 in 1993 dollars) and widely accepted by directly affected residents, wherever the appropriate criterion is "approached or exceeded". In practice, for purposes of environmental impact assessments and during the early stages of project design, mitigation is considered to be "potentially" warranted (subject to confirmation during detailed design) whenever predicted post-project levels are within 1.5 dBA of the relevant criterion.

Mitigation measures may include the traditional roadside noise barriers (walls), earth berms or berm-wall combinations. Depending on the situation, other less obvious approaches such as "quiet pavements" or the limitation of posted speeds, may be applied, either as "impact avoidance" measures (i.e. to prevent mitigation criteria from being exceeded) or as mitigation measures - either on their own or in combination with noise barriers.

Ultimately, the decision to mitigate is made by MoT project management upon weighing many factors including cost-effectiveness, impacts on other aspects of the project, community expectations and, in the case of B.C. Environmental Assessment Office and/or Canadian Environmental Assessment Agency reviewed projects, input from relevant federal and provincial agencies.

4. EFFECTIVENESS OF MoT POLICY

It may be argued that the effectiveness of a highway noise impact mitigation policy such as that developed by the B.C. MoT may best be judged by its track record in providing mitigation where significant project-related impacts have been forecast while limiting mitigation costs to what is affordable within the project budget. Since the revised MoT policy was adopted in 1993 there have been several major highway projects to which it has been applied. In general, it has been found that where growing traffic demands have warranted the construction of a entirely new controlled access highway or the undertaking of major highway improvements, the policy has resulted in mitigation measures being both warranted and carried out. With new highways, the potential for substantial noise increases is obvious and policy mitigation criteria have been frequently exceeded. Whether active mitigation (barriers) occurs, tends to depend on the density of populated areas along the new highway alignment and their proximity to it. With highway upgrading projects, it has been found that noise increases sufficient to warrant mitigation are generally forecast. These increases result from a combination of widening, increased traffic volumes over time (10 to 15 years) and, in some cases, increased average vehicle speed due to the freeing of congestion associated with insufficient highway capacity. Some examples are cited below.

4.1 Completed Highway Projects

- *Inland Island Highway* (new highway) - mitigation generally not warranted as numbers of residences were small and setbacks large,
- *Nanaimo Bypass* (new highway) – extensive use of noise walls since in places route passed close to existing communities on outskirts of Nanaimo,
- *Victoria Approaches* (upgrade) – highway widening and development of three new interchanges on TCH north of Victoria resulted in extensive use of barriers,
- *Westview Interchange* (upgrade) – last stoplight on TCH in Greater Vancouver area replaced with diamond interchange - resulted in extensive use of barriers,
- *Duke Point Highway* (new highway) –access to B.C. Ferry terminal located near heavy industrial park south of Nanaimo. Quiet pavement (OGA) and earth berms used to mitigate impacts at scattered rural residences.

4.2 Planned Highway Projects

Several major highway projects are now being planned to address increasing traffic congestion and improve U.S. border access for commercial traffic within the Greater Vancouver area. These include the South Fraser Perimeter Road (primarily a truck route) tying in with ferry and border crossing links, three "Border Infrastructure Program" projects involving Highways 10, 15 and 91/91A, the widening of the TCH between Surrey and Vancouver including the twinning of the Port Mann Bridge over the Fraser River, a New Fraser River Crossing between Surrey/Langley and Pitt Meadows/Maple Ridge. In addition, Highway 99 (Sea to Sky Highway) between Vancouver and Whistler is to be upgraded prior to the 2010 Winter Olympics. It is expected that all of these projects will include some noise impact mitigation measures based on assessment under the MoT noise policy.

REFERENCES

1. "Revised Policy for Mitigating the Effects of Traffic Noise from Freeways and Expressways", B.C. Ministry of Transportation & Highways, Highway Environment Branch, November 1993.
2. "Transit Noise and Vibration Impact Assessment", U.S. Dept. of Transportation, Federal Transit Admin., April 1995.

Environmental Assessments of Rail Projects - the Noise Factor

Bill Aird
Canadian Transportation Agency, Ottawa ON, K1A0N9 bill.aird@cta-otc.gc.ca

1. ENVIRONMENTAL ASSESSMENT

Under the *Canadian Environmental Assessment Act* (CEAA), a responsible authority must identify all potential adverse effects of the construction and operation of a proposed project and determine whether those effects are significant. Projects with insignificant effects are permitted to proceed while those with significant effects may require further study or may be prevented from proceeding at all.

2. RAIL INFRASTRUCTURE

There are two main types of rail infrastructure projects under the *Canada Transportation Act* (CTA) which require an (EA) - rail crossings and the construction and operation of new railway lines.

3. NOISE COMPLAINTS

The Canadian Transportation Agency (Agency) has received a variety of noise complaints over the past decade. Those have focussed on: idling diesels (steady drumming), car shunting (impulse), whistling (steady) and pass-by (steady). The Federal Court has ruled that the Agency has no authority, under it present legislation, to deal with noise complaints. Amendments to the CTA are being contemplated to given the Agency the authority to deal with noise complaints.

4. RAIL NOISE

The sensible noise level generated by idling locomotives is a combination of the noise of the diesel engines and the drumming effect of the multiple internal combustion engines running at the same time. A single idling diesel can generate as much as 93 dBA (at source) while with 2 diesels the noise rises to 101 dBA and with 16 diesels it can be 110 dBA.

Car shunting can involve noise as high as 100 dBAI while a train whistle must be 110 dBA, at source. Pass-by trains tend to generate 95 to 100 dBA of noise.

Idling locomotives are the predominant type of complaint received. The drumming contributes significantly to the annoyance. The worst case encountered by the author involved the storage of 32 idling locomotives over a Christmas season. In that case the predicted noise level was over 108 dBA in the complainant's yard. In most cases, the noise levels range from 80 to 90 dBA.

5. EFFECTS OF RAIL NOISE

Many rail yards are located within 100 m of sensitive areas. In some cases, a railway may have built or re-activated a yard near homes while in many cases, residential areas have encroached on rail yards. In other cases, a rail yard which was built over 1 km from homes, now has homes almost up to the yard fence.

Rail noise can interfere with sleep, awaken residents, cause fatigue, irritability and general stress. Impulse noises are novel, unheralded, or unexpectedly loud and thus can startle people or awaken them.

When examining the effects of rail noise one must look at the source type, the location of the source relative to the receiver, the time of day and the noise level at the receiver. Table 1 summarizes the possible effects of various rail noise sources.

Table 1. Effects of Rail Noise

Noise Source	Time of Day	Effect
idling	day	annoyance
	night	sleep prevention
shunting	day	startle, stress
	night	startle awakening
pass-by	day	minimal
	night	minimal

6. NOISE ASSESSMENT

When conducting an EA, one must determine whether an adverse effect is potential and whether it will be significant. Rail noise can definitely be an adverse effect and in some cases it can be significant.

Present noise standards fall into one of two types: emission or receiving. Table 2 summarizes the United States Environmental Protection Agency USEPA emission standards for railway operations:

Table 2 USEPA RAIL NOISE STANDARDS

Source	Level (dBA at 32 m)
idling locomotive	70 - 93
moving locomotive	90 - 96
idling switcher	70 - 87
moving switcher	90
coupling (impulse at receiver property line)	92 dBAI

Source rail noise levels are normally quoted at 32 m from the source. According to the laws of physics, noise tends to decay 6 dB per doubling of distance over bare ground or 3 dB per doubling over vegetated or rough surfaces. Given the attenuation of noise and the proximity of homes to rail yards, it is not surprising that rail operations can cause noise levels in the range of 60 to 110 dBA at nearby residences.

Health officials, on the other hand, consider the noise levels at the receiver. For residential areas, the generally accepted noise standards are: 55 dBA at night (2300 to 0700) and 45 or 50 dBA daytime (0700 to 2300).

If one uses the USEPA levels as typical rail noise source levels and seeks to achieve the residential noise standards, one quickly sees that rail operations should be separated 500 m to 1 km away from residential areas. If such an approach were to be applied in Canada, either the noise from many rail operations would have to be seriously curtailed or many residential areas should never have been built near rail operations in the first place. Thus neither the USEPA nor residential standards provide a workable model for assessing rail noise.

Regulators have turned to another approach when assessing airport noise. They examine the possibility of annoyance as reflected in the likelihood of complaints:

Noise Level (dBA)	**Annoyance**
less than 55	minimal
55-60	sporadic complaints
60-65	increasing complaints
65 plus	widespread complaints and increasing likelihood of legal action

When assessing the noise factor for a proposed rail project, one should consider the following information:

- land-use including sensitive areas (residential, hospitals, nursing homes, schools and day-cares)
- typical historical noise levels in sensitive areas
- predicted noise levels in sensitive areas during the construction and operation of the proposed project
- prediction of per centage of population in sensitive areas that will experience noise levels in excess of 65 dBA (DNL)

7. DISCUSSION

No single noise standard exists which can be used to effectively determine the significance of potential noise from rail operations.

- A predicted DNL above 65 dBA indicates that widespread complaints or legal action should be expected
- An increase os 10 dB above existing noise levels would be noticeable and could lead to complaints
- As the per centage of the population that will experience a DNL above 65 dB increases, one would expect increased complaints
- The EA community needs guidance to make noise assessment more easily understood by practitioners, proponents and the public at large.

NOTE: the views expressed in this paper are those of the author alone and do not reflect the policy of the Canadian Transportation Agency

Aircraft Noise Information for Public Consumption

Tom Lowrey
Environment and Land Use, Aerodromes and Air Navigation, Civil Aviation Directorate, Transport Canada
Ottawa Ontario K1A 0N8 email lowreyt,@tc.gc.ca

1. INTRODUCTION

Aircraft noise information has been traditionally presented to the public using tools that are not particularly suited for public consumption. These tools are mainly used to assist land use planners in designating the use of lands located near airports to achieve a level of use that is compatible with the noise from aviation activities at the airport. The principal tool is the Noise Exposure Forecast contouring system that produces contours that, when interpreted using social response information, predict public reaction to aircraft noise.

The Provinces of Ontario, Manitoba Alberta and the Northwest Territories have recognized the importance of protecting airports from incompatible land use and have legislatively adopted the requirement to pay due concern to an airport's noise contours when designating land uses near airports. Other provinces have encouraged local planners to pay heed to noise contours, some to a greater extent than others. Accordingly, in the whole of Canada, compatible land use near airports has been a relatively successful endeavour and this enables our aviation system to operate freely. Of course there are exceptions to the rule but nationally, the country is in good shape.

These logarithmally averaged sample traffic day contours have been successful for the purpose for which they were designed. Unfortunately they have are not designed to describe individual events or illustrate the noise effects of individual aircraft operations or indicate the effects of small numbers of flights that may be the cause of public concern.

The aviation noise management discipline has recognized the public information problem and is gradually coming to grips with this deficiency by developing a series of descriptors that are more easily understood by the non-expert.

The Australian government Department of Transportation and Regional Services has provided worldwide leadership in developing better methods of communicating aircraft noise information and this paper briefly describes their system and provides an insight into the public's reaction to it.

2. DISCUSSION

Land use planning contours have often been used by airport noise offices to discuss noise impacts of individual aircraft events. These contours are not suited for this purpose. In the Canadian noise model, the Noise Exposure Forecast (NEF) system calculations included a penalty for night operations such that each operation occurring between the hours of 10:00 PM and 7:00 am are multiplied by a factor of 16.66 to account for sensitivities during nighttime hours. These operations are then included in the total summation of noise at an airport. When discussing individual noise events or particular aircraft overflights with the public and using the aforementioned contours for illustrative purposes, it becomes evident that the wrong tool is being used.

A misconception that results is that a resident whose house is located outside of a noise contour will not hear aircraft. This is further compounded, in some cases, by residents living on property that is outside of a contour who do not even expect to see aircraft over their neighbourhoods.

Selected flights from an airport may be required to follow flight paths different from that of the majority of departures from a particular runway. The noise effects of these flights may not be of sufficient magnitude to affect a contour due to the averaging provisions of the NEF program. Accordingly, relatively few flights may become the source irritation to many people and yet not be reflected in the contour. This phenomena serves to decrease the public's confidence in the airport's noise management program.

When new runways are being planned for an airport, an environmental assessment is required. Past experience has shown that the public views noise contours with suspicion. Decision makers, in many cases, are presented with noise contours and the information may conflict with what they are being told by the public.

There is a need to have all participants in the aviation noise question talking with the same knowledge and understanding. This will facilitate effective discussions and no doubt lead to better decisions in the final analysis.

3. METHOD

The majority of complaints expressed by the public are triggered by one event, a particularly noisy overflight that an individual resident finds irritating. This concern usually leads a complainant to other events causing irritation. Accordingly, to better discuss the event and the cause, a tool has been developed that is separate from the noise contours discussed above. It has the capability to deal with single events.

The process starts with the delineation and presentation of the flight paths of individual arrival or departure operations at the airport. Secondly, a threshold noise level that is significant for public reaction, usually 70 dBa is chosen. The 70 dBa single event contour for each aircraft type in use at the airport and its flight path is plotted. The 70 dBa level has been chosen because it has been determined that this is the level at which outdoor noise affects indoor speech in a standard Australian house with windows open. The results of this are then depicted pictorially on an aerial photo of the airport and its environs. Finally, the number of events that correlate with the flight path and single event contour are entered on the picture.

The result is that the public can easily see the noise level and the areas that are subject to it with the number of times it will occur on a typical day.

A second variation of this picture can include the airport's noise contour and permit a more informed explanation of the relationship between individual events and the logarithmically average noise.

Finally the information package can be tailored to meet the needs of individual airports.

4. RESULTS

The implementation of this system has allowed government officials and noise management professionals to enter into meaningful discussions with the public that lead to better understanding of all aspects of the noise problem.

The system is being publicized and has met with enthusiastic response. This enthusiasm comes from both the public that finally can understand what the noise manager is talking about and the noise management community that can develop a degree on understanding with the public.

International airports in North America and Europe have become acquainted with the system and are in the process of or have adopted all or parts of the information package.

The information package is enjoying success and the principles and facets are becoming the subject of study at the world level through the International Civil Aviation Organization.

EVENT-RELATED POTENTIAL MEASURES OF SLEEP DISTURBANCE BY OBTRUSIVE ACOUSTIC STIMULI

Kenneth Campbell [1] and Alexandra Muller-Gass[1]
[1] School of Psychology, University of Ottawa, Canada K1N 6N5. kcampbel@uottawa.ca; amgass@uottawa.ca

1. AWARENESS OF AUDITORY INPUT

1.1 Active attention.

R. Näätänen, a prominent Finnish cognitive neuroscientist, has developed an elaborate model to explain how an individual becomes conscious of an acoustic sound (Näätänen, 1992). In the waking state, this usually requires active attention to the auditory source. The brain's attentional system largely inhibits the processing and consciousness of that which is not attended. Näätänen's model does provide two routes through which otherwise unattended acoustic input can however intrude into consciousness. In both routes, highly relevant auditory stimuli may cause awareness through passive attentional mechanisms.

1.2 Passive attention.

The first route involves the "transient feature detector system". All auditory input enters this system, whether attended or not. This system is, as its name implies, short-lasting and involved in basic feature detection of the auditory stimulus. The output of this system varies with the obtrusiveness of the stimulus. If the output reaches a certain, critical level, an interrupt is sent to the central executive (responsible for the allocation of attentional resources). This will result in an automatic switch of attention away from the task at hand and toward the obtrusive auditory stimulus, allowing the individual to become conscious of it. This is what is meant by "distraction".

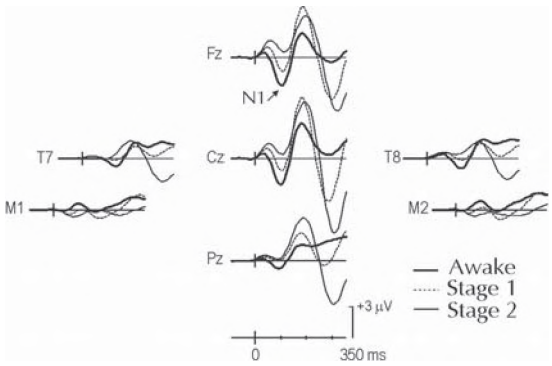
The second route involves the "change detector" system. Auditory stimuli that are not otherwise obtrusive can also result in a switch of attention, but only if they signal a change from the immediate acoustic past. It is assumed that auditory stimulus input is stored in a very brief duration echoic memory. If subsequent input matches this representation, further processing ceases and ongoing cognitive activity will not be disturbed. But, if the new input fails to match the representation, "change" is detected. If the extent of change is large enough, an interrupt will again be sent to the central executive and attention will be switched to the auditory channel.

2. EVENT-RELATED POTENTIALS

2.1 Waking auditory ERPs.

Measuring the extent of processing of unattended stimulus input presents many problems. Näätänen records the brain's minute response to the auditory stimuli, called the event-related potentials or "ERPs". ERPs allow the researcher to monitor the processing of auditory input even when the subject is not conscious that a stimulus had in fact been presented (Muller-Gass & Campbell (2002)

An auditory stimulus will elicit a series of long latency ERPs. In the figure below, a moderate intensity (80 dB SPL) brief auditory stimulus was presented every 1.5 s. This elicits a series of characteristic positive- and negative-going ERPs. Prominent among these is a negative wave, N1. In the waking state, N1 occurs at about 100 ms and is maximum over fronto-central (Fz/Cz) regions of the scalp. N1 is elicited automatically, independent of the attention and consciousness. Näätänen suggests that the amplitude of N1 varies directly with the output of the transient feature detector system. A particularly obtrusive stimulus will result in a switch of attention. The actual switching of attention is thought to be reflected by a later positive wave, often labeled the "P3a". P3a peaks from 250 to 300 ms after stimulus onset and is maximum over central areas of the scalp. Current evidence suggests that N1 is generated in the auditory cortex whereas the switching of attention associated with P3a may be a function of the executive control system of frontal lobe.



Acoustic change elicits a slightly later negativity called the Mismatch Negativity (MMN), peaking from 120 to 250 ms depending on the extent of acoustic change. The MMN is also thought to occur relatively independently of attention. A very large change in the auditory stimulus will elicit a switch of attention, observed again by the later positive wave, the P3a.

3. SLEEP ERPs

3.1 Sleep physiology

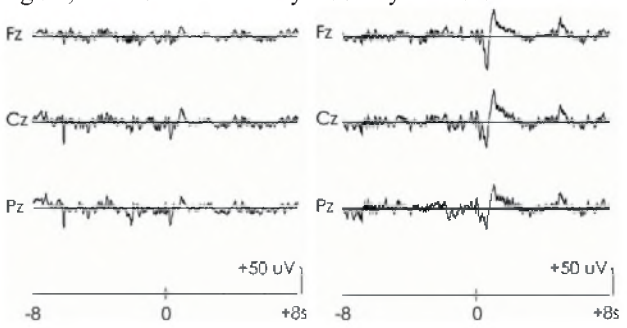
At sleep onset, consciousness (or "awareness") of the external environment and internal mental state is gradually lost. This is why sleep is called an "unconscious" state. Unlike other unconscious states, sleep can be reversed either through natural awakenings or through particularly obtrusive external stimuli. Sleep is not unique. It consists of two major sub-divisions or "stages", non-REM and REM sleep. Non-REM is further subdivided into stages, numbered 1 to 4, larger numbers corresponding to greater depth of sleep.

3.2 Loss of the waking ERPs.

In order to sleep, processing of external stimulus input must be inhibited. Nevertheless, highly relevant and obtrusive stimulus input must somehow be detected and if necessary, result in an awakening. Figure 1 indicates that activity of the transient detector system essentially ceases at sleep onset. N1 gradually declines in amplitude during the transition from a waking to a sleeping state (stage 1) and reaches baseline level during definitive stage 2 sleep (Campbell & Colrain, 2002). The MMN also gradually reaches baseline level during non-REM sleep. In the young adult, there is thus little evidence of consciousness of the auditory stimulus during non-REM sleep. During REM sleep, N1 returns to about 25% of its waking level. It is also possible to elicit P3-like ERPs (peaking from 300-400 ms) providing the stimulus is extremely obtrusive (it is loud and occurs extremely rarely, Cote et al., 2001) or if the stimulus is psychologically meaningful (perhaps the name of the subject). These late positivities, occurring from 250-400 ms following stimulus onset, may reflect a disturbance of the sleep state.

3.3 Sleep-Related ERPs: The K-Complex.

During non-REM sleep, even a highly obtrusive stimulus will usually fail to elicit the N1 or MMN potentials in young adults. It may however elicit the largest response known to occur in the human brain, the K-Complex (Bastien & Campbell, 1994). The K-Complex cannot be elicited in the waking state or in REM sleep. In the right portion of figure, a moderate intensity auditory stimulus elicits the K-Complex during non-REM sleep. The K-Complex consists of a series of components, the most prominent of which is a very large amplitude negative wave "N550"), peaking as it label suggests at about 550 ms after stimulus onset. Interestingly, identical stimulus parameters will sometimes elicit a K-Complex and sometimes will not (left-portion of figure).



In young adults, like the waking P3a, the K-Complex is most easily elicited by stimuli that occur very rarely, are very loud or very abrupt (Bastien & Campbell, 1994). There is some evidence that the subject's own name can also elicit the K-Complex. The function of the K-Complex remains disputed. It has been suggested that it may serve a protective role, protecting sleep from unwanted disturbances. It may act as a secondary inhibitory system following failure of the inhibition of either the transient or change detector systems that normally occurs in non-REM sleep. Thus, the K-Complex may be elicited by highly relevant stimuli, yet they are not so relevant that the sleeper needs to awaken. Interestingly, obtrusive stimuli that fail to elicit the K-Complex may produce sleep fragmentation. These are brief arousals that do not lead to awakenings yet affect performance on many tasks the next day.

It has been well-established that the elderly have difficulty in falling asleep and once asleep, remaining asleep. Recent studies have indicated that the inhibition of processing that is apparent in younger adults within sleep may often fail in the elderly (Crowley & Colrain, 2004). The P2 (or perhaps P3a) component of the auditory ERP increases in amplitude in the elderly during non-REM sleep. At the same time, K-Complexes are very difficult to elicit and when they are elicited, their amplitude is much reduced.

REFERENCES

Bastien C & Campbell K. (1994) Effects of rate of tone-pip stimulation on the evoked K-Complex. *J Sleep Res.*3, 65-72.

Campbell, K.B. & Colrain, I.M. (2002). Event-related potential measures of the inhibition of information processing: II. The sleep onset period. *Intl J Psychophys*, 46, 197-214.

Cote, K.A., Etienne, L., Campbell, K.B. (2001). Neurophysiological evidence for the detection of external stimuli during sleep. *Sleep* 24, 791-803.

Crowley K.E. & Colrain I.M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clin Neurophysiol.* 115, 732-44.

Muller-Gass, A. and Campbell, K.B. (2002)). I. The effects of selective attention on cognitive ERPs during the waking state. *Intl J Psychophys*, 46, 197-214.

Näätänen R., 1992. *Attention and Brain Function*, Erlbaum.

ACKNOWLEDGEMENTS

This research has been funded by Natural Science and Engineering Research Council (NSERC) of Canada and the Canadian Institutes of Health Research (CIHR).

AIRCRAFT NOISE AND SLEEP DISTURBANCE: A REVIEW OF FIELD STUDIES

David S. Michaud, Ph.D
Health Canada, Consumer and Clinical Radiation Protection Bureau, Product Safety Programme, Acoustics Division, 775 Brookfield Road, Ottawa, ON K1A 1C1 dmichaud@hc-sc.gc.ca

1. INTRODUCTION

According to the World Health Organization (WHO) (1), health should be regarded as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity." Adequate sleep is fundamental to physical and mental restoration. Thus, aircraft noise-induced sleep disturbance (AN-ISD) should be viewed as a potential health hazard, according to the WHO's definition of health.

AN-ISD includes circumstances where aircraft noise (i) interfered with the ability to fall asleep, (ii) curtailed sleep duration, (iii) decreased perceived quality of sleep, (iv) caused awakening(s) during sleep or (v) increased bodily movements during sleep (i.e. motility).

2. METHODS

Field research related to AN-ISD published in English since 1992 was identified using electronic databases, internet searches and conference proceedings. Electroencephalography, which measures electrical activity of the brain, is still considered to be the "gold standard" in sleep research; however it has some practical disadvantages that make the relative simplicity of monitoring motility and objective measures of awakenings a desirable alternative to sleep researchers. As such, most of the studies in this review assess sleep disturbance by monitoring motility and awakenings. Since available dose-response curves are based on behaviourally confirmed awakenings, the present paper focuses primarily on these results. Behaviourally confirmed awakenings are typically confirmed when the subject presses a button when awakened.

3. RESULTS

3.1 Heathrow, Gatwick, Stanstead and Manchester Airports study, United Kingdom

In 1992 a field study was conducted to assess AN-ISD in the UK (3). The focus was on the impact that aircraft noise had on awakenings and motility in 400 subjects between 20-70 years of age over 5,742 subject nights. Outdoor noise measurements were monitored to provide aircraft noise event (ANE) data. At outdoor maximum sound levels (Lmax) above 80 dBA, the probability of an awakening was 1 in 75, but below this level ANEs were unlikely to cause an awakening. Aircraft noise was deemed to minimally contribute to awakenings on average, but noise sensitive individuals were up to two times more likely to be awakened by an ANE than individuals classified as the least sensitive to noise. As reported in the US studies below, sleep in general became more disturbed as the night progressed, but this was not a result of ANEs *per se.*

3.2 U.S.A field studies on aircraft noise and sleep disturbance

Fidell et al. (4) assessed the effects of aircraft noise on behaviourally confirmed awakenings over approximately one month in 27 subjects living near a military airfield and 35 subjects living near the Los Angeles International Airport (LAX). An additional 23 subjects living without significant amounts of aircraft noise (mostly traffic) served as controls. Indoor Leq levels were recorded every 2-sec within the bedrooms of all participants and noise events (presumed to be aircraft) were logged every 0.5 sec. Across all 3 study locations, spontaneous awakenings were much more common than those resulting from an aircraft noise event (2.07 vs. 0.24/night, respectively). The mean indoor sound exposure level (SEL) of a noise event that awakened subjects was 80.6 dB, while events with an SEL of 74.1 dB failed to awaken participants. An increase of 10 dB SEL for indoor noise event levels was associated with a 1.6% increase in the percentage of people awakening. Awakenings were not affected when aircraft activity at the military airfield was reduced over weekends (from 53.5 to 47.7 Leq). As reported in the UK study, the probability of awakening was dependent upon the time since going to bed. In this study, for each 15 min since going to bed, the probability of awakening to a noise event increased by a factor of 1.06. Cumulative noise exposure over the entire night was unable to predict sleep disturbance.

Fidell et al. (1995) (5) assessed behaviourally confirmed awakenings and motility in 77 subjects over 2,717 subject nights around the Stapleton International Airport (DEN),

which was about to close and the Denver International Airport (DIA), which was scheduled to open. As a function of indoor noise events (presumably aircraft), the percentage of people awakening increased by 0.25% per 1 dB increase in indoor SEL (average threshold ~ 69 dBA). Ambient sound levels (Leq) within the bedrooms were inversely related to noise event associated sleep disturbance. Prior to opening DIA, there was an average of 1.71 behavioural awakenings per night that dropped to 1.13 per night after opening. Awakenings before and after the closing of DEN were not statistically different (1.8 vs. 1.64, respectively). The indoor SEL that caused motility ranged from 65-74dB and motility increased by 0.4-1.23% for each 1dB increase in SEL.

Using similar methodology, Fidell et al (2000) (6) measured indoor and outdoor sound levels in 12 homes (22 participants) around the DeKalb-Peachtree Airport (PDK) before, during and following the Atlanta Olympic Games. Average awakenings were higher before (1.8 per night), dropping slightly to 1.3 during and to 1.0 following the games. For each 10 dB increase in the indoor SEL there was an increase of 1.3% in the percentage of people awakened. At an outdoor SEL of 100dBA there was large variability in the prevalence of awakening (0 to 20%), which was less apparent at 60 dBA (0-2%).

3.3 Amsterdam Schiphol Airport (AMS), The Netherlands

The impact of aircraft noise from AMS on sleep disturbance has recently been reported in 418 subjects between 18-81 years of age (7). Aircraft noise was monitored within subject's bedrooms and at selected outdoor locations over an 11 day period. Behaviourally confirmed awakenings were statistically more likely to occur during an ANE compared to outside the event; but were not influenced by the indoor Lmax nor SEL of an ANE. The indoor Lmax of an ANE did predict motility. The threshold for the probability of motility was 32dBA indoor Lmax and increased as levels increased so that, at 68dBA, the probability of motility during aircraft noise was about 3 times greater than in the absence of aircraft noise. Subjects that were on average exposed to minimal ANEs at night, that yield a lower night time Leq, were more likely to respond to an ANE with an increase in motility compared to subjects exposed to higher night time Leq levels from aircraft noise. However, mean motility (averaged over 11 days) increased with higher indoor equivalent noise levels from aircraft.

4. DISCUSSION

These studies show that aircraft noise can be a significant source of sleep disturbance for a small percentage of a population exposed to nighttime aircraft noise. This may be a significant number of individuals, even though the majority of exposed people does not show measurable signs of disturbance. Thus, future research might best protect susceptible exposed individuals by elucidating the factors that contribute to the range in individual AN-ISD.

The studies reviewed here also suggest that, ideally, the AN-ISD element of environmental assessment guidelines should incorporate indoor sound levels based on single aircraft noise events. Indeed, a recent meta-analyses of field data produced a dose-response curve for aircraft noise and awakenings based on the indoor SEL:

$$\% \text{ Awakenings} = 0.58 + (4.30 * 10^{-8})\ SEL^{4.11} \quad (8)$$

Environmental assessment will undoubtedly improve if researchers are able to establish the cumulative health effects of AN-ISD and how to best assess sleep disturbance (motility, awakenings, stage change, complaint behaviour, etc..).

It is not known if the general conclusions from the studies reviewed here are in agreement with trends in complaint behaviour from people residing near airports. Guidance for health impact assessment of aircraft noise would be improved if research could identify the relationship between logged complaints and measures of sleep disturbance presented by noise researchers in peer review journals.

5. REFERENCES

1. WHO (Geneva) (1999) Guidelines for community noise.
2. Horne J., Pankhurst F., Reyner L., Hume K. and Diamond I. (1994) A field study of sleep disturbance: Effects of aircraft noise and other factors on 5,742 nights of actimetrically monitored sleep in a large subject sample. Sleep, 17: 146-159
3. Civil Aviation Authority (London) (1992) Report of a field study of aircraft noise and sleep disturbance.
4. Fidell S., Pearsons K., Tabachnick B.G., Howe R. and Silvati L. (1995) Field study of noise-induced sleep disturbance. J. Acoust. Soc. Am., 98: 1025-1033
5. NASA Langley Research Center (Hampton, VA) (1995) Noise-induced sleep disturbance in residences near two civil airports. NASA Contractor Report 198252
6. Fidell S., Pearsons K., Tabachnick B.G. and Howe R. (2000) Effects on sleep disturbance of changes in aircraft noise near three airports. J. Acoust. Soc. Am., 107: 2535-2547
7. Division Public Health (The Netherlands) (2002) Sleep disturbance and aircraft noise exposure: Exposure-effect relationships. TNO report no. 2002.027
8. Finegold LS and Elias B. (2002) A predictive model for noise induced awakenings from transportation noise sources. Published in *Proceedings of INTER-NOISE* 2002. I-INCE, Dearborn, MI.

MULTIBAND COMPRESSION AND CONTRAST-ENHANCING FREQUENCY SHAPING IN HEARING AIDS

Ian C. Bruce[1], Shahab U. Ansari, Harjeet S. Bajaj, and Kamran Mustafa
Department of Electrical and Computer Engineering, McMaster University,
1280 Main St. W., Hamilton, Ontario, Canada, L8S 4K1. [1]e-mail: ibruce@ieee.org

1. INTRODUCTION

Contrast-enhancing frequency shaping (CEFS) produces a better representation of formants in the auditory nerve (AN) response of an impaired ear than conventional amplification schemes (Miller et al., 1999; Sachs et al., 2002). CEFS compensates for the broadened tuning curves and elevated thresholds of an impaired ear by adjusting the relative amplitudes of the formants without distorting the spectral envelope between the formants. Multiband compression, on the other hand, has been utilized in hearing aids to compensate for the reduced dynamic range of the impaired ear.

We have previously shown that multiband compression and CEFS amplification can work together when used in series without counteracting one another (Bruce, 2004), in contrast to previous spectral enhancement schemes that are not compatible with multiband compression (Franck et al., 1999). In this paper we describe the combination of CEFS amplification and multiband compression in a single frequency-domain filterbank implementation, thus reducing the computational complexity and the signal delay. The CEFS gain-frequency response has also been improved to give a better neural representation of F2 and F3. This new multiband-CEFS (M-CEFS) algorithm is evaluated with models of the normal and impaired ears (Bruce et. al., 2003) and compared to linear amplification, multiband compression and CEFS without compression.

2. METHOD

As illustrated in Fig. 1, the M-CEFS algorithm utilizes a formant tracker (Mustafa and Bruce, 2004) to direct the gain-frequency response of a time-varying filter. The gain in each frequency band also depends on the short-term input signal energy in that band (to apply multiband compression) and the hearing-loss profile (to tailor the algorithm to the audiogram of a hearing aid user).

Compression was realized using a filterbank of 15 filters spaced at 1/3-octave, starting at 250 Hz. Filter bandwidths were approximately 2 equivalent rectangular bandwidths. A sampling frequency of 16 kHz was used. Details of the FFT-based implementation of the filterbank can be found in Bruce (2004).

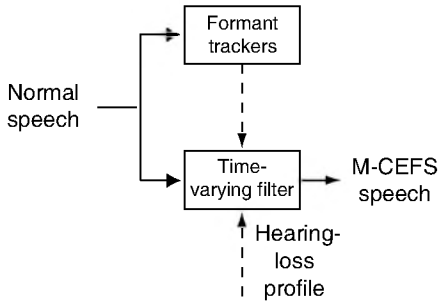


Fig. 1. Schematic of the M-CEFS amplification system.

Compression was applied independently in each frequency band of the filterbank based on the short-term input power in that band. The compression knee point was 40 dB SPL, above which a compression ratio of 2:1 was applied. The gain in each frequency band was adjusted dynamically to give a near-instantaneous attack and a release time of approximately 60 ms.

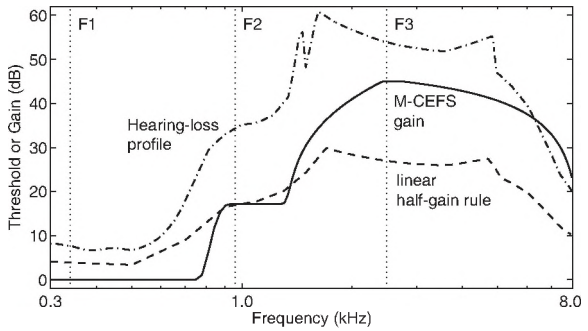


Fig. 2. Example gain-frequency response. The dot-dashed line gives the hearing-loss versus frequency profile used in the impaired auditory-periphery model. The dashed-line shows a linear gain-frequency response following the "half-gain rule." The solid line gives the M-CEFS gain-frequency response for the formant values F1, F2 and F3 indicated by the vertical dotted lines.

In addition to the gain adjustments made in each band to produce compression, M-CEFS amplification was applied using gain adjustments as a function of the current formant estimates and the hearing loss profile. An example gain-frequency profile is given in Fig. 2, where a comparison is made to a linear amplification scheme based on the "half-gain rule" (Dillon, 2001). The M-CEFS gain-frequency

response has the same gain as the linear scheme at the F2 frequency and is similar on average to the linear gain profile up to just below the F3 frequency. However, the M-CEFS gain is shaped to increase the contrast of the formants, and at F3 and above the M-CEFS gain is substantially higher than the linear gain. In contrast, the original CEFS gain-frequency response only applied contrast enhancement between F1 and F2 (Miller et al., 1999; Bruce, 2004).

The test speech signal used in this paper was the synthesized sentence 'Five women played basketball' (courtesy of R McGowan of Sensimetrics Corp, Somerville, MA). Using the auditory-periphery model of Bruce et. al. (2003), the neural representation of this sentence was evaluated via the short-term average discharge rate and the short-term synchronized rate versus time for a population of AN fibers.

3. RESULTS

Spectrograms of the original sentence, the linear-amplified sentence and the M-CEFS amplified sentence are show in Fig. 3. The linear amplification scheme applies the most gain in the F2 and F3 frequency region (~ 1–3 kHz), which helps prevent impaired model AN fibers (with elevated and broadened tuning curves) in this region from responding erroneously to F1. However, the contrast between the formants is not enhanced, and consequently the tonotopic representation of F2 and F3 are not correctly restored in the short-term average discharge rates and the synchronized rates of these impaired fibers.

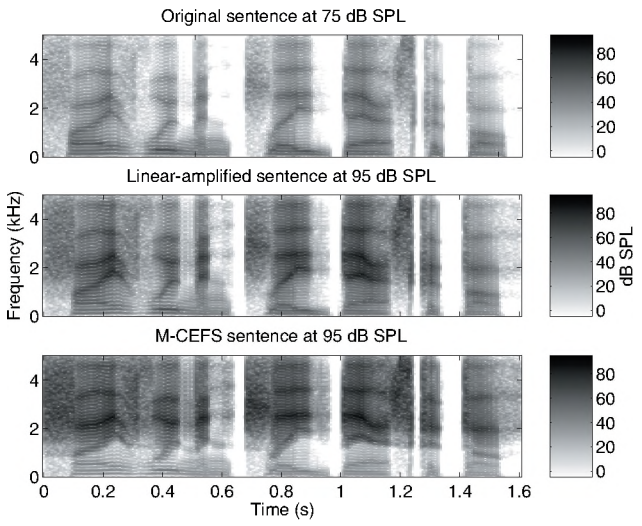


Fig. 3. Spectrograms of the original sentence (top panel), the linear-amplified sentence (middle panel) and the M-CEFS amplified sentence (bottom panel).

In comparison, the M-CEFS sentence (processed at an input level of 75 dB SPL) exhibits increased contrast between the formants. Particularly evident is the contrast at F2 as the formant tracker successfully adjusts the M-CEFS gain-frequency profile as the F2 frequency changes over time. Increased enhancement between F2 and F3 is also achieved. The resulting impaired model AN fiber responses have a more normal tonotopic representation of F2 and F3 than is produced by the linear-amplification gain-frequency response or the original CEFS scheme.

4. DISCUSSION

In this paper we have presented a hearing-aid amplification scheme that uses a formant-tracking algorithm and time-varying filter to produce improved contrast enhancement between formants. This contrast enhancement is not degraded by the application of multiband compression. The amplification scheme was incorporated into the same FFT-based filterbank implementation of the compression algorithm, thereby reducing the signal delay. The currently version has an average signal delay of 16 ms, which is somewhat larger than is desirable for a hearing aid speech processing scheme, so we will investigate if the delay can be reduced further without degrading performance. The modeling results indicate that M-CEFS should produce a better representation of formants in hearing aid users; human testing will be conducted to evaluate the actual performance.

REFERENCES

Bruce, I. C. (2004). Physiological assessment of contrast-enhancing frequency shaping and multiband compression in hearing aids. Physiol. Meas. 25, 945–956

Bruce, I. C., Sachs, M. B., and Young, E. D. (2003). An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses. J. Acoust. Soc. Am. 113, 369–388.

Dillon, H. (2001). Hearing Aids. New York: Thieme Medical Publishers.

Franck, B. A., van Kreveld-Bos, C. S., Dreschler, W. A., and Verschuure, H. (1999). Evaluation of spectral enhancement in hearing aids, combined with phonemic compression. J. Acoust. Soc. Am. 106, 1452–1464.

Miller, R. L., Calhoun, B. M., and Young, E. D. (1999). Contrast enhancement improves the representation of /ɛ/-like vowels in the hearing-impaired auditory nerve. J. Acoust. Soc. Am. 106, 2693–2708.

Mustafa, K., and Bruce, I. C. (2004). Robust formant tracking for continuous speech with speaker variability. Submitted to IEEE Trans. Speech Audio Processing.

Sachs, M. B., Bruce, I. C., Miller, R. L., and Young, E. D. (2002). Biological basis of hearing-aid design. Ann. Biomed. Eng. 30, 157–168.

ACKNOWLEDGEMENTS

The authors would like to thank Brent Edwards for supplying the code for the multiband-compression algorithm. This work was supported by NSERC Discovery Grant 261736 and the Barber-Gennum Chair Endowment.

AUTHOR NOTES

Kamran Mustafa is now with Nortel Networks, Ottawa, Ontario, Canada, K2H 8E9.

MACHINE LEARNING AND THE AUDITORY NERVE

Jeff Bondy[1], Ian C. Bruce, Sue Becker, and Simon Haykin
[1]Dept. of Electrical and Computer Engineering, McMaster University, 1280 Main St. W.,
Hamilton, Ontario, Canada, L8S 4K1. email: jeff@soma.ece.mcmaster.ca

1. INTRODUCTION

The goal to our research is to produce a hearing-aid algorithm that would enhance the intelligibility of the hearing impaired in noisy conditions. For this we have been using machine learning techniques to derive the important statistical quantities describing the differences between the normal and impaired auditory systems auditory nerves. This leads to a technique we call Neurocompensation because it is essentially trying to re-establish the neural response of the auditory system after hair cell loss.

The main advantage from using the neural coding of the auditory nerve is it is the closest physical variable after the impairment. By encompassing the impairment, theoretically, the resulting algorithms should be better because they are free from simplifying assumptions. For example, by basing hearing-aid processing design on the audiogram alone, an implied assumption is that the loss of cochlear gain is the only important variable. This does not encompass the large differences in temporal and spectral properties between the normal and impaired auditory system.

2. METHOD

To apply machine learning to the auditory system we have a four component model:

1. A model of the normal auditory system up to the auditory nerve.
2. A model of the impaired auditory system that encompasses the hearing impairment, in this case we specifically look at processing lost with hair cell damage.
3. A processing block to train, a surrogate attempting to replicate the missing processing of the damaged system.
4. An error metric that is an intelligibility predictor, based on distortions to the auditory nerve.

An acoustic input is processed by the normal model to come up with a control signal, while the processing block preprocesses the same input before being passed to the impaired block which comes up with a distorted auditory response signal. The control and distorted signal are then compared to calculate how intelligible the distorted signal is. By maximizing the intelligibility of the distorted signal by training the parameters of the preprocessing block we are really building a hearing-aid processor the restores what is important on the auditory nerve.

3. DISCUSSION

The two auditory models were provided by Bruce et al. (2003). The error metric was first suggested in Bruce et al. (2002), and then improved into its useful form in Bondy et al. (2004). The metric is largely based on deriving a neural equivalent to the Articulation Index (AI); in fact the validation directly parallels Steeneken (1992), one of the modernizers of the AI, who suggested the now widely adopted Speech Transmission Index (STI). The error metric we derived is called the Neural Articulation Index (NAI) and showed a deviation of empirical intelligibility from prediction of about 8% versus the standard AI's deviation of 10%. This error was on a nonsense syllable, rippled filtering condition test, which has historically proven very hard to predict.

The last piece of the puzzle was the processing block. Several possible blocks and their responses to hearing impairment are given in Bondy et al. (2004b). The largest success was in predicting linear hearing-aid strategies. The Neurocompensator strategy showed that the NAL-R (Dillon, 2001), a strategy very close to returning optimal intelligibility, is in fact minimizing the differences between the normal auditory nerve representation and the impaired one. We show the extremely close correlation of the NAL-R with the values calculated by the Neurocompensator approach. An example of optimizing the gain per dB of threshold shift through the Neurocompensator method is shown in Figure 1.

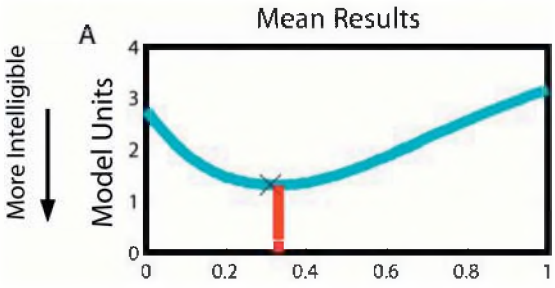


Fig. 1. The x-axis the gain in dB for every dB shift from an audiogram mimicking the NAL-R's gain ratio of 0.31. The X marks the NAL-R prescribed gain ratio, while the vertical bar is the Neurocompensator optimized value 0.32.

The Neurocompensator also has the added advantage of being able to be tailored to an individual, by maximizing the gains and shape for a specific hearing loss. It is also hoped that it could deal with the variance in intelligibility for people with similar audiograms, which can be derived from different ratios of inner and outer hair cell loss.

While the Neurocompensator did very well predicting normal, linear fitting strategies, it had problems when trying to derive the optimal non-linear parameters. We found no set of time constants, compression knees and rates or channel setups that produced a higher intelligibility prediction (Bondy et al. 2003). To try to come to terms with this we optimized the gain in different time windows of speech. The best possible values produced a scatter "function" versus input RMS; there was no connection between the input power and gain requirements, past the mean, "linear" response. Our initial trial listening tests produced quality and intelligibility deficits, subjects often complaining about odd artifacts. We decided to look closer at the differences between the normal and impaired response. In general we say the mode rate may be the same, but the shape of the activity was very different. Generally the mode discharge rate between the Normal auditory response and the impaired auditory response after NAL-R preprocessing looks similar. There are some points that the normal discharge rate is larger than the impaired response and some points when it is much smaller.

In describing the statistical differences between the normal and impaired auditory nerve responses we saw that there was a loss of contrast between different auditory landmarks. The well known lessening of the suppression effect in the impaired ear correlates neighbouring frequency bands, while the healthy cochlea produces a negative correlation. The loss of suppression reduces the peak-to-trough ratio of spectral information. Similarly, in the time domain, the loss of adaptation reduces the temporal contrast as shown in Figure 2.

The important peak-to-trough ratio of the impaired response is not as large as the normal response, it is not in the same place, cresting at different times over frequency and the adaptation response is much wider.

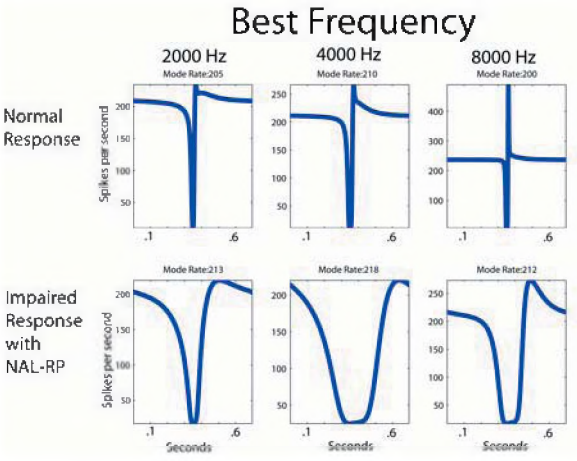


Fig. 2. Discharge rates for the normal and impaired auditory system with a 2 Hz 100% AM modulation applied to a tone at the best frequency.

That really defines an interesting formulation for the machine learning problem to address. Unlike our previous attempts which really were learning an average code and maybe not the important aspects of the AN responses, we are now looking at algorithms to reestablish suppression and onset/offset information in the impaired ear. We hope to address the fundamental question of segmentation enhancement for the hearing impaired. It is hoped that better segmentation will lead to more normal streaming, allowing the hearing-aid user the ability to unmask spectrally and temporally as well as a normal hearing person.

REFERENCES

Bondy, J., Bruce, I.C., Rong Dong, Becker, S., and Haykin, S., (2003) "Modeling intelligibility of hearing-aid compression circuits", Signals, Systems & Computers, 2003 The Thirty-Seventh Asilomar Conference on ,Volume: 1 , 720-724.

Bondy, J., Bruce, I.C., Rong Dong, Becker, S., and Haykin, S., (2004) "Predicting Speech Intelligibility from a Population of Neurons", in Advances in Neural Information Processing Systems 16, Sebastian Thrun and Lawrence Saul and Bernhard Schoelkopf eds., MIT Press.

Bondy, J., Becker, S., Bruce, I.C., Trainor, L., and Haykin, S., (2004b) "A novel signal-processing strategy for hearing-aid design: Neurocompensation", Signal Processing, **84**(7), 1239-1253.

Bruce, I.C., Bondy, J., Becker, S., and Haykin, S., (2002) "A Physiologically Based Predictor of Speech Intelligibility", International Conference on Hearing Aid Research, 2002.

Bruce, I. C., Sachs, M. B., and Young, E. D. (2003). An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses. J. Acoust. Soc. Am. 113, 369–388.

Dillon, H. (2001). Hearing Aids. New York: Thieme Medical Publishers.

Steeneken, H.J.M. (1992) "On measuring and predicting speech intelligibility", University of Amsterdam, 1992.

ONSET OF FEEDBACK HOWL IN HEARING AIDS

Michael R. Stinson and Gilles A. Daigle
National Research Council, Ottawa, Ontario, Canada, K1A 0R6, mike.stinson@nrc-cnrc.gc.ca

1. INTRODUCTION

Acoustical feedback howl in hearing aids limits the amount of gain that can be provided to a user. Feedback arises when some of the amplified sound leaks back, usually through the vent, to the microphone input port. When external conditions change, such as when a handset is brought near the pinna, there is an increased sensitivity of a hearing aid to feedback (Hellgren *et al.*, 1999). As a consequence, considerable effort is being devoted to feedback cancellation. According to the Nyquist criterion, the onset of feedback howl occurs when the overall feedback gain (i.e., the open-loop transfer function) is unity. Here, we report quantitative tests of this criterion, making using of hearing aids mounted in the pinna of a mannikin.

2. THE NYQUIST CRITERION

Consider the hearing aid shown in Figure 1. An incident sound pressure p_i is received at the microphone of the hearing aid. This signal is processed by the internal circuitry of the aid, producing an amplified signal in the ear canal. Some of this signal can leak out of the canal and get back to the microphone port, contributing a component p_{fb} to the total sound pressure $p_T = p_i + p_{fb}$ and creating a feedback condition. There are several possible feedback paths but the dominant one is usually through the hearing aid vent. The size of the feedback signal depends on the amplification settings of the hearing aid, the properties of the ear canal, and the transmission characteristics through the vent and out to the microphone. The overall gain G through the hearing aid system is given by $p_{fb} = G\, p_T$. It may then be shown that $p_T = p_i / (1\text{-}G)$. The total pressure gets very large as the gain approaches unity. According to

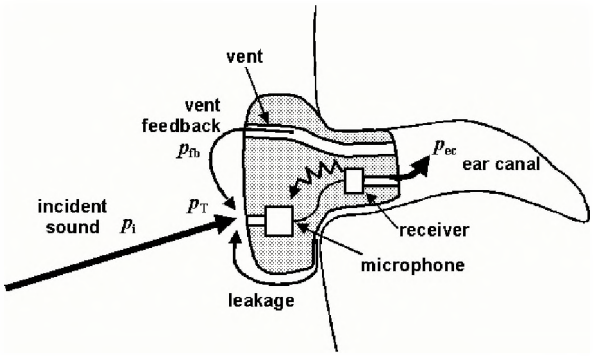


Fig. 1. Sketch of a hearing aid in an ear canal, showing several possible feedback paths.

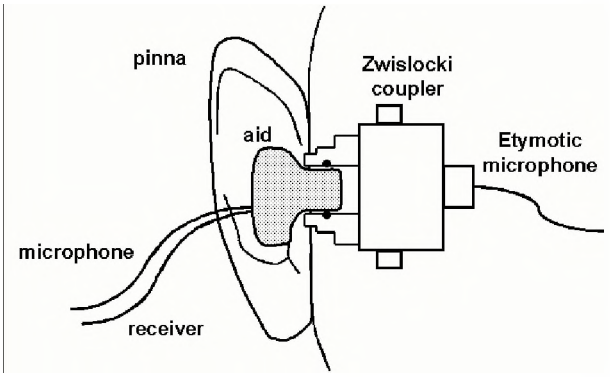


Fig. 2. Procedure for measurement of open-loop transfer function and ear canal sound pressure.

Nyquist (1932), feedback howl will occur when the magnitude of G exceeds unity and the phase of G is zero. Because G is measured in an open-loop condition, it is referred to as the *open-loop transfer function.*

3. EXPERIMENTAL PROCEDURE

Hearing aids were mounted in an artificial pinna on a KEMAR mannikin, as sketched in Figure 2. A Zwislocki coupler terminated the model ear canal and an Etymotic microphone monitored the "ear canal" sound pressure. The hearing aids have been modified, with the circuitry between microphone and receiver broken and breakout leads installed. To measure G, an electrical signal of known voltage is input into the receiver and the output from the microphone recorded; the ratio of these voltages is the open-loop transfer function (Stinson and Daigle, 2004). Shorting the two breakout leads restores normal function of the hearing aid. If conditions for feedback howl are met, the howl will be evident at the ear canal microphone. Measurements were made in an anechoic chamber. A telephone handset could be positioned near the pinna – its proximity to the pinna modified the acoustical feedback, typically increasing the open-loop transfer function at lower frequencies.

4. RESULTS

Figure 3 shows measured results for a Unitron Vista canal aid, a linear ITC aid. A telephone handset was in place just touching the pinna. The top two panels show the measured open-loop transfer function G, magnitude and phase. The magnitude is greater than unity (0 dB) between 600 Hz and 4000 Hz. Within this range, there are two zero crossings of

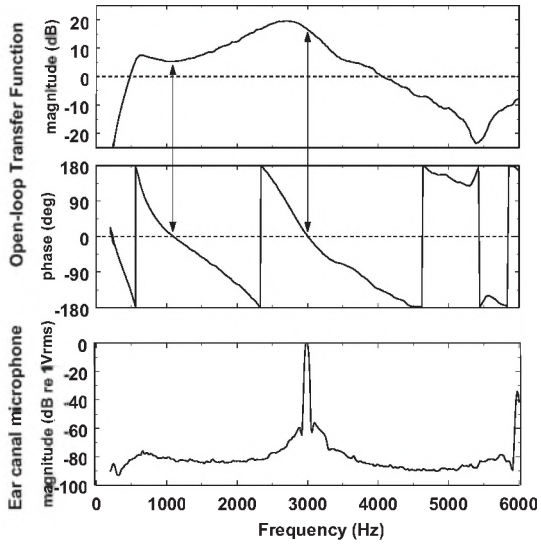


Fig. 3. Open-loop transfer function of a linear ITC hearing aid and its closed-loop response in the ear canal.

the phase component, i.e., 1100 Hz and 3000 Hz: these are the two possible frequencies of oscillation, according to the Nyquist criterion. When the breakout leads are shorted and normal operation restored, the hearing aid howls. The third panel shows the measured sound pressure inside the ear canal – the peak at 3000 Hz (the frequency having the largest open loop transfer function) is clearly evident.

The electroacoustic behaviour of this hearing aid was modified by introducing a low-pass filter (1.5 kHz corner) between the microphone and receiver components of the aid. The new open-loop transfer function is shown in the top two panels of Fig. 4. There are only two frequencies, 650 Hz and 1300 Hz, that satisfy the Nyquist criterion. In operation, the hearing aid howls at the latter of these. (Also seen are harmonics of this frequency.)

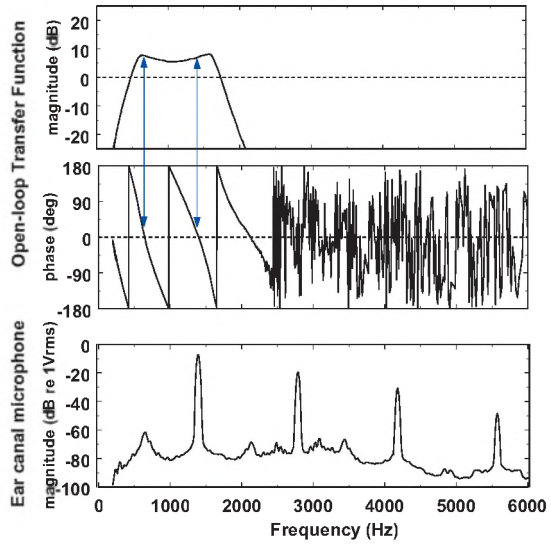


Fig. 4. Linear ITC hearing aid, with a low-pass filter inserted between microphone and receiver of hearing aid.

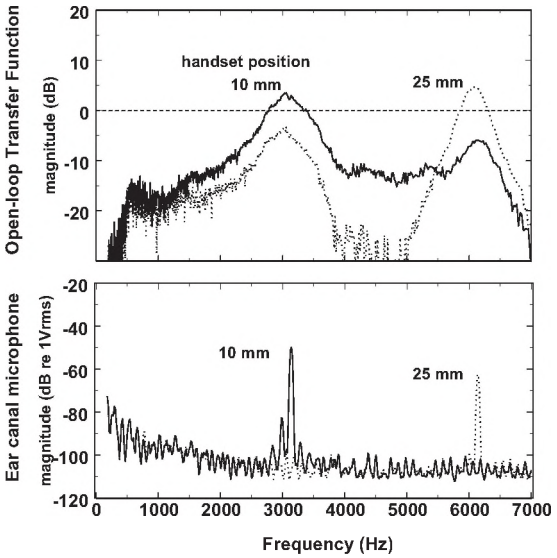


Fig. 5. Open-loop transfer function of a nonlinear ITE hearing aid and its closed-loop response, for two positions of a telephone handset.

A Unitron Conversa ITE hearing aid was also investigated. This hearing aid is inherently nonlinear as it incorporates compression and low-level expansion, so the open-loop transfer function will vary with input SPL level. There is a linear regime, nominally between 30 and 50 dB input SPL; the open-loop transfer function is constant and maximal here and measurements of it correspond to this region. For this hearing aid, the phase component of the open-loop transfer function changes rapidly, with zero-crossings occurring every 140 Hz. The magnitude component is shown in the top panel of Figure 5; the lower panel shows the closed-loop ear canal response. Two sets of results are shown corresponding to different positions of a telephone handset. For the solid curve, the handset was located 10 mm from the pinna. The open-loop transfer function exceeds unity for frequencies between 2700 Hz and 3200 Hz and feedback howl occurs at 3100 Hz. With the handset located 25 mm from the pinna, the open-loop transfer function exceeds unity between 5850 Hz and 6250 Hz and feedback howl occurs at 6100 Hz.

These tests, and others, confirm the applicability of the Nyquist criterion for predicting the onset of acoustic feedback in hearing aids.

REFERENCES

J. Hellgren, T. Lunner, and S. Arlinger, "Variations in the feedback of hearing aids", J. Acoust. Soc. Am. 106, 2821-2833 (1999).

H. Nyquist, "Regeneration theory", Bell Syst. Tech. J. 11, 126-147 (1932).

M.R. Stinson and G.A. Daigle, "Effect of handset proximity on hearing aid feedback", J. Acoust. Soc. Am. 115, 1147-1156 (2004).

REAL-TIME FEEDBACK CANCELLATION IN HEARING AIDS

Horst Arndt, Henry Luo
Unitron Hearing Ltd., 20 Beasley Drive, Kitchener, ON, N2G 4X1

1. INTRODUCTION

Feedback is a particularly debilitating problem in hearing aids. Feedback happens when the closed-loop system gain reaches values larger than unity and the phase of the feedback signal is 0^o or an integer multiple of 360^o (the Nyquist criterion). In digital systems and devices, solutions to the problem are implemented in the form of feedback reduction algorithms. [1,2,3]

The computational power in many hearing aids has been limited but cancellation methods are slowly becoming more effective. Adaptive anti-phase feedback canceling appears to be an optimal solution, although it requires significant dsp resources. The system continuously detects changes in the feedback path and, once feedback is detected, an anti-phase feedback signal is generated to cancel it. [4,5,6]

2. OVERVIEW

The fundamental issue with feedback management is the reliable and rapid detection of the onset of feedback. What is really desired is an effective method, which can detect feedback reliably and quickly with a minimum computational cost. Then a solution is required to adaptively cancel the feedback quickly while, at the same time, maintaining the input signal quality unaffected even for multiple feedback paths. By doing so, the wearer of the aids is not subjected to any annoying or upsetting feedback signals during a telephone call, or during meals and other daily activities requiring jaw movements.

In order to prevent any audible feedback reaching the hearing aid user's ear we need to detect the onset of the feedback build-up pattern and then to destroy the build-up very quickly. Most current feedback detection technologies, for example an anti-phase feedback canceller, need about 200 ms to detect a feedback path. Then it might need another 200 ms to eliminate feedback in the anti-phase feedback canceler. Therefore, a short burst of feedback is usually audible before it is canceled. This is perceived as annoying and can, at times, be uncomfortable.

Modern digital hearing aid amplifiers usually have multi-band filterbanks and feedback can occur in one or several of these bands. A new design of our digital hearing aids contains independent feedback detectors in each frequency band. These feedback detectors continuously monitor and detect feedback in real time independently and simultaneously in all individual frequency bands.

The signal strength in each band varies with time. In addition, the input signal changes its frequency components over time. An independent feedback detector analyzes the signal in each frequency band. Once feedback is detected in a frequency band, processing is applied to cancel the feedback instantaneously.

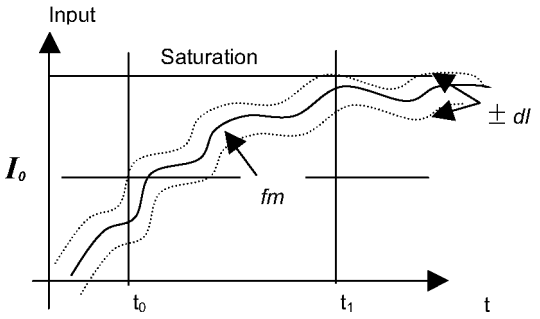
Our technique detects feedback during its build-up phase and destroys the feedback during the early build-up phase. In this way, the feedback burst duration is usually too short, or the feedback burst intensity too low, to be noticed. The method also uses multi-detectors and multi-cancellers operating independently in each frequency region in order to handle multiple feedback paths efficiently.

3. FEEDBACK DETECTION

Detection of feedback onset is a critical element of the method. Five criteria have been integrated together to rapidly detect feedback during its build-up phase. The methodology on which this detector is based is referred to as "tunnel within the frame".

The criteria that are integrated into the feedback detection module are listed here:

- Continuous and limited input signal level variation ($\pm dI$)
- Continuous modulation of input signal level (fm)
- Duration of a specific input signal ($dt=[t_1-t_0]$)
- The level of input signal (I_0)
- The difference between the current gain to the maximum deliverable gain.



Figdure 1: Feedback detection in a band

As shown in Figure 1, the input signal in a narrow frequency band has entered the narrow tunnel with a ***dI*** variation and a slow modulation frequency ***fm.*** If the signal persists over a certain duration ***dt*** and finally crosses over

the minimum threshold I_0, it is detected as a potential feedback signal in this frequency region. The typical feedback build-up time in a BTE hearing aid is about 200 ms. Feedback modulation can be affected by the compression characteristics and the actual acoustic feedback path. The final criterion integrated into our feedback detection is a comparison of the actual gain with the maximum gain (which is the maximum gain for the hearing aid set at time of fitting). Feedback is beginning to build up when the system gain is close to this maximum gain. Feedback starts when the Nyquist criterion is met. A telephone or a hand close to the hearing aid, or touching it, will cause a departure from the stable closed loop system gain that has been established at time of fitting.

4. FEEDBACK CANCELLATION

One example of such an adaptive feedback canceller is a feedback canceller with fixed feedback margin. When feedback is detected, the adaptive feedback canceler modifies the system gain around the knee-point T_1 by the fixed feedback margin shown in Fig. 2. This is done over a very narrow input signal range so that the overall dynamic input signal, such as speech and music, will not be significantly affected.

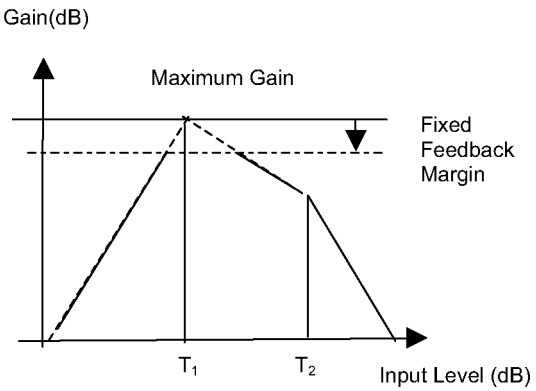


Fig. 2: Adaptive feedback canceller in a band

The output signal in the particular frequency band with a potential for feedback will be affected depending on how close the signal input is to the knee-point T_1. The maximum gain reduction occurs when the input signal coincides with the knee-point T_1. Experiments show that a 12 dB feedback margin provides the best performance in cancelling feedback without affecting signal quality. The farther the input signal level deviates from the knee-point T_1, the less will the gain be affected. The gain reduction is adaptively applied and adaptively removed in accordance with the continuous feedback detection decision.

The simplicity of using a fixed feedback margin is also its principal drawback, because it is not necessarily optimized to provide the best performance for all different hearing aids and different degrees of feedback. Sometimes, it might reduce gain more than required. At other times, the feedback canceler might under-react and not cancel the feedback completely. These considerations lead to the development of a Double Adaptive Feedback Canceller. Once feedback is detected during the build-up phase, or even when feedback is already established, this canceller will first apply a small feedback margin, such as 3dB. If the feedback is canceled immediately, it will stay with the small feedback margin. If the feedback still exists or continues to build up, the feedback margin is adaptively increased until the feedback is completely eliminated.

5. RESULTS

Two separate studies evaluated the real-time adaptive feedback canceller in a digital hearing aid during telephone use. The Unitron (UHL) study was a field trial with 24 hearing impaired subjects. A sepaerate study was carried out at the University of Buffalo (UB), with 25 patients. Hearing losses ranged from mild to severe and hearing aids included CIC ITC ITE and BTE styles. The results are shown in Figure 3.

	UHL	UB	
	% users	% users	% calls
Success	72	64	78
Problems	28	36	22

Figure 3. Field trial results

6. SUMMARY

A very fast and efficient multi-band feedback detection and cancellation method has been developed. Detection and cancellation of potential feedback takes place in in less than 100 ms, sometimes even before feedback is fully built up and becomes noticeble.

7. REFERENCES

1. D. Smriga; Problem Solving Through "Smart" Digital Technology, p58-60, V6, No 1, The Hearing Review, January 1999,.
2. B. W. Edwards et. al; Signal-processing algorithms for a new software-based digital hearing device,, V51, No. 9, Sep. 1998, The Hearing Journal.
3. T. Weidner; Method for Feedback Recognition in a Hearing Aid and a Hearing Aid Operating According to the Method; US Patent 6,404,895 B1; June 11, 2002.
4. Kates et al; Feedback Cancellation Apparatus and Methods Utilizing Adaptive Reference Filter Mechanisms; US Patent 6,434,247; August 2002.
5. Hays; Feedback Cancellation Improvements; US Patent 6,377,119; April 2002.
6. F. Kuk; Understanding Feedback and Digital feedback Cancellation Strategies, Feb., 2002, The Hearing Review.

A LOW-NOISE DIRECTIONAL MICROPHONE SYSTEM

Jim Ryan[1] and Brian Csermak[2]
[1]Gennum Corporation, 232 Herzberg Road, Suite 202, Kanata, Ontario, Canada, K2K 2A1 jim_r@gennum. com
[2] Gennum Corporation, 970 Fraser Drive, Burlington, Ontario, Canada, L7L 5P5 brian_c@gennum. com

1. INTRODUCTION

First-order directional microphones are commonly used in hearing aids. The conventional method for implementing a directional response using two non-directional microphones is depicted in Fig. 1. The microphones are arranged such that there is a front microphone and a rear microphone. A directional response pattern, with its main beam pointing toward the front microphone, is formed by subtracting the delayed rear microphone signal from the front microphone signal. The optional equalization filter (EQ) is provided to equalize the directional microphone on-axis frequency response to that of a single, omnidirectional microphone. A variety of directional patterns can be implemented by varying the delay [1].

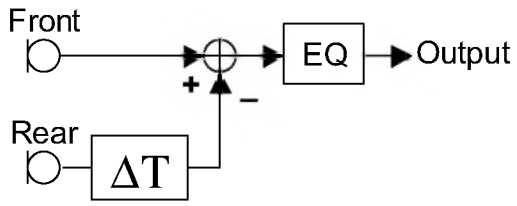


Fig. 1. Block diagram of a conventional directional microphone system.

While such microphones do provide significant directional gain and patient benefit it is a well-known fact that the unequalized frequency response possesses a first-order high-pass characteristic [2]. The reduction in signal level, particularly acute at low frequencies, leads to a reduction in signal-to-noise ratio (SNR) due to microphone self noise, preamp noise and wind. While the equalization filter overcomes the high-pass characteristic, it does not improve the SNR.

The increase in audible noise due to directional processing is considered to be one of the major drawbacks of a directional hearing instrument. In many cases it causes the wearer to deactivate the directional microphone.

Existing automatic directional systems handle the low-frequency noise problem in one of two ways: 1) they automatically switch the hearing aid to omnidirectional mode when the wearer enters a quiet environment, or 2) they automatically adjust the amount of low-frequency equalization applied to the microphone signal to reduce noise amplification. The first approach is disadvantageous in situations where a small amount of directional performance would benefit the wearer but the low-frequency noise is still a problem. By switching the hearing aid to omnidirectional mode, the wearer loses the benefit of the directional microphone. While the second approach does maintain the operation of the directional microphone, the lack of low-frequency equalization reduces the audio quality of the hearing aid and can impair audibility for wearers with a low-frequency hearing loss.

This paper describes a simple method for improving the SNR in directional microphone applications. The method provides the maximum directional gain for a prescribed allowable degradation in SNR. This is accomplished using a frequency-specific phase shift to create a controlled loss in directional gain over a frequency band of interest. In this way a minimal amount of directional performance is lost while maintaining a targeted amount of low frequency sensitivity or SNR. As will be shown, the signal processing requirements are simple enough to permit deployment in low-power digital hearing aids.

2. LOW-NOISE DIRECTIONAL SYSTEM

The proposed system limits the amount of noise amplification at low frequencies by replacing the pure time delay of the conventional system with a frequency-specific phase shift. This phase shift is implemented using specially designed filters to process both the front and rear microphone signals. The filter outputs are then summed to provide the directional microphone output. The equalization (EQ) filter is provided to equalize the directional microphone on-axis frequency response to that of a single, omnidirectional microphone. The block diagram is depicted in Fig. 2.

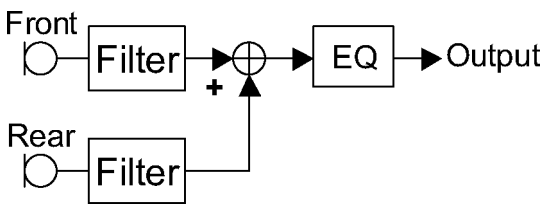


Fig. 2. Block diagram of a low-noise directional microphone system.

The phase shift required to implement a low-noise directional system is determined by the maximum SNR loss that is acceptable. If a high loss of SNR can be tolerated, then the phase shift approaches that of a conventional directional system. On the other hand, if there is a desire to limit SNR loss, then the phase shift deviates significantly from that of a conventional system.

As an example, consider the design of a directional microphone system with a front to rear microphone spacing of 10.7mm. The SNR loss as a function of frequency for a conventional system is shown in Fig. 3 (curve 'Conv'). Superimposed on this are curves representing maximum desired levels of SNR loss (20dB, 15dB, 10dB, 5dB and 0dB) for the same microphone. Fig. 4 shows the corresponding directivity index (DI). The main effect of the SNR loss limit is the decrease in low-frequency DI compared to a conventional system. The lower the desired SNR loss, the lower the resulting DI.

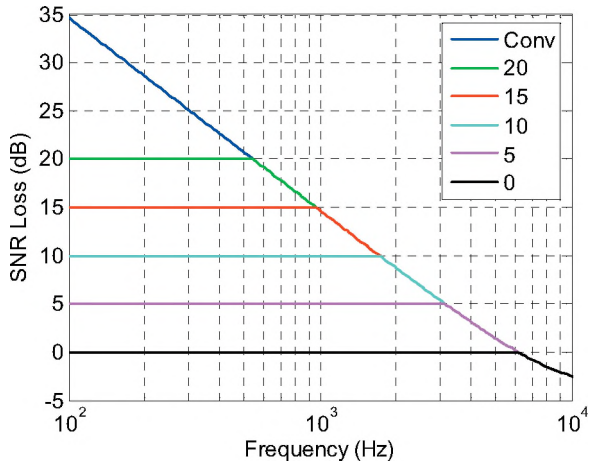


Fig. 3. SNR loss for directional microphone system with a front to rear microphone spacing of 10.7mm.

Since hearing aids are primarily intended for speech directional performance is often stated as AI-DI, which is a weighted average of the DI at frequencies 500,1000,2000 and 4000 Hz [3]. The points used in the AI-DI calculation are indicated by the circular markers on the curves in Fig. 4. The DI at the frequencies used in the AI-DI calculation is much less affected by the limit on SNR loss. For example, if the SNR loss is limited to 10dB, a reduction of 25 dB from the conventional system, the AI-DI only drops by 1.3 dB.

The inter-microphone phase shift required to implement the systems described in Fig. 3 and 4 is shown in Fig. 5. Strictly speaking, any type of filter could be used as long as the magnitude responses of both filters are identical. It is a much simpler design task, however, if allpass filters are used. By using allpass filters, the inter-microphone phase response from Fig. 5 is divided equally between the front and rear filters. This phase is then added to the phase of a known realizable allpass filter to obtain a realistic phase target. The required allpass filters can then be designed using any known allpass filter design procedure [4].

This approach has been used to obtain low-noise directional microphones using only one second-order allpass filter in each of the front and rear microphone signal paths.

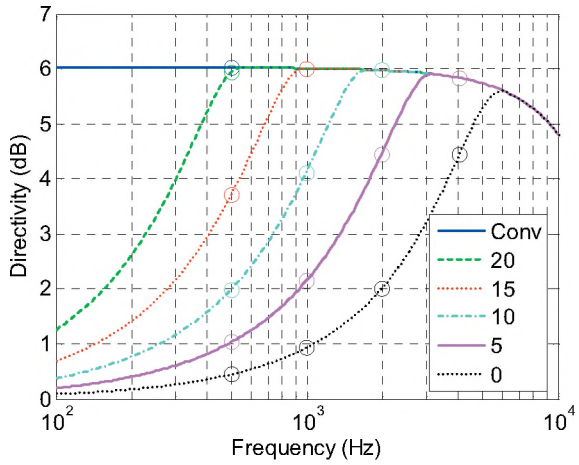


Fig. 4. Directivity Index achievable for the SNR loss curves shown in Fig. 3. The circles indicate the points used to calculate AI-DI. AI-DI values for these systems are: 'Conv' = 6 dB, '20' = 5.9 dB, '15' = 5.5dB, '10' = 4.7 dB, '5' = 3.6 dB and '0' = 2dB.

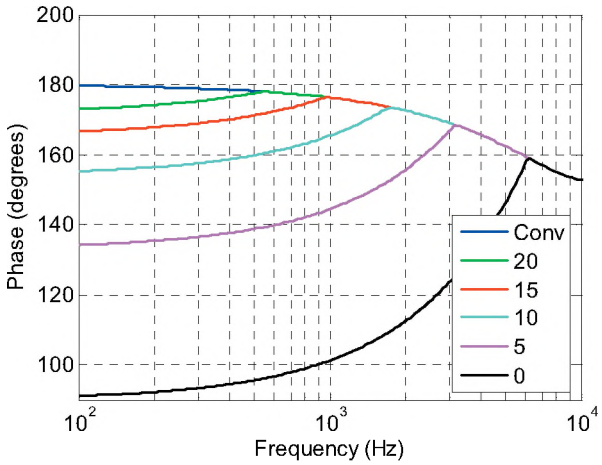


Fig. 5. Inter-microphone phase shift that is required to achieve the directional gains indicated in Fig. 4.

REFERENCES

[1] Csermak, B. (2000). A primer on a dual microphone directional system. *The Hearing Review*, vol. 7, pp.56-60.

[2] Elko, G.W. (2000). Superdirectional microphone arrays. In *Acoustic Signal Processing for Telecommunications*, pp.181-236.

[3] Roberts., M. & Schulein, R. (1997). Measurement and Intelligibility Optimization of Directional Microphones for Use in Hearing Aid Devices. *103rd AES Convention.*

[4] Kidambi, S.S. (1996). Weighted least-squares design of recursive allpass filters. *IEEE Trans. on Signal Processing*, Vol. 44, No. 6, pp. 1553-1557.

OBJECTIVE AND SUBJECTIVE EVALUATION OF NOISE REDUCTION ALGORITHMS FOR HEARING AIDS

Vijay Parsa[1,2], Gurjit Singh[1], Guo Chen[1] and Karthikeyan Umapathy[1]
[1]National Centre for Audiology, University of Western Ontario, London, Ontario, Canada N6G 1H1.
[2]Contact email address : parsa@nca.uwo.ca

1. INTRODUCTION

Speech is a highly redundant signal and because of this redundancy listeners with normal hearing are able to understand speech even in the presence of background noise. For listeners with sensorineural hearing-impairment however, there is a considerable loss of these redundant cues in the speech signal (Levitt, 2001). Thus, one of the most common complaints made by hearing-aid users is understanding speech in noise and consequently, one of the driving factors behind obtaining a hearing aid.

Recently, there has been an explosion in the number of digital hearing aids appearing on the market, with a number of these devices proffering noise reduction capabilities. Accurate and comprehensive evaluation of their noise reduction performance is important in order to quantify the relative benefits of these devices under a variety of listening conditions. In general, the noise reduction performance can be evaluated through objective electroacoustic measures and/or subjective listening tests. Subjective listening tests are preferred for their face validity, but they are often expensive, time-consuming and labour-intensive. Currently, there does not exist a validated or standardized electroacoustical procedure that allows clinical audiologists from assessing the relative benefits of various devices that offer similar, but not identical, noise reduction algorithms. The present study aims to address this issue through bridging electroacoustic and subjective measures of quality of speech processed through the noise reduction algorithms. The specific goals of this research were two-fold : 1) to evaluate the quality of speech processed through six different noise reduction algorithms with normal hearing and hearing impaired listeners, and 2) to identify instrumental measures of noise reduction performance that correlate best with the behavioural data.

2. METHOD

Six candidate noise reduction algorithms were evaluated in this study (Umapathy and Parsa, 2003). Algorithm 1 (EPH) is a technique based on minimizing the Mean Square Error (MMSE) of the Short Time Spectral Amplitude (STSA) estimator as proposed by Ephraim and Malah (1984). Algorithm 2 (WOL) is also based on the STSA estimator, however instead of the MMSE based amplitude estimator, the criterion used in this algorithm is the MMSE power spectrum estimator. Algorithm 3 (ES) is based on the subspace projection technique where noisy signals are decomposed into signal and noise subspaces, with the assumption that the signal is present only in the signal subspace whereas noise spans both subspaces (Klein and Kabal, 2002). Algorithm 4 (WV) uses wavelet packet decomposition and auditory masking properties (Lu and Wang, 2003). Both algorithms 5 (MP) and 6 (MP2) are based on the matching pursuit algorithm where time-frequency atoms from a Gaussian dictionary of time-frequency functions are adaptively moulded to fit the speech signal (Mallat and Zhang, 1993). The MP algorithm used a varying number of time-frequency functions to reconstruct the enhanced signal by applying a threshold on the slope of the rate of energy capture curve, while the MP2 algorithm used a fixed number of time-frequency atoms irrespective of the SNR values.

Ten adult participants with normal hearing (pure tone thresholds $\leq$ 20 dB HL at 1, 2, and 4 kHz) and 10 adult participants with hearing loss (mild to profound sensorineural hearing impairment) were paid to participate in this study. Participants rated the improvements in sound quality on 11-point scales, ranging from -5 to 5, on five dimensions: clarity, listening comfort, listening effort, background noise, and overall quality. Sound quality ratings were obtained using two sentences from the Hearing In Noise Test (HINT) database (Nilsson et al., 1994) which were corrupted by either speech-shaped noise (SSN) or multi-talker babble (MTB) at SNRs ranging between –4dB to +12 dB. In parallel, several instrumental measures of speech quality were computed from the speech stimuli processed by the noise reduction algorithms. These included the Perceptual Evaluation of Speech Quality (PESQ) (ITU, 2001), Perceptual Speech Quality Measure (PSQM) (ITU, 1998), Measuring Normalizing Blocks (MNB) (ITU, 1998), and Measuring Perceptual Spectral Density Distribution (MPSDD) (Chen and Parsa, 2004).

3. RESULTS

Figures 1a and 1b display the subjective quality ratings for different noise reduction algorithms from normal and hearing impaired listeners respectively. Algorithms based on spectral subtraction techniques (EPH and WOL)

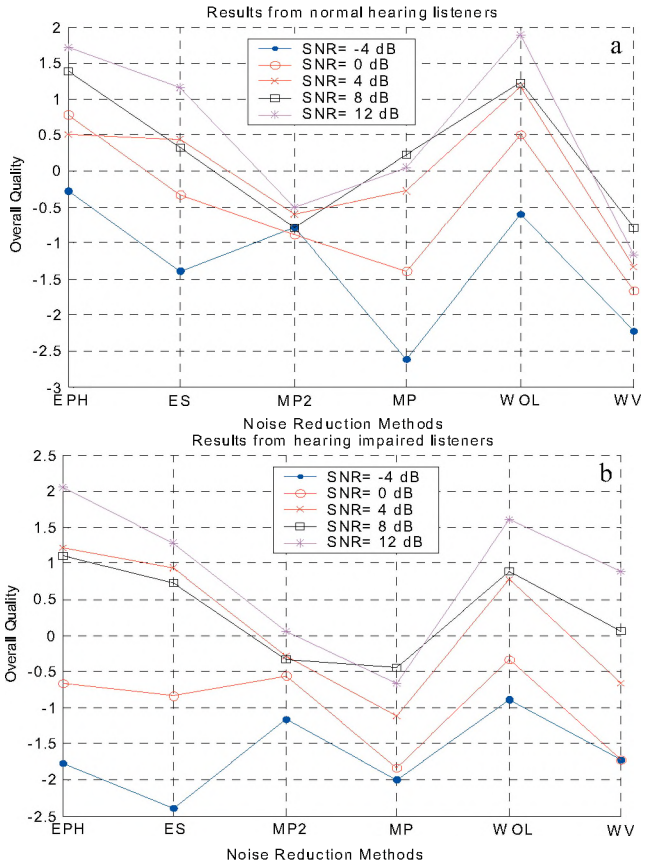


Fig. 1. Sound quality ratings from normal and hearing impaired listeners. Ratings were obtained from speech stimuli corrupted by the speech shaped noise at different SNRs and processed by different noise reduction algorithms.

can be seen to improve perceptual sound quality ratings by normal hearing listeners in conditions where the SNR is ≥ 0 dB. For hearing impaired listeners, the same algorithms were judged to improve the sound quality when the SNR is ≥ 4 dB. Table 1 displays the correlation between the instrumental measures of speech quality and the subjective ratings. These correlations were computed separately for perceptual data collected from normal hearing and hearing impaired (HI) subjects for the two noise conditions. It is evident from this table that the PESQ parameter correlated best with perceptual ratings under the speech-shaped noise condition, and its performance is similar to the PSQM parameter under the multi-talker babble condition.

Table 1: Correlations between instrumental and subjective measures of speech quality.

Measure	SSN		MTB	
	Normal	HI	Normal	HI
PESQ	0.79	0.83	0.71	0.85
PSQM	0.63	0.71	0.73	0.88
MNB	0.30	0.62	0.44	0.70
MPSDD	0.74	0.77	0.74	0.77

4. CONCLUSIONS

This paper investigated the performance of six noise reduction algorithms under a variety of listening conditions using instrumental measures and subjective ratings of speech quality. Ratings obtained from normal and hearing impaired listeners showed improvements in speech quality at positive SNRs for algorithms based on short-time spectral amplitude estimation. Instrumental measures such as the PESQ measure exhibited a good degree of correlation with quality ratings from normal and hearing impaired listeners. These results show that the PESQ measure can potentially be of use in the development and optimization of noise reduction algorithms for hearing impaired listeners.

REFERENCES

Chen, G. & Parsa, V. (2004). Output-based speech quality evaluation by measuring perceptual spectral density distribution, IEE Electronics Letters, 40, 783-785.

Ephraim, Y. & Malah, D. (1984). Speech enhancement using a minimum mean-square error short-time amplitude spectral estimator. IEEE Transactions on Acoustics, Speech, and Signal Processing, 32, 1109-1121.

International Telecommunication Union (1998). Perceptual Speech Quality Measure , ITU-T recommendation P.861, Geneva, Switzerland.

International Telecommunication Union (2001). Perceptual Evaluation of Speech Quality , ITU-T recommendation P.862, Geneva, Switzerland.

Klein, M. & Kabal, P. (2002). Signal subspace speech enhancement with perceptual post-filtering. Proceedings of the IEEE International Conference on ICASSP, 1, 537-540.

Levitt, H. (2001). Noise reduction in hearing aids: a review. Journal of Rehabilitative Research and Development, 38, 111-121.

Lu, C.T. & Wang, H.C. (2003). Enhancement of single channel speech based on masking property and wavelet transform. Speech Communication, 41, 409-427.

Mallat, S.G. & Zhang, Z. (1993). Matching pursuits with time-frequency dictionaries. IEEE Transactions on Acoustics, Speech, and Signal Processing, 41, 3397-3415.

Nilsson, M., Soli, S.D., & Sullivan, J.A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. Journal of the Acoustical Society of America, 95, 1085-1099.

Umapathy, K. & Parsa, V. (October, 2003). Objective evaluation of noise reduction algorithms in speech applications, Paper #Z6-2, 115th Audio Engineering Society Convention, New York, USA.

ACKNOWLEDGEMENTS

Financial support from the Oticon Foundation, Denmark, and the Ontario Rehabilitation Technology Consortium is gratefully acknowledged.

ADVANCED MEASURES OF BONE ANCHORED HEARING AIDS: DO THEY CORRELATE WITH PERCEPTUAL JUDGMENTS?

W. E. Hodgetts[1], G. Chen[2], and V. Parsa[2]
[1]Dept. of Speech Pathology and Audiology, University of Alberta, Edmonton, Alberta, Canada, T6G 2G4
[2]National Centre for Audiology, University of Western Ontario, London, Ontario, Canada, N6G 1H1
Contact email address: bill.hodgetts@ualberta.ca

1. INTRODUCTION

It can be argued that, as the Bone-Anchored Hearing Aid (BAHA) continues to become the standard of care for patients with conductive hearing loss, comparisons between BAHA and traditional Bone-Conduction Hearing Aids (e.g. those held in place by headband) have ceded their importance to comparisons between different types of BAHAs.

Currently, there are two ear-level BAHAs (the Compact and Classic 300) and a body-worn processor for more severe hearing losses (the Cordelle). The two ear level devices are linear but limit maximum output differently. Specifically, the Classic 300 saturates like a peak clipper, while the Compact limits output by compression. The Cordelle II uses a K-amp circuit but also saturates at high levels.

Neither subjective nor instrumental comparisons of sound quality have been formally evaluated with the current BAHA devices. Subjective ratings from patients are often confounded by a number of biases. For example, patients are not blind to the BAHA under test and may have several thousand dollars invested in the device. Rather than using patients for the subjective ratings, normal hearing listeners can be used to evaluate the sound quality of recordings made from each device. We were also interested in whether the subjective sound quality ratings correlated with instrumental measures of sound quality such as the Perceptual Evaluation of Speech Quality (PESQ; ITU, 2001) and the Perceptual Evaluation of Audio Quality (PEAQ; ITU, 2001).

The following research questions were of interest: 1) Are there perceptual differences between the different BAHA recordings, and 2) Do these perceptual judgments correlate with objective measures of sound quality?

2. METHOD

Fourteen adults with normal hearing participated in this study. Their task was to listen to the BAHA recordings and provide a mark on a visual analogue scale (VAS) between 1 and 10 that best represented their overall impression of the sound quality of each recording.

Three frequency responses (F1=Manufacturer's default setting at full volume, F2= Manufacturer's default setting at volume 2, F3=Potentiometers adjusted to opposite of Manufacturer's default setting at full volume) and two input levels (65 & 75 dB SPL) were chosen for the three BAHAs for a total of 18 listening conditions. The "carrot passage" from the Connected Speech Test (CST; Cox et al., 1987) spoken by a male talker was used as the input stimulus to each BAHA on a skull simulator (coupler). Recordings were digitized through an Aardvark Q10 audio interface and stored as .wav files. These .wav files were then RMS equalized.

The two objective measures of sound quality (PESQ and PEAQ) were obtained from the input and output BAHA recordings using MATLAB.

3. RESULTS & DISCUSSION

Mean and standard deviation VAS scores for each recording condition are presented in Table 1.

Table 1. Mean and standard deviation values ordered from highest to lowest mean score for all 18 recording conditions.

Condition	Mean VAS	SD
F1_Compact_65	7.21	1.52
F2_Compact_65	6.93	1.91
F3_Compact_65	6.44	1.69
F2_Cordelle_65	6.42	1.80
F2_Classic_65	6.39	1.42
F3_Classic_65	6.33	1.35
F1_Classic_65	6.14	1.36
F2_Compact_75	5.54	2.12
F2_Cordelle_75	5.39	2.18
F1_Cordelle_65	4.81	2.12
F3_Cordelle_65	4.77	2.49
F1_Compact_75	4.69	1.48
F2_Classic_75	4.18	2.04
F3_Compact_75	4.03	1.41
F1_Cordelle_75	3.18	1.61
F3_Classic_75	2.75	1.41
F1_Classic_75	2.30	1.52
F3_Cordelle_75	2.26	1.09

A 3 (Device) x 3 (Frequency Response) x 2 (Input) repeated measures ANOVA was computed. Results are plotted in Figures 1 and 2.

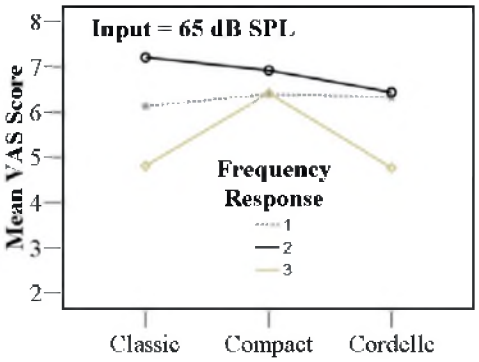


Fig. 1. Mean Frequency Response x Device VAS scores for a 65 dB SPL input to the BAHA.

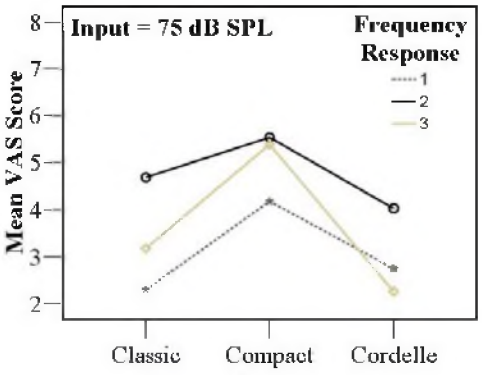


Fig. 2. Mean Frequency Response x Device VAS scores for a 75 dB SPL input to the BAHA.

A significant main effect was found for "device" with the Compact receiving the highest sound quality ratings. There was also a clear preference for F2 regardless of type of device or input SPL. This was a logical result as the F2 frequency response was obtained for all devices set to volume 2 (out of 3). Significant interactions were also investigated. It was expected that the Classic and Cordelle would saturate at full volume (F1 and F3) and that ratings would be lower as the input level increased. Indeed, this was the case. Conversely, since the Compact employs output compression, ratings were expected to remain high for all conditions regardless of input level. This was true with the exception of F1 at an input of 75 dB SPL. Unfortunately, F1 represents the "default" frequency response for the Compact. From a sound quality perspective, it would appear that F3 would be a better option for patients wearing the Compact at volume 3 in the presence of loud speech.

Next, the subjective scores were correlated with the PESQ and PEAQ instrumental measures. Results for the PESQ are plotted in Figure 3. Similar results for the PEAQ are not shown due to space limitations. The scores for the subjective ratings show greater dispersion compared to the instrumental measure. This is reflected in the relatively low correlation (r = 0.51) between the two measures.

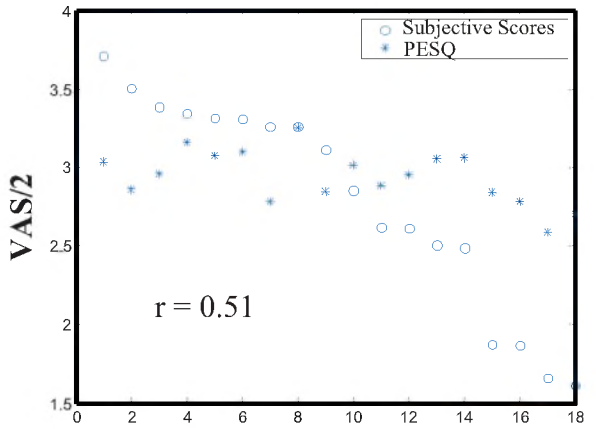


Fig 3. Mean VAS and PESQ scores by recording condition. VAS scores were divided by 2 so that data could be plotted on the same 5 point scale as the PESQ data.

4. CONCLUSION

This paper investigated whether normal hearing listeners were able to differentiate the sound quality of different BAHAs in different processing conditions. A predictable advantage was found for the Compact with output compression as the input speech level increased. However, for the Compact, a consequence of output compression is reduced MPO. It is not clear that BAHA users would still prefer the sound quality of the Compact if the compression compromised the audibility of speech. More research is needed to answer this question.

A second goal of this study was to determine if the perceptual data correlated with instrumental measures of sound quality used in the telecommunications industry. Correlations were modest for both the PESQ and PEAQ, suggesting further modifications are required for the instrumental measures to match with the perceptual data.

REFERENCES

Cox, R. M., Alexander, G. C., & Gilmore, C. (1987). Development of the Connected Speech Test (CST). Ear & Hearing., 8, Suppl-126S.

International Telecommunication Union (2001). Perceptual Evaluation of Speech Quality, ITU-T recommendation P.862, Geneva, Switzerland.

International Telecommunication Union (2001). Perceptual Evaluation of Audio Quality, ITU-R recommendation BS-1387, Geneva, Switzerland.

HEARING LOSS PREVENTION IN THE MILITARY ENVIRONMENT

C. Giguère & C. Laroche
Programme d'audiologie et d'orthophonie, Université d'Ottawa, 451 Smyth Road, Ottawa, Ontario K1H 8M5

1. INTRODUCTION

In the military, noise can be particularly noxious to hearing. The personnel face a wide range of noise-hazardous situations, many of which are seldom encountered in other work environments. High noise levels are associated with the operation of small and large calibre weapons, combat vehicles, aircraft, ships, vessels, and industrial equipment. Exposure to such noise can cause hearing loss, compromise speech communication, localization of sound sources and detection of warning sounds and thus, can jeopardize life or safety of the military and civilian personnel.

The ultimate goal of a hearing-loss prevention (HLP) program is to preserve hearing health as well as all hearing abilities necessary for effective operations. This paper reviews the essential elements of a prevention program proposed for the Canadian Armed Forces (CF). It has been designed to meet the Canadian Occupational Safety and Health (COSH) Regulations [1]. Additional measures and limits beyond COSH are also included to address issues specific to the military environment. A draft policy based on this proposal is currently under review by the CF.

2. PROGRAM ELEMENTS

2.1 Hazard Assessment and Identification

An effective HLP program is based on accurate and up-to-date sound level measurements for all noise-hazardous areas, facilities and operational equipment.

For steady-state or fluctuating noise, detailed surveys are to be conducted in all environments requiring the initiation of hearing-loss prevention procedures under COSH [1] (> 84 dBA). The A-weighted sound levels and the duration of exposure must be reported. For impulse noise with high peak levels, the A-weighted sound exposure level (SEL) per single impulse must be reported as well as the total number of rounds fired in a work shift or day. The noise survey data from all operational military equipment (ships, aircraft and vehicles) and weapons systems should be included in a central noise database to be maintained and updated for access by all CF personnel involved in the implementation or evaluation of the prevention program.

In assessing hazard, all sources of noise must be included in the calculation of exposure. Under COSH [1], the maximum 8-hr noise exposure limit from all sources is 87 dBA. In addition, the critical SEL limits set in a recent North Atlantic Treaty Organization (NATO) study [2] for single impulses from weapons systems are not to be exceeded (see Sec. 2.4). The new NATO data generally indicate that the risk from small calibre weapons (or short impulse duration) are under-estimated using current damage-risk criteria based on CHABA [3], while the risk from large calibre weapons (or long impulse duration) may be over-estimated.

Occupational noise regulations, like COSH, are based on a typical workday of about 8 hrs followed by a long rest period. In the military, sustained exposure largely exceeding an 8-hour workday can occur on a regular or irregular basis. For exposures lasting 12 hrs or more, a rest period at least as long as the exposure duration is recommended [4]. In all cases, the rest period should be sufficiently long to ensure that the temporary threshold shift (TTS) induced by the exposure has decreased to a value 2.5 dB or less, which is the residual TTS expected after an exposure to 87 dBA for 8 hrs and 16 hrs of rest. Data in [5] can be used to estimate such a minimum rest period, given exposure duration and level. The rest environment should be lower than 74 dBA.

2.2 Engineering Noise Control Measures

Engineering control measures is the preferred method to reduce exposure to safe levels. No other prevention method can match the long-term health, safety and workplace communication efficiency benefits of a quieter environment.

The best time for initiating engineering noise control measures is during the procurement process. The documentation for all new or retrofitted equipment and facilities should include noise specifications. If technically feasible, noise levels for all operators will not exceed COSH regulatory limit of 87 dBA during normal use. Otherwise, the specifications should ensure that all state-of-the-art engineering control measures be considered to deliver the quietest possible products, and that the noise levels from all sources at each operator position be specified upon delivery.

2.3 Administrative Controls

Administrative controls refer to measures used to inform personnel of potentially noise-hazardous area, and to staffing procedures used to further limit the duration and level of noise exposure once all engineering controls have been implemented. COSH regulations require informing all personnel of the potential risk to hearing whenever noise exposure is likely to exceed 84 dBA. In addition, there is the mandatory installation of visible and permanent warning signs to identify noise-hazardous areas, and the supplying of hearing protectors when noise levels exceed 87 dBA.

2.4 Personal Hearing Protection

Personal hearing protection devices (HPDs) are to be used to reduce noise exposure only once all engineering and administrative control measures have been exhausted. In a working environment as complex as the military (e.g. variable work schedules), the guiding rule to ensure that the daily noise exposure does not exceed the regulatory limit of 87 dBA is to require that exposure at each noise-hazardous site be below 87 dBA, through proper use of HPDs, irrespective of the duration of the exposure.

For steady-state or fluctuating noise, selection of HPDs will be made to ensure exposure is below the 87 dBA limit, preferably in the range 77-82 dBA to avoid over-protection. A central database of HPDs should be integrated with the noise database to identify proper devices in each setting.

HPDs are to be used on firing ranges by all personnel in the vicinity of the weapons systems. HPD selection will be restricted to a range of devices tested or approved by the CF for each weapons system. The maximum daily number of rounds allowable will be calculated according to:

$$\text{MAX. DAILY ROUNDS} = 28880 \times 10^{-(\text{SEL-ATT-87})/10}$$

where 87 dBA is COSH noise exposure limit, SEL is the free-field sound exposure level in dBA per single impulse, and ATT is the attenuation (dB) of the HPD for the particular weapons impulse. When the SEL is measured under the protector, this value must be used instead of SEL-ATT in the equation above. For protection against small-calibre weapons, a minimum HPD attenuation is also necessary to ensure the critical SEL limit of 116 dBA for an unprotected single impulse [2] is not exceeded, as follows:

$$\text{ATT} \geq \text{SEL} - 116 \text{ dBA} \quad \text{(small-calibre)}$$

When the SEL is measured under the protector, then the protected SEL per single impulse will not exceed 116 dBA.

2.5 Monitoring Audiometry

Audiometric monitoring of the CF personnel at risk is needed to (1) identify and document the hearing status of individuals with hearing loss, (2) provide proper care, protection, employment follow-up for those who incur hearing loss, and (3) monitor the general effectiveness of the HLP program. It important to note, however, that audiometric testing is not in itself a prevention method if there is no effective intervention to limit noise exposure. There are also reliability issues associated with the use of audiograms in occupational settings.

Audiograms should be recorded with automatic audiometers to standardize the measurement process across CF facilities. A computerized record keeping system should be put in place to automatically identify hearing conditions requiring follow-up. Periodic audiograms should be performed at any time during the work shift, and preferably late in the shift. The standard threshold shift due to noise is to be defined as a change from the baseline audiogram of 15 dB or more at 500, 1000, 2000, 3000, 4000 or 6000 Hz, in either ear. The baseline audiogram should be taken on all personnel entering the CF, within 30 days after initial noise exposure.

2.6 Education

An educational component is required (1) to ensure the CF personnel is aware of the effects of noise on health and safety, and (2) to explain the advantages and limits of each element contained in the HLP program. Training should be provided at least annually to all personnel working in areas where noise levels can exceed 84 dBA. In the military environment, a major challenge is to ensure continuity in the training process.

2.7 Program Evaluation

The objective of program evaluation is to assess or monitor the effectiveness of the HLP program in preventing hearing damage in the CF personnel. The use of general program evaluation tools based on audiometric databases is questionable in the military environment. Instead, specific activities can include but should not be limited to (1) the identification of high-risk tasks or military occupations, (2) the field evaluation of the attenuation of hearing protectors, and (3) the validation of impulse noise damage risk-criteria and prevention measures.

2.8 Documentation

The critical documents (acoustical standards, regulations, etc.) necessary to conduct the daily procedures contained in the HLP program should be easily accessible by the responsible personnel.

3. CONCLUSIONS

In the military, the importance of accurate noise surveying, engineering and administrative noise controls, proper fit of hearing protection and regular audiometric monitoring of the hearing of exposed personnel cannot be over-emphasized. It is only through the utilization of all available methods that the hearing of the personnel will be protected.

[Work carried out under a contract from the Canadian Forces Medical Services].

REFERENCES

[1] Canadian Occupational Safety and Health (COSH) Regulations Part VII: Levels of Sound (SOR/2002-208, 30 May 2002).

[2] NATO (2003). Reconsideration of the Effects of Impulse Noise (RTO TR-017/HFM-022).

[3] Forshaw, S. (1970). Guide to Noise-Hazard Evaluation. Defence Research Establishment Toronto, Review Paper No. 771.

[4]Shaw, E.A.G. (1985). Occupational Noise Exposure and Noise-Induced Hearing Loss: Scientific Issues, Technical Arguments and Practical Recommendations (APS707; NRCC/CRNC No 25051).

[5] Shaw, E.A.G. (1983). "On the Growth and Decay of Asymptotic Threshold Shift in Human Subjects," 4th Int. Cong. on Noise as a Public Health Problem, Vol. 1: 297-308 (Turin, Italy).

FUTURE DIRECTIONS FOR RESEARCH ON THE COMBINED EFFECTS OF NOISE AND VIBRATION ON COGNITION AND COMMUNICATION

Ann M. Nakashima and Sharon M. Abel
Defence Research and Development Canada – Toronto, P.O. Box 2000, 1133 Sheppard Ave West, Toronto, ON, M3M 3B9 E-mail: ann.nakashima@drdc-rddc.gc.ca

1. INTRODUCTION

Exposure to noise and vibration in military vehicles is still an unsolved problem. High noise levels have been known to cause hearing loss and difficulties in using intercommunication systems. Whole-body vibration (WBV) is a frequently reported source of annoyance that has been linked to incidences of back pain, but effects on performance are not well documented. For military personnel, clear communication and sharp situational awareness are essential for effective performance and, ultimately, survival. In order to gain a better understanding of how to optimize the performance of personnel operating military vehicles, it is important to consider the effects of realistic levels of noise and vibration. This paper presents an overview of previous studies of communication and cognitive performance in noise and vibration, and gives suggestions for future research on the combined effects of noise and vibration with regard to crew performance in military vehicles.

2. SPEECH INTELLIGIBILITY

The effect of noise on speech intelligibility has been studied extensively. Different types of hearing protectors have been considered. Kryter [7], for example, found that the use of earplugs does not negatively affect speech intelligibility for noise levels above 80 dB; in fact, it tended to increase the intelligibility. However, conventional hearing protectors decrease intelligibility for individuals who have hearing loss [1]. In military vehicles, operators use communication headsets that are sometimes modified with active noise reduction (ANR) technology or worn in combination with earplugs to improve ambient noise attenuation. Speech intelligibility tests using ANR headsets and headsets combined with communications earplugs (CEP) have shown that the noise reduction technology did not have a negative effect on intelligibility scores, for both normal hearing and hearing-loss individuals [10].

The issue of speech understanding in noise for individuals with hearing loss is important because 1) many military personnel suffer from noise-induced hearing loss and 2) it has been found that exposure to combined noise and vibration, causes temporary threshold shift (TTS). For example, Seidel *et al.* [11] found that the combination of noise and vibration can cause greater TTS than exposure to either noise or vibration alone. This included a significant amount of TTS at the 4 kHz octave band, which is known to be crucial to speech understanding [2].

3. COGNITIVE TASK PERFORMANCE

Previous studies of cognitive performance in noise and combined noise and vibration have utilized vigilance, short-term memory and counting tasks. The results of the studies are difficult to generalize because of the differences in the types of noise and vibration used (e.g. white noise vs. recorded noise, intermittent vs. continuous noise, sinusoidal vs. stochastic vibration). Two indicators of performance that have been used are reaction time and accuracy.

In an extensive study by Manninen [9], subjects were exposed to combinations of complex noise, WBV, temperatures and psychic load. One of several responses that were measured was reaction time. The effects of psychic load were observed by comparing the results of a competitive group, who had a monetary incentive to perform well, to those of a non-competitive group. Subject reaction times during a simple light cancellation task were observed. Reaction times increased when the non-competitive subjects were exposed to noise, but a combined effect of noise and vibration on reaction time was not found.

In a study that tested visual vigilance in the presence of intermittent noise, it was found that subjects who were exposed to noise tended to have a faster response time but poorer accuracy than subjects who were tested in quiet [4]. Exposure to WBV alone (without significant background noise) has been shown to cause slower reaction times and, in the case of 1.0 m/s^2 vibration only, increase response errors [12]. In another study of performance on a memory task, for different intensities of combined noise and vibration, significant changes in reaction time were not found across the three intensity levels [8]. However, since the effects of noise and vibration were not tested independently in the study, no conclusions could be drawn about the effects of vibration.

4. FUTURE RESEARCH DIRECTIONS

To date, most experiments of human response to combined noise and vibration have focused on subjective annoyance or discomfort, motor control and TTS. Fewer studies have looked at the effects on cognitive performance, and, to the best of the authors' knowledge, the effects on speech intelligibility have not been studied. Suggestions for consideration in the design of noise and vibration experiments are given below.

4.1 Noise and Vibration Signals

When performing noise exposure experiments using human test subjects, it is generally considered that the appropriate jurisdictional exposure guidelines for workplace noise should be followed (e.g. 85 dBA for 8 hours by Canadian standards [3]). There are currently no widely accepted exposure limits for WBV exposure in Canada. International standard ISO 13090-1 [6] gives guidelines for vibration experiments in terms of the magnitude of vibration that would require the attendance of a physician or medical doctor. These values are listed in Table 1.

Table 1. Exposure to vibration and repeated shock requiring the attendance of a physician or medical doctor [6].

Duration of exposure in any one 24hr period	16 min	1 h	4 h	8 h
Acceleration magnitude, m/s^2 (frequency-weighted rms acceleration)	2.2	1.6	1.1	0.9

Spectra of the noise and vibration signals must also be considered. The noise signals can be broadband, or recordings of actual vehicle/aircraft noise. In previous vibration experiments, sinusoidal vibration and stochastic vibration (typically between 2 and 16 Hz) have been used. These signals do not represent realistic vibration spectra, which often have dominant spectral components. In addition, previous experiments have used only vertical (z-axis) vibration. While vibration is dominantly in the vertical direction for many vehicle seats, vibration in the horizontal axes (left-right and back-front) and the seat back may also be significant.

4.2 Speech Intelligibility

Helmet-mounted communication systems used in military vehicles are required to provide adequate hearing protection while enhancing speech intelligibility in high levels of background noise. The finding from previous experiments that exposure to combined noise and vibration tends to exacerbate noise-induced TTS suggests that vibration should be considered when testing the speech intelligibility of communication systems. Normal hearing and hearing impaired individuals loss should be used as test subjects, and different types of hearing protection (ANR headsets, headsets combined with earplugs, etc.) should be considered.

4.3 Cognitive Tasks

Reaction time and accuracy have been used to evaluate the effect of noise and vibration on cognitive performance. Since previous studies have shown that exposure to noise tends to decrease both reaction time and response accuracy, while vibration tends to increase reaction time and decrease accuracy (for a particular vibration magnitude), it is of interest to investigate their combined effects. This would involve observing the performance of test subjects who are exposed to each stressor alone, and in combination to study the interactive effects of noise and vibration. For application to military vehicles, performance on vigilance tasks should be tested to investigate the possibility of attention lapses and mental fatigue.

REFERENCES

1. Abel SM, Alberti PW, Haythornthwaite C, Riko K (1982). Speech intelligibility in noise: Effects of fluency and protector type. J Acoust Soc Am 71:3, 708-715.
2. Abel SM, Krever EM, Alberti, PW (1990). Auditory detection, discrimination and speech processing in aging, noise-sensitive and hearing-impaired listeners. Scand Audiol 19, 43-54.
3. CSA Z94.2-02 (2002). Hearing protection devices – Performance, selection, care and use. Canadian Standards Association, Rexdale, ON.
4. Carter NL, Beh HC (1987). The effect of intermittent noise on vigilance performance. J Acoust Soc Am 82:4 1334-41.
5. Harris CS, Shoenberger RW (1980). Combined effects of broadband noise and complex waveform vibration on cognitive performance. Aviat Space Environ Med 51:1, 1-5.
6. ISO 13090-1 (1998). Mechanical vibration and shock – Guidance on safety aspects of tests and experiments with people – Part 1: Exposure to whole-body mechanical vibration and repeated shock. International Organization for Standardization, Geneva, Switzerland.
7. Kryter KD (1946). Effects of ear protective devices on the intelligibility of speech in noise. J Acoust Soc Am 18:3, 413-417.
8. Ljungberg J, Neely G, Lundström R (2004). Cognitive performance and subjective experience during combined exposures to whole-body vibration and noise. Int Arch Occup Environ Health 77:217-221.
9. Manninen O (1985). Cardiovascular changes and hearing threshold shifts in men under complex exposures to noise, whole body vibrations, temperatures and competition-type psychic load. Int Arch Occup Environ Health 56:251-274.
10. Riberia J, Mozo B, Murphy B (2004). Speech intelligibility with helicopter noise: Tests of three helmet-mounted communications systems. Aviat Space Environ Med 75:2,132-137.
11. Seidel H *et al.* (1988). Isolated and combined effects of prolonged exposures to noise and whole-body vibration on hearing, vision and strain. Int Arch Occup Environ Health 61:95-106.
12. Sherwood N, Griffin MJ (1990). Effects of whole-body vibration on short-term memory. Aviat Space Environ Med 61:12, 1092-1097.

Noise Exposure Assessment – A New ISO Standard

Alberto Behar
IBBME, University of Toronto, Toronto, ON, Canada,M5S 3G9.
alberto.behar@utoronto.ca

1. INTRODUCTION

It is universally accepted that the hazard of occupational noise induced hearing loss has to be assessed measuring the noise exposure of the exposed person and not just by measuring the noise level the person is exposed to. This is because hearing loss from occupational noise is the result of long-term (years?) exposure to noise. Therefore there is a need for the assessment of this long-term exposure.

There is no lack of documents that describes the procedures for this kind of measurements. In Canada the Canadian Standard Association has produced the standard CSA Z107.56 (R1999) Procedures for the Measurement of Occupational Noise Exposure. Although it is under revision, the eventual modifications are not expected to be substantial. In the United States, the American National Standard Institution has produced the standard ANSI S12.19-1996 Measurement of Occupational Noise Exposure. The International Organization for Standardization, ISO, has two standards related to this subject. One is the ISO 1999:1990 Acoustics – Determination of occupational noise exposure and estimation of noise-induced hearing loss. It does not provide to many details with regard to the measurement itself.

The only specific standard among the ISO documents is the ISO 9612:1997: "Acoustics -- Guidelines for the measurement and assessment of exposure to noise in a working environment". New developments in the field of noise exposure measurement techniques have made the ISO 9612:1997 obsolete, to the point that it was decided to write a new standard rather than review the old one. To that effect, the ISO/TC 43/SC 1, "Noise" has requested that WG 53 (Working Group 53) be responsible for preparing the new document. (There are two Canadian members in the WG: Tim Kelsall of Hatch Association and this author). The WG held its first meeting in January 2004, where the basics of the standard where discussed. Substantial

progress was achieved at a second meeting in June. A new meeting is planed for September, when the first draft is expected to be finished. (Because of the deadline for submitting this summary, the results of the September meeting are not included).

2. THE NEW DRAFT

The complete name for the new document is: ISO/WD2 9612 Acoustics – "Measurement and calculation of occupational noise exposure – Engineering method". (WD = Working Draft). Some of the sections will be described in the following paragraphs.

2.1 Sect 1:Scope

The purpose of the standard is to provide a method for the measurement of the noise exposure in the workplace. It is explicitly stated that the following issues are beyond the scope of the document: environment noise levels (noise maps), masking of communication, infrasound and ultrasound, extra-auditory effects as well as estimation of the noise exposure when hearing protectors are worn.

The standard does not deal with measurements required for noise reduction either.

2.2 Sect 6: Instrumentation

The standard recognizes the use of handheld sound level meters as well as dosimeters (called in the document "personal sound level meters". A table included in the document provides guidance on when each of those instruments is the best choice for the task.

2.3 Sect 7: Identifying equal noise exposure groups or a noise exposed worker.

This is a process aimed at reducing the number of required measurements. By grouping the workers on the basis of their exposures, there will be no need of measuring each one of them. This, by the way, is a common Industrial Hygiene technique, used for exposures to other hazardous substances chemical of physical.

Two approaches are recommended: one based on the tasks being performed and the second by the function. The second is especially useful when a worker performs the same task at different locations (e.g., maintenance operators).

2.4 Sect 8: Description of representative working day

This is a task necessary when characterizing a noise exposure situation of a worker in the entity under measurement. The standard includes a specific procedure for the overview and understanding of all the different tasks

that may influence the noise exposure. This has to be done in consultation with the worker in question and his supervisor.

There are provisions for a short-term evaluation to ascertain that the description of the measurement period is representative. However, another evaluation, under the heading of "long-term" requires that the measurements be repeated in another occasion(s) to confirm the previous findings.

2.5 Sect 10: Measurements

Probably the most useful section of the document provides guidance on how the measurement should be performed. It specifies that the result of the test should be presented as L_{eq}, L_{EX} and L_{Cpeak}.

Several sampling strategies are included and a table presents a selection of the strategies that can be used and the one that are recommended. Details are also included regarding how the instrument should be used, calibration procedures for the same, number and duration of the measurements, etc.

2.5 Sect 11: Evaluation of the uncertainty

For the last couple of years, it has been a requirement for each new standard to evaluate the uncertainty. This is to eliminate the rule of tumb presently used, stating that measurement results are true within +/- 2 dBA.

3.0 CONCLUSION

As mentioned above, the work is still in progress and there are many issues to be dealt with, trying to get consensus within the Working Group. We hope that most of the problem areas will be cleared at the September meeting and that by the CAA 2005, we will have a finished ISO standard.

ÉTUDE DE L'EXPOSITION PROFESSIONNELLE AU BRUIT DES CONDUCTEURS D'AUTOBUS SCOLAIRES

Pierre Marcotte, Paul-Émile Boileau et Jérôme Boutin
Institut de Recherche Robert-Sauvé en Santé et en Sécurité du Travail
505, boul. de Maisonneuve Ouest, Montréal, QC, H3A 3C2 marcotte.pierre@irsst.qc.ca

1. INTRODUCTION

L'exposition professionnelle au bruit des conducteurs d'autobus scolaires constitue une problématique mise en lumière au cours des dernières années. En effet, il semble que plusieurs conducteurs se soient plaints du niveau élevé de bruit relié à leur travail. Une étude effectuée en 1993 par le service de santé au travail du CLSC de l'Estuaire (Rimouski) [1] a mesuré l'exposition professionnelle au bruit des conducteurs d'autobus scolaires. Les mesures d'exposition (L_{eq}, facteur de bissection de 3 dB) variaient de 69,8 à 84,8 dB(A) avec un niveau sonore moyen de 82 dB(A) pour les grands véhicules de 72 passagers et de 78 dB(A) pour les petits véhicules de 48 passagers et moins. Par ailleurs, cette étude a démontré que, de façon générale, les enfants sont plus bruyants en après-midi.

Les objectifs de la présente étude étaient de mesurer l'exposition au bruit des conducteurs et d'estimer la dose quotidienne d'exposition au bruit, d'évaluer l'importance de l'exposition par rapport à des valeurs de référence, d'identifier les sources de bruit pouvant contribuer le plus aux niveaux de bruit auxquels sont exposés les conducteurs, de déterminer si des différences importantes existent entre les différent modèles de véhicules ainsi que de suggérer, au besoin, des correctifs à apporter pour réduire le bruit à l'intérieur des autobus scolaires.

2. MÉTHODE

Des mesures de l'exposition au bruit ont été réalisées dans des autobus scolaires appartenant à deux entreprises possédant des marques et modèles d'autobus scolaires offrant une bonne représentation de ceux utilisés au Québec. Un échantillon de 10 véhicules de 72 passagers a été sélectionné pour réaliser les mesures. Le tableau 1 identifie l'ensemble des 10 véhicules ainsi que le type de moteur pour chacun des véhicules. Pour chacun des autobus retenus, les mesures de l'exposition au bruit ont été effectuées sur des parcours réels avec et sans étudiants. Afin de simplifier le processus de collecte de données, l'exposition a été évaluée sur une demi-journée de travail (pendant l'après-midi), puisque c'est l'après-midi que les passagers ont tendance à être le plus bruyant [1].

Tableau 1. Véhicules sélectionnés pour les mesures de bruit.

MARQUE	MODÈLE	ANNÉE	MOTEUR
Blue Bird	TC-2000	1999	Cummins
Bue Bird	TC-2000	1999	Cummins
Blue Bird	TC-2000	2003	Cummins
Blue Bird	TC-2000	2002	Cummins
Blue Bird	TC-2000	1998	Cummins
Blue Bird	TC-2000	2003	Cummins
Thomas		2001	Cummins
Blue Bird	GMC	2001	Caterpilar
Freightliner	FS65	2004	Mercedes
International	30S	2003	DT466

Des spectres en fréquence de bruit ont été enregistrés à l'aide d'un microphone (BK 4165 avec préamplificateur BK 2669) relié à un analyseur de signal portatif (BK *Pulse*). Un dosimètre de bruit portatif (Larson-Davis Spark™ 706) a été positionné près de l'oreille droite des conducteurs pour mesurer les doses de bruit auxquelles ils sont exposés. Les niveaux de bruit équivalents (L_{eq}) ont été mesurés par le dosimètre selon la norme internationale ISO 1999:1990 [2] (pas de seuil et facteur de bissection de 3 dB) ainsi que selon le *Règlement sur la santé et la sécurité du travail* [3] (seuil de 85 dB(A) et facteur de bissection de 5 dB).

3. RÉSULTATS

3.1. Caractérisation des différentes sources de bruit

Les niveaux de bruit équivalents (L_{eq}, facteur de bissection de 3 dB) ont été mesurés à l'aide du microphone relié à l'analyseur de spectre, sur des périodes de temps variant de quelques secondes à plusieurs minutes. Une analyse des résultats a démontré que les écarts entre le véhicule le plus silencieux et le véhicule le plus bruyant sont de 8,6 dB(A) et 7,8 dB(A) pour respectivement les conditions de véhicule au repos et de véhicule en déplacement [4]. Les types de parcours empruntés par les véhicules ont été divisés en plusieurs catégories : urbain, rural, route (80 ou 90 km/h) et autoroute. Par la suite, le bruit provenant du véhicule a été moyenné sur l'ensemble des autobus, en fonction du type de parcours, et est présenté dans le tableau 2. Ces résultats suggèrent que le bruit

provenant du véhicule est directement relié à la vitesse de celui-ci, étant donné qu'une vitesse plus élevée est associée à un régime de moteur plus élevé ainsi qu'à une augmentation du bruit provenant du contact pneu-chaussée.

Tableau 2. Bruit du véhicule en fonction du type de parcours.

Type de parcours	Niveau de bruit moyen
Au repos	66,3 dB(A)
Urbain	73,4 dB(A)
Rural	74,2 dB(A)
Route	75,7 dB(A)
Autoroute	78,5 dB(A)

La contribution des autres sources de bruit a été évaluée en soustrayant la contribution du moteur lorsque celui-ci était en marche. Pour chacune des sources de bruit retenues, le niveau de bruit a été moyenné sur l'ensemble des mesures. Le niveau de bruit moyen ainsi obtenu ainsi que le niveau de bruit maximum sont rapportés dans le tableau 3. Il apparaît que les sources de bruit autres que le véhicule ont une contribution non négligeable à l'environnement sonore du véhicule, surpassant même le bruit moyen du véhicule. Par ailleurs, même si la source de bruit « SRG (CB) » (service de radio général) atteint un niveau moyen de 83,4 dB(A), sa contribution à la dose de bruit reçue par les conducteurs est moindre étant donnée sa nature intermittente et limitée dans le temps. Par ailleurs, il semble que la contribution de la source « Équipement hiver » soit non négligeable, avec un niveau de 77,4 dB(A), ce qui contribuerait à rendre le véhicule plus bruyant durant la saison hivernale.

Tableau 3. Contribution des autres sources de bruit.

Source	Niveau moyen	Niveau max.
Élèves primaires	78,1 dB(A)	83,5 dB(A)
Élèves secondaires	74,3 dB(A)	78,1 dB(A)
SRG (CB)	83,4 dB(A)	87,3 dB(A)
Équipement hiver	77,4 dB(A)	---

3.2. Dosimétrie

L'étude dosimétrique effectuée sur les conducteurs a révélé que ceux-ci sont exposés à des niveaux équivalents quotidiens ($L_{EX, 8h}$, ISO 1999) variant de 74,6 à 85,2 dB(A), et à des doses de bruit variant de 1,0 % à 17,2 % selon le *Règlement sur la santé et la sécurité du travail* (RSST), correspondant à des niveaux équivalents quotidiens ($L_{EX, 8h}$, RSST) allant de 57,3 à 77,3 dB(A). Par la suite, en utilisant la norme ISO 1999, il a été estimé que, pour une exposition quotidienne de 85,2 dB(A), un individu se situant au 95^{ième} percentile sur une distribution normale de susceptibilité aux effets nocifs du bruit aurait un déficit auditif calculé, selon le *Règlement annoté sur le barème des dommages corporels* de la CSST [5], de 1,0 dB après une exposition de 1 an, et de 3,3 dB après une exposition de 40 ans [4].

4. CONCLUSION

La caractérisation des différentes sources de bruit de 10 autobus scolaires montre que le bruit du véhicule est, en moyenne sur l'ensemble des véhicules, de 66,3 dB(A) lorsque le véhicule est au repos, de 75 dB(A) lorsque le véhicule est en mouvement et atteint 78,5 dB(A) lorsque le véhicule se déplace sur l'autoroute. En plus du bruit provenant du véhicule, la contribution des autres sources de bruit a été évaluée et s'élève, en moyenne, à 78,1 dB(A) pour les élèves du niveau primaire, à 74,3 dB(A) pour les élèves du niveau secondaire, à 83,4 dB(A) pour le SRG (CB) qui fonctionne de façon intermittente et à 77,4 dB(A) pour l'équipement d'hiver.

L'étude dosimétrique effectuée sur les conducteurs a révélé que ceux-ci sont exposés à des niveaux équivalents quotidiens ($L_{EX, 8h}$, ISO 1999) variant de 74,6 à 85,2 dB(A). En utilisant la norme ISO 1999, il a été estimé que, pour une exposition quotidienne de 85,2 dB(A), un individu se situant au 95^{ième} percentile sur une distribution normale de susceptibilité aux effets nocifs du bruit aurait un déficit auditif calculé, selon le *Règlement annoté sur le barème des dommages corporels* de la CSST, de 1,0 dB après une exposition de 1 an, et de 3,3 dB après une exposition de 40 ans.

BIBLIOGRAPHIE

[1] Daoust, C. (avril 1993). Transport scolaire - Exposition au bruit des conducteurs. Service de santé au travail de Rimouski.

[2] Norme internationale ISO 1999 (1990). Acoustique – Détermination de l'exposition au bruit en milieu professionnel et estimation du dommage auditif induit par le bruit.

[3] Gouvernement du Québec (18 juillet 2001). Règlement sur la santé et la sécurité du travail. Gazette officielle du Québec, Partie 2, Lois et règlements, Volume 133, Numéro 29, pages 5020-5133.

[4] Marcotte, P., Boileau, P.-É. et Boutin, J. (mars 2004). Étude de l'exposition professionnelle au bruit des conducteurs d'autobus scolaires. IRSST / Rapport R-364.

[5] Commission de la santé et de la sécurité du travail (2000). Règlement annoté sur le barème des dommages corporels.

MUSIC TEACHERS' NOISE EXPOSURE

Dejan Zivkovic[1], **Peter Pityn**[1]
[1]OSHTECH Incorporated, 400 York Street, London, ON, N6B 3N2, dejanz@oshtechinc.com

1. INTRODUCTION

Music teachers are exposed to "noise" during the course of their activity. The size and activities of the classes vary, as well as the sound levels. Noise exposure can be classified as "occupational" and evaluated as such, because exposures are the result of the work environment, which some may dispute.

A survey was conducted to determine music teachers' professional noise exposure in secondary schools and to assess their risk of hearing loss. Noise exposures of 6 music teachers at six different secondary schools within the same board of education were measured.

The findings are part of an extensive survey, currently underway, which will continue into the coming school year.

2. METHOD

In each school two noise dosimetry measurements were obtained. One personal sample was collected from a music teacher and an area noise dosimeter was posted on a music stand (close the teacher, in front of the class). All music teachers were willing participants, informed about the objectives of survey. The test period was variable depending upon the individual teaching schedules (lasting from 1 to 5 hours).

Noise dosimetry samples were collected in accordance with the CSA Standard to quantify the time-varying noise levels. Dosimeters collected samples of A-weighted noise, over a range of 50-146 decibel (dB). The prescribed settings were followed - upper limit of 115 dB, slow response time constant, 90 dB criterion level, with 5 dB exchange rate and no lower threshold. This is the maximum permissible level of noise exposure set by the Ontario Ministry of Labour (MOL) established in Occupational Health and Safety Act, Regulation for Industrial Establishments.

Data were also collected as per other widely accepted criterion with an 85 dB exposure limit and 3 dB exchange rate. This criterion is recommended by National Institute for Occupational Safety and Health (NIOSH), as well as the American Conference of Governmental Industrial Hygienists (ACGIH), and is regarded by many health and safety professionals as a guideline for best practice.

Each dosimeter was pre- and post-calibrated with a Quest Electronics, Model CA-32, Permissible Sound Calibrator, which produces a 110 dB pure tone at 1000 Hertz.

During the survey information was collected about hearing protection practices, hearing protection equipment, teacher's concern about noise exposure, as well as individual history related to audiometry. Students played their instruments (flutes, trumpets, clarinets, saxophones, guitars, trombones). The teachers conducted music, instructed class and helped individual students. Often teachers needed to position themselves close to individual students, placing themselves in the middle of the instrument section, directly exposed to the sound from these sources.

Averaged noise levels (L_{avg}, L_{eq}) during the measurement period were recorded. The L_{avg} represents the average noise level based on a 5 dB exchange rate, and the L_{eq} is the equivalent noise level where the average is based on a 3 dB exchange rate. The normalized 8-hour exposure, termed time weighted average exposure (8 hr TWA), was determined to evaluate compliance with MOL requirements. The normalized 8 hr. noise exposure level (L_{ex}) was computed for comparison with the best practices guideline. In the laboratory the logged data, as well as the information from the discussion were downloaded into a database.

Band practice, which is typically held three times per week, also could significantly contribute to individual exposure. Band practice was not a part of this survey. Band practice is a voluntary extra – curricular activity and it may not fit the definition of "occupational exposure", even occurs in the workplace

3. RESULTS

The results of noise dosimetry are shown in Table 1. The findings indicate differences in noise dosimetry results obtained from personal and area samples and differences between teachers due to teaching schedules. In all cases the TWA exposure of the teachers complied with the Ministry of Labour Limit of 90 dB(A). Only half complied with the NIOSH standard of best practice - 85 dB(A) - for noise exposure averaged over 8 hours (personal samples at Schools 1, 2, 4, and 5).

Table 1. Noise Dosimetry Results

Sampling Description		Noise Exposure			
Sample	Duration (min)	L_{avg}	L_{eq}	TWA	L_{ex}
School #1					
Music teacher #1	208	89.0	91.5	86.3	87.9
Area sample #1	209	84.8	87.7	81.2	84.1
School #2					
Music teacher #2	248	92.6	97.8	89.7	95.0
Area sample #2	245	91.5	93.3	88.6	90.4
School #3					
Music teacher #3	71	87.9	90.2	79.6	81.9
Area sample #3	69	84.6	87.2	76.2	78.8
School #4					
Music teacher #4	255	88.5	91.1	85.8	88.3
Area sample #4	255	83.4	86.4	80.7	83.7
School #5					
Music teacher #5	310	89.9	93.1	88.0	91.2
Area sample #5	307	88.1	91.4	86.2	89.5
School #6					
Music teacher #6	245	82.0	84.9	79.1	82.0
Area sample #6	245	76.5	81.3	74.6	78.4

The samples from school #6 were not entirely representative, because of the classroom activities devoted to preparation for final exams.

As the noise in music room is usually intermittent then L_{avg} is several decibels less than L_{eq}. Specifically, the increase in L_{eq} levels is determined by the amount of high intensity "bursts" of noise. When classes play musical compositions in unison the flow of music determines the undulating characteristics of the sound, but when students are allowed to practice individually there is opportunity for more random "noise".

It can be concluded that noise exposure is influenced by the type of performed music activity and teaching approach.

4. DISCUSSION

The results show that there is a potential risk of hearing loss for music teachers. Exposures usually complied with the MOL limit of 90 dB(A), but not with the NIOSH standard of best practice - 85 dB(A). If high noise levels during band practice are included (usually three times per week), then it can be predicted that the risks would increase considerably. This may be controversial, because band practice is a voluntary activity. In other words, the noise exposure occurs in the workplace, but is not "part of the job"...or is it?

Measures should be implemented for reducing noise exposures. The key elements are: training, appropriate usage of protective equipment, audiometric hearing tests, periodic reviews, and written documentation (including record keeping).

Two specialized types of hearing protectors are available for music teachers and professional musicians: custom fitted earplugs (musician's plugs "ER" with 9, 15 or 25 dB NRR) and "ER 20" Hi-Fi earplugs. They have almost flat attenuation properties and they are designed to let the music teacher hear the full range of music. Training on the appropriate usage and care of the protective equipments is important.

Students also can be exposed to high noise levels while playing instruments, but the duration of their exposure is much shorter than the teacher's. It is important to point out that the music teacher is a role model and can set a good example for students by demonstrating appropriate usage of protective equipment.

Audiometry is an objective method for determining whether hearing loss is being prevented and for identifying individuals with established hearing loss. In this regard, audiometric records may be of significant importance pre-employment, as well as on a routine basis for monitoring the risks of hearing loss in music teachers. Periodic audiometry, noise dosimetry and reviews of hearing protection practices are important for effective hearing conservation programs.

Implementation of engineering noise controls (acoustical treatment of music rooms, teachers' offices and practice rooms) also can reduce noise and improve room characteristics. Cost-effectiveness is an issue.

5. REFERENCES:

1. Behar, A., MacDonald, E., Lee, J., Cui, J., Kunov, H., Wong, W. (2004). Noise Exposure of Music Teachers. J. of Occup. and Environ. Hyg., 1: 243-247.
2. Canadian Standards Associations (CSA) (1994). Proc. for the Measurement of Occupational Noise Exposure (Z107.56-94). Mississauga, Ontario, Canada: CSA.
3. Ontario Ministry of Labor (1998). Ontario Occupational Health and Safety Act; Regulations for Industrial Establishments, Section 139. Ontario, Canada. Occupational Health and Safety Division.
4. National Institute for Occupational Safety and Health (NIOSH) (1998). Criteria for a Recommended Standard: Occupational Noise Exposure. Publication No. 98-126, Cincinati, Ohio, NIOSH.
5. American Conference of Governmental Industrial Hygienists. 2004 Threshold Limit Values and Biological Exposure Indices. Table 1, p. 104. Cincinnati, Ohio, ACGIH.

NEW METHOD AND DEVICE FOR CUSTOMIZING IN SITU A HEARING PROTECTOR

Jérémie Voix and Frédéric Laville

École de technologie supérieure, 1100 Notre-Dame Ouest, Montréal (QC) H3C-1K3, jeremie.voix@etsmtl.ca

1. INTRODUCTION

To address the issues of discomfort (from both a physical and perceptual perspective) and unknown performance of existing hearing protectors, a new concept of a re-usable earplug has been developed [1,2]. For physical comfort this earplug is instantly custom-fitted with soft biocompatible silicone rubber [3]. A sound bore through the earplug is used two ways: first, for the measurement of the sound pressure level inside the earplug (see [2] and [4] for the assessment of the earplug attenuation) and, second, for the filtering of the earplug with passive acoustical dampers (see [1] for the customization of the attenuation to limit the speech and warning signals degradation). This paper will present how all those features have been integrated over the last 2 years into a field solution for customizing, in situ, a hearing protector.

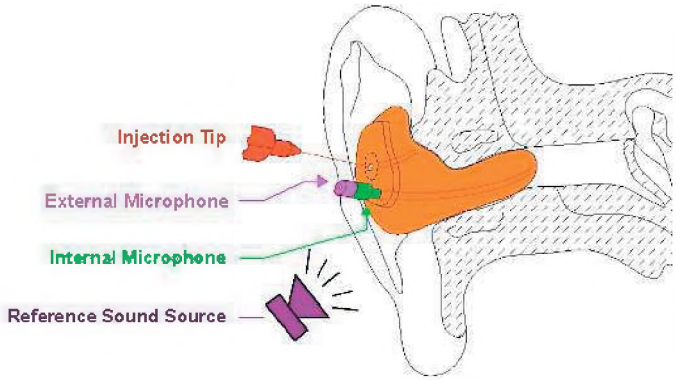


Fig. 1. Schematic Overview of the Custom Earplug and the Measurement Device

2. IN-SITU ATTENUATION MEASUREMENT DEVICE

Field Calibration of the Device

Daily Calibration Check

This check is performed prior to any use of the device. The dual microphone probe is clipped in the middle of the Reference Sound Source speaker grid (acoustic near field), and a Transfer Function (TF) measurement is performed between the External and Internal microphones while the sound source is generating a moderate pink noise level. This check ensures that the Sound Source is functioning properly (since the Sound Pressure Level measured at the reference microphone must be within a given range) and that both microphones are working correctly (since the measured TF must respect a given template). The TF is stored for later usage.

Microphone Probe Length Correction

The acoustical length of the microphone probe tube is increased when the probe microphone is inserted inside the earplug sound bore. This added length is merely a function of the earplug size and the associated level correction can be stored per octave band in a table to be used later on.

In-Situ Attenuation Measurement

Noise Reduction Measurement

Once the fitted earplug has cured, a Noise Reduction (NR) measurement is performed on the earplug: a loud pink noise is generated with the reference sound source and the Transfer Function between the 2 microphones is computed including the corrections from the daily calibration TF and the sound bore length correction. A method presented by the authors [2] is used to predict the REAT (Real Ear Attenuation at Threshold) by octave band from this NR measurement.

Predicted Personal Attenuation Rating

From this REAT prediction, a new indicator has also been proposed [5]: the P-PAR (Predicted Personal Attenuation Rating). The P-PAR is comparable to an NRR (*Noise Reduction Rating,* a single number value representing the attenuation [6]) but it is obtained from an objective measurement (not from a subjective evaluation), on a particular user (not on a population sample), under realistic wearing conditions for the hearing protector (not under laboratory conditions).

3. CUSTOMIZING IN SITU THE EARPLUG

Adaptation of the Earplug Attenuation

Attenuation of the Filtered Earplug

The use of an acoustical damper creates a second sound path that has been quantified using REAT measurements with and without the damper in place following ANSI standards [7]. The prediction of the *filtered* earplug attenuation can then be obtained by combining the predicted attenuation of the *unfiltered* earplug with the quantified attenuation of the damper. This latter value is essentially function of its acoustical resistance which is guaranteed by the manufacturer to be very stable between samples The attenuation of the filtered earplug can consequently be reliably predicted per octave band enabling filter selection which is described in the following paragraph.

Filter Selection

Following existing recommendations [8,9], the filter that leads to a protected exposure level ranking between 70 and 85 dBA will be selected among a set of 5 available filters as presented by the authors [1].

Performance Tests

Respectively from the rough NR measurement (see "Noise Reduction Measurement" section), the PPAR computation (see "Predicted Personal Attenuation Rating" section) and the Protected Exposure Level (see "Attenuation of the Filtered Earplug" section), three performance tests are undertaken: the first is an *Acoustical Seal Test* that checks if the earplug provides a proper and comfortable fit; the second is a *Rating Test* that checks if the earplug is efficient and provides at least the NRR it has been rated for (thereby ensuring that the earplug does not need to be "de-rated"), and the third is a *Protection Test* that checks that the filtered earplug offers an amount of protection that matches the user's noise exposure, therefore avoiding overprotection and the resulting voice and warning signals intelligibility degradation. These performance tests are undertaken a first time right after the earplug has been fitted (*Post-Curing Test*) and a second time after the earplug has been removed and replaced by the user himself (*Subject-Fit Test*); the comparison between the two not only ensures the robust prediction of field performance but is also an effective indicator that helps training the user to properly insert the earplug.

4. CONCLUSION

The recent developments of a method and the associated hardware/software system for a custom-fitted earplug have been integrated into a comprehensive in situ approach that includes in situ attenuation measurements, adaptation of the attenuation with an acoustical damper, and several performance tests. This is a first step towards meeting the goal established by NIOSH (National Institute for Occupational Safety and Health) of finding a way for workers to be individually fitted and to offer them increased comfort and the ability to hear speech and warning signals [10] while being protected.

ACKNOWLEDGEMENTS

The support of Sonomax Hearing Healthcare Inc, IRSST (*Quebec Occupational Health and Safety Research Institute*) and NSERC (*Natural Sciences and engineering Research Council of Canada*) is gratefully acknowledged.

REFERENCES

1. Voix, J., F. Laville, and J. Zeidan. *Filter Selection to Adapt Earplug Performances to Sound Exposure*. in *Acoustics Week in Canada PEI 2002*. 2002. Province of Prince Edward Island (PEI).
2. Voix, J. and F. Laville. *Expandable Earplug With Smart Custom Fitting Capabilities*. in *InterNOISE 2002*. 2002. Dearborn, MI, USA.
3. McIntosh, I. and R. Saulce, *Expandable in-ear device*. 2004, Sonomax Hearing HealthCare Inc.: United States Patent.
4. Voix, J. and F. Laville, *Method and apparatus for determining in situ the acoustic seal provided by an in-ear device*. 2004, Sonomax Hearing Healthcare Inc: United States Patent.
5. Voix, J. and F. Laville. *Problématiques associées au développement d'un bouchon d'oreille "intelligent" (Problematics associated with the development of a "smart" earplug)*. in *RRSSTQ-Congrès annuel de l'ACFAS*. 2004. Montréal,QC, Canada.
6. EPA, *Noise Labelling Requirements for Hearing Protectors*, in *Fed. Regist. 44(190), 40CFR Part 211*. 1979.
7. ANSI, *American National Standard Method for the Measurement of Real-Ear Protection of Hearing Protectors and Physical Attenuation of Earmuffs*. 1974, American National Standard Institute: ANSI S3.19-1974 (R 1990).
8. CSA, *Hearing Protection Devices - Performance, Selection, Care, and Use*. 2002, Canadian Standards Association: Z94.2-02.
9. EN458:1993, *Hearing Protectors - Recommendations for selection, use, care and maintenance - Guidance Document EN 458, CEN (European Committee for Standardization)*. 1993.
10. NIOSH, *Criteria for a Recommended Standard-Occupational Noise Exposure*. 1998, (National Institute for Occupationnal Safety and Health).

LABORATORY ACCREDITATION OF THE ACOUSTICAL STANDARDS PROGRAM AT THE INSTITUTE FOR NATIONAL MEASUREMENT STANDARDS

George S. K. Wong, Lixue Wu and Won-Suk Ohm
Institute for National Measurement Standards, National Research Council Canada,
Building M-36, Montreal Road, Ottawa, Ontario K1A 0R6

1. INTRODUCTION

National Metrology Institutes (NMIs), such as NIST in the USA and NRC in Canada and similar national laboratories in other countries should have a Quality System that defines how business (calibration & measurements) is conducted. The ISO/IEC 17025: 1999 "General Requirements for the Competence of Testing and Calibration Laboratories" is the current "Bible" for laboratory assessments.

The question come to mind is that "Why do we need laboratory assessments?" In the old days "Trust us" was the answer, but nowadays, the client may ask "Is the laboratory accredited?" Internationally, there is a higher body: Consultative Committee on Acoustic Ultrasound and Vibration (CCAUV), under BIPM, locates in Sèvres cedex, France, that dictates the behaviour of its memberships, and most NMIs made substantial contributions to maintain their memberships. One of the requirements of CCAUV is that national laboratories must support their claims of calibration and measurements capabilities (CMC) by inter-comparison and with assessment and accreditation of their laboratory calibration procedures.

2. QUALITY SYSTEM FOR A NATIONAL METROLOGY INSTITUTE (NMI)

ISO/IEC 1705 specifies a quality system that covers from "cradle to grave" for calibration laboratories. It may start from the technical expert telephone first contact with a client, the arrival of the devices to be calibrated to the return of the devices to the client. It is essential to have "paper trails" of nearly everything. The Institute for National Measurement Standards has standard procedures that cover the relatively mundane part of the requirements, such as how to handle the instruments when they arrive, but the technical experts of the Institute must ensure that technical requirements are satisified.

3. INSTRUCTIONS TO THE STAFF

One must instruct the staff that spending lots of time in preparing for the accreditation is essential for international recognition but admits that these are "non productive" work. It is not even recognised as professional development or as excuses for "promotion".

The procedures developed for calibration require updating from time to time, but do not assume that these work can be devoted as one's career just to service the procedure manuals.

4 PREPARING THE MEASUREMENT PROCEDURES

The best person to draft the calibration procedure is the person who is going to perform the calibration. The Acoustical Standards Program included some brief theory behind the method used. This will also be a training manual for new employees in future, and to inform the technical assessor who may not be too familiar with the laboratory practice. After the draft and the revised drafts are written, one should have a "dry run" to confirm the steps of calibration. One will be surprised on what one may turn up. If the calibration procedures include some "checking" requirements such as to ensure the FFT analyser is functioning, these "checking" must be documented and noted in the procedures.

5. VALIDATION OF A BLACK–BOX INSTRUMENT

The question on how to validate a black-box such as an FFT analyzer is very difficult to answer. To satisfy the assessment it is necessary to find a simple test that the solution is known. For a FFT analyzer, use a known wave-form such as a square wave with harmonics easily calculated.

6. REDUCTION OF PAPER WORK FOR FUTURE YEARS

It is mandated to update the calibration procedures when there are changes. For example: When new equipment

(a microphone or a DVM) is pressed into service, one must issue a new procedure to reflect the change in the equipment list. It is reasonable to include a statement such as "If the new equipment has similar specifications, the Program Leader can approve the use of the new equipment". This will save a lot of paper work in the future.

7. STAFF TRAINING RECORDS

It is a requirement to have an up-to-date record on staff training that includes specifics on which staff has been trained to perform which calibration. The importance of this can be illustrated with an example:

During the assessment, a question was asked on how the Program Leader, who signed calibration reports, was sure that the calibration results were correct. The initial answer was based on scientific facts: There were four sets of repeated data that were very similar, and the curves plotted were smooth and normal, etc. Not acceptable. The same question was repeated, and the second attempt to answer was that the Program Leader could repeat the measurements himself but he might make mistakes too. Not acceptable.

It turned out that the correct answer should be: There is a training record, signed by the Program Leader, for that particular staff member, and therefore the above calibration results should be correct.

On reflection on the above, one may wonder why scientific evidence is inferior to a signature on paper.

8. SOME SALIENT POINTS THAT MAY BE HELPFUL

If computer programs are used in a calibration, be sure to have an answer on how to verify that the computer programs are correct.

When that program is loaded from a computer file, how can one be sure that it is the correct version of the program? The problem was solved by loading the computer program from a disc every time, and not from the list of programs that reside in the computer.

If the data during calibration are stored in a computer, be sure to use a unique file name to identify that calibration. For example: file names with serial numbers, month & day are not good enough.

9. CONCLUSIONS

A short summary of an external accreditation is presented. Things to remember include paper trails, documentation, and understand how the assessor interprets ISO/IEC 17025. It is known that assessors may have different interpretations. Based on our experience, one may need to issue 30 % to 40 % more documents than our original set of prepared procedures. One final suggestion for those laboratories contemplating ISO/IEC 17025 accreditation: Do not depart from the final focus on accreditation. To resist is futile, comply.

REFERENCE

[1] IEC/ISO 17025, General requirements for the competence of testing and calibration laboratories, 1999.

ACOUSTIC AUDIT OF ENVIRONMENTAL NOISE IMPACT ASSESSMENT

C. A. Krajewski
Ontario Ministry of the Environment, 2 St. Clair Avenue West, 12th Fl. Toronto, Ontario, M4V 1L5
krajewch@ene.gov.on.ca

1. INTRODUCTION

Acoustic audit of environmental noise impact is an investigative procedure consisting of measurements or a combination of measurements and acoustic modelling of noise emissions due to operation of the audited sources (equipment/facilities), carried out, primarily, to compare the measured/modelled noise emissions with the design objectives.
However, the purpose of an acoustic audit may be multifold as outlined below:

a) to establish background sound level which is then used as a criterion of acceptability in the certification process for equipment/facilities that contain significant sources of noise emission;
b) to assess the magnitude of noise impact due to noise source(s) operation;
c) to verify if the level of noise emissions is not exceeding the applicable performance limit;
d) to evaluate effectiveness of noise abatement programs for existing noise sources, and
e) to determine if noise control measures applied to proposed, new noise source(s) are adequate to ensure design objectives.

In the course of environmental noise impact assessment the acoustic audit is typically required when equipment/facilities contain complex noise sources displaying a wide range of temporal and spectral variation in the noise emission pattern, when various empirical models are available for prediction of sound emission levels of a particular noise source and when the outcome of noise impact analysis indicates marginal compliance with the design objective (performance limit).

The term Aaudit@, in the context of environmental noise impact assessment, signifies a follow-up examination or verification process. For the purpose of equipment/facilities certification such verification should be carried out by an impartial, neutral investigator (a qualified acoustic consultant) having no prior association with the project being audited and, specifically, not involved in the original assessment of noise emissions and/or the design/implementation of the required noise control measures.

- calculation of sound level due to noise source(s) subject to the acoustic audit (excluding the contribution of the

Findings of the acoustic audit are documented in an acoustic audit report, which may be required as a condition of certification, and the audit performed in accordance with the methods and procedures accepted by a regulatory agency involved in the certification process.

2. METHODS

A number of factors may determine the selection of a suitable method for performance of the acoustic audit. The predominant factors are; operating conditions of noise source(s), the nature of acoustic environment at point(s) of reception, availability of measurement instrumentation and the range of measurement instrumentation capability. The following four methods are typically applied for the performance of the acoustic audit:
- direct measurements at point(s) of reception;
- source(s) sound pressure level measurements combined with acoustic modelling;
- sound intensity measurement combined with acoustic modelling, and
- long-term monitoring and data analysis of sound levels at point(s) of reception.

All four methods outlined below focus on verification of compliance with the applicable performance limit. A brief description of each method, including applicability, instrumentation and reporting requirements as well as limitations is provided:

2.1 Direct Measurements

This method can be applied when contribution of continuous, steady noise emissions from noise source(s) in a facility is dominant, and the effect of extraneous noise sources other than vehicular traffic is insignificant at the point(s) of reception.

Depending on characteristics of noise source(s) emissions, Class 1 or Class 2 integrating-averaging sound level meter with impulse and frequency analysis capability is required.
Application of this method requires:
- measurement of the overall noise emissions with noise source(s) in full operation at selected point(s) of reception [1]

established background sound level determined in accordance with) [2].

2.2 Source SPL Measurement and Acoustic Modelling

This method, combining SPL measurements at close distance to significant noise source(s) and acoustic modelling of propagation path from the intermediate measurement location (close to the source) to selected point(s) of reception, is typically applied when there is a significant contribution of extraneous sources to acoustic environment, source(s) operation is intermittent or cyclical and noise source(s) subject to acoustic audit are in continuous operation. Application of this method requires:

- measurement of individual source(s) noise emissions (SPL) at close distance following a standardized measurement procedure (such as one described in reference [3]). The measurement distance to each source must be selected outside the acoustic near field and must also ensure that noise emissions of the measured source are dominant over other noise sources in the vicinity.
- calculation of sound power level of the source(s) based on the results of SPL measurements.
- acoustic modelling of propagation path from the measurement location to selected point(s) of reception) for each noise source. [4]
- calculation of an aggregate contribution from all significant sources in a facility to the overall SPL at point(s) of reception.

2.3 Sound Intensity Measurement and Acoustic Modelling

In this method, sound power level of noise source(s) is determined through measurement of sound intensity of noise radiating surfaces for significant noise source(s).[5] The remaining steps required to determine if noise emissions of the audited noise source(s) conform with the applicable performance limit are identical to those outlined in 2.2 and in the Summary.

The main advantage in using this method is that it enables quantifying the PWL of sources in-situ in close proximity of other noise sources.

Due to complexities involved in sound intensity systems data acquisition and signal processing, application of this method is limited to experienced users fully trained in the technology and operation of the system, capable of interpreting measurements to provide meaningful results.

2.4 Long-term Monitoring and Data Analysis of Sound Levels at Point(s) of Reception

This diagnostic method may be applied when noise emissions due to audited noise source(s) are steady and continuous in a background of non-steady, intermittent extraneous sources including other industrial/commercial facilities and road/rail traffic. It requires a detailed analysis of acoustic data collected from a long-term monitoring survey, carried out simultaneously at close distance to a facility containing audited noise sources and at selected point(s) of reception. Class 1 or Class 2 sound level meter capable of continuous monitoring and data logging of the equivalent sound level $L_{eq}$and ninety percentile sound level L_{90} values is required. Implementation of this method is based on:

- analysis of the monitored values of the hourly equivalent sound level L_{eq} and the nineties percentile sound level L_{90} distribution..
- acoustic modelling of propagation path from the monitoring location(s) close to the facility to selected point(s) of reception [4].
- comparison of the calculated (through acoustic modelling) sound levels with the monitored L_{eq} and L_{90} data at points(s) of reception.
- establishment of a pattern (consistency) between the modelled and monitored L_{90} values. The results of comparison analysis must clearly demonstrate that the monitored levels are representative of contribution of noise emissions from the facility.

3. SUMMARY

Application of each method requires an assessment (monitoring or prediction) of the background sound levels in accordance with an accepted procedure [1][2], adjustment of the measured levels for special quality of sound, if applicable, and a comparison of the calculated resultant sound level due to the operation of the audited noise sources with the applicable performance limit.

Each of the above outlined methods may be suitable for a particular noise environment. In all cases it is important to demonstrate why a particular method was applied and to describe how it was implemented.

REFERENCES

[1] ISO 1996-2, "Acoustics - Description, assessment and measurement of environmental noise, Part 2 Determination of environmental noise levels."

[2] ORNAMENT, Ontario Road Noise Analysis Method for Environment and Transportation, Technical Document, Ontario Ministry of Environment, ISBN 0-7729-6376, 1989.

[3] ISO 3744:1994, "Acoustics - Determination of sound power levels of noise sources using sound pressure - Engineering method in an essentially free field over a reflecting plane."

[4] ISO 9613-2:1996, "Acoustics - Attenuation of sound during propagation outdoors - Part 2: General method of calculation."

[5] ISO 9614-2:1996, "Acoustics -Determination of sound power levels of noise sources using sound intensity - Part 2: Measurement by scanning."

MEASURING WITH SOUND LEVEL METERS: THE REAL DIFFERENCE BETWEEN TYPE 1 AND TYPE 2

Richard J. Peppin
Scantek, Inc., 7060 Oakland Mills Rd #L, Columbia, MD 20146, PeppinR@asme.org

1. INTRODUCTION

I will briefly discuss the differences between Type 1 and Type 2 meters, based on ANSI ands IEC standards. However, knowing the differences in instrumentation tolerances, we really do not know how they translate into differences in acquired data. I will discuss the differences in results for two meters, exactly the same except for Type designation.

There are differences between the parameters that instruments: essentially they lie in the differences in instrumentation tolerances between Type 1 and Type 2. The output depends on the spectrum and frequency response of meter. This paper reports the result of a "theoretical study," based on numbers alone, and of a simple measurement test.

So some issues include: What does "frequency range" mean and what is the effect of the frequency range on sound level meter accuracy and it's ramifications for measurement? And what does the "Type" of sound level meter have to do with accuracy of measurements?

2. FREQUENCY RESPONSE OF SOME METERS

To give some idea as to the frequency response reported in the literature of some manufacturers, Table 1 shows data gleaned from product data sheets:

Meter	Type	Range (Hz)
CESVA SC-160	2	31-16k (OB)
RION NA-26	2	20-8k
RION NL-22	2	20-8k
RION NL-32	1	20-20k
Norsonic N-118	1	6.3-20k
Norsonic N-121	1	0.1-20k
LD 812, 820	1	?
LD-824	1	2.5-20k (1/3rd OB)

Table 1 Claimed frequency ranges of some meters

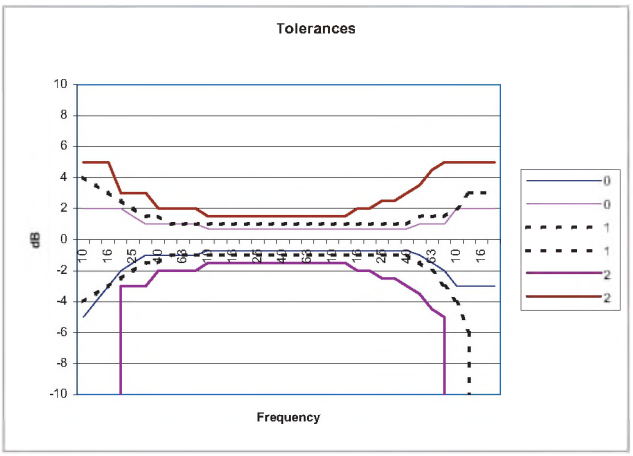


Figure 1 Tolerances on meters per ANSI

When discussing tolerances, the meter and the microphone together must meet specs. And the most inaccurate component is probably the microphone because it is electromechanical

The A-weighting filter has tolerances associated with it. The frequency response is

10 - 20k Hz in S1.4-1983
10 - 100k Hz in S1.4A-1985

The sum of all energy in the band, and weighted per the curve, is the range for which the "dBA" is meant. Any other range (like 50 to 800 Hz) is not the same. It is a partial weighting. That is for a sound measurement any differences between the A-weighted levels measured with two meters depends on the spectrum and the meter response. A-weighted value is the energy modified by the A-weighted curve. This brings up a really basic, if exaggerated, question: What does it mean to measure a smaller frequency range? Suppose you measure only between1000 Hz and 1250 Hz 1/3rd OB data. Is that somehow an A-weighted number?

3. COMPARISON OF METERS

I did a theoretical modeling of noise and a field test with three different meters to evaluate differences.

2.1 Modeling Procedure

I put random noise into meter simulated by random numbers into a mathematical model of an A-weighting filter of different bandwidths. The energy sum, Correct for A-wtd. Was calculated and I computed the difference between highest and lowest values of A-wtd data for a given (random) spectrum. This was done 30 times. Figure 2 shows the results of a series of broad band spectra for different meter types. Figure 3 shows, for 30 measurements, the means, standard deviations, as well as a maximum range for the results.

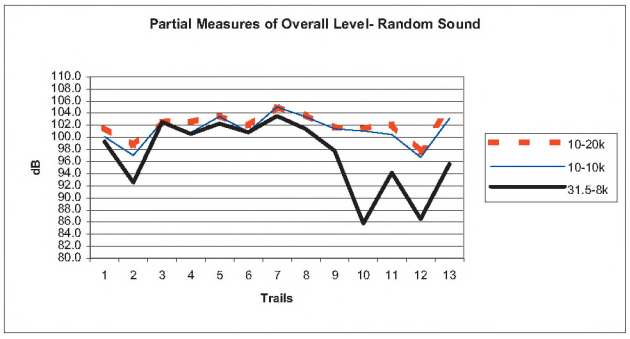


Figure 2 Partial measures of Overall level

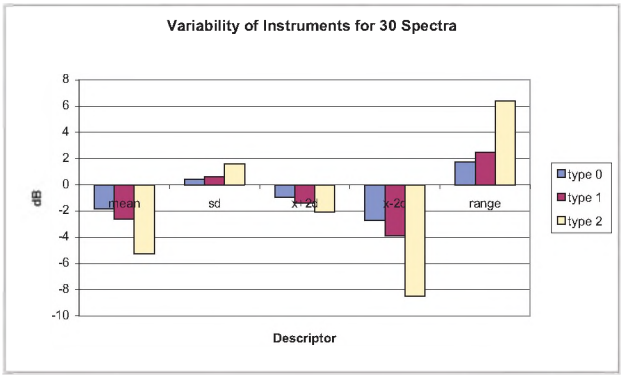


Figure 3 Difference between Types for 30 measurements

2.2. Field test:

I used three integrating-averaging sound level meters of identical sizes, identical preamps, presumed almost identical electronics, different microphone, A-weighted and measured 10 Min Leq with Autostore. The meters were: RION NL-21, Type 2, RION NL-22, Type 2, and the RION NL-32, Type 1. Placed about 12 cm apart and 1-m from ground, I measured traffic noise about 15 m from the right-of-way. Table 2 shows the results of six, 10-min samples.

	NL-21	NL-22	NL-32
1	66.3	66.5	66.4
2	66.6	66.7	66.6
3	67.3	67.5	67.3
4	65.8	66.0	66.0
5	67.2	67.4	67.3
6	68.0	68.2	68.1
Hourly Average	66.9	67.1	67.0

Table 2 ten-minute Average sound level for three meters.

4. DISCUSSION AND CONCLUSIONS

4.1 Discussion

The oral presentation contains more details on both modeling and field tests. The field tests included attempting to determine the effects of microphone windscreens. There is a lot of uncertainty present, some we know, from the results of a traceable calibration and some we don't because we have no apriori and quantitative information about the sound and environment.

4.2 Conclusions

The errors in sound level meters are based on more that we normally think: The tolerance by the manufacturer, the bandwidth of the signal (which may be different for the same type of instrument, the actual signal in the sound, and the orientation of the microphone compared to the sound wave, if indeed we even know where the sound is coming from (discussed a paper long ago.) Also,

Type 2 is the just about the same as a Type 1
- –For this level of uncertainty
- –For this type of noise
- –For this type of sound level meter
- –For the usefulness of the results

Type 2 is better than type 1
- –Less costly
- –Less expensive to break
- –Calibrator less expensive
- –Microphone is less expensive

Type 1 is better than type 2
- –For low noise environments (<20 dB v. < 30 dB or more)
- –For accurate measurements of low frequency or high frequency noise (where tolerances can give differences of + 4)
- –A Type 1 meter is a Type 2 meter

ANGULAR RESPONSE OF A SOUND INTENSITY PROBE AT HIGH FREQUENCIES

S.E. Keith [1], and V. Chiu[2]
[1,2]Health Canada, Consumer and Clinical Radiation Protection Bureau, 775 Brookfield Rd. 6301B, Ottawa, ON, K1A 1C1
[1]skeith@hc-sc.gc.ca, [2]sound@hc-sc.gc.ca

1. INTRODUCTION

A-weighting is used in the assessment of noise related to human health. The dominant frequencies in this weighting are from 630 Hz to 10 kHz. Yet, in most current standards, intensity measurements are restricted to frequencies below 6.3 kHz due to concerns related to finite difference approximation errors. Using a numerical simulation, Jacobsen[1] has shown that measurements are possible up to 10 kHz with a ½ inch p-p intensity probe and 12 mm spacer, with diffraction effects compensating for finite difference approximation errors. These findings were also demonstrated experimentally by Keith[2, 3].

In this study, the polar response of a ½ inch p-p intensity probe and 12 mm spacer was measured in 1/3-octave band frequencies. In agreement with Jacobsen's conclusions, the angular response of the probe showed no evidence of finite difference approximation errors. At frequencies up to 10 kHz, the intensity approximated the ideal cosine dependence on angle.

The results also showed an intensity probe could provide superior directivity compared to a ½ inch microphone.

2. METHOD AND APPARATUS

Measurements were made in the large (13 x 9 x 8 m^3) hemi-anechoic chamber at the Consumer and Clinical Radiation Protection Bureau. The measurements were made under hemi-anechoic conditions above a 9 x 13 m^2 hard floor constructed from 3 cm thick concrete tiles.

Measurements were made using GRAS 26AA preamplifiers and either ½ inch Brüel & Kjær (B&K) 4181 or GRAS 40AI 12 mm microphones with 12 mm spacer. Supplementary measurements used a single B&K 4165 free field microphone. A B&K type 2133 1/3-octave band Class 1 frequency analyzer was used for direct intensity and mean pressure measurement up to 10 kHz. A Hewlett Packard 35670A FFT analyzer was used for simultaneous measurements up to 25.6 kHz with 400-line resolution. Intensity was calculated from the cross spectrum [4], and mean pressure was calculated from the two channel auto spectra and inter-microphone phase. Bursts of 32 second duration pseudo random pink noise from the B&K analyzer were reproduced by a 25 mm diameter dome tweeter flush mounted in the concrete floor.

Calibration levels were checked using a B&K 3541 intensity coupler with pistonphone. This coupler was also used with a pink noise source to check the pressure-residual intensity index, δ_{pI0}, of the instrumentation[5]. Checks were made before and after measurements. In 1/3-octave bands, the minimum measured δ_{pI0} was greater than 18 dB for 1/3-octave bands ranging from 1 kHz to 10 kHz.

Estimation of δ_{pI0} from the FFT measurements relied on values obtained when the probe was oriented at 90° to the incident sound[6]. The minimum measured δ_{pI0} of the instrumentation was 14 dB for 1/3-octave bands from 1000 Hz to 12.5 kHz. These δ_{pI0} values were lower than above because the measurement chain was arranged to compensate for a known phase mismatch in the B&K analyzer.

Figure 1 shows the probe configuration with two ½ inch microphones on either side of the centre spacer (black), one 6 mm diameter preamplifier on the left, and another preamplifier on the right seamlessly joined to the 6 mm support rod. The probe was mounted 2.44 m above the tweeter sound source in a frame made of 6 mm diameter steel rods. Probe angle was varied from 0° to 95° in 1° increments using a B&K 9664 five axis microphone-positioning robot attached via cables. Maximum angular errors were less 2°. The alignment was checked i) at 0° by reflecting a laser off the microphone diaphragm, ii) at 90° by comparison with the δ_{pI0} calibration [6], and iii) at other angles by measuring the time delay between the probe and an additional microphone mounted on the rotating part of the frame.

3. RESULTS AND DISCUSSION

Comparison of Figures 1 and 2 shows small differences in the directivity of the B&K and GRAS microphone pairs. This suggests that the directivity may depend on small differences in probe configuration. Preliminary experiments also suggested similar differences due to preamplifier position and microphone body length.

At 10 kHz the intensity response is close to ideal. At 5 kHz, however, Figures 1 and 2 show that for angles exceeding 60° the intensity drops 1 dB relative to the ideal response. Angles over 60° are associated with half the sound power radiated from an omnidirectional source in half space. This suggests that sound power measurements at 5 kHz using large planar measurement surfaces could give lower results

than obtained with the same measurement over a hemispherical measurement surface.

Above 60° at 12.5 kHz the differences from the ideal intensity responses are 3 to 4 dB in Figures 1 and 2. The worst-case measurement error in the intensity would also be 3 to 4 dB, and could occur with a highly directional source and planar measurement surface (ignoring effects of background noise). At 12.5 kHz, a 3 to 4 dB difference is also found in Figure 3, which shows the response of a ½ inch B&K 4165 free field microphone [7]. This would produce a similar magnitude worst-case error in the sound pressure level, (for a highly directional source and planar measurement surface in a free field).

Compared to a ½ inch pressure microphone, below 10 kHz, the mean pressure response of the intensity probe is more omnidirectional. Juhl and Jacobsen [8] have shown that the mean pressure directivity can be improved to 10 kHz using a weighted response from each microphone, i.e., $p_{mean} = x \cdot p_A + (1-x)p_B$ (where p_{mean} is the mean pressure, p_A, and p_B are the two time domain microphone pressure signals, and x is the weighting). Mean pressure is typically calculated using x =0.5 as shown in the left side plot in Figure 4. The right hand plot shows significant improvements are possible up to 16 kHz with x = 0.7.

4. CONCLUSIONS

This study has shown a ½ inch p-p intensity probe can approximate the ideal intensity response up to 10 kHz. The same probe can also have a more omnidirectional response than a ½ inch pressure microphone up to 16 kHz.

REFERENCES

1. F. Jacobsen, V. Cutanda and P.M. Juhl, A Numerical and Experimental Investigation of the Performance of Sound Intensity Probes at High Frequencies, JASA, 103(2), 953-961, (1998).
2. S.E. Keith, G. Krishnappa, Sound power of noise sources using sound pressure derived from sound intensity measurements, Internoise 2002, paper 267, Detroit, (2002).
3. S.E. Keith, G. Krishnappa, V. Chiu, The Potential for Extending the Frequency Range of Sound Intensity Standards, Euronoise 2003, Naples, paper 169 (2003).
4. F.J. Fahy, Sound Intensity, Elsevier Applied Science, London (1989).
5. ISO 9614-1 Acoustics - Determination of sound power levels of noise sources using sound intensity - Part 1: Measurement at discrete points, (1993).
6. S. Keith, G. Krishnappa, V. Chiu, Dynamic Capability of Sound Intensity Measurements at High Frequencies, Internoise 2003, paper N94, (2003).
7. Condenser Microphones and Microphone Preamplifiers for Acoustic Measurements, Brüel & Kjær Data Handbook, (1982)
8. P. Juhl and F. Jacobsen, Sound Pressure Measurements With Sound Intensity Probes, ICA 2004, paper We4.G.2, Kyoto (2004).

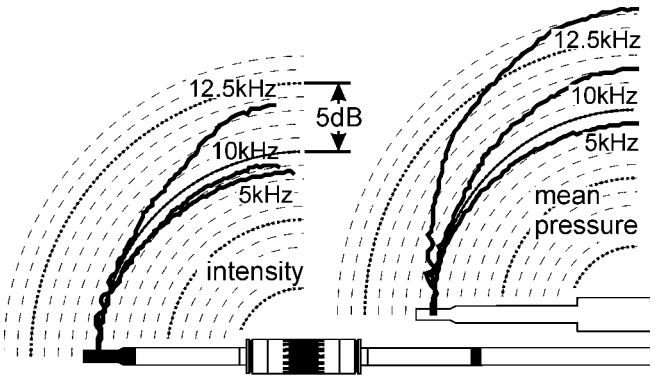


Fig. 1. Difference from ideal response for the intensity and mean pressure of the B&K probe (normalized to 0°, on axis, response).

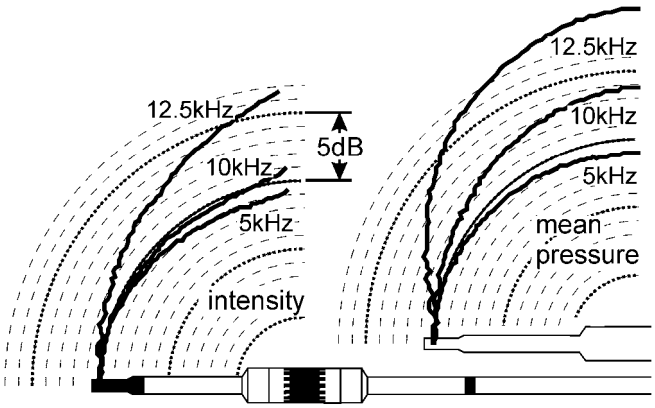


Fig. 2. Difference from ideal response for the intensity and mean pressure of the GRAS probe.

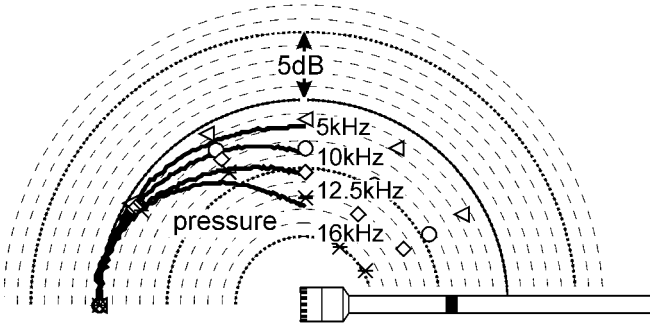


Fig. 3. Difference from ideal omnidirectional response for a free field microphone. Symbols are manufacturer's data.

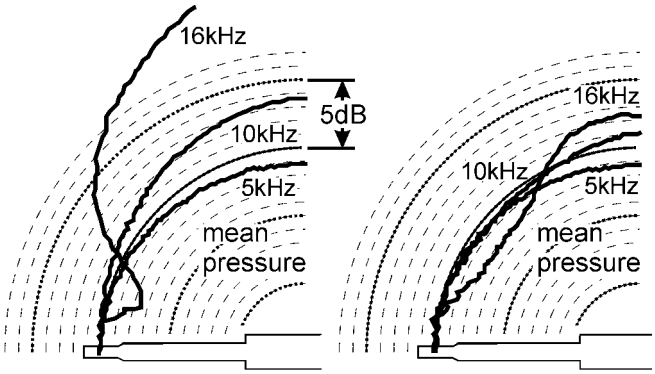


Fig. 4. B&K probe mean pressure response using weighted microphone signals: left side, x=0.5; right side, x=0.7.

THE EFFECT OF MASS LOADING ON THE SENSITIVITY OF SHOCK ACCELEROMETERS

Lixue Wu, George S. K. Wong, Peter Hanes, and Won-Suk Ohm
Institute for National Measurement Standards, National Research Council Canada
Building M-36, 1200 Montreal Road, Ottawa, Ontario K1A 0R6

1. INTRODUCTION

Mass loading affects the sensitivity of accelerometers. The mass loading effect can be corrected using mass loading correction curves published by manufacturers [1]. These curves, however, are only applicable to the sinusoidal acceleration below 500 m/s^2. The mass loading effect on the sensitivity of accelerometer Endevco 2270 in shock calibration, from 500 m/s^2 and up, is investigated using a vibrometer. The output waveforms of the accelerometer and vibrometer are compared in this study. Since the output waveform of the vibrometer is an integration of the acceleration waveform, small changes in the peak value of the acceleration waveform are masked by the acceleration waveform shape changes between measurements. Thus, the mass loading effect cannot be determined from the direct output of the vibrometer. Traditional differentiators that can convert the velocity output to acceleration suffer from noise. To overcome the noise problem, digital signal processing techniques are used. Existing filtering algorithms [2], [3] are not suitable for this application because of the rapid change of acceleration in shock calibration and the requirement for precision measurement.

The aim of this paper to present a new method for conversion of a velocity signal to an acceleration signal. With this method, the sensitivities of an Endevco 2270 accelerometer for two different mass loads at a shock level of 10,000 m/s^2 are measured. The mass loading effect in shock calibration for this accelerometer is then obtained.

2. METHOD

Signal quantization can represent a significant limiting factor in converting velocity signals to acceleration signals. Additional conversion errors are introduced when the velocity signal is subject to added noise prior to quantization. Reduction of these effects is the main task in determining the mass loading effect using the vibrometer output.

2.1 Data Acquisition

A data acquiring program is implemented using Visual Basic 6. Raw sample data from a digital oscilloscope Tektronix TDS 240L as shown in Figure 1 are transferred to a PC via GPIB and saved in a file that Matlab can read. The sample values are integers that correspond to the output of the A/D converter. The sample values are kept in integer format through rounding during the data processing to eliminate the numerical error introduced by calculations.

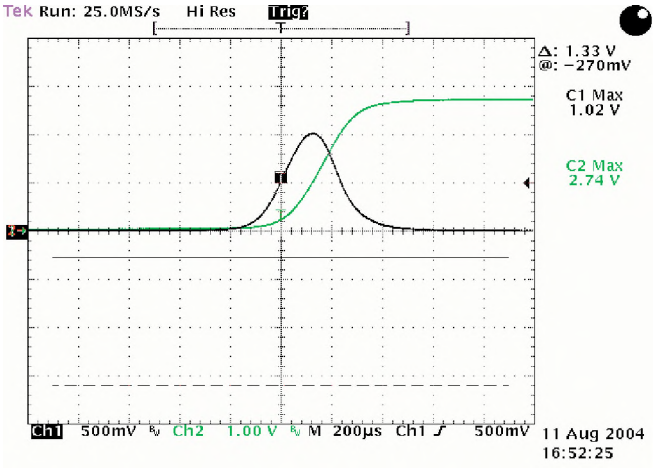


Figure 1: Acceleration (Ch1) and velocity (Ch2) signals.

2.2 Data Filtering

The velocity signal is first filtered through an analog low-pass filter (cut off, 5 kHz) prior to quantization. The quantized velocity signal (as an integer sequence) is then filtered through a median filter of variable length.

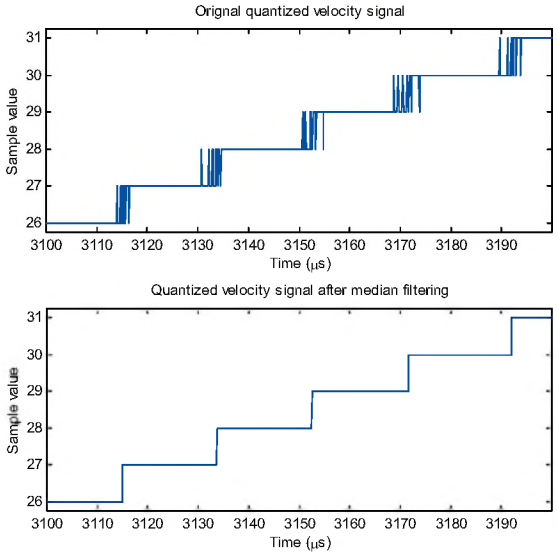


Figure 2: The effect of median filtering of a signal.

The length of the median filter is adaptively determined so that the power of impulsive noise in the filtered signal is minimized. Figure 2 shows a portion of the quantized velocity signal before and after median filtering.

2.3 Rate Estimation

The filtered velocity signal is close to the ideal output of an A/D converter: a stepwise function in time. The differentiation of such a signal creates zeros and impulses in the resulting acceleration signal. The stepwise signal could be smoothed before differentiating by either low-pass filtering or least squares curve fitting. This process, however, would result in a distorted acceleration signal and would generates a large error in acceleration peak estimation. A new method is proposed here, based on the assumption that the change rate (acceleration) of velocity within adjacent quantization levels is constant. That is, the stepwise signal can be replaced by a piecewise signal consisting of line segments. Each line segment connects the midpoints of two adjacent steps in the stepwise signal (velocity signal). The midpoint for a quantization level is estimated by finding a step at that level that has the greatest length. Other steps at that level are not considered, so that the noise has little effect on the midpoint estimation. As an example, the estimated change rates for a velocity signal using this method are shown in Figure 3 by crosses.

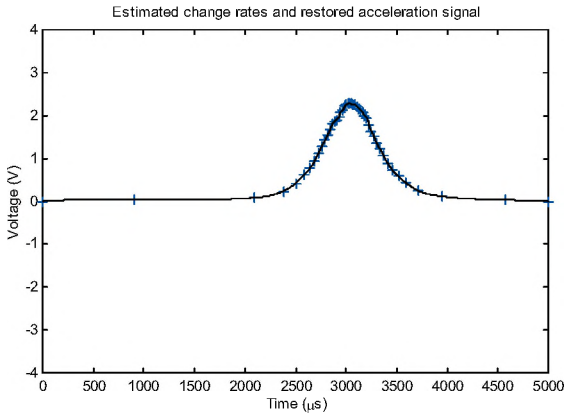


Figure 3: An example of a restored acceleration signal.

2.4 Acceleration Restoration

The acceleration signal is restored using cubic spline interpolation [4] from the discrete change rates estimated from the stepwise velocity signal. The restored acceleration signal using the estimated change rates is shown in Figure 3 by the solid line. The peak acceleration is then estimated from the restored acceleration signal.

3. RESULTS

Using the proposed method for converting velocity signals to acceleration signals, the sensitivities of an Endevco 2270 accelerometer (S/N CB18) for two different mass loads at are measured. The accelerometer is mounted on an anvil of a shock generator (Endevco 2925 POP) with the dummy mass loads on top of it. A vibrometer (Polytec OFV 050) is used to measure the velocity of the accelerometer. The output of the accelerometer, which is affected by mass loading, is compared with the acceleration signal converted from the velocity output of the vibrometer. The measured sensitivity of the accelerometer, normalized to the peak acceleration determined using the vibrometer, is shown in Figure 4 for 0 g and 200 g mass loads at accelerations of 10,000 m/s^2 ± 300 m/s^2. The dashed lines in Figure 4 indicate mean values for ten repeated measurements. The difference due to mass loading for this case is about 1.5%.

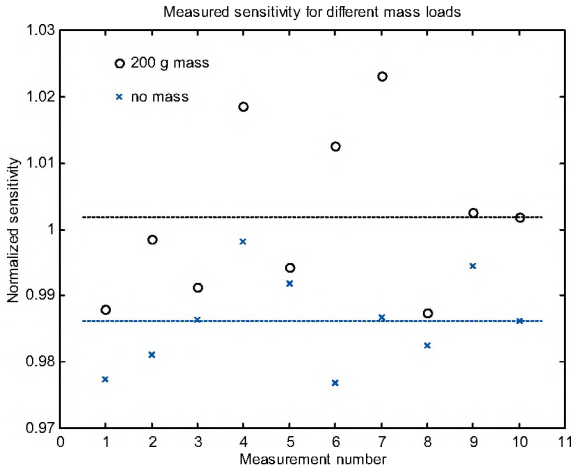


Figure 4: Measured sensitivity for 0 g and 200 g mass loads.

4. DISCUSSION

Mass loading affects the sensitivity of an accelerometer in high shock calibration. The variance of the measured sensitivity is mainly due to the resolution limitation (8-bit) of the A/D converter in the digital oscilloscope. Therefore, fine tuning of the proposed conversion algorithm may not lead to any significant improvement. The next task is to determine the mass loading effect at a number of different mass loads and shock levels.

REFERENCES

[1] Sill, Robert D. "Mass loading in back-to-back reference accelerometers," *Technical Paper 310*, Endevco Corporation.

[2] Marble, V. C., M. McIntyre, R. Hastings-James, and W. Hor, (1981). "A comparison of digital algorithms used in computing the derivative of left ventricular pressure," *IEEE Trans. on Biomedical Engineering*, vol. BME-28, no. 7, pp. 524-529.

[3] Vainio, O., M. Renfors, and T. Saramäki, (1997). "Recursive implementation of FIR differentiators with optimum noise attenuation," *IEEE Trans. Instrum. Meas.*, vol. 46, pp. 1202-1207.

[4] de Boor, C. (1978). *A Practical Guide to Splines*, Springer-Verlag.

SIMPLE *A PRIORI* SELECTION OF SOUND POWER MEASUREMENT STANDARDS

S.E. Keith
Health Canada, Consumer and Clinical Radiation Protection Bureau, 775 Brookfield Rd. 6301B, Ottawa, ON, K1A 1C1
skeith@hc-sc.gc.ca

1. INTRODUCTION

Workplace noise is a major occupational health problem resulting in thousands of disability claims per year. Noise emission declarations assist in the purchase of quieter machinery by enabling manufacturers to formally provide sound level data in an agreed format.

Canadian standard CSA Z107.58-2[1] contains guidance on provision and verification of noise emission declarations. The guidance is based on standards from the International Organization for Standardization (ISO), and is consistent with European Union (EU) Directives.

There are 10 basic ISO standards for measurement of sound power and 5 standards for measurement of emission sound pressure. The CSA document summarizes trade offs in measurement situations to provide the optimum solution for a given situation.

In this paper an excerpt of the CSA standard is explained, based on the relative dimensions of the machine and the measurement environment.

2. METHOD

The relative sizes of a noise source and measurement environment are often the most significant influence on the ability to make an accurate measurement. Figure 1 shows the required size of the room as a function of the absorption coefficient, α, for four ISO standards, and also includes background noise requirements. The accuracy of these standards is engineering grade with a standard deviation of ±1.5 dB in the A-weighted sound power level. The CSA guidelines refer to additional standards and provide details for characterizing room α values (used on the abscissa of Figure 1). The derivation of Figure 1 follows.

2.1 Free Field Environments

The normal intensity through an enclosing measurement surface is the primary quantity required for determination of sound power. Pressure measurements in free field sound power standards[2] overestimate this quantity due to background noise and room reflections. Background noise is discussed in clause 2.5. The magnitude of the remaining overestimate is given by K_2:

$$K_2 = 10\log\left(1 + 4S/S_V\alpha\right) \text{ dB} \qquad (1)$$

where S is the area of the measurement surface, S_V is the total area of the room surfaces, and α is the average sound absorption coefficient in the room.

It is difficult to visualize surface areas, so equation (1) was simplified using the average dimension of the room, L_{Vavg}, or source, L_{avg}. These values are simply the average of the length, width and height of the room or source. For the measurement surface, the minimum possible area is $S = 5L_{avg}^2$, and the absorbing area of the room boundaries is $S_V = 6L_{Vavg}^2$. By substituting in equation (1) we obtain:

$$L_{Vavg}/L_{avg} = \sqrt{10/3\alpha\left(10^{0.1K_2} - 1\right)} \qquad (2)$$

Equations (1) and (2) are strictly only applicable to uncluttered rooms with uniform absorption. The rooms should also be roughly cubical in shape, (i.e., the ratio of largest to smallest dimension should be less than 4). If these conditions are not met, L_{Vavg} will overestimate the size of the room. In this case K_2 should be obtained by measurements using either a reference sound source, or the reverberation time. To compensate for deviations from ideal conditions, Figure 1 uses equation (2) with K_2 values reduced by 10%.

ISO3744[2] allows measurements if K_2 is less than 2 dB. This requirement is shown as a function of α in Figure 1. In this figure L_{avg} actually defines the measurement surface dimension, so that the smallest surface lies on the surface of the source. This makes the number of measurement points prohibitively large. To reduce the number of measurement points to 40, reevaluate Figure 1 with L_{avg} increased by 1.5 times. For 10 measurement points reevaluate Figure 1 with L_{avg} increased by 3 times (note also that ISO3744 requires the measurement surface to be at least 0.25m from the source).

2.2 Reverberant Environments

The reverberant room standard, ISO3743-1[3], requires the room volume to be at least 40 times the volume of the source and α must be less than 0.2.

Changes to ISO3747 in 2000[4] were not incorporated in the CSA guidelines. In this standard[4], accuracy is determined by examination of measurement results. A conservative estimation of requirements was obtained in Figure 1 by comparison with ISO3743-1 and ISO3744.

2.3 Intensity standards

For the ISO 9614 series of intensity standards [5,6], K_2 approximates the pI index (if background noise is treated separately as in clause 2.4). The average pI index (and thus K_2) cannot exceed the measuring equipment dynamic capability, which is typically limited to 10-13 dB.

2.4 Background Noise

In a reverberation room, measurements are made far from the source, and the source must dominate background noise everywhere in the room. For free field measurements the source need only dominate the background noise on the measurement surface. In intensity measurements the allowable background noise is limited by the dynamic capability of the measuring equipment and at high levels can reduce the allowable K_2. Background noise must be evaluated by measurement since allowable levels can exceed the noise produced by the source.

3. CONCLUSIONS

Figure 1 can be used to estimate which standards can be used in a given measurement situation. The Figure also shows how a measurement can be improved. For example, by changing rooms, or by increasing the absorption in a room, it may be possible to significantly improve a measurement. Additional information can be found in the CSA guidelines.

REFERENCES

1. CSA Z107.58-02:2002, Noise emission declarations for machinery
2. ISO3744:1994, Acoustics -- Determination of sound power levels of noise sources using sound pressure -- Engineering method in an essentially free field over a reflecting plane.
3. ISO3743-1: 1994, Acoustics -- Determination of sound power levels of noise sources -- Engineering methods for small, movable sources in reverberant fields -- Part 1: Comparison method for hard-walled test rooms.
4. ISO3747:2000, Acoustics -- Determination of sound power levels of noise sources using sound pressure -- Comparison method in situ.
5. ISO9614-1:1993, Acoustics -- Determination of sound power levels of noise sources using sound intensity -- Part 1: Measurement at discrete points.
6. ISO9614-2:1996, Acoustics -- Determination of sound power levels of noise sources using sound intensity -- Part 2: Measurement by scanning.

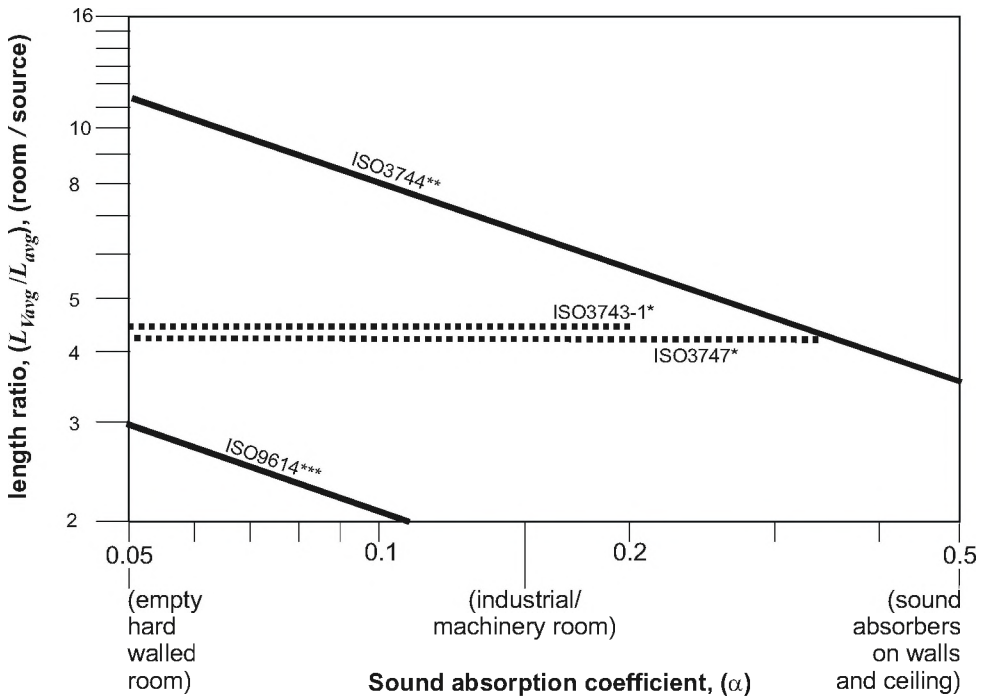


Fig. 1. Minimum average room length to average source length ratio, L_{Vavg}/L_{avg}, as a function of the room's mean sound absorption coefficient, α for engineering grade standards. Background noise limitations are: *source dominates background noise at all points in the room; **source dominates background noise on measurement surface, (i.e., near source); ***background noise dominates source (source must be audible on measurement surface)

USE OF SOUND QUALITY METRICS FOR THE ANALYSIS OF AUTOMOTIVE INTAKE NOISE

Colin Novak, Helen Ule and Robert Gaspar
Dept. of Mechanical, Automotive and Materials Engineering, University of Windsor
401Sunset Ave., Windsor, Ontario, Canada, N9B 3P4, novak1@uwindsor.ca

1. INTRODUCTION

Due to consumer demands, the need to attenuate the various sources associated with the operation of today's automobiles has become paramount. One such source of noise often overlooked from a consumer's perspective, is engine induction noise. This intake noise is the audible result of pressure fluctuations propagating through the intake system ducting as a result from the valve action of the engine.

Automotive intake noise is most often attenuated through the application of passive control techniques. These techniques are the simplest and least expensive form of attenuation but do not always yield the best results. A common example of a passive noise control technique used in automotive applications is the Helmholtz resonator. Experimentation with active noise control techniques has thus far also shown promise. This method, however, has not yet realized full commercial viability.

Independent of which method of attenuation is used to control intake noise, quantification of its success is paramount. However, from a consumer's point of view, the perceived quality of noise emissions takes precedence over what traditional acoustical analysis techniques may imply. Given this, sound quality metrics can be an important analysis tool for the evaluation of the various noise control techniques. The present study investigates the validity of using several psychoacoustic metrics for the analysis of automotive intake noise.

2. PSYCHOACOUSTICS

In the evaluation of the acoustic comfort of a sound, fundamental quantities such as the acoustic sound pressure level are not adequate to truly represent the actual hearing sensations. The science of psychoacoustics involves the quantitative evaluation of these subjective sensations using sound quality metrics. Application of sound quality metrics allow for the visualization of the complicated relationship that exists between the physical and perceptual acoustic quantities.

For this investigation, several sound quality metrics were used to evaluate automotive induction noise. These metrics included loudness, sharpness, roughness, fluctuation strength and articulation index. Before a discussion of any results using these metrics can be presented, a description of each is warranted.

Loudness is a standardized metric that describes the human perception of how loud a source is perceived as opposed to a simply reported sound pressure level. Loudness, using the units of sones, takes into account the temporal processing and audiological masking effects of sounds across the frequency range.

Sharpness, which has units of acum, describes the high frequency annoyance of noise by applying a weighting factor on sounds above 2 kHz. This overall measurement is most often applied to broadband sources such as rushing air noise, gear meshing noise and grinding sounds.

Roughness and fluctuation strength are two metrics used to describe the annoyance of modulating sounds. Fluctuation strength, with units of vacil, focuses on sounds which modulate at frequencies between 0.5 Hz and 20 Hz, with a 4 Hz fluctuation being the most annoying. Roughness, using units of asper, focuses on modulating noise at frequencies between 20 Hz and 300 Hz. The most annoying modulation occurs at 70 Hz. A siren is an example of a modulating sound.

Articulation index is a calculated measure of the intelligibility of speech in the presence of a sound. This is accomplished through consideration of the levels of the masking signal in the frequency bands most critical to the understanding of speech. A calculated AI of 0.1 indicates poor speech intelligibility whereas an AI greater than 0.6 is considered good with a value of 1.0, or 100%, being the best.

3. PROCEDURE

Induction noise was measured from a 16 valve 4 cylinder Toyota 4A-GE engine. The engine, which was located in a semi-anechoic room, was motored on a DC dynamometer at engine speeds from 1000 to 6500 rpm in increments of 500 rpm. The microphone was located at a distance of 100 mm in front of the intake orifice and was connected to a 01dB Orchestra data acquisition system which allowed recording to a PC hard disk. This data was later post-processed to determine, among other parameters, the psychoacoustic metric results.

4. RESULTS

Figure 1 shows the measured results for A-weighted pressure level and loudness. These two metrics are expected to show similar data trending since they are both designed to compensate for the human perception of sound amplitude at various frequencies. The difference is that loudness also compensates for other effects such as temporal processing. Given these similarities, it can be surmised that loudness is just as useful, if not better, a tool for induction noise analysis.

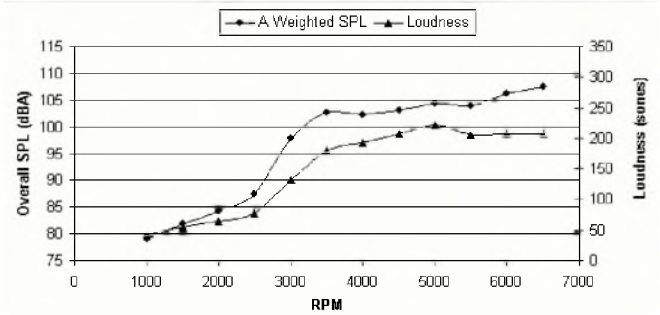


Fig. 1. Results of A-Weighted Sound Pressure Level and Loudness with respect to Engine Speed illustrate good correlation.

It was assumed that sharpness might be an appropriate metric for the evaluation of intake noise given that a high frequency component of intake noise is created by the intake air traveling across the intake valve seat at a high velocity. Examination of figure 2, however, shows that very little sharpness is realized. This is most likely due to a reduction of high frequency content in the noise resulting from influences of the manifold ducting and air cleaner.

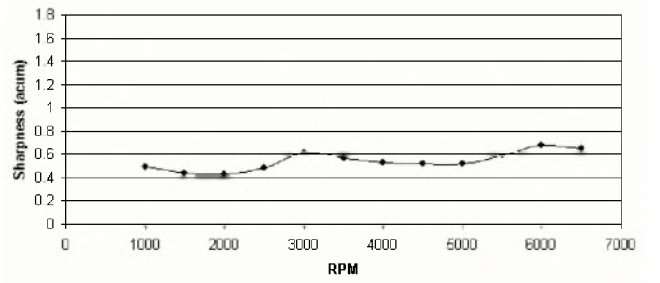


Fig. 2. Results of Sharpness versus Engine Speed show very little high frequency content in the induction noise signal.

Figure 3 illustrates the calculated roughness and fluctuation strength values across the reported engines speed range. Values for roughness are relatively low indicating that for this type of noise source, very little low frequency modulation exists. The reported values for fluctuation strength are similarly low with a couple of peaks occurring at specific engine rpm's which are the result of harmonics occurring at these engine speeds. This can be controlled through appropriate tuning of the intake ducting. These observations have been reinforced by a parallel study that used various frequency analysis tools.

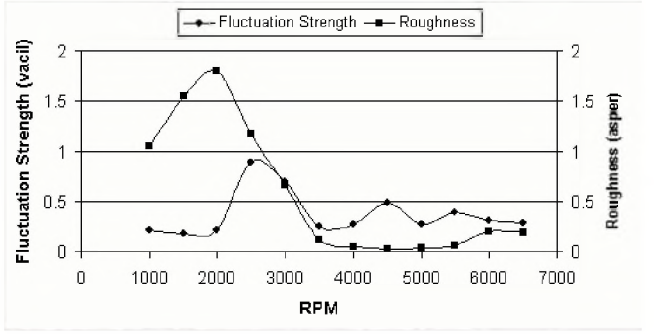


Fig. 3. Results of Fluctuation Strength and Roughness with respect to Engine Speed show very little low frequency and moderate mid frequency modulation of the induction noise data.

Figure 4 shows the results of an articulation index analysis. It should be noted that the location of these measurements was not the most appropriate. It would have been better to conduct these measurements inside the passenger cabin of the vehicle. This would have not only resulted in higher values for speech intelligibility but would have also been more representative of true human exposure. The merits of this metric, however, are still apparent.

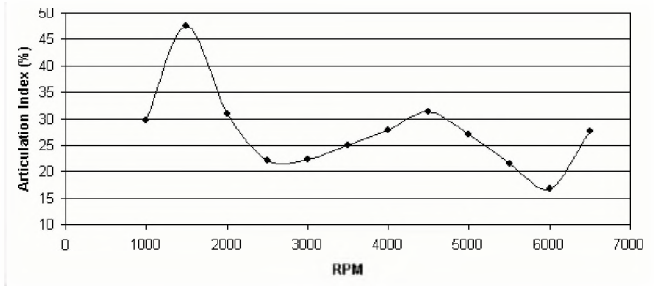


Fig. 4. Results of Articulation Index with respect to Engine Speed.

5. CONCLUSION

The focus of this investigation was to investigate the applicability of using several sound quality metrics for the analysis of automotive induction noise. It was seen that loudness correlated well with an A-weighted SPL analysis, thus demonstrating it usefulness. Sharpness, however, was not shown to be an appropriate metric in this case due to the lack of high frequency content in the acquired noise. The roughness analysis was able to demonstrate a lack of low frequency modulation while a more apparent fluctuation strength illustrated the presence of harmonics. This information could prove useful in a tuning exercise of these ducts. While the articulation index analysis provided good results, the noise data for this analysis should not have been collected in the under hood environment. Instead, this noise data needs to be collected at a more representative location such as the passenger compartment of the vehicle.

While some of the metrics employed in this study proved more applicable than others, the overall merit of using psychoacoustic tools for intake noise was demonstrated.

COMPARISON OF EXPERIMENTAL AND MODELED INSERTION LOSS OF A COMPLEX MULTI-CHAMBER MUFFLER WITH TEMPERATURE AND FLOW EFFECTS

Helen Ule[1], Colin Novak[1], Tony Spadafora[1], Ramani Ramakrishnan[2] and Robert Gaspar[1]
1. Dept. of Mechanical, Automotive and Materials Engineering, University of Windsor, 401Sunset Ave., Windsor, Ontario
2. Department of Architectural Science, Ryerson University, 35 Victoria St., Toronto, Ontario

1. INTRODUCTION

Due to both legislative and consumer demands, the need to attenuate noise emissions associated with the operation of today's automobiles has become paramount. Further to this, improvements in engine performance capabilities have resulted in the generation of greater noise propagation from the engine. Exhaust system engineers, therefore, must design mufflers that are capable of attenuating this higher amplitude noise propagation while at the same time not restricting engine performance by increasing flow resistance.

The complex multi-chamber muffler is the most common noise control filter used for automotive exhaust systems today. The design of these filters has always been a specialized field that involved a degree of experience and intuition along with an understanding of the fundamental design equations. With this, a prototype would be constructed, tested and improved as part of a trial and error process. This is both costly and time consuming. As a result, the design of such systems is now done with the aid of computer modeling programs. Care must still be taken in the implementation of such an approach to ensure good correlation between modeled and actual results.

For this investigation, a computer simulation program, Ricardo WAVE, is used to predict the insertion loss of an "off the shelf" multi-chamber muffler. Included in the investigation are the effects of both temperature and flow. These theoretical results are compared to experimental measurements of the same muffler design that was used in the computational model.

2. METHOD

The insertion loss of the muffler was determined both experimentally on an engine motored on a dynamometer as well as analytically using WAVE. The following is a description of these two approaches.

2.1 Experimental Approach

To experimentally determine the insertion loss of the muffler, it was attached to a Toyota 4A-GE engine motored on a DC dynamometer within a semi-anechoic environment. A cutaway view of the muffler showing the multiple chambers is illustrated in Figure 1. This filter is classified as a reactive muffler. Here, the multiple pipes and chambers provide an impedance mismatch for the acoustic energy traveling through it. This impedance mismatch causes some of the acoustic energy to reflect back to the source thus preventing some of the noise from transmitting through the muffler.

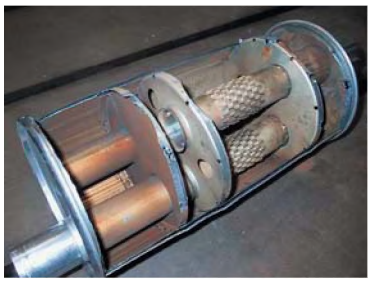


Fig. 1. Cutaway view of the muffler used showing the multiple passages and chambers that make up this reactive filter.

The engine was operated at engine speed from 100 to 4000 rpm in increments of 500 rpm. A microphone 100 mm from the exit of the exhaust pipe measured the resulting noise. Flow temperature and velocity data was also collected 0.5 metres before and after the muffler. The insertion loss was determined by comparing these results to similar measurements made with the muffler replaced by a straight section of pipe. The difference in realized sound pressure level represents the insertion loss of the muffler.

2.2 Analytical Approach

Using Ricardo WAVE, implementing a one-dimensional finite-difference formulation, the realized sound pressure level at the exit of the exhaust pipe was predicted with both the muffler and straight pipe in place. In order to accomplish this, both the Toyota engine that was used as the dynamic noise source and the muffler needed to be modeled. Figure 2 is an illustration of the WAVE model schematic showing the input building blocks of the muffler only.

3. RESULTS

Figure 3 shows the realized sound pressure level with and without the muffler installed for both the

experimental and modeled analysis. It can be seen that an insertion loss of up to 20 dB is realized with implementation of the muffler in the experimental exercise with the maximum occurring at the higher engine speeds. While a positive insertion loss also occurred in the modeled results, the degree of insertion loss is much less. It is felt that the higher noise levels for the unmufflered experimental results were due to wind noise, which is not realized in the modeled results.

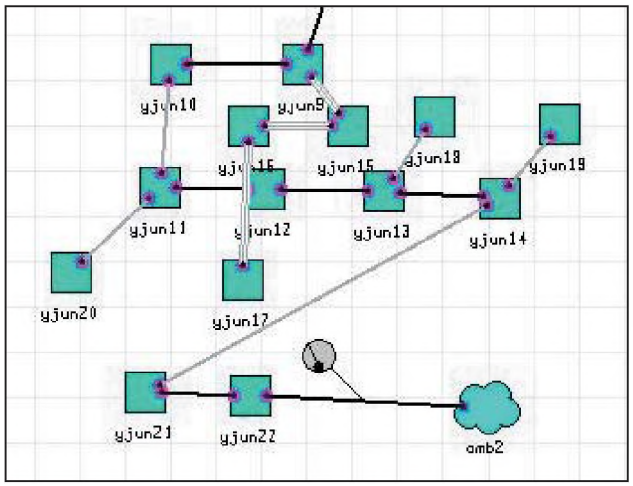


Fig. 2. Schematic of Ricardo WAVE model for the muffler used in this investigation illustrating the various input building blocks of the model.

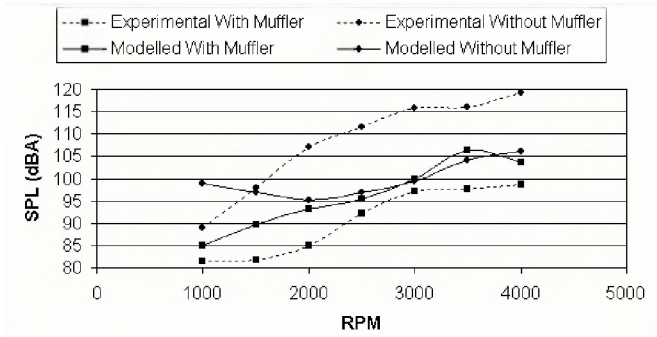


Fig. 3. Sound pressure levels with and without muffler for experimental and modeled engines illustrating realized insertion loss.

Figure 4 shows the experimental and modeled flow temperature determined 0.5 metres after the muffler exit location with and without the muffler installed. It can be seen that while the temperatures for both experimental results are approximately the same, they do increase with an increase in engine speed. It is assumed that this increase is due to frictional effects of the motored engine. This assumption is reinforced through examination of the modeled results. Here, the predicted temperatures not only remain constant, but also show no significant difference between the case with and without the muffler. This is due to the inability of the computational model to include any frictional effects

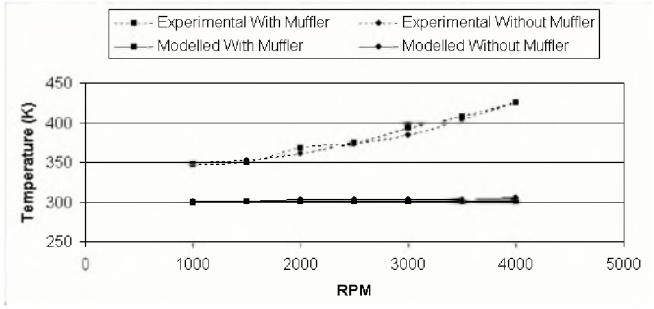


Fig. 4. Flow temperatures with and without muffler for experimental and modeled engines.

Figure 5 shows the gas flow velocity for the experimental and modeled flows determined 0.5 metres after the muffler exit location with and without the muffler installed. It can be seen for the most part that the flow velocity was higher for the experimental results in the absence of the muffler. This is to be expected since the presence of the muffler acts as a dampening reservoir. The results for the modeled results, however, do not follow this same trend. It was found that the modeled flow velocities without the muffler were unexpectedly low for engine speeds of 1000 and 3000 rpm.

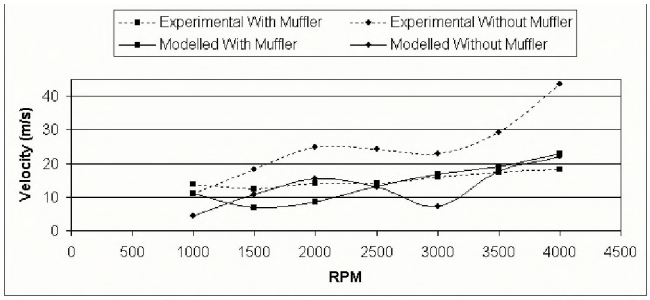


Fig. 5. Gas flow velocity with and without muffler for experimental and modeled engines.

4. CONCLUSION

The focus of this investigation was to investigate the difference in realized insertion loss of a motored engine using both experimental and modeled results. Also included were temperature and flow effects. It was felt that the experimental noise measurements were contaminated by wind noise over the microphone and that the temperature differences were due only to frictional effects of the motored engine. Good insertion loss results, however, were obtained by the computational model. As expected, no temperature changes resulted in the modeled output. It was further determined that the analytical model was not able to accurately predict the flow velocity for the case without the muffler installed.

ACOUSTIC ANALYSIS OF MRI SCANNERS

W. Shao and C. K. Mechefske
Department of Mechanical and Materials Engineering, Queen's University,
Kingston, Ontario, Canada K7L 3N6 Email: shao@me.queensu.ca

1. INTRODUCTION

Acoustic noise generated by MRI scanners has a tendency to raise stress levels in patients undertaking the scanning. MRI acoustic noise may even lead to temporary or permanent shifts in the hearing threshold for patients [1]. This acoustic noise is mainly caused by the vibration of the gradient coil system due to the Lorentz forces generated by the interaction of the magnetic field around the conductors in the gradient coil and the main static magnetic field. Noise levels as high as 120 –130 dB have been reported in some MRI scanners [2]. The increase in acoustic noise levels in recent years is primarily due to the trend toward the use of high static magnetic field strengths and high-performance switching gradients with high maximum amplitudes and slew rates.

Measures have been taken to control the acoustic noise generated by MRI scanners recently. The technique of Active Noise Control (ANC) for the reduction of MRI noise has been studied by Mechefske *et al.* [3], but it is normally effective at low frequency and less effective at high frequency. Mansfield *et al.*[4] proposed a new technique called active acoustic screening for quiet gradient coil design. Unfortunately, this acoustic screening inevitably reduces the gradient strength. Yoshida *et al.*[5] used "independent suspension" of the coil to dampen solid vibration; while "vacuum vessel enclosure" of the coil shields transmission of residual vibration through the air. However, these methods are not suitable for all kinds of MRI scanners.

It is obvious that optimizing the design of gradient coil systems will eliminate the root cause of the noise. Thus, a good understanding on the characteristics of acoustic radiation of the gradient coil system is necessary for the design of quiet MRI scanners. Due to the geometrical symmetry of the gradient coil system, an analytical model of finite cylindrical ducts with infinite flanges is used to investigate its acoustic radiation characteristics (see Figure 1) in this paper. The radiation impedances will be calculated for a finite cylindrical duct with rigid and absorptive walls. Based on these results, the inside sound field generated by the vibrating wall will be simulated.

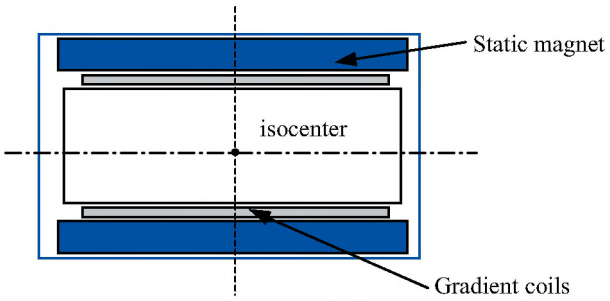


Figure 1. Schematic cross-sectional view of a MRI scanner

2. THEORY

The sound field inside an infinite cylindrical duct can be expressed as a sum of modal solutions (a time factor $e^{i\omega t}$ is understood throughout this paper)

$$p(r,\theta,x)=\sum_{m=-\infty}^{\infty}\sum_{n=1}^{\infty} J_m(\alpha_r^{mn} r)e^{-im\theta}[A_{mn}e^{-i\alpha_x^{mn}x}+B_{mn}e^{i\alpha_x^{mn}x}] \quad (1)$$

where x is the coordinate in the axial direction of the duct, r is in the radial direction and θ is in the circumferential direction. $J_m(\alpha_r^{mn} r)$ are the Bessel functions of the first kind of circumferential order m, n is the radial mode number. α_x and α_r are the wavenumbers in the axial and the radial direction respectively. A and B are the modal coefficients of the forward-propagating and backward-propagating acoustic wave modes respectively.

For the duct wall with a finite acoustic impedance, the boundary condition should satisfy the equation:

$$\alpha_r^{mn} J'_m(\alpha_r^{mn} a) = i\beta k J_m(\alpha_r^{mn} a) \quad (2)$$

where a is the radius of the duct and β is a specific acoustic admittance of the wall. The acoustic pressure and velocity amplitudes at the open ends of the duct can be expressed in terms of the acoustic modes in radial r and circumferential θ directions as

$$p(r,\theta,x)=\sum_{m=-\infty}^{\infty} e^{-im\theta}\sum_{n=1}^{\infty} P_{mn} J_m(\alpha_r^{mn} r) \quad (3)$$

$$u_x(r,\theta,x) = \frac{1}{\rho c}\sum_{m=-\infty}^{\infty} e^{-im\theta}\sum_{n=1}^{\infty} V_{mn} J_m(\alpha_r^{mn} r) \quad (4)$$

where P_{mn} and V_{mn} are the modal coefficients for the pressure and velocity respectively. The relation between the modal pressure and velocity amplitudes can be expressed by the generalized radiation impedance z at the open end:

$$P_{mn} = \sum_{l=1}^{\infty} Z_{mnl} V_{ml} \quad (5)$$

The impedance z can be calculated by the continuity of pressure and velocity at the open ends. The velocity distribution of the vibrating wall can be written in the form of a Fourier series:

$$u_r(a,\theta,x) = \frac{1}{4\pi^2}\sum_{m=-\infty}^{\infty} e^{-im\theta}\int_{-\infty}^{\infty} \tilde{U}_m e^{-i\alpha_x x} d\alpha_x \quad (6)$$

Therefore the total sound pressure inside the duct can be calculated by solving the equation (1) by satisfying the boundary conditions at the wall (vibrating and finite impedance) and the open ends (radiation impedance).

3. RESULTS

To verify the validity of the above mathematical model, the sound pressures in the isocenter of a finite cylindrical duct will be calculated. Assuming different wall acoustic impedances: rigid and $\beta = 0.1$, the cylindrical duct with 0.3 m radius and 1.2 m length will be used for the simulations. These results will be compared with the data calculated by the commercial code LMS SYSNOISE, which is based on the Boundary Element Method (BEM).

A uniform velocity distribution with an amplitude 0.0001 *m* is used to move the duct wall. The sound pressures at the isocenter calculated by both the analytical model and BEM model with rigid wall and absorptive wall ($\beta = 0.1$) are shown in Figures 2 and 3 respectively. Comparing these two figures, It can been seen that the overall general shape of all the curves is the same. This suggests good agreement at all frequencies between the BEM results and the analytical results.

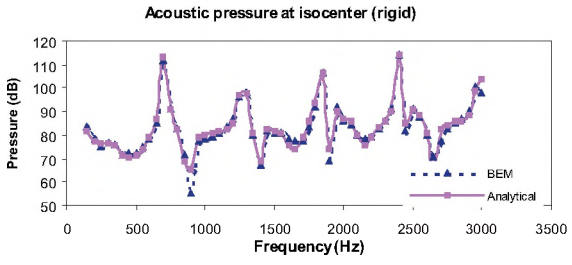


Figure 2. Acoustic pressures at the isocenter (rigid wall)

It can also be seen that an absorptive wall can significantly reduce the noise inside the duct. It is also obvious that there are five peaks in these frequency spectra. This is due to the fact that the wave reflections at the open ends reach their maximum value at the cut-off frequency, therefore causing acoustic resonance. The cut-off frequencies are dependent on the geometrical dimension (radius) of the duct.

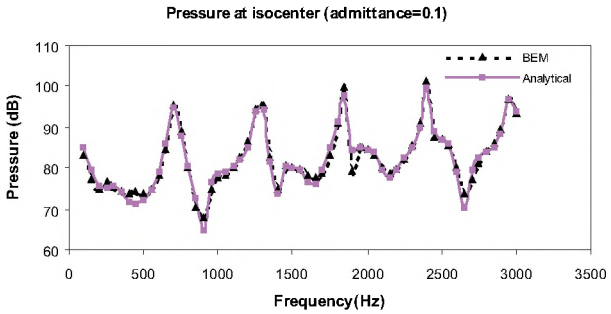


Figure 3. Acoustic pressures at the isocenter ($\beta = 0.1$)

4. CONCLUSIONS

An analytical model that predicts the acoustic noise radiation from gradient coils in MRI scanners has been presented. The acoustic response of the analytical model was found to be in good agreement with the results obtained using a BEM model. Compared with BEM, the most important feature of analytical methods is that they can generally show the dominant parameters for the modeled problems more directly and yield more physical insight. In addition, they are always much more computationally efficient (For instance, it normally takes about 2 weeks for calculating a model with absorptive walls from 100 to 3000 Hz using the BEM in a computer while the same calculation required only two or three hours for the analytical model in the same computer).

REFERENCES:

1. Brummett RE, Talbot GM, Charuhas P. (1988). Potential hearing loss resulting from MR imaging." Radiology, **169**:539-540.
2. Counter SA, Olofsson A, Borg E. (2000). Analysis of magnetic resonance imaging acoustic noise generated by a 4.7 T experimental system. Acta Otolaryngol **120**: 739-743.
3. Mechefske CK and Geris R. (2002). Active noise control for use inside a magnetic resonance imaging machine. Ninth International Congress on Sound and Vibration.
4. Mansfield P, Glover P and Bowtel R, (1994) Active acoustic screening: design principles for quiet gradient coils in MRI. Meas. Sci. Technol. **5**, 1021-1025.
5. Yoshida T, Takamori H and Katsunuma A, (2001) Excelart™ MRI system with revolutionary Pianssimo™ noise-reduction technology. Med. Rev. **71,**1-4.
6. A. Kuijpers, S. W. Rienstra, G. Verbeek and J. W. Verheij. (1998). The acoustic radiation of baffled finite ducts with vibrating walls. J. Sound Vib. **216**, 461-493.

THE INEXPENSIVE NOISE CONTROL METHOD – A CASE STUDY

Ramani Ramakrishnan
Associate Professor, Dept. of Architectural Science, Ryerson University
350 Victoria St., Toronto, Ontario, Canada, M5B 2K3, rramakri@ryerson.ca

1. INTRODUCTION

Stack exhausts are one of the worst radiators of noise levels from major industrial sites, due to their height as well as the loss of directivity corrections due to the large propagation distances. In addition, oven stacks from major process operations suffer from strong load variations such as temperature and flow changes. The complexity of the stack noise generation and its attenuation is discussed through one atypical example. The stack under investigation is located downstream of a pollution control systems used to control the effluent from four ovens of a major industrial facility. Unlike typical exhausts that produce steady sound levels around its azimuth, strong variations were identified. Implementing a typical stack silencer would not have guaranteed noise attenuation success. The details of this case study and the inexpensive modifications are presented in this paper.

2. SITE DESCRIPTION

Most major industrial plants such as paper and pulp mills, process plants, auto manufacturers, auto-parts manufacturers, and chemical plants have large exhaust stacks on their roof. They contain mostly fans that are used to exhaust effluents at large heights above the plant roof. Most of these stacks generate (mostly fans) broadband hum. If the adjacent sensitive receptors are more than 300 m or more away from the plant, the resulting noise levels are insignificant. However, there are exceptions to the general separation guideline applied in many locations. In Ontario, as contained in the Municipal Act, the separation distance between major industrial sites and residential receptors is of the order of 300 m.

The plant under study is a large industrial facility with a myriad of exhaust stacks on its roof. New paint facilities were added near one corner of the plant. Residential receptors were at a distance of about 300 from the plant. The present study was instituted as part of the audit noise measurements requested by the regulatory agencies. The plant layout is shown in Figure 1. The line of stacks is shown in Figure 2.

Four of the exhaust stacks from the new facility were connected to a pollution control system. The output from the pollution control system was exhausted through a single stack. The stack, around 110 cm in diameter was 10 m above the top roof of the new extension. The audit measurement program indicated the major concern is the noise output from the single stack. The details of the measurement program and the resulting analysis are described below.

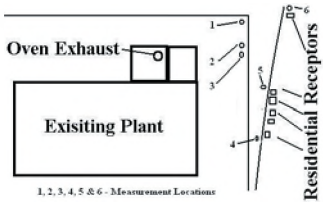


Figure 1. General Layout of the Plant and the Oven Exhaust.

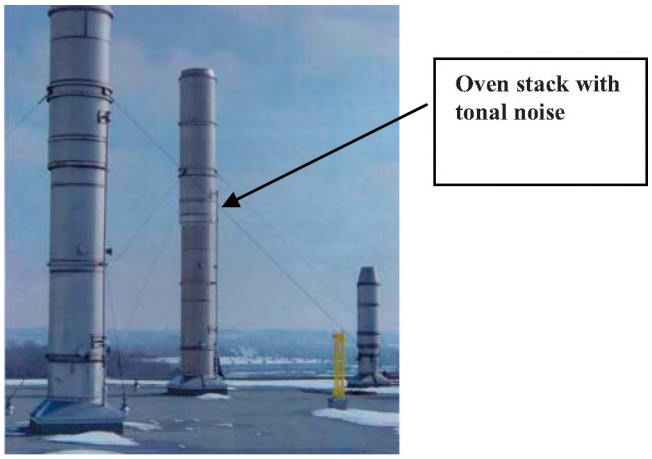


Figure 2. General Layout of the Plant and the Oven Exhaust.

3. PLANT NOISE AND MITIGATION

The new extension had to satisfy the noise emission limits. Since the plant is located in an urban setting, the ambient noise is set by local traffic and any other industry. As per the Provincial guidelines the noise limit to be satisfied by the extension is 45 dBA or the combined noise from the ambient road traffic and the existing plant, which ever is higher. In addition, if the source has any special characteristics such as a tone, a penalty of 5 dB must be added to the source level, before evaluating the noise impact. These levels are expressed as

one-hour energy averaged sound level, L_{eq}, in dBA. The existing ambient levels were evaluated through measurements. The lowest noise level was measured to be 48 dBA.

The plant, as part of its approval conditions, had to perform audit noise measurements of the extension. During the measurement program, a strong tonal noise was noticed. Narrow band analysis showed that the tone was centred at 266.5 Hz. The noise signature exhibited strong variation, both temporally and spatially. The measured noise levels are summarized in Table 1 below. The offending tone was barely audible at Location 5 (Figure 1 and Table 1).

Location	Before		After	
	250 Hz Band	dBA	250 Hz Band	dBA
1	42/64	49/59	44	48
2	62/70	57/63	57	55
3	57/50	56/49	57	54
4	-	-	48	49*
5	44/43	50/51*	45	48*
6	-	-	47	50*

*** - Dominant Noise - Road Traffic.**

Table 1. Summary of Measured Noise Levels.

Measurements were also conducted close to the sources of the new extension on two roof levels. The spectra of the rood top measurements are shown in Figure 3 below. These measurements indicated the exhaust stack of the pollution control system generated the offending tonal noise. Table 1 and Figure 3 shown results conducted at a number of locations as well as at different time periods. The 'before' measurements show the results, before any modifications to the stack were implemented. The results indicate strong variations of the order of 20 dB of the 266.5 Hz tone. The spatial variations can be seen from Figure 3 and Table 1.

The temporal variations can be seen from the results at Locations 1, 2 and 3 (Table 1). The residential receptor noise level, including the tonal penalty, was estimated as 57 dBA at the 2nd Storey bedroom window. The noise impact was of the order of 9 dB. A simple solution would have been to install a circular silencer with a minimum insertion loss of 10 dB in the 250 Hz band. However, due to the strong spatial and temporal variations, the successful performance of the silencer was in doubt. Any mitigation measure can be useful, only after the reasons for the variations can be established. Closer inspection of the stack revealed that the top 3 m of the pipe was slightly larger than the bottom 7 m pipe, thus creating a short annulus. The above prevents any condensed fluid from trickling back to the fan. The annulus creates an entrained flow, which provides tuning of the 266.5 Hz tone of the fan. The plant decided to close the annulus. The modification is described as well as highlighted in Figure 4. The results are shown as 'After' in Table 1 and Figure 3. The strong reduction of the tone level, between 8 and 10 dB, can be seen from Figure 3.

Finally, the residential noise levels showed the plant was within 1- 2 dB of the guideline limits.

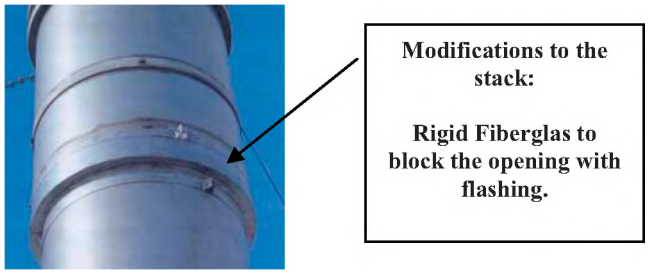


Figure 4. General Layout of the Plant and the Oven Exhaust.

REFERENCES

Model Municipal Noise Control By-Law, Ministry of the Environment, Ontario, Canada, 1978.

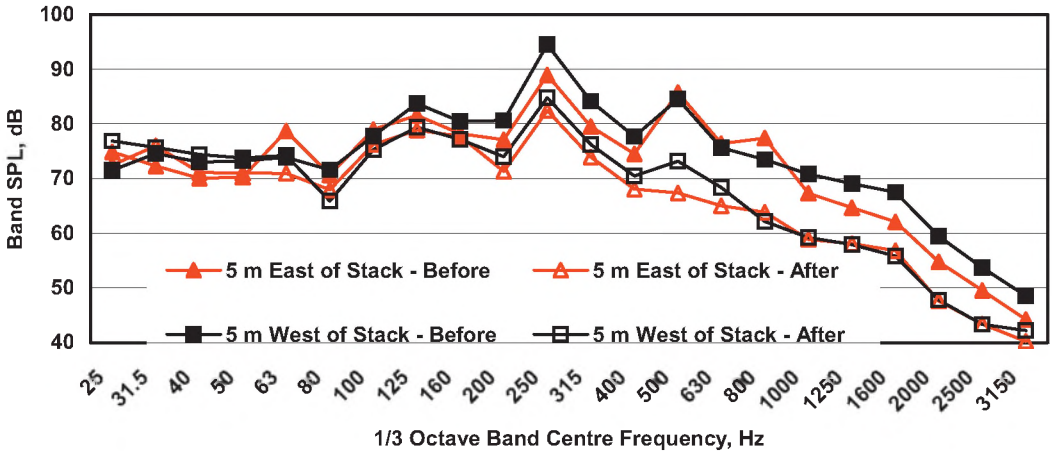


Figure 3. Noise Spectra of the stack exhaust, on the rooftop.

EXPÉRIENCES SUR LE NIVEAU DE BRUIT DES VENTILATEURS DOMESTIQUES PERSONNELS

Annie Ross[1], Marie-Josée Boudreau[2]
[1]Dép. de génie mécanique, École Polytechnique de Montréal, CP6079 succ. Centre-ville, Montréal, Québec, Canada H3C 3A7, Annie.Ross@polymtl.ca
[2]Les Éditions Protégez-Vous, 5199, rue Sherbrooke est, bureau 3699, Montréal, Québec, Canada, H1T 3X2

1. INTRODUCTION

En été, les ventilateurs personnels sont couramment utilisés pour rafraîchir les gens et leur permettre notamment de mieux dormir. Très souvent, on a recours à des appareils amovibles : ventilateurs sur pied ou de table. Or, ces appareils sont souvent bruyants, ce qui ne favorise pas le sommeil.

Le laboratoire d'acoustique de l'École Polytechnique a mesuré et comparé le niveau de bruit de 29 ventilateurs personnels sur le marché en 2004. Ce nombre comprenait des appareils à débit axial et des appareils centrifuges. Les niveaux sonores ont été évalués en bandes d'octave, pour chacune des trois vitesses d'opération des ventilateurs. Le présent article fait état des résultats de cette étude.

2. MÉTHODE EXPÉRIMENTALE

Les expériences entreprises n'avaient pas pour but de déterminer la puissance acoustique des ventilateurs, mais bien d'établir un classement des appareils entre eux. Bien qu'il existe des normes régissant l'évaluation sonore des ventilateurs domestiques (e.g. ACMA 300-96), leurs exigences ont été jugées trop sévères pour les besoins de cette étude. Ainsi, une procédure plus simple a été établie pour ces expériences.

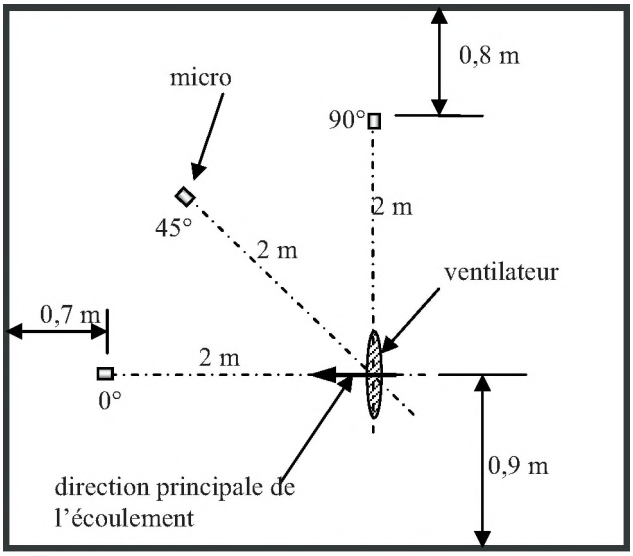


Fig. 1. Disposition de l'équipement en chambre anéchoïque.

2.1 Détail des mesures

Les essais ont été effectués en chambre anéchoïque, ce qui a permis d'obtenir des niveaux de pression sonore (Lp) en champ libre à 2m des ventilateurs, sur 8 bandes d'octave de 63 Hz à 8000 Hz. Les mesures on été prises simultanément en 3 positions (0°, 45° et 90°) dans le plan horizontal passant par l'axe principal de l'écoulement, tel qu'illustré à la figure 1 (vue de dessus). Les ventilateurs de table étaient posés sur un tabouret de bois, alors que les ventilateurs sur pied reposaient sur le grillage qui forme le plancher de la chambre.

Trois micros électrets 130A étaient alimentés par un conditionneur 514A de PCB. Chaque micro était maintenu par un pied métallique mince, de façon à ne pas osciller sous l'effet du vent. L'acquisition et les calculs étaient effectués à l'aide d'un analyseur spectral Oros OR38.

La pression acoustique était échantillonnée à 32,8 kHz et soumise à un filtre anti-repliement dont la fréquence de coupure était de 12,8 kHz. Dans chaque cas, l'échantillonnage couvrait une période de 5 minutes, ce qui permettait d'obtenir une moyenne stable. La résolution fréquentielle des données était de 2 Hz.

2.2 Bruit de fond

Le bruit de fond de la chambre a été évalué à six reprises durant les essais et se sont avérés très constants. Le tableau 1 indique les niveaux moyens mesurés lorsque les ventilateurs étaient éteints, ainsi que les écarts maximaux obtenus. Dans les figures ci-après, les bandes d'octave sont identifiées par les numéros attribués au tableau 1.

Tableau 1. Niveau moyen et variation du bruit de fond.

fréquence centrale	no. de l'octave	moyenne	écarts maximums	
63 Hz	1	30.5	-0.2	+ 0.9
125 Hz	2	23.6	-0.8	+ 0.8
250 Hz	3	26.5	-0.3	+ 0.6
500 Hz	4	21.6	-0.7	+ 1.0
1000 Hz	5	27.4	-0.3	+ 0.5
2000 Hz	6	26.4	-0.8	+ 1.3
4000 Hz	7	28.2	-0.7	+ 1.0
8000 Hz	8	30.8	-0.3	+ 0.4
global	---	36.8	-0.3	+ 0.4

3. RÉSULTATS ET OBSERVATIONS

Bien que les spectres de bruit aient aussi été observés à partir des données en bandes fines, l'analyse présentée ici porte sur les résultats en bandes d'octave.

3.1 Ventilateurs bruyants et silencieux

Les figures 2 et 3 représentent les cas extrêmes d'un appareil très bruyant et d'un appareil silencieux. Il s'agit de ventilateurs axiaux de performance moyenne en terme de déplacement d'air. Dans les deux cas, les mesures ont été obtenues à la vitesse maximum d'opération.

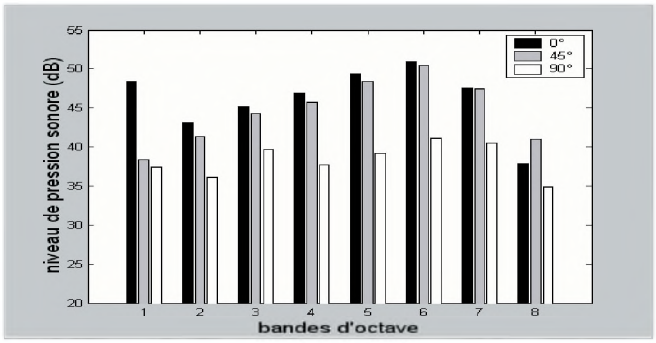


Fig. 2. Niveaux Lp d'un ventilateur « bruyant » à vitesse élevée.

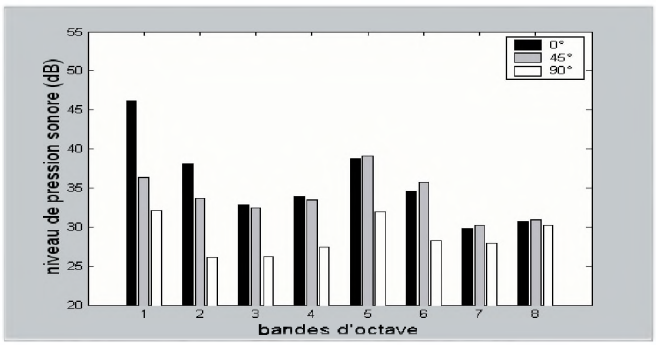


Fig. 3. Niveaux Lp d'un ventilateur « silencieux » à vitesse élevée.

La distribution en fréquence (de même que les spectres) est compatible avec les spectres obtenus par Wu et Su. En plus d'un maximum dans la bande de 63 Hz, on y distingue une courbe douce qui culmine dans le voisinage de 1000 à 2000 Hz. (La position exacte varie d'un appareil à l'autre.) Malgré ces similitudes, on constate que le niveau de bruit est très variable d'un appareil à l'autre. Le niveau global (toutes fréquences confondues) diffère de 8 à 11 dB, selon la position angulaire, entre les deux appareils présentés ici. À titre indicatif, le critère de bruit (NC – « noise criterion ») a été déterminé pour la chambre anéchoïque avec les ventilateurs en fonctionnement. Il passe de NC35 à NC55, selon le ventilateur. Il va sans dire qu'un niveau NC55 est difficilement admissible dans une résidence.

3.2 Directivité

Pour chaque ventilateur, le bruit mesuré à 0° et à 45° est de niveau comparable. (Ceci est valable dans toutes les bandes, sauf dans l'octave centrée sur 63 Hz où le niveau Lp est plus élevé à 0°.) Par contre, le niveau à 90° est généralement plus faible (jusqu'à 6 dB). Incidemment, cette position se trouve en dehors du cône de ventilation (Boudreau). Le bruit se propage donc principalement dans la direction où se trouve l'utilisateur du ventilateur.

3.3 Vitesse d'opération

Il n'est pas étonnant de constater à la figure 4 que le niveau de bruit du ventilateur augmente avec la vitesse d'opération. Cette augmentation diffère selon la bande de fréquences. (On voit dans le cas ci-dessous que les bandes de 4000 et 8000 Hz sont très peu influencées par le régime.) La distribution varie aussi d'un appareil à l'autre. (Dans d'autres cas, les niveaux Lp varient grandement aux hautes fréquences et peu aux basses fréquences.) Pour cette raison, le classement global des ventilateurs demeure difficile.

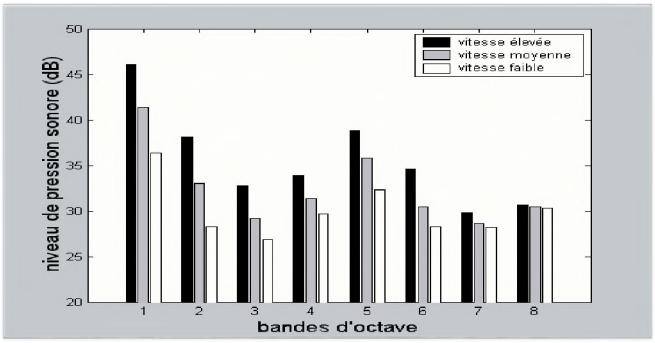


Fig. 4. Niveaux Lp d'un ventilateur en fonction de la vitesse.

3.4 Ventilateurs axiaux et centrifuges

Bien que les ventilateurs centrifuges soient reconnus pour être silencieux (Alton Everest), cette étude a révélé de pauvres performances acoustiques pour ce type d'appareils. Leur niveau de bruit n'est comparable à celui des ventilateurs axiaux que pour environ un tiers (1/3) de la capacité de ventilation des appareils axiaux.

RÉFÉRENCES

ACMA, norme 300-96 (1996). Reverberant room method for sound testing of fans.

Alton Everest, F. (2001). Master Handbook of Acoustics, McGraw-Hill, 4th ed.

Boudreau, M. (2004). V'là l'bon vent. Protégez-Vous, juillet 2004, 16-19.

Wu, S.F., Su, S.G. (1997). Modelling of the noise spectra of axial flow fans in a free field. Journal of Sound and Vibration, 200(4) 379-399.

REMERCIEMENTS ET NOTE

Cette étude a été commandée et subventionnée par les Éditions Protégez-Vous (projet C.D.T. P3006). Pour des raisons de confidentialité, la marque et le modèle des ventilateurs étudiés ne sont pas divulgués.

TIRE NOISE ASSESSMENT OF ASPHALT RUBBER CRUMB PAVEMENT

Steven Bilawchuk, M.Sc., P.Eng.
aci Acoustical Consultants Inc., Suite 107, 9920 – 63 Ave, Edmonton, AB, T6E 0G9
stevenb@aciacoustical.com

1. INTRODUCTION

With the ever increasing traffic volumes and prevalent desire to minimize residential noise levels, various noise mitigating methods are commonly employed. The use of Asphalt Rubber Crumb (ARC) pavement is widespread in the southern United States, and has a proven track record of performance [1]. Use of ARC in Canada, however, has been limited mainly to pilot projects covering relatively short sections of road. The purpose of this paper is to discuss measured noise level results obtained during a pilot ARC paving project conducted in and around Edmonton, Alberta in 2003.

2. PAVEMENT DESCRIPTION

Typical conventional asphalt pavement is comprised of aggregate (small rock) and a binder of 5% to 6% conventional asphalt cement by total weight [2]. The ARC mix used for the study contained approximately 7.5% to 8.5% asphalt rubber binder by total weight. The asphalt rubber binder itself contained approximately 19% rubber crumb by weight, thus about 1.4% to 1.6% of the total ARC pavement contained the rubber crumb. The rubber crumb typically comes from recycled vehicle tires. For this study the primary source was large truck tires.

In production, the asphalt mix is heated and the rubber crumb is added resulting in a gel-like material that surrounds and bonds with the asphalt cement. Application is exactly the same as conventional asphalt.

3. MEASUREMENT DESCRIPTION

Various road sections in and around the Edmonton area were paved with ARC as part of the 2003 pilot project. At one highway location (single lane in each direction and a posted speed limit of 100 km/hr), a 7 km stretch of old conventional pavement was re-surfaced with ARC pavement, and an adjacent 14 km stretch was re-surfaced with new conventional pavement. As such, a direct comparison could be made with *before* and *after* conditions for both ARC and conventional pavement.

3.1. LONG TERM NOISE MONITORING

A 26-hour environmental noise monitoring was conducted at both the ARC and conventional pavement locations both *before* and *after* the application of the new surface. The noise monitoring was conducted using a 30-second L_{eq} time period in both broadband (linear and A-weighted) and 1/3 octave band spectral analysis. In each case, the noise monitor was located approximately 20m from the centerline of the road. The key was to maintain the same location for both the *before* and *after* measurements to minimize the effects of distance attenuation, ground absorption, air absorption, and surface reflections. The 26-hour time was used so that a 2-hour observation period could be used at the same time on two consecutive days to document traffic conditions, determine consistency of traffic, and record subjective observations.

3.2. SHORT TERM MAXIMUM SOUND LEVELS

While on site for the 2-hour observation periods, the short term maximum sound levels obtained with specific vehicle pass-by's were noted. These maximum sound levels were collected and analyzed statistically to further determine the consistency in traffic conditions for each observation period, as well as give another measure of the amount of noise reduction.

3.3. CONTROLLED VEHICLE TESTING

The final measurement involved the use of a specific vehicle for controlled drive-by testing. A 2002 Dodge Grand Caravan (a very common vehicle type) was driven by the sound level meter at a specific distance (10 m) and speed (100 km/hr) in each road direction with engine on and off. The sound level meter was set to measure with 1-second L_{eq} sound levels in both broadband (linear and A-weighted) and 1/3 octave band spectra.

4.1. RESULTS: LONG TERM NOISE MONITORING

The results of the long term noise monitoring are presented in Table 1. There was a reduction in sound levels with the application of both the ARC and conventional asphalts. The amount of reduction with the ARC, however, was greater. Note the two key external factors that affected the measured sound levels (listed below the table). It is estimated that the day-time sound levels at the conventional location *before* would have been approximately 2-3 dBA lower than those measured resulting in less of a reduction from *before* to *after*. As well, the daytime sound levels at the ARC section *after* would have been approximately 1-2 dBA lower, resulting in a larger reduction in the sound levels from *before* to *after*. Both of these factors would have resulted in L_{eq}Day and L_{eq}24 sound reductions of approximately 6 dBA for the ARC section and 2 dBA for the conventional section.

In addition, there was a very small amount of traffic during the night-time. As a result, even small changes in vehicle counts for the night-time period will result in large changes to the L_{eq}Night. Due to this, the L_{eq}Night is not particularly useful for comparison at this location.

Table 1. Long Term Noise Monitoring Results

	Leq24 (dBA)	LeqDay* (dBA)	LeqNight (dBA)
Before (ARC)	57.8	59.2	52.9
After (ARC)	53.1**	54.9**	44.8
Difference (ARC)	**-4.7**	**-4.4**	**-8.1**
Before (Conventional)	58.7***	60.0***	54.8
After (Conventional)	54.5	56.3	46.0
Difference (Conventional)	**-4.2**	**-3.7**	**-8.8**

* Day-time hours are 07:00 – 22:00; night-time hours are 22:00 – 07:00

** Farm machinery operating in nearby field during day-time (results estimated to be 1-2 dBA higher than normal)

*** Abnormally high volume of dump-trucks during day-time (results estimated to be 2-3 dBA higher than normal)

The typical 1/3 octave band spectral results are shown in Fig. 1. It can be seen that both surfaces resulted in lower levels near 800Hz, but the ARC was consistently lower at frequencies beyond this, while the conventional resulted in no difference.

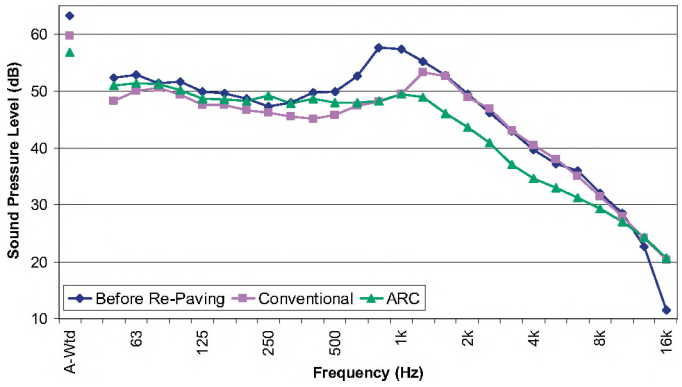


Figure 1. Typical 1/3 Octave Band Results

4.2. RESULTS: SHORT TERM MAX SOUND LEVELS

For light autos, the sound level reductions were much greater with the ARC section than the conventional pavement section. There was an increased amount of reduction with larger vehicles as well, but the difference between the ARC and conventional was not as great. This gives evidence that less of the total noise emanating from the larger vehicles is associated with tire noise than compared to the light autos (as expected).

4.3. RESULTS: CONTROLLED VEHICLE TESTING

The spectral results matched very will with those of the long term monitoring. This test also illustrated the increase in the slope for the rise and fall of the sound levels resulting from the passing vehicle. Thus, in conjunction with reduced maximum sound levels, the ARC pavement also reduced the length of time during which the higher vehicle pass-by sound levels occurred. The net effect is that residents in proximity to the roadway would experience both lowered maximum sound levels and shorter exposure times (both of which affect the L_{eq} sound levels).

4.4. RESULTS: SUBJECTIVE OBSERVATIONS

In general, it was noted that the use of ARC resulted in lower overall noise levels, as well as a substantially notable reduction in the mid to high frequencies. Essentially, it sounded as if the tire noise was somewhat "muffled" compared to both the old and new conventional pavement. The new conventional pavement was noted to have a slightly noticeable reduction in noise levels, but the frequency content of the noise did not change. In addition, newer light vehicles could be heard as far as 1 – 2 km away with the new conventional asphalt but only 300 – 400 m away with the ARC.

5.5. FUTURE WORK

There are still many important unknowns which should be addressed. Of prime importance for most regions within Canada is the effect of one or several freeze/thaw cycles. Road surface conditions such as partial snow or dirt/mud coverage and varying stages of road repair could have an impact on the noise levels. Also, variable mixtures of ARC could be investigated to find an optimal mixture for noise reduction. Finally, other vehicle related aspects such as different vehicle speeds could be investigated to determine the relative reduction levels for highway conditions compared to urban roads with slower speeds.

6.0. CONCLUSION

The use of asphalt rubber crumb pavement as a road surface material has been quantitatively and subjectively noted to reduce tire noise levels compared to conventional asphalt pavement. The various measurement techniques used to quantify the level of reduction all achieved similar results and the measured data corroborated well with subjective observations. Further work is also required to determine the longevity of the noise reduction benefits.

REFERENCES

1. Rubber Pavements Association, 1801 South Jentlilly Lane, Suite A-2, Tempe, AZ 85281 USA, Web: www.rubberpavements.org
2. Alberta Rubber Asphalt Project Report, Prepared for the Consulting Engineers of Alberta by EBA Engineering Consultants Ltd. of Edmonton.

DEVELOPMENT OF A SOUND MASKING SYSTEM FOR ROAD CONSTRUCTION NOISE

Tony Leroux[1], Jean-Pierre Gagné[1], Pierre André[2] and Line Gamache[3]
[1]École d'orthophonie et d'audiologie, Université de Montréal, C.P. 6128, succ. Centre-ville, Montréal, Québec, H3C 3J7
[2]Département de géographie, Université de Montréal
[3]Ministère des Transports du Québec, Direction Île-de-Montréal, 440, René-Lévesque ouest, Montréal, Québec, H2Z 2A6

1. INTRODUCTION

A large proportion of Montreal' highway network was built during the 60'. Today, many overpasses need major renovation. To avoid traffic jams, the road works are made mostly at night while the highway can be closed. A fairly large numbers of highways are located in densely populated areas and inhabitants living near by are complaining about the noise related to the construction works. The main objective of this project was to examine the effectiveness and feasibility of a noise masking technique used to reduce the population's annoyance related to emerging construction noises. The current research is a replication of Leroux et al. (2004).

The overpass Roi-René above the Metropolitain highway (A-40) has been crumbled down and rebuilt during the summer of 2004. The project was divided into three phases. The *first phase* measured the level of emerging noises at different distances from the construction sites. Surveys were also carried out among inhabitants to obtain their opinion and perception before and during the road construction. The *second phase* led to the evaluation, in laboratory, of the perception of the noise emerging from the construction site Subjects were asked to comment individually on the disturbing dimension of the noise generated by the construction work, presented against 12 different background noises designed to reduce the annoyance. Focus groups were also organized to get a closer look at the relation between psychoacoustics dimensions and emotions. A *third phase* examined, directly within the affected areas, how a system broadcasting various background noises selected from the second phase, was able to reduce the annoyance expressed by the inhabitants. This paper provides the results of the first two phases.

2. METHOD

2.1 Environmental analysis and survey

A sound level meter (Larson-Davis model LD 824) was used to measure the level of ambient noise before and during the construction noise in the residential neighbourhood ($L_{Aeq\text{-}15m}$). Samples of noise were recorded at 17 different locations onto a digital audio tape (Sony MZ-R55). In parallel, one hundred residents were asked to provide a rating of the level of annoyance at night, under two conditions: (1) while the highway was open to regular traffic at night and, (2) when it was closed for refection purposes.

2.2 Listening sessions and focus groups

Three groups of 20 subjects took part in the laboratory experiment: (1) young normal adults, (2) adult residents from the neighbourhood where the field work was conducted, and (3) adults that were matched to the subjects in group 2 except that they did not reside in the neighbourhood. The subjects were tested individually during three test sessions. They were seated in a quiet room (Hoeg et al., 1997) and they were instructed to rate the level of masking efficacy (using a 100 point scale arranged along five adjectives) of complex signals that were presented to them thru loudspeakers positioned in front of them. An adapted version of SEAQ (CRC, 2003) was used to present the signals and to record the responses. During the first session the subjects rated 12 different background noises that were played in isolation. During the following test session the subjects rated the same 12 background noises, presented simultaneously with a recording of the construction noise. Each masking signal was presented at a S/N ratio of + 5 dB. Finally, the subjects were asked to rate their emotional feelings when listening to the background noises. Six scales (7 point each) were used to assess the pleasantness, relaxing, restfulness, soothing and happiness dimensions.

Six (3 resident and 3 non-resident) focus groups were organized with each six subjects. The subjects were first asked to listen to series of artificial sounds manipulated according to 1) frequency, 2) amplitude modulation and 3) frequency/amplitude dimensions. The sound were obtained from a white noise digitally filtered using the above mentioned dimension. Subjects were then asked to rate the pleasantness (7 point scale) and provide a general appreciation of each sound using their own words. A discussion was then conducted asking subjects to provide their opinion about the best/worst sound for relaxing, sleeping, concentration and subjective stress response.

3. RESULTS

3.1 Environmental analysis and survey

The ambient noise measured during normal night operation of the A-40 highway show levels of noise varying from 65.2 to 68.9 dBA at the closet street to 50.2 to 53.4 dBA for farther away row of houses. The level of ambient noise when the highway is closed varies at the same

locations from 55.5 to 50.5 dBA. The noises generated by the construction itself are intermittent and the most frequent ones seem to be vehicle reverse alarms and jack hammers.

The responses revealed that the level of annoyance was greater when A-40 was closed than when it was open to regular traffic (see Figure 1). These responses were obtained even though the overall noise level was 10 dB lower when the highway was closed for refectory work. The residents identified the reverse alarm signals and the noise of the jack hammers as being the most annoying components of the noise they perceived.

3.2 Listening sessions and focus groups

The data were submitted to a Friedman rank sum technique (Mack & Skillings, 1980). The analysis made it possible to identify the masking noises that were the least annoying (i.e., most pleasant) and the more efficient for the participants. The ranking obtained for each of the 12 masking sounds are provided in Table 1. The results revealed that the four least annoying and more efficient maskers were: (1) continuous large waterfall, (2) babbling brook, (3) ocean waves and (4) sea. When looking at the subjects' emotional responses provided by Table 2, the ocean waves sound received the more positive emotional responses for all of four studied dimensions. As a basis of comparison the underwater bubbles sound evoked the more negative emotional responses.

The focus groups provided identification of desirable (least annoying) acoustic components in terms of frequency and amplitude modulation. As expected a low frequency spectrum was less annoying than a high frequency content. However, there is a lower boundary where low frequency evoked negative emotions such as fear. The mean for the most desirable frequency content is around 1754 Hz. For the amplitude modulation, a fast modulation was perceived as highly annoying while a slow modulation was less annoying. Again, a low boundary was found, a very slow modulation evoked exasperation and irritation. The mean for the most desirable amplitude modulation is around 2.8 Hz.

4. DISCUSSION

Many investigators have reported that indices of equivalent noise level. (i.e., L_{eq}) do not provide accurate predictions of annoyance for intermittent noise sources. The results of the present investigation are consistent with this observation. The study of the acoustic characteristics of masking signals that reduce the level of annoyance revealed optima for frequency content and amplitude modulation. Moreover, the non acoustic load/meaning characteristics of sounds seemed to be even more important to the receivers than acoustic factors.

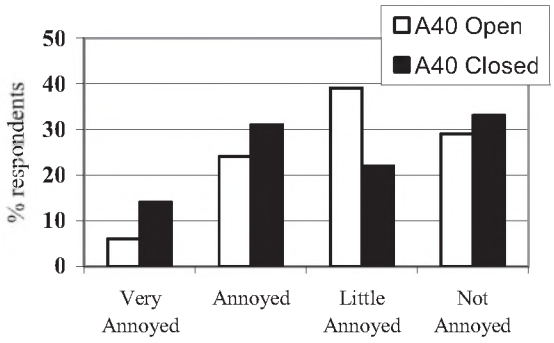


Figure 1: Level of annoyance

Table 1: Friedman's rank analysis for pleasantness and masking efficacy

Masking signals	Rank
Continuous large waterfalls	8,93
Babbling brook	8,84
Ocean waves	8,54
Sea	7,88
Heavy rain	7,23
Rainfall in traffic noise of low density	6,43
Heavy rain with reverberation	5,49
Whispering leaves	5,30
Heart beats	5,12
Frogs in a pond	5,04
Cat purr	5,02
Underwater bubbles	4,18

Table 2: Emotional subjective appreciation

Masking signals	Relax	Restful	Sooth	Happy
Ocean waves	5.55	5,42	5.48	5.40
Babbling brook	5.50	5.32	5.48	5.38
Waterfalls	5.45	5.30	5.50	5.12
Sea	4.77	4.98	5.05	4.97
Heavy rain	4.60	4.22	4.43	4.15
Underwater bubbles	*2.53*	*2.63*	*2.68*	*2.67*

REFERENCES

Communications Research Centre. (2003) (SEAQ) System for the Evaluation of Audio Quality, Ottawa.

Hoeg, W. et al. (1997). Évaluation subjective de la qualité du son - Systèmes et méthodes de l'UER. Revue Technique de l'Union européenne de radiodiffusion, 40-50.

Mack, G. A. & Skillings, J. H. (1980). A Friedman-type rank test for main effects in a two-factor ANOVA". J. Amer. Statist. Assoc., Vol. 75, 947-951.

Leroux, T., Gagné, J.-P., André, P. & Gamache, L. (2004) Discomfort reduction of a population exposed to noise related to a major urban highway refection. 18th ICA, IV-2639-2642.

ACKNOWLEDGEMENTS

Work supported by the Quebec Ministry of Transport.

ESTIMATING COMMUNITY NOISE IN A STANDARD WAY: A DISCUSSION ABOUT AND THE STATUS OF ISO 9613-2

Richard J. Peppin
Scantek, Inc., 7060 Oakland Mills Rd #L, Columbia, MD 20146 www.scantekinc.com, PeppinR@asme.org

1. INTRODUCTION

This summarizes the status of the work of a large working group of Canadians and Iraq occupiers who, on and off, have worked on developing a national, if not North American standard that is comparable to the international standard dealing with prediction of outdoor noise levels.

The ISO 9613-2, and subsequently, we hope, the new Draft ANSI S12.62 deals with the propagation of noise outdoors accounting for various environmental factors: barriers, foliage, berms, ground effects, etc. Essentially the Standard is a collection of algorithms that can be used with hand calculations or in computer programs; they can help estimate the noise from one or more sound sources. The method described in this standard is general in the sense that it may be applied to a wide variety of noise sources, and covers most of the major mechanisms of attenuation.

ISO 9613-2 is a much used standard in most countries in the world and is much maligned, sometimes deservedly so. This paper reviews the existing ISO 9613-2 and indicates issues the ANSI working group has with adopting the standard as written several months ago.

This standard specifies an engineering method for calculating the attenuation of sound produced by one or more sound sources during downwind propagation or, equivalently, for propagation under a well-developed moderate ground-based temperature inversion, such as commonly occurs at night outdoors, in order to predict the levels of environmental noise at a distance from a variety of sources. The method predicts the equivalent continuous A-weighted sound pressure level under meteorological conditions favorable to propagation from sources of known sound emission (sound power level).

The method also predicts a long-term average A-weighted sound pressure level. The long-term average A-weighted sound pressure level encompasses levels for a wide variety of meteorological conditions, and seasons of the year.

The calculation method consists specifically of octave-band algorithms (with nominal midband frequencies from 63 Hz to 8 kHz) for calculating the attenuation of sound, which originates from a point sound source, or an assembly of point sources. The source (or sources) may be moving or stationary.

This method is applicable in practice to a great variety of noise -fixed base installations, surface transportation sources, industrial noise sources, construction activities, and many other ground-based noise sources. It does not apply to sound from aircraft in flight, or to blast waves from mining, military, or similar operations.

Specific terms are provided in the algorithms for the following physical effects:

- Geometrical divergence,
- Atmospheric absorption,
- Ground effect,
- Reflection from surfaces,
- Screening by obstacles,
- Propagation around one or more buildings (housing),
- Foliage, and
- Industrial sites

The basic equation is

$$L_{fT}(DW) = 10 \lg \left\{ \frac{\left[\left(1/T\right) \int_0^T p^2{}_f(t)\, dt \right]}{p_0{}^2} \right\} dB$$

Where

$L_{fT}(DW)$ = downwind average octave band level, f is octave band mid frequency, p = sound pressure, p_0 = 20 μPa, and T= integrating time

This standard assumes that the wind blows from the source to the receiver. If this standard is used either to predict at multiple receivers or to draw noise contours, then it implicitly assumes that the wind is omni-directional, and simultaneously blows from the source to each receiver.

The equations to be used are for the attenuation of sound from point sources and their image sources from reflecting objects other than the ground. Extended noise sources therefore, such as road and rail traffic or an industrial site (which may include several installations or plants, together with

traffic moving on the site) are represented by a set of sections, each having a certain sound power and directivity. Attenuation calculated for sound from a representative point within a section is used to represent the attenuation of sound from the entire section. A line source may be divided into line sections, an area source into area sections, each represented by a point source at its center.

$$L_{fT}(DW) = L_W + D_c - A$$

where

L_W is the octave-band sound power level, in decibels, produced by the point sound source relative to a reference sound power of one picowatt (1 pW):

D_c is the directivity correction, in decibels, that describes the extent by which the equivalent continuous sound pressure level from the point sound source deviates in a specified direction from the level of an omnidirectional point sound source producing sound power level L_W. D_c equals the directivity index D1 of the point sound source plus an index $D\Omega$ that accounts for sound propagation into solid angles less than 4π steradians; for an omnidirectional point sound source radiating into free space, $D_c = 0$ dB;

A is the octave-band attenuation, in decibels, that occurs during propagation from the point sound source to the receiver.

These values of "A" are determined from the algorithms.

2. STATUS TO DATE (August 2004)

The working group, composed of perhaps 40 people of which ten or so are active, originally was planning to help produce an ANSI Nationally Adopted International Standard (NAIS) that would essentially adopt all technical requirements of the ISO 9613-2 (which it had to) but also contain editorial changes and an annex that provided notes and commentary to help interpret sections that were a problem. For example, the barrier equation includes a non-zero term for a small path length difference which implies even a roadway curb will give 3-5 dB attenuation as a barrier.

As a result of financial issues, and after all, for developers as apposed to volunteers, standards is a moneymaking proposition, ASA will not develop any more NAIS documents. This gives the working group an opportunity to start fresh and develop a standard set of algorithms, like ISO 9613-2 or not, that are based on agreed upon ways of doing things. It implies that all who use the standard will estimate the same propagation losses.

A big issue is the relationship between the actual propagation and those that estimated with this standard. Many users of software expect the software to predict levels measured. I submit that is hard or impossible to do because:

1- Little is known about micro atmospheric conditions on any day, much less based on some weather prediction

2- Ground impedance is not well characterized and its relationship to propagation losses is not well developed. In any location with hard/soft/ conditions, chances are good that only a few people, if any, have good data at all with which to input to a program.

3- Foliage data are almost none existent and the data available suggest a very large uncertainty.

4- Sound power data of machines are either unknown or with high uncertainty because the data do not reflect actual operating conditions, added appurtenances (like conveyor belts or duct attenuators the field performance of which is unknown), and so on.

5- Sound data are often transient.

6- Configurations of propagation paths are often unknowable or so complex that data to describe them are not feasible to obtain.

With all of this, it remains for the working group to determine how to resolve this. I strongly believe, though, that every algorithm must include an estimate of uncertainty, which I suspect will be high, to indicate to laypersons and to professional that the narrow contours on a piece of paper

REFERENCES

ISO-9613-2, 1996 "Acoustics — Estimation of outdoor sound propagation by calculation"
Bies, D.A. and Hansen, C.H., Engineering Noise Control, 2nd Ed.

Woodyard Sounds in the Community

Cameron W. Sherry
President Enviro-RISQUE Inc., P.O. Box 190, Howick, Quebec, J0S 1G0

INTRODUCTION

The pulp and paper mill is composed of several major components spread over a large area. One of these components is the North Woodyard that is located north of the parking lot used by the employees and covers an area of about three blocks by two blocks. To the east of the yard are a railway right of way and another major industrial site that has east of it a major highway. To the north is a street that has commercial and manufacturing sites on either side of it. On the west side is a residential area where the critical sound receptors are located. The closest residences are about 130 metres from the property line. Also adjacent to the yard on the west side is the Municipal water filtration plant and the Bus Transit garage. In between the west residential area and the company property are another railway right of way and a natural gas line right of way.

The North Woodyard has been in operation for many years. The major operations are:

1. Log-handling yard
2. Woodroom (debarking and chipping plant)
3. Chip-handling yard
4. Screenroom (chip-blowing and screening system)

The community around the mill is familiar with the mill sound levels and there are no known complaints from the area. The sound levels at the critical receptor locations are in excess of those given in the publications of the Ontario Ministry of the Environment. The mill North Woodyard sounds that can be heard at the east end of the streets prior to the modifications made in September and October of 2001 were:

1. Pay loaders
2. Chip blower
3. Chip line

In addition, there is a significant noise contribution from the buses at the Transit garage, the nearby water purification plant, a plant to the north, plant noise from those portions of the mill to the south of yard, road traffic on the surrounding streets and railway traffic.

RESULTS

In the fall of 2001 the mill decided that it would improve the quality of its chip stock by installing a new chip handling system. This initiated the requirement of obtaining a certificate of approval from the Ontario Ministry of the Environment. Initially it was proposed that the new system would not make any more noise at the receptor locations. This was not acceptable to the ministry as the sound levels at any of the three time periods was in excess of the guidelines of 45 dBA (23:00 to 07:00), 47 dBA (18:00 to 23:00) and 50 dBA (07:00 to 18:00).

A predictive model was developed was developed based on the sound levels of the existing equipment and the new equipment. It was very obvious that the ministry guidelines could not be met without carrying some mitigation steps. The first step was building of a five-metre high earth berm around the west side of the property joining up with an existing three-metre high earth berm on the north side.

Twenty-four hour measurements showed that the yard was not in compliance. The berm was effective in reducing the paylift yard noise at the northern receptor, however, the sounds from the silencers on the blowers was excessive. The next step was to construct a Durisol wall around the silencers.

Twenty-four hour measurements showed that the northern receptor was for all practical purposes in compliance when there was a gentle west breeze. The southern receptor because the berm was not wrapped around was not in compliance. The next step was to complete the construction of the berm along the south side up to the office.

The measurements in the spring of this year appear to indicate that the sound level at each receptor is no longer affected by the sounds from the yard. However, it appears that since this project has started the background sound levels have increased and the ministry guidelines are still being exceeded. A time period is now being looked for when the background sound levels can be re-measured while the yard and the mill are down.

ACOUSTIC SIGNAL PROCESSING DEVELOPMENTS FOR NON-INVASIVE MONITORING OF VITAL SIGNS AND ULTRASOUND INTRACRANIAL SYSTEMS

Stergios Stergiopoulos
CANAMET Inc., 1120, Finch Ave West, Suite 201, Toronto ON, M3J-3H7, stergios@canamet.com

1. BACKGROUND

Current system concepts for non-invasive monitoring of vital signs are limited in providing estimates for blood pressure, ECG, pulse oxymetry and tympanic core temperature. Although this information is considered to be sufficient in most emergency and search and rescue operations, the lack of accuracy of these vital signs measurements makes the relevant system concepts unattractive to medical practitioners.

For the specific case of monitoring the blood pressure, the traditional system concepts are based on the oscillometric technique rather that the auscultatory method that monitors the sounds as blood flows through the brachial artery (i.e. Korotkoff sounds) in the same way as the medical practitioner employs a stethoscope along with a mercury sphygmomanometer. In noisy and vibration intensive environments the pressure fluctuations caused by these disturbances are sizable compared to the pressure fluctuations that need to be detected for the proper operation of the technique, thereby reducing the accuracy. Furthermore, in noisy environments, such as onboard (or near) a rescue helicopter or ambulance, the environmental noise frequently overwhelms the acoustic signal of interest, making it impossible to make a measurement using the conventional auscultatory method.

2. SYSTEM CONCEPT AND IMPLEMENTATION

Canamet Corporation has overcome these issues by using adaptive signal processing techniques that include adaptive interference cancellation, band pass filtering, and peak discrimination (pattern recognition) algorithms.[1,2]. The use of these advanced signal processing techniques leads to a feasible system for providing vital sign measurements in challenging noisy environments, such as ambulances and helicopters. The block diagram in Figure 1 shows the block diagram layout for the blood pressure measurement system. More specifically, Canamet's Piesometer MK-1 portable blood pressure monitor has a primary acoustic sensor integrated with the cuff and placed on the brachial artery to collect the Korotkoff sounds and any environmental noise. A secondary acoustic sensor (on the back of the patient's arm) collects only the environmental noise, as shown in Figure 1. The adaptive interference cancellation algorithm is a non-linear filter that removes the interference measured by the secondary sensor from the desired signal received by the primary sensor that was corrupted by environmental noise. The band pass filter removes any noise outside the frequency of interest. Finally, the peak discrimination algorithm extracts valid peaks from the Korotkoff sounds in the acoustic signal that result from heartbeats. Peaks that do not satisfy these constraints are discarded; however a degree of arrhythmia is accounted for during this process. In summary, Canamet's Piesometer MK-1 system is based the auscultatory technique that emulates the operations of the medical practitioner using the mercury sphygmomanometer and the stethoscope. This system has proven with clinical trials to have accuracy equivalent to that of a medical practitioner for patients at resting position [1,2].

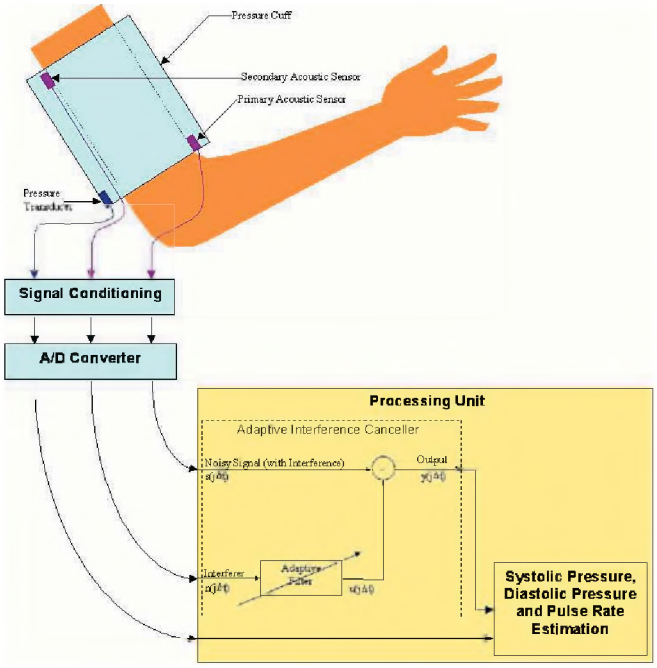


Fig 1: System Concept of Adaptive Noise Cancellation for Blood Pressure Systems based on the auscultatory method

However, advanced adaptive algorithms are extremely computationally intensive, requiring the execution of billions of operations per second (GOPS), and relying heavily on the use of Fast Fourier Transforms (FFTs) for frequency domain operations, such as time delays and spectral analysis. Conventional DSP architectures are not up to the task. Fixed-point computing architectures do not possess the appropriate processing capabilities to efficiently execute these computationally intensive algorithms and tend to introduce inaccuracies in the weight vector calculation that actually increase the noise in the system. Floating-point arithmetic enables much more accurate calculations and provides faster development cycles because the C-code does not have to be translated to a fixed point format.

In addition to the computationally intensive requirements, the need for miniaturization and portability with the

requirement to include into the system's architecture a real time clock, microcontrollers, I/O analog/digital peripherals, telemedicine functionality through serial port/USB and graphic interfaces for user friendly operations, makes the existing floating point DSP processors not very attractive for this kind of advanced and portable medical system applications. An ideal DSP processor architecture should include all the above peripherals and functionalities. Canamet's Adaptive blood pressure system is designed to be part of an open system. This type of open system design is shown in Figure 2. It allows for the integration of a number of modular designs into the complete system, based on need.

An addition to the system shown in Figure 2 is the non-invasive monitoring of the density variations of the brain due to changes of temperature or pressure, stroke, heat stroke, head injuries, hemorrhage, variations of blood flow in the skull due to drug effects, variations in metabolism and stress [3]. This is a new system innovation, called intracranial ultrasound technology. The design of this system mirrors the design of the blood pressure monitoring system as a monitoring intracranial vital signs system. Furthermore, a portable 3D ultrasound system technology [4,5] can be integrated as a modular unit into the above open system design.

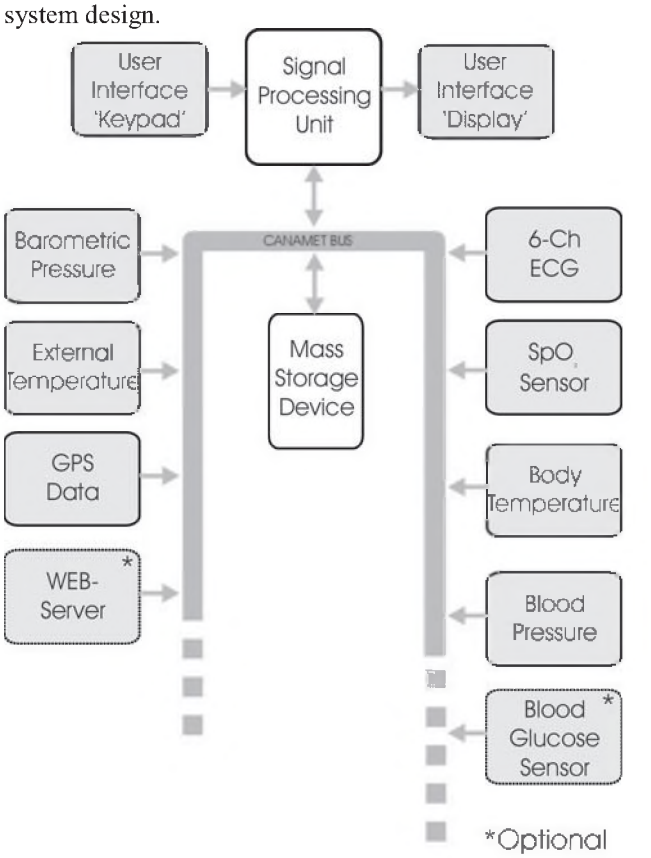


Fig. 2: Canamet's Open Modular System Design for Monitoring Vital Signs.

3 CONCLUSION

This article describes the evolution of innovative new acoustic signal processing algorithms and DSP architectures implemented in a wide variety of new medical electronics applications that form the EMERGING TRENDS IN THE FIELD OF NON-INVASIVE MEDICAL DIAGNOSTIC TECHNOLOGIES.

The advance signal processing structure of the present development and its implementation into a system computing architecture has demonstrated successful performance in obtaining:

- Blood Pressure Monitoring of systolic - diastolic pressure in noise intense environments, such as helicopters, ambulances, emergency rooms, etc.
- Implementation of an open system architecture with a common system bus to allow the design through a generic system design.
- Non-invasive monitoring of the density variations of the brain due by intracranial ultrasound technology.
- Continuous 24-hour monitoring of vital signs, such as pulse oxymetry, six-electrode (vector) ECG with high sampling rate (1KHz) to allow diagnosis of a wide spectrum heart diseases, tympanic ear thermometer.

It is anticipated that the present development would address most of the medical requirements for non-invasive monitoring of vital signs for Home Care, in Hospital Emergency Departments, Ambulances, Hospital Intensive & After Care, and applications for Family Medical Clinics, Insurances and Old-Age Nursing Homes.

REFERENCES

[1] Pinto L., Dhanantwari A., Wong W. and Stergiopoulos S., "Blood Pressure Measurements in Noise Intense Environments Using Adaptive Interference Cancellation", Annals of Biomedical Engineering, 30, 657-670, 2002.

[2] Stergios Stergiopoulos and Amar Dhanantwari, "Method and Device for Measuring Systolic and Diastolic Blood Pressure and Heart Rate in an Environment with Extreme Levels of Noise and Vibrations", US Patent 6,520,918, issued 18 February 2003.

[3] Stergiopoulos S. and Wrobel M. "Non-Invasive Diagnostic Ultrasound System Monitoring Brain Abnormalities", US Patent application (DND File No. 1416-01/007 USA), Dec-2001.

[4] A. Dhanantwari, S. Stergiopoulos, C. Parodi, F. Bertora, P. Pellegretti, A. Questa, "An efficient 3D beamformer implementation for real-time 4D ultrasound systems deploying planar array probes", Proceedinds of the IEEE UFFC'04 Symposium, Montreal, Canada, Aug 2004.

[5] Stergiopoulos S. and Dhanantwari A., "High Resolution 3D Ultrasound Imaging System Deploying A Multi-Dimensional Array of Sensors and Method for Multi-Dimensional Beamforming Sensor Signals", United States Patent 6,482,160, 19 Nov-2002.

PERFORMANCE OF ULTRASONIC IMAGING WITH FREQUENCY-DOMAIN SAFT

Daniel Lévesque, Alain Blouin, Christian Néron, and Jean-Pierre Monchalin
Industrial Materials Institute, National Research Council Canada,
75 de Mortagne, Boucherville, Québec, Canada, J4B 6Y4 daniel.levesque@cnrc-nrc.gc.ca

1. INTRODUCTION

Ultrasonic testing is widely used for detecting, locating and sizing flaws in many applications. Lateral resolution and signal-to-noise ratio (SNR) can be improved by focusing the acoustic field with lenses or curved transducers or by using a computational technique that performs the focusing numerically. The last method is known as the Synthetic Aperture Focusing Technique [1]. Originally developed in the time domain, SAFT can be advantageously implemented in the frequency domain (F-SAFT). F-SAFT is based on the plane wave decomposition of the measured ultrasonic field [2] and provides an accurate and time-efficient algorithm for 3-D reconstruction. Also, laser-ultrasonics has brought practical solutions to a variety of NDT problems that cannot be solved by using conventional ultrasonic techniques [3, 4]. It uses two lasers, one with a short pulse for the generation of ultrasound and another one, long pulse or continuous, coupled to an optical interferometer for detection. Recently, we have introduced several improvements to the F-SAFT algorithm and we have coupled it to laser-ultrasonics [5, 6]. In this paper, we demonstrate the capability of this technique for detecting and characterizing flaws in structural materials.

2. F-SAFT METHOD

We consider that the generation and detection beams are focused at the same location onto the surface. Following acquisition of a grid of signals over the surface of the tested part, traditional SAFT identifies a defect by summing all the signals within a certain area (the aperture) after giving proper time-delays. This data processing approach, while straightforward in its principle and implementation, is very computation intensive. An alternative method is F-SAFT whereby the data processing is performed in the 3-D Fourier space using the angular spectrum method [2]. Starting from the measured ultrasonic field at the sample surface, a 3-D Fourier transformation is performed. Then, the transformed field is backpropagated to any depth and an inverse 2-D Fourier transformation yields the sub-surface image.

Recent improvements to the F-SAFT algorithm include temporal deconvolution to enhance both axial and lateral resolutions, control of the aperture to improve signal-to-noise ratio, as well as spatial interpolation in each sub-surface plane. The control of the aperture also allows reduction in the sampling requirements therefore making the technique more attractive for industrial use. Usually, the amplitude of laser-generated longitudinal (L) waves decreases away from the normal to the surface and one should limit the aperture to a maximum value to avoid the addition of noise. Also, laser-generated shear (S) waves have an angled emission pattern and one should use an annular aperture for the same reason. Introducing a lower limit could also be useful for L-waves to reduce direct contribution from the backwall and improve contrast in the image. By considering the direction cosines of the wave components, it follows a condition on the temporal frequencies to be included in the summations [6]. For L-waves, it has been shown that the SNR rapidly increases with aperture angle, reaches a maximum for about 30^{o} and then progressively decreases by at least 6 dB, as a result of including components that contributes more to noise than to coherent signal. In addition, the processing time is further reduced since less data points are included in the summations. Other improvements in the implementation of F-SAFT can be found in Ref. [5].

3. RESULTS AND DISCUSSION

A first application to illustrate the capability of F-SAFT with laser-ultrasonics is the detection of inclusions in steel slabs. The presence of small defects in cast slabs (voids and particular inclusions), mainly found within a depth of about 10 mm below the surface, is hard to avoid and could lead to a variety of problems down the processing line. Samples taken from an industrial steel slab and having flat-bottom holes at different depths were tested. The inspection was performed in a scanned area of 28 X 28 mm^2 with a step size of 0.2 mm. The F-SAFT reconstruction was made with L-waves and results on a descaled sample are presented in Fig. 1. The figure includes the amplitude C-scans (top) gated at the depth of the holes, and B-scans (bottom) across the holes. In Fig. 1a, the presence of the deepest 1 mm hole is seen in the C-scan (at the intersection of the two cursors), with a depth estimated from the B-scan to be 3.6 mm. The B-scan also shows the 2 mm hole at a slightly different depth. In Fig. 1b, the C-scan shows the presence of the 1 mm hole near the surface and the depth estimated is about 0.9 mm by considering the strongest reflections in the B-scan. It is found that a step size of nearly 1 mm would be sufficient to detect the 1 mm holes. A statistical approach where randomly selected areas on the steel slab are inspected can be considered in practice.

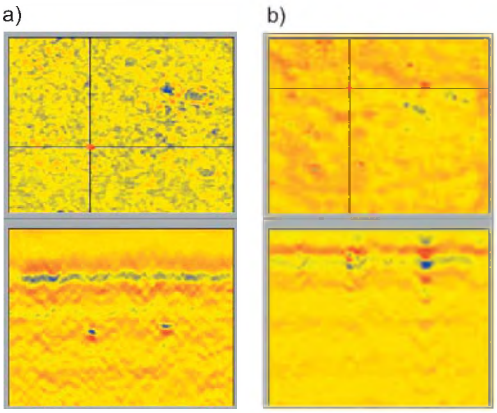


Fig. 1. C-scan (top) and B-scan (bottom) of a descaled sample, with the 1 mm holes at a depth of 3.6 mm in a), and near the surface at 0.9 mm in b).

A second application is the visualization of stress corrosion cracks (SCCs) using laser-ultrasonics. Tested samples were made from stainless steel and contained a few SCCs, each typically less than 0.05 mm wide and from 0.5 to 5 mm deep. The samples were interrogated from a surface opposite to cracking. The step size of the scan was 0.2 mm and the scanned area was about 40 x 40 mm^2. The aperture was from 5° to 25° for L-wave reconstruction and from 35° to 55° for S-wave reconstruction. C-scans obtained using L-waves and S-waves are compared to the images by liquid penetrant testing (PT) in Fig. 2. Each C-scan was obtained by selecting the maximum amplitude in a narrow gate at depths corresponding to the cracking surface. F-SAFT provides very fine images of the surface which contains SCCs with a detectability that can exceed that of PT, being the most sensitive method to visualize fine cracks. When comparing images, the spatial resolution of F-SAFT with S-waves is higher. Indeed, B-scan images reconstructed with L-waves show the crack roots as a lack of signal having a width not well defined. The image with S-waves shows a feature having a cross-like shape ("X" shape), the center of which has the maximum amplitude and is located at the position of the crack root on the surface. The "X" shape is related to the existence of a corner near the crack root and further investigations have shown that the amplitude at the cross point is indicative of the crack depth. Notice that the presence of the "X" shape is a particular feature that only laser-generated S-waves can produce.

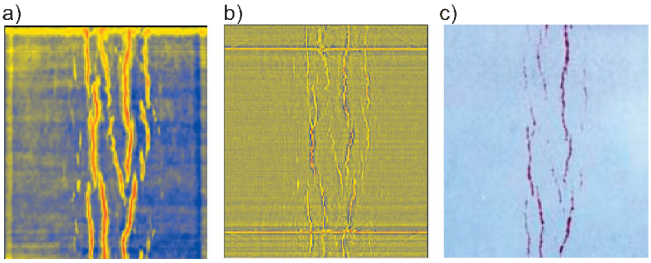


Fig. 2. C-scans of the cracking surface on a SCCs sample using a) L-waves, b) S-waves and c) PT.

A third application is visualization of delaminations along curved interfaces. An example is Mg/Al casting, composed of an Al core enclosed in a Mg shell and investigated by an automotive manufacturer for engine parts. In the case of a Mg/Al interface not parallel to the top Mg surface, the ultrasonic wave can be reflected obliquely from the interface to any location on the Mg surface. F-SAFT allows synchronization of the ultrasonic signals scattered back in the different directions. A testpiece of Mg/Al casting was made with flat-bottom holes drilled in the Al part down to the interface of the two alloys. The inspection was from the Mg surface in an area of 40 x 15 mm^2 with a step size of 0.2 mm. Fig. 3a and Fig. 3b are C-scans taken at depths of 8.2 and 9.7 mm respectively. Fig. 3c is a B-scan showing the curved Mg/Al interface as well as the depth of the two C-scans. In Fig. 3b, two flat-bottom holes of diameters 10 and 5 mm are observed at a depth of 9.7 mm where the interface is almost flat and parallel to the scanned surface. In Fig 3a, corresponding to a depth of 8.2 mm, a horizontal line of variable amplitude is observed. This line is the intersection of the curved Mg/Al interface with the observation plane and the darker amplitudes are indicative of delaminations.

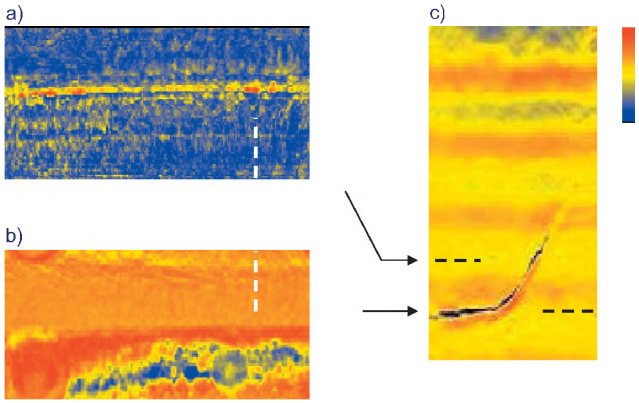


Fig. 3. SAFT images of a non-parallel Mg/Al interface. The C-scans a) and b) have depths indicated in the B-scan c). The B-scan is at position indicated in a) and b).

REFERENCES

[1] Doctor, S. R. et al. (1986), "SAFT - the Evolution of a Signal Processing Technology for Ultrasonic Testing," *NDT-International* 19, pp. 163-167.

[2] Busse, L. J. (1992), "Three-Dimensional Imaging Using a Frequency-Domain Synthetic Aperture Focusing Technique," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 39, pp. 174-179.

[3] Scruby, C. B. and L. E. Drain (1990), "Laser-ultrasonics: Techniques and Applications," Bristol, Adam Hilger.

[4] Monchalin, J.-P. et al. (1998), "Laser-ultrasonics: from the Laboratory to the Shop Floor," *Advanced performance materials* 5, pp. 7-23.

[5] Lévesque, D. et al. (2002), "Performance of Laser-Ultrasonic F-SAFT Imaging," *Ultrasonics* 40, pp. 1057-1063.

[6] Ochiai M. et al. (2004), "Detection and characterization of discontinuities in stainless steel by the laser ultrasonic synthetic aperture focusing technique", *Mater. Eval* .62, pp. 450-459.

ANALYSIS AND SYNTHESIS OF THE GUQIN -A CHINESE TRADITIONAL INSTRUMENT

Ying Ding and David Gerhard
Department of Computer Science, University of Regina,
3737 Wascana Pkwy, Regina, SK S4S 0A2.
Ding200y@uregina.ca

1. BACKGROUND

The GuQin, a seven-stringed zither, is China's oldest stringed instrument, with a history of some 3000 years. In Chinese, "Gu" means old and "Qin" means musical instrument. Historically, GuQin was referred to as " Qin " in most ancient Chinese writings. Owing to its long history, it has been widely called GuQin during the last 100 years. This summer in China, we conducted an interview with Fengyun Lee, a renowned master of the GuQin. She introduced the characteristic of GuQin and showed some of typical playing techniques to help us to understand the acoustics and techniques related to the GuQin.

Throughout the history, the GuQin has been viewed as a symbol of Chinese high culture. Qin music has been regarded as the essence of highly complicated aesthetics, philosophy, and musical temperament. It was the popular instrument of the Chinese literati, who played it for self-cultivation and personal enjoyment. In ancient China, a well-educated scholar was expected to be skilled in Four Treasures: Qin (the GuQin), Qi (Chinese chess), Shu (calligraphy) and Hua (painting).

In the long history of Chinese music, GuQin is most able to express the essence of Chinese music. More than 100 harmonics can be played on the GuQin, allowing the instrument to produce a large number of overtones. Undoubtedly, the GuQin is a part of world's heritage, but today fewer than two thousand people can play it and only about 70 pieces have been transcribed into modern western music notation. In order to preserve world heritage, we propose using physical modeling techniques to simulate the acoustic GuQin. This paper will introduce these models and how they apply to the GuQin.

2. PHYSICAL MODELLING

Physical modeling of musical instrument is an exciting topic in digital sound synthesis, and it has been a very popular research domain since the 1980s. By using physical modeling techniques that simulate a real instrument, the sound of virtual instrument can be more realistic and controllable, because the sound created is based on real-world information and parameters. To model the process of plucking a GuQin string to create a synthesized "virtual GuQin", the simplest algorithm to use is the Karplus-Strong algorithm.

A typical elementary GuQin tone consists of a short fast attack and long smooth decay. The Karplus-Strong algorithm can be used to model this plucked-string behaviour.

The original Karplus-Strong algorithm is based on wavetable synthesis. In wavetable synthesis, a buffer is filled with samples that are then read by a pointer that circulates from the beginning of the table to the end and then starts back at the beginning, thus creating a periodic sound. The initial conditions of the wavetable determine the resulting timbre and its length determines the periodicity of the sound, and hence the pitch. Kevin Karplus and Alex Strong [4], two researchers at Stanford University, used similar idea in the early 1980s. Karplus and Strong suggested using a delay line to represent the waves on the string. The waveguide is initially filled with random values or white noise to represent the pluck effect, and the wavetable is low-pass filtered to simulate the decay in the higher-frequency harmonics over time.

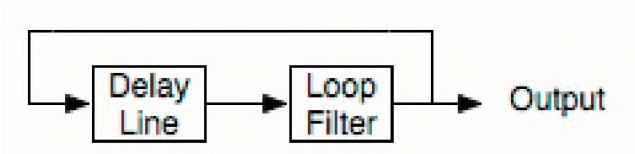


Figure 1. The Karplus-Strong Algorithm

Karplus and Strong have explored different variations and refinements of their algorithm. Jaffe and Smith [1] also proposed a number of ideas that improved the Karplus-Strong algorithm in its usefulness and accuracy. They not only solved some of the original Karplus-Strong algorithm problems, they also proposed a dynamic control system. Meanwhile, Sullivan suggested a number of methods based on Karplus-Strong. He suggests using two delay lines, one to represent waves travelling in one direction and another to represent the opposite direction. By this method an electric guitar's pickup and its output were represented. Smith [5] extended the ideas of physical modelling and model-based synthesis further, resulting in the digital waveguide technique, an important principle for discrete-time modelling of acoustic system. Later, Karjalainen *et al* [3] presented a detailed derivation of reduced string models.

This has shown how the two-directional digital model can be reduced to a more efficient signal delay loop (SDL) model. They also mentioned the idea of dual polarization of the guitar string. The system is different from the Karplus-Strong algorithm in that the excitation is provided by an impulse instead of random noise.

3. SIMULATION OF THE ACOUSTIC GUQIN

From the description of playing techniques, we now move to analysis and synthesis of the GuQin. Basically, the GuQin is able to produce three types of sound: *san yin* (open notes); *fan yin* (harmonics), and *an yin* (stopped notes). Some of the most important and typical analysis and simulation are present in below.

It should be noted that since GuQin music is usually used to express the player's feeling and personality, sometimes a harder attack might be required than is available in current physical modelling techniques. This is also a special part of GuQin.

3.1 Tuning

Open notes simply require that the correct string is plucked. By using Karplus-Strong synthesis, the strong and clear sound for the important notes of GuQin can be produced. More tasteful sounds could possibly be produced by extended the original Karplus-Strong algorithm. Therefore, a further refinement is achieved by applying other types of filtering. When a GuQin string is plucked, the high pitch decays faster than the low pitch. The low frequencies will last longer in sometimes non-harmonic ways. To entirely control the decay-time, we should know how to shorten or stretch the decay. These are solved by introducing a loss factor and a strength factor for modifying note duration. When playing GuQin using the fleshy part of the finger, the loss factor is set close to 1 to simulate a string being damped by soft materials.

3.2 Simulation of pluck location

When playing GuQin, pluck location is an essential element to decide tone. An effective solution of simulating pick position is to introduce comb filter. This can fairly accurately mimic the effect of plucking a string at varying distances from the bridge of GuQin.

3.3 Simulation of Sliding Effect

One of the characteristics of GuQin music is the number of glides. Indeed some authorities prefer to call stopped notes. It is also called sliding sounds. To play stopped notes, a left hand finger presses the string firmly while the right plucks, then may slide to other notes or create many different vibratos or ornaments, to achieve delicate and expressive sounds. By changing the delay buffer length over time, a rapid alternation of ascending and descending pitch changes is possible, offering a sliding sound.

3.4 Simulation of Style of Attacks

Attack pattern is an important element in perception of timbre, and using a different degree of finger pressure or pluck strength will create a different timbre as well. The inwards and outwards movements of different fingers on the string can change the attack character. A realistic simulation of attack style has been created by using a one-pole, lowpass filter.

3.5 Sympathetic String Simulation

A sympathetic vibration will be produced by a string close to the string that is plucked, which has a resonant frequency close to the frequency of the note played on the plucked string. Here we can define that as in the basic algorithm the input of random value is indicated as the plucked string, and algorithm is excited by plucked string is indicated as the sympathetic string. In the original Karplus-Strong, each string and resonance has the same pitch. To present sympathetic string effect, it needs a copy of string by a small percentage of the output from the plucked string to create a different pitch.

The sympathetic string can be seen to act as a bandpass filter. Frequencies close to the resonances of the partials of the plucked string will be reinforced, and all other partials will die away. To prevent an overflow problem with harmonic reinforcement, a loss factor will be involved in this modification.

References

[1] D.A. Jaffe and J.O.Smith. "Extension of the Karplus-Strong Plucked-String Algorithm". Computer Music Journal 7(2): 56–69, 1983.

[2] T. John. GuQin Silk String Zither. [online] http://www.silkqin.com/, accessed 15 May 2004

[3] M. Karjalainen, V. Välimäki, and T. Tolonen, "Plucked string models: from Karplus-Strong algorithm to digital waveguides and beyond," Computer Music Journal, 22(3): 17–32, 1998.

[4] K. Karplus and A. Strong. "Digital synthesis of plucked string and drum timbres," Computer Music Journal 7(2): 43–55, 1983

[5] J. O. Smith. "Physical modeling using digital waveguides." Computer Music Journal, 16(4): 74–91, 1992.

PANOCAM: COMBINING PANORAMIC VIDEO WITH ACOUSTIC BEAMFORMING FOR VIDEOCONFERENCING[1]

David Green and Mark Fiala
Institute for Information Technology, National Research Council of Canada, M-50, 1200 Montreal Rd, Ottawa ON Canada K1A 0R6, dave.green@nrc-cnrc.gc.ca

1. INTRODUCTION

Videoconferencing systems in use today, typically rely either on fixed or pan/tilt/zoom cameras for image acquisition, and close-talking microphones for good quality audio capture. These sensors are unsuitable for scenarios involving multiple users seated at a meeting table, or non-stationary users. In these situations, the focus of attention should change from one talker to the next or should track a moving talker. This paper describes an experimental, combined panoramic video camera and microphone array which is placed at the centre of a meeting table and which can detect and track in real-time the talkers seated around a table.

Kapralos [1] uses a panoramic camera and a simple microphone array for videoconferencing. This work discusses pointing accuracy but does not address talker tracking. Cutler [2] describes a panoramic system composed of multiple cameras and a beamforming microphone array used to archive meetings. The system that we describe uses video to locate potential talkers, and a circular beamforming microphone array to continuously search the meeting space for sound cues. Only rudimentary calibration is required. In real-time, we can select the talker candidates in the video image using a combination of cues while directional audio is used to choose the active talker. Finally, we have integrated the above functionality into an application that is compatible with Microsoft NetMeeting (http://www.microsoft.com/windows/netmeeting/) using OpenH.323 (www.openh323.org/).

2. PANORAMIC VIDEO SYSTEM

The video system is composed of a Pixelink digital video colour camera (http://www.pixelink.com/)) fitted with a Remote Reality NetVision Assembly B panoramic lens/mirror assembly (http://www.remotereality.com/) (Figure 1.). It captures a color image of 1280x1024 pixels of which an annular region of 800 pixels diameter contains the panoramic image.

A first transformation warps the useful pixels in the raw image into a standardized panorama accounting for all device specific parameters such as focal length, radial profile, etc. A second transformation produces a final image with correct perspective. A number of cues were investigated to detect candidate talkers in the image. These included skin colour, motion, "face" [3] based on OpenCV (www.intel.com/research/mrl/research/opencv/overview.htm), and "marker" using ARToolkit (http://mtd.fh-hagenberg.at/depot/graphics/artoolkit/). We have found motion detection to be the most robust. It is used to generate a set of azimuths that point to candidate talkers. As each candidate is detected, a video sub-image with correct perspective is produced, Figure 2. Audio information is then used to select one of these sub-images for output.

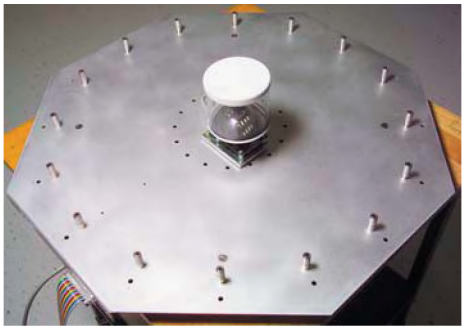


Fig. 1. Panoramic Audio/Video Sensor. A video camera with a panoramic lens is surrounded by a 16-element microphone array.

3. CIRCULAR BEAMFORMING MICROPHONE ARRAY

The circular microphone array is composed of 16 omnidirectional sensors uniformly distributed on a diameter of 57 cm, Figure 1.

The raw microphone signals are pre-amplified and filtered at 4.8 kHz before being digitized by a 16-channel ADC sampling at 11.025 kHz per channel. A 200 MHz TMS320C6201 DSP PCI card (http://www.innovative-dsp.com/) performs the delay-and-sum computations.

The DSP computes 16 "look directions" in each sample interval, Figure 2. One direction is specified for audio capture and is used for analog output via a DAC (steered audio). The remaining 15 look directions which are equally

[1] NRC No. 47171

distributed, are used to estimate speech power in each direction [4].

4. VIDEOCONFERENCING OVERVIEW

While the system software runs on a desktop PC, the microphone array controller runs on the DSP. The DSP receives steering commands from the main application, and returns the directional power estimates. Steered audio is available as an analog output signal. An experimental "Face Server" maintains a database of faces that have been previously captured. As a new face is detected, its position is noted and a matching algorithm selects the best candidate and labels the image [3]. The optional "Marker" module uses ARToolKit to analyze the image to detect unique markers carried by the users. This information can be used to locate the user and to annotate the display. The H.323 module manages communication over the Internet. The system has been demonstrated communicating remotely with other users using NetMeeting. The remote user sees a rectangular image window aimed at the talker, and hears the steered audio.

5. TALKER TRACKING

A key requirement is to be able to automatically select the current talker in real-time. We address this by combining information from candidate sub-images with the audio search results.

When motion is detected in the panoramic image, a "potential talker" is created, and will persist according to some simple rules. A candidate persists for 120 seconds without motion. A non-moving candidate persists for 30 seconds after being the loudest audio direction. The system will track several potential talkers, Figure 2. The vision processing maintains a list of potential talkers, of which one is selected based on audio information. Every 11.6 ms, the DSP provides a frame of audio intensity as a function of heading angle, Figure 2. The final output viewing direction is chosen by selecting the potential talker with the highest audio response. A perspective view is warped in the direction of the selected talker. The selected view and the steered audio are transferred to the H.323 module. In this way sensory fusion is achieved; audio information chooses the general direction and video processing fine-tunes the steering direction.

Figure 2 shows an example with 3 candidate talkers participating in a discussion. The figure illustrates how video sub-images are selected using motion cues, and how the microphone beamformer is used to determine the instantaneous direction to the current talker.

6. DISCUSSION

The image quality of the prototype is limited by two factors. First, with the current optics, the raw image only occupies about 36% of the camera pixels. Sub-images therefore are of low pixel resolution. Second, the average light level determines the camera exposure. Therefore a very bright region on the raw image will produce poor contrast in other regions. We are exploring solutions to these problems.

7. CONCLUSIONS

We have described a panoramic audio/video sensor for videoconferencing applications. Problems related to low audio beam resolution and to reverberation are mitigated by the use of video-based face detection. The system has been demonstrated to dynamically detect and track the active talker amongst a group participating in a round-table discussion. We have experimented with face detection and identification, and marker detection for image annotation. The system includes H.323 functionality and so is compatible with Microsoft NetMeeting.

Fig. 2. Panocam. The raw video image appears in the lower left. The initial corrected panoramic image appears at the top. The real-time audio search response shows an active talker near 12:00 in the left centre frame and three virtual images are shown in the centre frame. The audio and corresponding virtual video (lower image, "6-10") is selected for output.

REFERENCES

[1] Kapralos, B., Jenkin, M. and Milios, E. (2003). Audiovisual localization of multiple speakers in a video teleconferencing setting. Intl. Jour. Imaging Systems and Technology, 13(1): 95-105.

[2] Cutler, R., Rui, Y., Gupta, A., Cadiz, J.J., Tashev, I., He, L-W., Colburn, A., Zhang, Z., Liu, Z. and Silverberg, S. (2002) Distributed meetings: a meeting capture and broadcasting system. Proc. 10th ACM Intl. Conf. On Multmedia, 503-512.

[3] Gorodnichy, D. (2003) Facial Recognition in Video. Proc. of IAPR Int. Conf. on Audio- and Video-Based Biometric Person Authentication (AVBPA'03), LNCS 2688, 505-514.

[4] Flanagan, J.L., Johnston, J.D. Zahn, R. and Elko, G.W. (1985) Computer-steered microphone arrays for sound transduction in large rooms. J. Acoust. Soc. Am. 78(5), 1508-1518

AUDIO FOR COMMITTEE ROOMS

Jeff Bamford
Engineering Harmonics Inc., 29A Leslie Street, Toronto, ON, Canada, M4M 3C3, JBamford@EngineeringHarmonics.com

1. INTRODUCTION

Audio services for committee rooms are paramount to the success of discussions. For governments, the successful distribution of the spoken word is also very important. Committees are used to discuss bills, government policy and party strategies. To properly implement a long-term solution for integrating technology into committee rooms, a prototype committee room was built. This unique opportunity allowed many facets of the room to be tested and configured. This was especially important for determining the accommodation requirements for multimedia devices in the room, such as television cameras and large-screen projection.

Of particular interest was testing a digital delegate system. Existing infrastructure for committee rooms uses technology from the late 1970s. The prototype would allow testing of current technology and determining if it would be appropriate. The prototype room would also allow for testing sound reinforcement in a broadcast environment. Traditionally, larger committee rooms require sound reinforcement. However, as many committees are televised, it is important to ensure that any sound reinforcement in the room does not have a negative affect on televising.

2. GENERAL LAYOUT

Figure 1 - Room Layout

Figure 1 shows the general layout of the room. A Simultaneous Interpretation Booth, a Multimedia Control Room and a Multimedia Equipment Room are located at the front of the room. A pair of projection screens is located near the front of the room on the left and right sides. Six cameras are used to capture the proceedings. A large table is in the centre of the room; public seating is at the rear. Each position in the room is equipped with a listening station; this is required to provide simultaneous interpretation. At the large table, the listening station is integrated with a delegate station. The delegate station includes a microphone and small loudspeaker for sound reinforcement.

3. DELEGATE SYSTEM

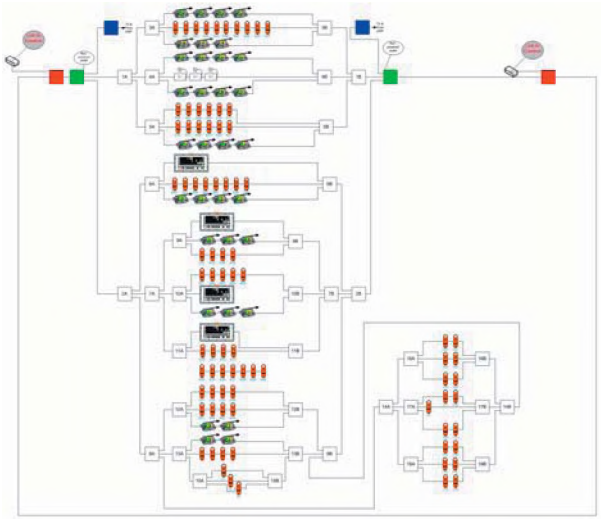


Figure 2 - Digital Delegate System

Figure 2 shows the block diagram of the delegate system. Of interest in the design of this system is the redundant loop. This is one of the main requirements of this system – it must be able to withstand failures and continue operating. In this case, there are redundant controllers, power supplies, interface boxes (for access to external audio systems) and cabling paths. The failure of any single device will have no affect on the overall system – it must continue to operate. This design philosophy allows proceedings to continue while hardware is repaired or replaced.

Access to other audio systems was provided via an analog interface device. This allowed any of the system channels to be delivered to external systems. This was used to

provide audio for conferencing, sound reinforcement (both ceiling loudspeaker and infrared) and television. Inputs to the system were also used to provide audio from conferencing, presentation materials and ambient room sound for the interpreters. These signals were delivered via the DSP system.

4. DSP SYSTEM

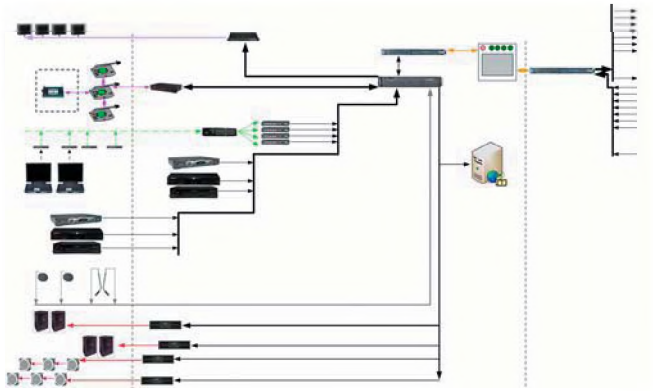


Figure 3 - Audio Block Diagram

The DSP system provided two main functions: routing and processing; as can be seen in **Figure 3**. As can be seen in the diagram, the major purpose of this system is the routing of signals within the room and to the central control area. Of interest is the "stove" like device near the upper right corner of the diagram. This is used to "lock-out" signals leaving the room; in certain situations, it is imperative that discussions be held in confidence. In order to ensure that electronic eavesdropping is not possible, all signals leaving the room are severed. In the prototype, this was done by de-powering Ethernet hubs which were carrying Cobranet traffic. This removed the committee room from the central control area. Other systems were also locked-out using similar techniques.

One interesting result from the prototype was the discovery of a loss of synchronization between the video and audio from the room, with the video being delayed to the audio. It was determined that there was a six-frame delay of the video. This was perceptible to the trained eye and was particularly noticeable when viewing a video conference. It was thought that the delay was caused by cascading of several video format conversions. In the case of a video conference, the far-end video signal passed through more processing devices than the committee room camera feeds, thus increasing the overall delay.

Each signal conversion adds one or two frames of delay (33 or 66 ms) to the video signal. By comparison, the audio delay is very short; latency for audio conversions is approximately 10 ms. In a traditional broadcast plant environment, the audio is usually bundled with the video in an SDI chain – which will aid in synchronizing the two signals. In the prototype, audio was handled separately as there are numerous channels. Also, the video signals from the cameras pass through the committee room to a central control area, where all camera switching is done. Operators in this location do not control all aspects of the room audio, thus it is not possible to include the audio feeds with the video feeds.

It is a relatively simple matter to introduce additional delay through the DSP to resynchronize the audio and video.

5. RESULTS

The prototype systems were demonstrated to the various stakeholders. Each group was brought in separately and a run through of the various systems was done. Once all of the groups were brought in, their input was used to change aspects of the system.

In general, all groups were pleased with the room and were happy with the results. One area of particular concern was with ambient sound in the multimedia control room. Two operators work in this room, one of whom controls the microphones in the room. Traditionally, they work inside a committee room. In the prototype, they worked behind glass and required ambient sound. The initial results indicated that there was not enough a sense of direction from the playback system. Thus, an additional microphone pair was tested on the front wall of the room. This augmented the existing pair in the ceiling above the middle of the table. The results were then demonstrated and a difference was noted. The pair over the table offered better intelligibility at the expense of a sense of direction. The pair on the front wall offered better direction, but did not achieve the same level of fidelity. It is likely that both will be installed in future rooms and the operator will be able to choose which pair to use for monitoring.

6. CONCLUSION

In conclusion, the prototype offered a unique opportunity to test multimedia systems and integration techniques. A consensus was achieved and the task of designing systems will be much easier now that a baseline has been established for multimedia systems and integration techniques.

Figure 4 - The Prototype, looking toward the front of the room

HUMAN-CENTERED DESIGN OF ACOUSTIC AND VIBRATORY COMPONENTS FOR MULTIMODAL DISPLAY SYSTEMS

William L. Martens
Faculty of Music, McGill University, 555 Sherbrooke Street W., Montreal, CQ, Canada H3A 1E3
wlm@music.mcgill.ca

1. INTRODUCTION

It is well established that human tolerance for latency between audio and video reproduction for teleconferencing is an important design consideration. Deployment of the multimodal display system described here required the coordination of signals for three sensory modalities, auditory, visual, and vibratory. Though a great deal of research has been done investigating audio/visual interaction, relatively little is known about interaction between reproduced acoustic and vibratory components of remotely captured events. Therefore, a study was undertaken to determine the intermodal delay required for brief acoustic and structural vibrations to be perceived as synchronous. The structure-borne component of recorded impact event was presented via a motion platform on which the observer was seated. The air-borne component of the event was presented via a multichannel loudspeaker array, with simulated indirect sound arriving from all around the observer. By varying the relative level and intermodal delay of the vibratory (structure-borne) components and the acoustic (air-borne) components, conditions allowing successful time order judgment (TOJ) were estimated using a two-alternative, forced-choice (2AFC) tracking procedure. Then, in order to avoid sequential response biases in the tracking procedure, the method of constant stimuli was used to determine the range of intermodal delay values associated with observers' reports of perceived simultaneity as a function of the relative level of the vibratory stimulus. Since the results of this investigation provide a basis for deployment of multimodal display technology that is generated through perceptual experimentation with relative levels and intermodal delay values, they are said to enable human centered design (Martens, 1999). In contrast to displays developed using conventional engineering approaches, then, these results may lead to the creation of more satisfying and convincing virtual environments for applications such as teleconferencing and realistic reproduction of remote musical performances.

1.1 Multimodal display technology

Multimodal display technology that is used to reproduce a remotely captured and/or recorded event is most effective when the transmitted and reproduced stimulation is synchronized with minimal intermodal delay (Barfield, et al., 1995). Such coordinated display of visual, auditory, tactile, and kinesthetic information can produce for an observer a strong sense of presence in a reproduced environment when asynchrony is below threshold for human detection (Miner & Caudell, 1998), but even when asynchrony is detectable, there is useful variation in human experience within the tolerable range of asynchrony (Martens & Woszczyk, 2004). Other recent work (Woszczyk & Martens, 2004) has focused upon asynchrony between acoustic and vibratory display components in an attempt to quantify their multimodal integration in isolation from other display modalities. First hand experience with such bimodal display of these events suggested that physical synchrony between display components was not necessarily required to produce a subjective experience of simultaneity. The novel aspect of the research reported here is that it examined the impact of the relative level of vibration upon the perceived realism and naturalness of remotely reproduced impact events.

2. METHOD

This section describes both the stimulus generation methods and the experimental methods used in the experimental tests. First, an overview of the employed multimodal display system is presented, along with a description of the selected experimental stimuli.

2.1 Acoustic component display

The acoustic component was presented via a spherical loudspeaker array consisting of 5 low-frequency drivers (ranging from 25 to 400 Hz) and 32 high-frequency drivers (ranging from 300 to well over 20,000 Hz). The low-frequency drivers were "Mini-Mammoth" subwoofers manufactured by the Quebec-based company D-BOX Technology, and these were placed at standard locations for the 5 main speakers in surround sound reproduction (the speaker angles in degrees relative to the median plane were –110, -30, 0, 30, and 110). The high-frequency drivers were dipole radiating, full range transducers featuring the "Planar Focus Technology" of Level 9 Sound Designs, Inc. of British Columbia, and these 32 loudspeaker panels were placed pairwise in 16 locations lying on the surface of an imaginary sphere of 2-meter radius.

The stimuli were selected as the most representative from a number of transient sound sources that were recorded in a rectangular shaped music hall (Redpath Concert Hall) at McGill University using a Schoeps CCM 21H wide-cardioid microphone pointing at the stage. The most satisfying recording was that made by dropping a stack of 3

telephone books from above the stage onto the floor, at a distance of 2 meters from the microphone. For the current study, the level of the acoustic stimulus was held constant at 82 dB(A).

2.2 Vibratory component display

Only vibration along the vertical axis was presented in this study, although the employed vibration transducer was capable of generating multidimensional vibration stimulation, providing users with motion having three Degrees of Freedom (3DOF) in a home theater setting (Paillard, et al., 2002). The motion was generated by the Odyssée™ system, a commercially available motion platform manufactured by D-BOX Technology. The Odyssée™ system uses four coordinated actuators to displace the observer linearly upwards of downwards, with a very quick response and with considerable force (the feedback-corrected linear system frequency response is flat to 50 Hz). The vibratory stimulus was generated by gating to a 30 ms duration the initial portion of the audio signal (which was a highly reverberant recording a the impact of a phone book on a wooden stage), and then applying a lowpass filter with a cutoff frequency of 50. For the current study, the maximum vertical acceleration RMS value was 1.3 m/sec^2, measured at the observer's foot position (using a B&K Type 4500 accelerometer). The vibration level was attenuated from this maximum RMS value in 7 steps, each of -3 dB, to cover a vibration range of 18 dB. The vibratory stimulus was also delayed relative to the acoustic stimulus in 7 steps of 10 ms, reaching a maximum of 40 ms, but also leading the acoustic stimulus in two cases. Observers made time order judgments, and also reported when the two components seemed to occur simultaneously.

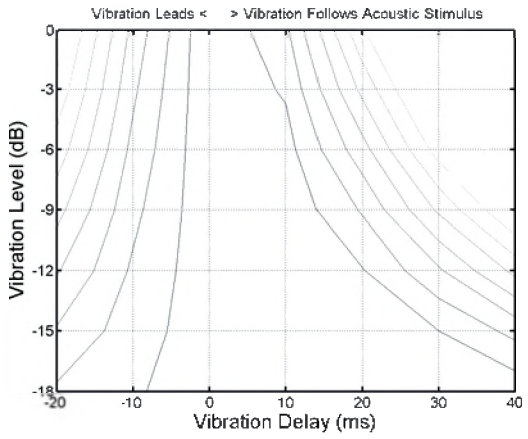


Fig. 1. Contour plot showing acceptability of a multimodal reproduction as a function of two parameters of the vibration display component, delay and level (see text).

3. RESULTS AND CONCLUSIONS

The results of this investigation can be summarized using the single contour plot acceptability of the multimodal reproduction shown in Figure 1. The darkest contours in the plot surround the region of highest acceptability on the plane defined by the 7 by 7 factorial combination of presented vibration delay and level values. At the lowest vibration level presented, the range of most acceptable vibration delay values extends from –10 ms (vibration leading the acoustic stimulus) to the maximum tested delay of 40 ms. As the vibration level was increased, the acceptable range of delay decreased, so that only true physical synchronization produced a reliable response of perceived simultaneity. When criteria other than strict simultaneity were employed, such as "naturalness," the range of acceptable delay values grew to include longer vibration delay values, but for the most natural impression, the vibration could not lead the acoustic stimulus. Also worth noting is the variation in subjective intensity of the impact event. For example, at the lowest vibration level presented, one observer reported that the impact event seemed more "powerful" when the vibration followed the acoustic stimulus by 20 ms, even though this was combination was not associated with the greatest sense of simultaneity.

REFERENCES

Barfield, W., Hendrix, C., Bjorneseth, O., Kaczmarek, K. A., and Lotens, W. (1995) "Comparison of Human Sensory Capabilities with Technical Specifications of Virtual Environment Equipment," Presence: Teleoperators and Virtual Environments, 4(4).

Martens, W. L., & Woszczyk, W. (2004) "Perceived synchrony in a bimodal display: Optimum intermodal delay values for coordinated auditory and haptic reproduction," Proceedings of the 10th International Conference on Auditory Display, Sydney, Australia, July.

Martens, W. L., (1999) "Human-centered design of spatial media display systems", Proceedings of the 1st International Workshop on Spatial Media, University of Aizu, Aizu-Wakamatsu, Japan.

Miner, N., & Caudell, T. (1998) "Computational Requirements and Synchronization Issues for Virtual Acoustic Displays," Presence: Teleoperators and Virtual Environments, 7 (4), 396-409.

Woszczyk, W., & Martens, W. L., (2004) "Intermodal delay required for perceived synchrony between acoustic and structural vibratory events," To appear in: Proceedings of the 11th International Congress on Sound and Vibration, St. Petersburg, Russia, July.

Paillard, B., Roy, P., Vittecoq, P., & Panneton, R. (2002) "Odyssée: A new kinetic actuator for use in the home entertainment environment." Proceedings of DSPFest 2000, Texas Instruments, Houston, Texas, and July, 2000.

ACKNOWLEDGEMENTS

The author would like to thank Dr. Bruno Paillard of D-BOX Technologies for many helpful discussions. This research was supported by Valorisation-Recherche Québec.

EXTRACTING SEMANTICALLY-COHERENT KEYPHRASES FROM SPEECH

Diana Inkpen[1], and Alain Désilets[2]
[1]School of Information Technology and Eng., University of Ottawa, 800 King Edward St., Ottawa, ON, Canada, K1N6H5
diana@site.uottawa.ca
[2]Institute for Information Technology, National Research Council, 1200 Montreal Rd., Ottawa, ON, Canada, K1AOR6
Alain.Desilets@nrc-cnrc.gc.ca

1. INTRODUCTION

Browsing through large volumes of spoken audio is known to be a challenging task for end users. One way to facilitate this task is to provide keyphrases extracted from the audio, thus allowing users to quickly get the gist of the audio document or sections of it.

Previous methods for extracting keyphrases from spoken audio have used text-based summarisation techniques on automatic speech transcription. The method of Désilets et.al. (2000) was found to produce accurate keyphrases for transcriptions with Word Error Rates (WER) of the order of 25%, but performance was less than ideal for transcripts with WERs of the order of 60%. With such transcripts, a large proportion of the extracted keyphrases included serious transcription errors.

In this paper, we extend those previous methods by taking advantage of the fact that the mistranscribed keyphrases tend to have a low semantic coherence with the correctly transcribed ones. For each pair of extracted keywords, we determine their semantic coherence by computing a Pointwise Mutual Information (PMI) score based on a very large web corpus. We then use those semantic coherence scores to identify semantic outliers and filter them from the set of extracted keyphrases. The effect of the method on the accuracy of the extracted keyphrases is evaluated. We also use the same approach to filter semantic outliers in the speech on transcripts, before extracting keyphrases from it.

1.1 Data

We used a subset of the ABC and PRI stories of the TDT2 English Audio data that had correct transcripts generated by humans. We conducted experiments with two types of automatically-generated speech transcripts. The first ones were generated by the NIST/BBN time-adaptive speech recogniser and have a moderate WER (27.6%), which is representative of what can be obtained with a state of the art SR system tuned for the Broadcast News domain. See an example of a transcribed paragraph in Fig.1. The second set of transcripts was obtained using the Dragon NaturallySpeaking speaker dependant recogniser. Their WER (62.3%) was much higher because the voice model was not trained for speaker independent broadcast quality audio, in order to approximate the type of high WER seen in more casual less-than-broadcast quality audio.

1.2 Extracting keyphrases

Our approach to extracting keyphrases from spoken audio is based on the Extractor system developed for text by Turney (2000). Extractor uses a supervised learning approach to maximise overlap between machine extracted and human extracted keyphrases and it was estimated to be approximately 80% accurate. A keyphrase consist of one, two, or three keywords. See Fig.1 for some examples of keyphrases, extracted from the manual transcripts and from the BBN transcripts.

2. METHOD

Our algorithm detects the semantic outliers to be filtered out from keyphrases. It declares as outliers all the keywords with low similarity to the other keywords.

For a set of keyphrases containing the keywords (w_1, w_2, ... w_n), the algorithm has the following steps:
1. Compute semantic similarity scores $S(w_i, w_j)$ between all the pairs w_i, w_j, for all $1 <= i, j <= n$, $i \neq j$, using PMI.
2. For each keyword w_i, compute its semantic coherence score $SC(w_i)$ by summing up all $S(w_i, w_j)$, $1 <= j <= n$, $i \neq j$.
3. Compute the average score of all keywords.
4. Declare as outliers the keywords with score $SC(w_i) < K\%$ of the average score. The value of K in the threshold is chosen empirically, as shown in Section 3.

The **semantic similarity score between two words** w_1 and w_2 is their pointwise mutual information score, defined as the probability of seeing the two words together over the probability of each word separately. $PMI(w_1, w_2) = \log P(w_1, w_2) / (P(w_1) \cdot P(w_2)) = \log C(w_1, w_2) \cdot N / (C(w_1) \cdot C(w_2))$, where $C(w_1, w_2)$, $C(w_1)$, $C(w_2)$ are frequency counts, and N is the total number of words in the corpus. The scores were computed using the Waterloo Multitext system with a very large corpus of Web data (Clarke and Terra 2003).

A variant of this algorithm detects each keyword with low similarity to its closest semantic neighbour, by using the maximum score in Step2, instead of the sum. The threshold is chosen differently, as a function of the minimum $SC(w_i)$.

Manual transcript: Time now for our geography quiz today. We're traveling down the Volga river to a city that, like many Russian cities, has had several names. But this one stands out as the scene of an epic battle in world war two in which the Nazis were annihilated.
Keyphrases:
- Russian cities --> (22.752942)
- city --> (22.752942)
- Volga river --> (22.752942)
- Nazis --> (11.376471)
- war --> (11.376471)
- epic battle --> (11.376471)
- scene --> (11.376471)

NIST/BBN transcript: time now for a geography was they were traveling down river to a city that like many russian cities has had several names but this one stanza is the scene of ethnic and national and world war two in which the nazis were nine elated
Keyphrases:
- russian cities --> (22.752942)
- city --> (22.752942)
- river --> (22.752942)
- elated --> (11.376471)
- nazis --> (11.376471)
- war --> (11.376471)
- scene --> (11.376471)
- stanza --> (11.376471)

Detected outlier keywords: stanza, elated
Lost keywords: --none--

Fig.1. Fragment of a manual transcript and the extracted keyphrases; the BBN transcript, the extracted keyphrases, and the detected outliers.

3. RESULTS

Table 1 shows the results of our outlier detection algorithm on the keyphrases extracted from the BBN transcripts and from the Dragon Naturally Speaking transcripts. The second column shows the WER in the speech transcripts, measured with the standard NIST tool as a function of the number of insertions, deletions, and substitutions. The third column shows the word error rate in the keyphrases extracted by Extractor, and the last column shows the error rate after the outliers were eliminated. The word error rate in the keyphrases (kWER) is measured as the number of words that are in the keyphrases but not in the manual speech transcript. Table 1 shows that the number of wrong keywords caused by recognition errors reduces by almost half when the outliers are eliminated (for K=80%).

The variation of the error rate in keyphrases with K is shown in Fig.2. The higher the threshold, the more outliers are eliminated, but some good keywords can also be lost. Fig 2 shows the percent of lost keywords, computed as the percent of the keywords that are were wrongly declared as outliers (they are considered good keywords because they are in the keyphrases extracted from the manual transcripts). The variant of the algorithm that uses maximum scores in Step 2 produces better reduction in kWER, but higher loss of good keywords. Its results are not shown here because of space limitations.

Table 1. Word error rate in the transcripts, in the initial keyphrases, and in the filtered keyphrases (plus % lost keywords, for K=80%).

Tran-scripts	WER transcripts	kWER initially	Filtered keyphrases % Lost k.	kWER
BBN	27.6%	10.6%	14.5%	**5.4%**
Dragon	62.3%	43.3%	4.3%	**27.6%**

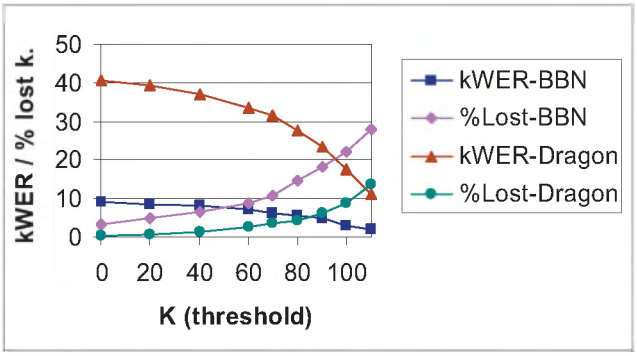


Fig.2. The variation of the kWER and percent of lost keywords for the two sets of data, in function of the threshold K% of the average similarity of a set of keyphrases.

4. CONCLUSION

We presented a method for filtering semantic outliers from keyphrases that summarize speech. Future work includes experimenting with other methods for computing semantic outliers (Jarmasz and Barrière 2004), and building small domain models from reliable keywords in order to detect the outliers relative to them (Turney 2003). We also plan to run the outlier detection algorithm directly on the speech transcripts. In this case the input to the algorithm is all the content words in the transcript.

REFERENCES

Clarke, C. and Terra, E. (2003). Passage retrieval vs. document retrieval for factoid question answering. ACM SIGIR'03, 327-328.
Désilets, A. and de Brujin, B. and Martin. J. (2001). Extracting keyphrases from spoken audio documents. SIGIR Workshop on Information Retrieval Techniques for Speech Applications, 36-50.
Jarmasz, M. and Barrière, C. (2004). Keyphrase extraction: enhancing lists. Proceedings of CLINE'04.
Turney, P. D. (2003). Coherent keyphrase extraction via Web mining. Proceedings of IJCAI'03, 434-439.
Turney, P.D. (2000). Learning algorithms for keyphrase extraction, *Information Retrieval*, 2 (4), 303-336.

ACKNOWLEDGEMENTS

We wish to thank Peter Turney and Gerald Penn for their useful feedback and discussions. We thank Egidio Terra, and Charlie Clarke for allowing us to use the Multitext System, the NRC copy.

MAINTAINING SPEECH INTELLIGIBILITY IN COMMUNICATION HEADSETS EQUIPPED WITH ACTIVE NOISE CONTROL

A. J. Brammer[1,2], D. R. Peterson[1], M. G. Cherniack[1], S. Gullapalli[1] and R. B. Crabtree[3]
[1]Ergonomic Technology Center, University of Connecticut Health Center, Farmington, Connecticut, U.S.A., 06030-2017
[2]Institute for Microstructural Sciences, National Research Council, Ottawa, Ontario, Canada, K1A 0R6
[3]Defence Research and Development Canada – Toronto, Toronto, Ontario, Canada, M3M 3B9

1. INTRODUCTION

The primary requirement for a communication headset is to maintain speech intelligibility under all conditions of use. For headsets equipped with active noise reduction (ANR), the performance of the control system may influence the communication signal reaching the ear. Conversely, the communication signal may perturb the operation of the ANR system. The potential for interaction between the communication and control signals would appear to depend primarily on the control structure.

The intelligibility of speech reproduced at the ear of persons wearing a headset equipped with ANR has been reported in several studies. The influence of the control structure on the speech intelligibility, however, has not been frequently investigated [1,2].

In this paper, the performance of two circumaural headsets is compared, one employing feedback control with a fixed filter and analog signal processing, an approach commonly used in commercial devices, and one employing adaptive digital feed-forward control. Both devices attempt to control low-frequency noise, as the passive attenuation of the earmuff substantially reduces noise at frequencies above 500 Hz. The headset with the feedback control system was chosen to match, as close as possible, the ANR of the headset with the feed-forward control system.

2. METHODS

2.1 Sound Fields and Measurements

The passive, and total, noise reductions of the headsets were measured when they were worn by a human subject, or by a manikin (Bruel & Kjaer head and torso simulator, HATS). In the former case, a miniature microphone recorded the sound pressure under the earmuff when the control system was, and was not, operating. A separate measurement of passive noise reduction was performed with a miniature microphone positioned in the concha (i.e., headset doffed and donned). In the latter case, the built-in microphone within the external ear simulator of HATS was used to record the sound pressure. The measurements were conducted in a reverberation room using band-limited white noise.

2.2 Speech Transmission Index

The influence of ANR on speech intelligibility was determined using the Speech Transmission Index (STI), which is a figure of merit for a communication link that varies from zero (no intelligibility) to unity (ideal intelligibility). The A-weighted sound level at the ear produced by the STI signal fed to the earphone was set to 70 dB. A noise spectrum shaped to approximate that of the long-term average of speech was established at the subject position within the reverberation room, and the sound level adjusted to produce a range of STI values. In each case, a miniature microphone within the volume enclosed by the earmuff was used to record the combination of the STI test signal and the confounding noise both with, and without, active control. For all measurements, the A-weighted sound level of the noise was adjusted to be the same for both headsets when the ANR was not operating.

3. RESULTS

3.1 Passive and Total Noise Reduction

The differences between the passive, and between the total, noise reductions of the two headsets are shown as a function of frequency in Fig. 1. The results are plotted for the headset with the feed-forward control system – headset with the feedback control system. The passive attenuation of the headset with the feed-forward control system can be seen to exceed that of the other headset at frequencies from 200 to 1500 Hz, and at frequencies above 3.5 kHz, while the opposite was observed at other frequencies. The difference in the total noise reduction shows the same pattern at mid and high frequencies, but displays a different pattern at frequencies below 1 kHz, reflecting the contributions of the active control systems. The similarity between the two curves in Fig. 1 at frequencies below 200 Hz implies that the ANR of the two control systems is similar at these frequencies. At frequencies from 200 to 400 Hz, the ANR of the feedback system can be seen to exceed that of the feed-forward system. The extent to which the control systems fulfilled the selection criterion may thus be deduced from the similarities and differences between the two curves in Fig 1 at frequencies below 1 kHz.

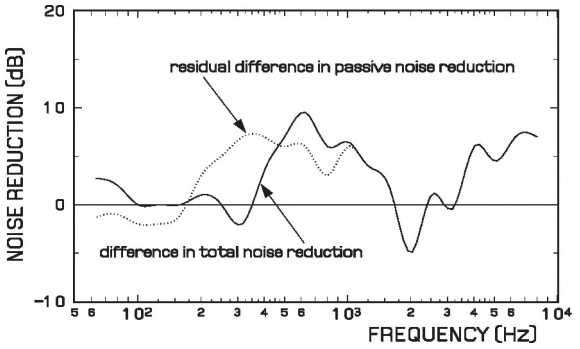


Figure 1: Difference between passive, and total, noise reduction

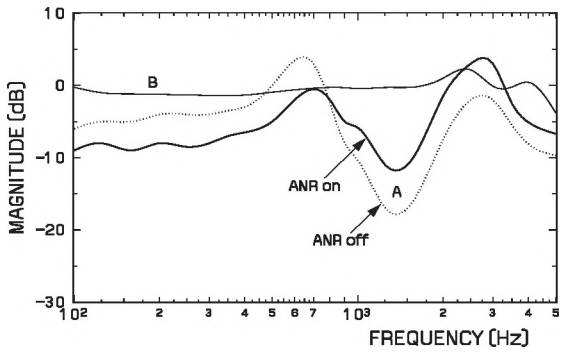


Figure 2: Frequency response of sound reproduction sub-systems

3.2 Speech Reproduction

The frequency responses of the speech reproduction sub-systems are shown in Fig. 2. The results were obtained when the headsets were mounted on HATS with cushions sealed. It can be seen from Fig. 2 that the speech reproduction sub-system of the headset with the feed-forward control system possessed little dependence on frequency from 100 to 4000 Hz (curve B). In contrast, the frequency response of the speech reproduction sub-system of the headset with the feedback control system displayed large frequency-dependent variations in amplitude (curves A, dashed line). Similar, but not identical, variations were observed when this control system was operating (curves B, continuous line). There was no change in sound reproduction by the other headset when its control system was operating. It should be noted that the earphone selected for a feedback control system is usually a compromise between the need to maintain stability of the feedback loop and for communication fidelity: frequency-dependent amplification of the communication signal may also be employed.

3.3 Speech Transmission Index

The STIs of the two headsets for various speech-to-noise (S/N) ratios at the ear are shown in Table 1. It can be seen from the Table that the different S/N ratios produced a range of STI values, as expected, with the largest STI values being obtained with no interfering noise, and the smallest values with the most intense noise (S/N = 2.5 dB). While the STI recorded in noise by the feed-forward control system was greater than that recorded by the feedback system (viz: 0.69 versus 0.55, and 0.85 versus 0.73), it can be seen from the Table that the increase in STI with active control was much greater for the headset with feedback control than for the other headset.

4. DISCUSSION

In a feedback ANR system, the microphone providing the input to the controller is positioned at the ear, under the earmuff, and so senses both the residual noise and the reproduced speech sounds. The controller will attempt to null this "error" signal, i.e., cancel both noise *and* speech. While there are strategies to mitigate the cancellation of speech, it will always occur with this control structure. In contrast, the error signal does not provide the input to a feed-forward controller – the input is taken from a reference microphone outside the earmuff, and so does not contain speech unless there is a substantial air leak in the seal between the earmuff and the head. Thus, the control signal does not perturb the speech reproduced by a circumaural headset with a feed-forward control structure, resulting in improved STI in noise. The disproportionate increase in the STI of the headset with the feedback control system when the ANR system was operating suggests that there may be a factor other than the S/N ratio to consider, as this would be approximately the same for both headsets. Reference to Fig. 2 shows that the frequency response of sound reproduction changed when this controller was operating, presumably as a consequence of the speech component of the error signal, to the benefit of speech intelligibility.

5. REFERENCES

1. A. J. Brammer, R. B. Crabtree, D. R. Peterson and M. G. Cherniack, "Intelligibility in active communication headsets," *Proc ICBEN 2003*, edited by R. de Jong, T. Houtgast, E. A. M. Franssen and W. F. Hofman, pp. 58-64, 2003.

2. A. J. Brammer, R. B. Crabtree, D. R. Peterson, M. G. Cherniack, and S. Gullapalli, "Active headsets: Influence of control structure on communication signals and noise reduction," (to appear).

Table 1: Mean STI for various speech-to-noise (S/N) ratios, for two control structures (from Ref. 2)

Control Structure	Speech Transmission Index		
	S/N=2.5 dB	S/N=8.5 dB	No Noise
fixed-filter, feedback, ANR off	0.42	0.6	0.98
fixed-filter, feedback, ANR on	0.55	0.73	0.97
adaptive filter, feed-forward, ANR off	0.67	0.83	0.99
adaptive filter, feed-forward, ANR on	0.69	0.85	0.98

SPEAKER RECOGNITION IN REVERBERANT ENVIRONMENTS

Joseph Gammal, Rafik Goubran
Dept. of Systems and Computer Engineering, Carleton University, Colonel By Dr., Ottawa, Canada, K1S 5B6
{jgammal,goubran}@sce.carleton.ca

1. INTRODUCTION

Closed set speaker recognition is the task of, given a speech utterance, correctly selecting the identity of an unknown speaker from a limited set of speakers. Different methods are used for this purpose and each can employ a variety of feature vectors extracted from the speech. The objective of this paper is to objectively compare these methods and parameterizations when there is a mismatch between training and test conditions caused by the existence of reverberation.

Three classes of speaker recognition algorithms were used: Gaussian mixture models (GMM), covariance models and AR-vector models. Each of the last two classes employs two different speaker recognition measures.

Two different feature vectors were extracted from the speech, these are LPC cepstral (LPCC) and Mel-warped cepstral (MFCC) vectors [1]. The effect of the addition of delta-cepstral vectors was investigated.

2. METHOD

2.1 Reverberation Models

Reverberation was simulated using the image-method [2]. Impulse responses were generated for two rooms. The first had dimensions 3.6x4.2x3m with reflection coefficients for the walls as 0.9 and floors as 0.7, and the second had dimensions 3x6x2.5m with reflection coefficients of 0.93 and 0.8. Microphone to speaker separation in the first room was 0.75m and in the second was 0.54m. The first configuration is characterized as minor reverb., the second as major reverb. The impulse responses were generated using a sampling frequency of 8Khz. The lengths of the impulse responses were 1 and 2 seconds respectively.

2.2 Database and Signal Processing

The KING speaker recognition database was used. The speech files in the 51-speaker database were band limited to the 300-3400Hz telephone band and μ-law coded. A sampling frequency of 8Khz was used. Reverberated versions of the database were produced by convolving the original database with either of the impulse responses before further processing. 20ms frames were extracted every 10ms after silence removal. Either MFCC or LPCC vectors were extracted. For MFCC vectors the filter bank contained 19 filters. When delta-cepstral (Δ) vectors were used, they were produced using a 5-frame first order orthogonal polynomial fit. Cepstral mean subtraction (CMS) was applied to all vectors.

2.3 Recognition Methods

The GMM was produced using the method outlined in [3] and using diagonal covariance matrices. Both the sphericity (SM) and divergence shape (DS) [4] were used for the covariance-based methods. The DS and SM, which are distances between a claimant model and test utterance are calculated as follows [4]:

$$DS_{1,2} = DS(C_1, C_2) = \frac{1}{2} trace[(C_1 - C_2)(C_2^{-1} - C_1^{-1})] \quad (1)$$

$$SM_{1,2} = SM(C_1, C_2) = \frac{1}{2} trace(C_1 C_2^{-1}) trace(C_2 C_1^{-1}) \quad (2)$$

Here C_1 is the covariance matrix of the training speech and C_2 is that of the test utterance. For both the second-order AR-vector methods the training method specified in [6] was used to train the models. For a set of p-dimensional training vectors $\{\vec{x}_t\}_{1 \le t \le N}$ with mean $\vec{\mu}$, the 2nd order AR-vector model whose solution as follows [6]:

$$\sum_{i=0}^{2} A_i(\vec{x}_{t-i} - \vec{\mu}) = \vec{e}_t \text{ with } A_0 = I_p \quad (3)$$

The lagged covariance matrices are defined as follows [6]:

$$\chi_k = \frac{1}{N} \sum_{t=k+1}^{N} (\vec{x}_t - \vec{\mu})(\vec{x}_{t-k} - \vec{\mu})^T \text{ with } k = 0..2 \quad (4)$$

The Toeplitz matrix X is defined as follows:

$$X = \begin{bmatrix} \chi_0 & \chi_1^T & \chi_2^T \\ \chi_1 & \chi_0 & \chi_1^T \\ \chi_2 & \chi_1 & \chi_0 \end{bmatrix} \quad (5)$$

With A= [A_0 A_1 A_2] let $E_X^{(A)} = AXA^T$. A set of vectors from the test utterance $\{\vec{y}_t\}_{1\le t\le M}$ has model B. Let $E_Y^{(B)} = BYB^T$, $E_X^{(B)} = BXB^T$ and $E_Y^{(A)} = AYA^T$ [6]. The distance measure referred to in this paper as AR1 is as follows [5]:

$$AR1_{1,2} = \frac{1}{2}\log[trace(\frac{E_Y^{(A)}}{E_Y^{(B)}})\ x\ trace(\frac{E_X^{(B)}}{E_X^{(A)}})] \qquad (6)$$

AR2 requires that the vectors be sorted randomly prior to training and testing. The distance measure is as follows [6]:

$$AR2_{1,2} = \log\left[\frac{\frac{1}{p}trace(E_X^{(A)^{-\frac{1}{2}}}E_Y^{(A)}E_X^{(A)^{-\frac{1}{2}}})}{[\det(E_X^{(A)^{-\frac{1}{2}}}E_Y^{(A)}E_X^{(A)^{-\frac{1}{2}}})]^{\frac{1}{p}}}\right] \qquad (7)$$

3. RESULTS

Table 1. Recognition accuracy.

Method & feature	Recognition accuracy (%)		
	No reverb	Minor reverb	Major reverb
GMM64 LPCC	93.4	79.0	70.6
GMM64 LPCC+Δ	96.3	72.0	62.0
GMM64 MFCC	93.1	77.5	67.1
GMM64 MFCC+Δ	94.2	76.4	66.3
AR1 LPCC	91.9	40.3	30.8
AR1 LPCC+Δ	92.2	50.1	38.9
AR1 MFCC	90.5	33.7	29.1
AR1 MFCC+Δ	85.9	44.4	40.9
AR2 LPCC	95.1	80.7	75.8
AR2 LPCC+Δ	96.3	68.6	63.4
AR2 MFCC	91.4	53.3	46.4
AR2 MFCC+Δ	94.2	35.2	29.4
SM LPCC	94.8	73.2	67.4
SM LPCC+Δ	94.2	57.6	52.2
SM MFCC	89.0	58.5	50.7
SM MFCC+Δ	93.9	43.5	40.9
DS LPCC	94.2	76.1	69.5
DS LPCC+Δ	94.2	44.4	33.4
DS MFCC	91.9	70.3	67.1
DS MFCC+Δ	93.9	48.4	41.8

Each recognition method was trained using sessions 1- 3, and tested using 30s segments from sessions 4-10.

4. DISCUSSION

The results reveal that performance for all methods degrades under reverberation. Delta-cepstral coefficients degrade recognition performance considerably under conditions of reverberation for all methods except AR1, where in the case of AR1 they enhance recognition performance. When delta-cepstral coefficients are used the GMM is the most robust to reverberation. When delta-cepstral coefficients are not used AR2 using LPCC vectors is the most robust to reverberation followed by the GMM. LPCC and MFCC vectors are affected differently by reverberation. When delta-cepstral features are not used, LPCC vectors are more robust to reverberation than MFCC vectors for all methods except the GMM.

It was found that when training was performed with reverberant speech prior to testing with minor reverb or major reverb that recognition accuracy improved for all methods. This was found regardless of whether the impulse response used during training was the same as that used during testing. It was also found that training with reverberant speech and testing with non-reverberant speech gave better results than when training was performed with non-reverberant speech and testing was performed with reverberant speech.

REFERENCES

[1] S. Davis, P. Mermelstein (1980) "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences", IEEE Transactions on Signal Processing, vol. 28, pp. 357-366

[2] J. B. Allen, D. A. Berkley, (1979) "Image method for efficiently simulating small-room acoustics", Journal of the Acoustical Society of America, vol. 65. pp. 943-950

[3] D. Reynolds, R. Rose, (1995) "Robust text-independent speaker identification using Gaussian mixture speaker models", IEEE Transactions on Speech and Audio Processing, vol. 3, Issue: 1, pp. 72-83

[4] R. Zilca, (2001): "Using second order statistics for text independent speaker verification", in ODYSSEY-2001, pp. 45-49.

[5] C. de Lima, D. da Silva, A. Alcaim, J. Apolinario Jr., (2002) "AR-Vector using CMS for robust text independent speaker verification", in Proc. 14th International Conference on DSP, vol 2, pp 1073-1076

[6] I. Magrin-Chagnolleau, J. Wilke, F. Bimbot (1996), "A further investigation on AR-vector models for text-independent speaker identification", in Proc. ICASSP'96 pp. 401-404

A RECURSIVE LEAST-SQUARES EXTENSION OF THE NATURAL GRADIENT ALGORITHM FOR BLIND SIGNAL SEPARATION OF AUDIO MIXTURES

M. Elsabrouty, T. Aboulnasr and M. Bouchard
School of Information Technology and Engineering, University of Ottawa, 800 King Edward,
Ottawa, Ontario, K1N 6N5 {melsabro, aboulnas, bouchard}@site.uottawa.ca

1. INTRODUCTION

Blind Signal Separation (BSS) refers to the operation of recovering a set of sources that are as independent as possible from another set of observed linear or non-linear mixtures. The term blind indicates that this operation is done without knowing the mixing coefficients. Although this problem is more difficult than the classical filtering problem, a solution is still feasible provided that some information about the original independent components is provided. Major attention in the literature has been focused on different criteria to be used as objective functions for BSS. However, most of the existing on-line methods can be categorized as using a stochastic gradient. The need for second-order based algorithms for BSS can be easily revealed, as the fast convergence rate of the Recursive-Least-Squares (RLS) based algorithms can be advantageous for many applications. Among the many existing objective functions for Blind Signal Separation, Maximum Likelihood and Negentropy stand as strong criteria which are well justified, as they minimize the mutual information between the original independent components [1],[2]. Maximum likelihood proved especially suitable for heavy tailed distribution signals, such as audio data [3].

In the case of pre-whitened inputs, the separating de-mixing matrix is constrained to be orthogonal, lying on a Stiefel manifold [4]. It is widely recognized that applying the constraints of the Stiefel manifold to an optimization problem leads to a better performance of the algorithm. Following the Stiefel geometry leads to a modified gradient rather than the ordinary gradient of the objective function [4]. The new constrained gradient is referred to as the natural gradient, for the pre-whitened case [4],[5]. A natural question would be how to extend and combine the performance of the natural gradient with second-order based RLS algorithm. It is the objective of this paper.

2. EXISTING MAXIMUM-LIKELIHOOD BASED ALGORITHMS

Let $s_i(n), i = 1,2,\ldots,N$ be scalar inputs (or sources) to the blind signal separation model at a time n. For simplicity, it is assumed that the mixing is linear and that the mixing matrix is square, i.e. the number of inputs N is equal to the number of mixtures $x_i(n), i = 1,2,\ldots,N$. Therefore, the mixing matrix A is a square matrix of size $N \times N$. The mixing model can be expressed as:

$$\mathbf{x}(n) = \mathbf{A} \times \mathbf{s}(n) \quad (1).$$

The mixture $\mathbf{x}$ is then applied to a whitening matrix $\mathbf{V}$. The resulting whitened mixtures in $\mathbf{z}$ are expressed as:

$$\mathbf{z}(n) = \mathbf{V} \times \mathbf{x}(n) = \mathbf{V} \times \mathbf{A} \times \mathbf{s}(n) = \mathbf{B} \times \mathbf{s}(n) \quad (2),$$

where $\mathbf{B}$ is the resulting mixing matrix after the whitening stage. The purpose of the blind signal separation algorithms is to estimate a matrix $\mathbf{W}$ such that $\mathbf{W} \times \mathbf{B} = \mathbf{I}_{N \times N}$, where $\mathbf{I}$ is an identity matrix. Then the outputs of the separation process referred to as $y_i(n)$ would be identical to the source inputs $s_i(n)$. Maximum Likelihood targets a separation via increasing the likelihood between the outputs $y_i(n)$ and the inputs $s_i(n)$ [5]. In the case of pre-whitened inputs, the cost function of the log-likelihood $L(\mathbf{W})$ of the de-mixing matrix $\mathbf{W}$ can be expressed as:

$$L(\mathbf{W}) = E\left\{ \sum_{i=1}^{N} \log p_i(\mathbf{w}_i \mathbf{z}) \right\} \quad (3),$$

where $E\{\ \}$ refers to the expected value, $\mathbf{w}_i$ is the i^{th} row of the matrix $\mathbf{W}$ and $p_i(\ )$ is a probability density function. The above cost function has the gradient $\nabla L(\mathbf{W})$ as:

$$\nabla L(\mathbf{W}) = E\left\{ \sum_{i=1}^{N} \log \mathrm{g}(\mathbf{y}) \mathbf{z}^T \right\} \quad (4),$$

where $g(y_i) = \dfrac{p_i'}{p_i}$ and is usually set to $2\tanh(y_i)$ for super-gaussian data, such as audio data. Pre-whitening also constrains the matrix $\mathbf{W}$ to be orthogonal, meaning that $\mathbf{W}\mathbf{W}^T = \mathbf{I}_{N \times N}$. This constraint places the optimization of the cost function on a Stiefel manifold, where the knowledge of the differential geometry can be used to adjust the original gradient $\nabla L(\mathbf{W})$ to the natural gradient $\tilde{\nabla} L(\mathbf{W})$ which follows the geometry of the manifold [6],[7]:

$$\begin{aligned} \tilde{\nabla} L(\mathbf{W}) &= \nabla L(\mathbf{W}) - \mathbf{W} \nabla L(\mathbf{W})\ \mathbf{W} \\ &= (\mathrm{g}(\mathbf{y})\mathbf{y}^T - \mathbf{y}\,\mathrm{g}(\mathbf{y})^T)\ \mathbf{W} \end{aligned} \quad (5).$$

The natural gradient algorithm is based on taking the instantaneous value of the update for the above gradient, so that the update equation becomes:

$$\mathbf{W}_{new} = \mathbf{W}_{old} + \mu (\mathbf{y}\, g(\mathbf{y})^T - g(\mathbf{y})\, \mathbf{y}^T)\, \mathbf{W}_{old} \quad (6),$$

where μ is the step size adequately set to 0.0005.

3. THE QUASI-RLS STIEFEL ALGORITHM

A time averaged approximation of the Hessian of the cost function is (except for a scaling factor of $1-\lambda$):

$$\mathbf{R}_{new} = \lambda \times \mathbf{R}_{old} + \tilde{\nabla} L(\mathbf{W})_{vec} \times \tilde{\nabla} L(\mathbf{W})^T{}_{vec} \} \quad (7),$$

where $\tilde{\nabla} L(\mathbf{W})_{vec}$ is a re-arrangement of the $N \times N$ matrix $\tilde{\nabla} L(\mathbf{W})$ as a vector of size $N \times 1$, via column stacking. The resulting $\mathbf{R}$ is thus a matrix of size $N^2 \times N^2$. λ is a forgetting factor close to 1, common in RLS algorithms. The update can thus be calculated as:

$$\Delta \mathbf{W}_{vec} = \mathbf{R}^{-1} \times \tilde{\nabla} L(\mathbf{W})_{vec} \quad (8).$$

The inverse of the quasi-Hessian matrix $\mathbf{R}$, which is required for the above update, can be calculated recursively using the matrix inversion lemma as:

$$\mathbf{R}^{-1}{}_{new} = \frac{1}{\lambda}\left[\mathbf{R}^{-1}{}_{old} - \frac{\mathbf{R}^{-1}{}_{old}\, \tilde{\nabla} L(\mathbf{W}) \tilde{\nabla} L(\mathbf{W})^T\, \mathbf{R}^{-1}{}_{old}}{\lambda + \tilde{\nabla} L(\mathbf{W})^T\, \mathbf{R}^{-1}{}_{old} \tilde{\nabla} L(\mathbf{W})}\right] (9).$$

To enhance the robustness of the algorithm, the update calculated in (8) is projected on the Stiefel manifold as:

$$\tilde{\Delta} \mathbf{W}_{vec} = \mathbf{W}\mathbf{W}^T \Delta \mathbf{W}_{vec} - \mathbf{W} \Delta \mathbf{W}_{vec} \mathbf{W} \quad (10).$$

The new update is rearranged into a matrix $\tilde{\Delta}\mathbf{W}$ of the original size $N \times N$ to be added to the current estimate of the matrix $\mathbf{W}$:

$$\mathbf{W}_{new} = \mathbf{W}_{old} + \mu_{RLS}\, \tilde{\Delta} \mathbf{W} \quad (11).$$

4. SIMULATION RESULTS AND PERFORMANCE COMPARISON

To compare the different BSS algorithms, tests were performed on a mixture of audio data files (speech) sampled at 8 kHz and of duration 3.7 sec. There were 4 sources considered, two male files and two female files. The choice of the speech files duration was made short to emphasize on the fast convergence property of the new algorithm. The mixing matrix chosen for this aural scene is rather harsh. The comparison of the proposed algorithm is performed with the Natural Gradient based on Maximum Likelihood (NAG, step size μ set to 0.0005 [8]), and with the RLS-modified Natural Gradient (RLS-NAG). The RLS-NAG proposed in [8] suggests an RLS-update of the Natural Gradient algorithm by modifying the update from

$$\Delta \mathbf{W} = \mu \left[\mathbf{I} - g(\mathbf{y})\, \mathbf{y}^T\right] \mathbf{W} \quad (12)$$

to

$$\Delta \mathbf{W} = \mu_r \left[\mathbf{I} - g(\mathbf{y})\, \mathbf{y}^T\right] \mathbf{Q}^{-1} \mathbf{W} \quad (13)$$

with $\mathbf{Q} = g(\mathbf{y})\, \mathbf{y}^T$ and has the form of a covariance matrix. This update is applied on each element of $\mathbf{W}$ individually and convergence is obtained when $\mathbf{Q}$ is a diagonal matrix, i.e. when the mutual information between $g(\mathbf{y})$ and $\mathbf{y}$ is minimized. This algorithm works efficiently and provides in most cases a better performance than the ordinary Natural Gradient algorithm [8]. For the RLS-NAG algorithm μ_r is set to 0.008 while the forgetting factor is set to 0.991. For the new proposed algorithm, the step size μ_{RLS} is set to 0.12 and the forgetting factor λ is set to be time varying starting at 0.9993 at time n=1 and ending at 0.9996 at n=10000. To evaluate the effectiveness of the separation algorithms, the PESQ scores (from ITU-T P.862 [9]) of the separated outputs $y_i(n)$ were computed. PESQ scores have values varying between -0.5 to 4.5, and higher values indicate a higher speech quality. The results of the above test are provided in Table 1. From this table, it can be seen that the proposed algorithm converges much faster than the other algorithms, as shown by the significant difference in the PESQ scores achieved by the different algorithms.

5. CONCLUSION

This paper presented a new algorithm named the Quasi-RLS Stiefel. This algorithm combines the principles of natural-gradient on differential manifolds (Stiefel manifold in our case) and RLS-based algorithms. Simulation results quantified the good on-line convergence speed of the proposed algorithm and proved that the algorithm is very suitable for real-time Blind Signal Separation.

Table 1. PESQ scores for a mixture of 4 speech files

File	New algorithm	NAG	NAG-RLS
Female1	3.244	2.311	1.293
Female2	3.310	1.732	1.641
Male1	3.643	1.807	2.288
Male2	3.201	1.753	1.475

REFERENCES

[1] Amari, S. (1998). Natural gradient works efficiently in learning. Neural Computation, vol. 10, 251-276.

[2] Hyvarinen, A. , Karhunen, J. and Oja, E. (2001). In: Independent Component Analysis, John Wiley & Sons.

[3] Kidmose, P. (2001). Blind Separation of Heavy Tail Signals. Ph.D. thesis, Technical University of Denmark.

[4] Amari, S. (1999) Natural gradient learning for over- and under-complete bases in ICA. Neural Comput., vol.11, 1875–1883.

[5] Cichocki, A. and Amari, S. (2002) Adaptive Blind and Image Processing, John Wiley & Sons.

[6] Douglas, S.C. (2000). Self-stabilized gradient algorithms for blind source separation with orthogonality constraints. IEEE Trans. Neural Networks, vol. 11, 1490–1497.

[7] Edelman, A., Arias, T. and Smith, S. T. (1998). The geometry of algorithms with orthogonality constraints, SIAM J. Matrix Anal. Appl., vol. 20, 303–353.

[8] Benesty, J. (2000). An Introduction to Blind Source Separation of Speech Signals. In: Acoustic Signal Processing for Telecommunication, Kluwer Academic Publisher, 13-61.

[9] ITU-T P.862 (2000). Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone network and speech codecs. International Telecommunication Union.

A KALMAN FILTER WITH A PERCEPTUAL POST-FILTER TO ENHANCE SPEECH DEGRADED BY COLORED NOISE

Ning Ma [1], Martin Bouchard [1], and Rafik A. Goubran [2]
[1] School of Information Technology and Engineering, University of Ottawa, 800 King Edward, Ottawa, Ontario, Canada, K1N 6N5 email: {nma, bouchard} @site.uottawa.ca
[2] Department of Systems and Computer Engineering, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6, Canada, email: Rafik.Goubran@sce.carleton.ca

1. INTRODUCTION

Speech enhancement algorithms have been employed successfully in many areas such as VoIP, automatic speech recognition and speaker verification. Some of the methods assume that the environmental noise is white noise. However, when used in colored noise environments, those methods will produce a weaker performance. Approaches for colored noise have also been previously proposed, however those previous methods have to detect non-speech frames for the noise covariance estimation. This paper proposes a method for colored noise speech enhancement based on a Kalman filter combined with a post-filter using masking properties of human auditory systems. No detection of non-speech frames is needed in the proposed method.

2. THE PERCEPTUAL KALMAN FILTER ALGORITHM

A clean speech signal $s(n)$ can be modeled as:

$$s(n) = \sum_{i=1}^{p} a_i s(n-i) + u(n) \quad (1)$$

where $s(n)$ is the n-th sample of the clean speech signal, $u(n)$ is a zero mean white Gaussian process with variance σ_u^2 and a_i is the i-th autoregressive (AR) model parameter. The n-th sample of the noisy speech signal $y(n)$ is:

$$y(n) = s(n) + v(n) \quad (2)$$

where $v(n)$ is a colored measurement noise process with covariance matrix $\mathbf{R}$. Using a vector Kalman filter as in [1], a state-space model can be expressed as

$$\mathbf{x}(n) = \mathbf{F}\mathbf{x}(n-1) + \mathbf{G}u(n) \quad (3)$$

$$\mathbf{y}(n) = \mathbf{H}\mathbf{x}(n) + \mathbf{v}(n) \quad (4)$$

where

$$\mathbf{F} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ a_p & a_{p-1} & a_{p-2} & \cdots & a_1 \end{bmatrix} \quad (5)$$

$$\mathbf{x}(n) = [s(n-p+1)\cdots s(n)]^T \quad (6)$$

$$\mathbf{y}(n) = [y(n-p+1)\cdots y(n)]^T \quad (7)$$

$$\mathbf{v}(n) = [v(n-p+1)\cdots v(n)]^T \quad (8)$$

$$\mathbf{G} = [0\ 0\ \cdots\ 1]^T \quad (9)$$

and $\mathbf{H}$ is a p-th order identity matrix. Thus the Kalman filter estimation and updating equations are as follows:

$$\mathbf{e}(n) = \mathbf{y}(n) - \hat{\mathbf{x}}(n \mid n-1) \quad (10)$$

$$\mathbf{K}(n) = \mathbf{P}(n \mid n-1) \times (\mathbf{P}(n \mid n-1) + \mathbf{R})^{-1} \quad (11)$$

$$\hat{\mathbf{x}}(n \mid n) = \hat{\mathbf{x}}(n \mid n-1) + \mathbf{K}(n) \times \mathbf{e}(n) \quad (12)$$

$$\mathbf{P}(n \mid n) = (\mathbf{I} - \mathbf{K}(n)) \times \mathbf{P}(n \mid n-1) \quad (13)$$

$$\hat{\mathbf{x}}(n+1 \mid n) = \mathbf{F}\hat{\mathbf{x}}(n \mid n) \quad (14)$$

$$\mathbf{P}(n+1 \mid n) = \mathbf{F}\mathbf{P}(n \mid n)\mathbf{F}^T + \mathbf{G}\mathbf{G}^T\sigma_u^2 \quad (15)$$

where $\mathbf{e}(n)$ is the innovation vector, $\mathbf{K}(n)$ is the Kalman gain, $\hat{\mathbf{x}}(n \mid n)$ represents the filtered estimate of the state vector $\mathbf{x}(n)$, $\hat{\mathbf{x}}(n \mid n-1)$ is the minimum mean-square estimate of the state vector $\mathbf{x}(n)$ given the past observations $y(1), \ldots, y(n\text{-}1)$, $\mathbf{P}(n \mid n)$ is the filtered state error covariance matrix; and $\mathbf{P}(n \mid n-1)$ is the *a priori* error covariance matrix. The last element of $\hat{\mathbf{x}}(n \mid n)$, $\hat{s}(n)$, is the output of the Kalman filter. The computation of the model noise process statistics (variance σ_u^2) and the measurement colored noise statistics (covariance matrix $\mathbf{R}$) can be done with a covariance matching method, as in [1].

The Kalman filtered speech signal is then processed by a post-filter, on frame-by-frame basis, with the frame length defined as L. Both time domain forward masking effects and frequency domain simultaneous masking properties are considered in the proposed post-filter. The psychoacoustic time domain forward masking effects are modeled as a psychoacoustic specific loudness versus critical-band rate and time. The total loudness Q, defined as the sum of the output specific loudness in all critical bands, is used as an estimate of the time domain forward masking level [2]. The total masking level is determined by integrating the frequency-domain simultaneous masking effect and the time-domain forward masking effect [2]. The total masking level of the i^{th} critical band ($i = 1, 2, \cdots, 18$) at time t is:

$$M_t(i) = \max\{M_s(i),\ M_t(i)^* \cdot \exp^{-\Delta t/(\tau(i)\cdot Q)}\} \qquad (16)$$

where $M_t(i)$ and $M_t(i)^*$ are the total masking levels of the current frame and the previous frame, respectively; $M_s(i)$ is the masking level computed from the simultaneous frequency domain masking model [3], Δt is the time difference between two frames, $\tau(i)$ is the maximum decay time constant in each critical band [2], and Q is the total loudness level computed as in [2]. The post-filter performs thresholding on its input signal based on the computed total masking level $M_t(i)$. To perform the thresholding, the following procedure is used:

(1) Mapping the total masking level $M_t(i)$ in each critical band to frequency domain (FFT bins) to obtain $T(\omega_j)$ ($j = 0, 1, \cdots, 255$).

(2) From previously computed filtered state error covariance matrices ($\mathbf{P}(n-i \mid n-i)$ $i = 0,1,2,\cdots$) , estimate a covariance function for the filtered state error signal. The power spectrum density (PSD) $P_e(\omega_i)$ of the filtered state error signal is then computed by a 256-points FFT of the covariance vector, i.e., for $\omega_j = 2\pi/256 \cdot j$ ($j = 0,\cdots,255$). Then the thresholding is performed on the speech spectrum:

$$\left|\tilde{\hat{S}}(\omega_i)\right| = \begin{cases} \left|\hat{S}(\omega_i)\right| \times \alpha^{P_e(\omega_i)/T(\omega_i)} & \text{if } P_e(\omega_i) < T(\omega_i) \\ \left|\hat{S}(\omega_i)\right| \times \alpha \times (1+\alpha^{P_e(\omega_i)/T(\omega_i)}) & \text{otherwise} \end{cases} \qquad (17),$$

where α ($0 < \alpha < 1$) is a tonality coefficient [3].

(3) Doing an IFFT using $\left|\tilde{\hat{S}}(\omega_j)\right|$ and the phase of $\hat{S}(\omega_j)$, keeping the last L values of the size-256 IFFT outputs to obtain the frame of improved speech signal.

From (17), the masking properties and the characteristics of the speech frame are taken into account. If $P_e(\omega_i) < T(\omega_i)$, most of the power in the Kalman filtered signal can be kept, to reduce distortion. If $P_e(\omega_i) > T(\omega_i)$, whether to increase the power of the Kalman filtered signal or to reduce it depends on the value of α and $P_e(\omega_i)$. For $\alpha > (\sqrt{5}-1)/2$ (tone-like frame) and $P_e(\omega_i) < T(\omega_i)\ln(1/\alpha - 1)/\ln\alpha$, the power of the Kalman filtered signal is increased. In any other cases, it is reduced.

3. SIMULATION RESULTS

Four different speech sentences of 4.75 seconds spoken by 2 females and 2 males were used in the simulations. A colored noise $v(n)$ was used as the noise source, and it was obtained by running a white noise signal through a 8th order AR filter. The frame size was 80 samples (L=80), i.e. 10 ms frames. The AR prediction order p was set to be 10, and the AR coefficients were updated for every frame. The data length used to compute the AR parameters was 160 samples (the current noisy frame and one previous enhanced frame). The performance index used was ITU-T P.862 PESQ scores, in order to have a close match with subjective speech quality scores. In the P.862 standard, the lowest PESQ score is −0.5 and the highest score is 4.5. High scores stand for good speech quality. The PESQ scores obtained by using the proposed approach and other recent algorithms [3,4] for colored noise under various noisy speech signal-to-noise ratios (SNR) are shown in Table 1. The simulation results show that in the view of the PESQ scores, the new proposed method has the best performance for any input noisy speech SNR value.

4. CONCLUSION

In this paper, a total masking threshold including frequency domain simultaneous masking effects and time domain forward masking effects was applied as a post-filter to a Kalman filtered signal, to further enhance it in a perceptual sense. A thresholding procedure suitable for colored noise based on the computed masking level was proposed. Simulation results have shown that the new method leads to very promising results. No speech versus noise detection (i.e. VAD) is required in the proposed method.

REFERENCES

[1] Ma, N., Bouchard M. and Goubran, R. (2004). Perceptual Kalman filtering for speech enhancement in colored noise, Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Proc. (ICASSP)

[2] Huang, Y.-H. and Chiueh, T.-D. (2002). A New Audio Coding Scheme Using a Forward Masking Model and Perceptually Weighted Vector Quantization, IEEE Trans. Speech and Audio Proc., vol.10, 325-335.

[3] Virag, N. (1999). Single Channel Speech Enhancement Based on Masking Properties of the Human Auditory System, IEEE Trans. Speech Audio Proc., vol. 7, 126-137

[4] Popescu, D. C. and Zeljkovic, I. (1998). Kalman Filtering of Colored Noise for Speech Enhancement, Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Proc. (ICASSP), vol.2, 997-1000.

Table 1. PESQ scores obtained by different algorithms

	PESQ Scores				
Input SNR (dB)	Original Noisy Speech	Spectral Subtraction	Method from [3]	Method from [4]	Proposed Method
-5	1.293	1.266	1.296	1.371	**1.577**
0	1.605	1.636	1.673	1.692	**1.906**
5	1.882	1.950	1.981	1.977	**2.227**
10	2.182	2.259	2.304	2.292	**2.549**
15	2.502	2.585	2.627	2.603	**2.873**

IMPROVED PACKET LOSS CONCEALMENT FOR PCM VOIP

Q. Li, M. Elsabrouty and M. Bouchard
School of Information Technology and Engineering, University of Ottawa, 800 King Edward,
Ottawa, Ontario, K1N 6N5 {qili, melsabro, bouchard}@site.uottawa.ca

1. INTRODUCTION

Voice-over-IP (VoIP), the transmission of packetized voice over IP networks, is gaining much attention as a possible alternative to conventional public switched telephone networks (PSTN). However, impairments present on IP networks, namely jitter, delay and channel errors can lead to the loss of packets at the receiving end. This packet loss degrades the speech quality. Model-based speech coders, such as the International Telecommunication Union (ITU-T) G.729A and G.723.1 standards, have been extensively used for speech coding over IP networks because of their low bit rates requirements (5.3 to 6.4 kbit/s for G.723.1 and 8 kbit/s for G.729A). Also, they have an inherent ability to recover from erasure: using their model-based structure they include a built-in packet loss concealment scheme which makes their quality drop slowly with increasing amount of packet loss. However, their model requires a few frames to adapt to the transition from a concealed state to a correct state. Thus, model-based speech coders actually tend to corrupt a few good packets before recovery, as a result of a phenomenon known as "State Error"[1].

On the other hand, speech coded with Pulse Code Modulation (PCM, ITU-T G.711, 64 kbit/s), although having a higher quality compared to G.729A and G.723.1 in the periods of normal operation, does not have the built-in ability to conceal erasure. This results in a serious drop in speech quality during loss periods. Yet, PCM-based coders can recover from packet loss faster than model-based coders, since the first speech sample in the first good packet restores the speech to its original quality. With a suitable packet loss concealment scheme, PCM would make a very viable alternative to G.729A or G.723.1 for VoIP, considering its low complexity, its superior quality under normal conditions and its good performance in tandem coding. This paper introduces a high-performance concealment algorithm for PCM-coded speech.

2. A NEW PCM PACKET LOSS CONCEALMENT ALGORITHM

The new linear prediction based concealment technique is using a linear prediction with a fairly large order filter to model the speech:

$$S(n) = \sum_{i=1}^{P} \left(a_i \times S(n-i)\right) + b(n) \qquad (1),$$

where $S(n)$ is the n^{th} speech sample, P is the prediction order (set to 50), a_i are the linear prediction coefficients, and $b(n)$ is the residual signal of the prediction. As can be seen from (1), the current speech sample $S(n)$ is composed of two components. The first component is the predictable part carrying the information of the vocal tract along with the correlation between the current sample and the 50 previous ones. The second component is the residual signal $b(n)$ that contains the current unpredictable excitation.

In the case of a lost packet, the previous speech samples are available and thus the predictable term in (1) can be computed. However, the residual signal $b(n)$ is unknown to the receiver side. In this case, a good choice can be to use a small percentage of a pitch-predicted signal as $b(n)$, or in other words as the input excitation for the synthesis system. Here, the pitch-predicted signal refers to a Reverse-Order Pitch Period Replication (RORPP) of the lost frame, estimated in a manner similar to the concealment algorithm implemented in the ITU-T G.711-Annex A [2]. Thus, using a small percentage of the pitch-predicted signal, equation (1) can be re-written as:

$$S(n) = \sum_{i=1}^{P} \left(a_i \times S(n-i)\right) + \hat{S}(n) \times G \qquad (2),$$

where $S(n)$ denotes the concealed speech sample and $\hat{S}(n)$ is the pitch-predicted signal obtained from the ITU-T G.711-A (RORPP) concealment scheme. A value of $G = 0.01$ was found to give the best results in practice [3].

In "best effort" IP networks, future packets often arrive early. Thus it becomes possible to use future packets from the jitter buffer to predict a lost packet. A Hamming window can be applied to combine the prediction from past packets and the prediction from a future packet (or more, if available). Also, to provide a better approximation of the original signal, the concealment algorithm can be modified to perform a weighted summation of the speech predicted by linear prediction (i.e. equation (2) for the case where

future samples are not considered) and the pitch-based prediction, using a Voiced/Unvoiced (V/UV) speech classification to tune the parameters. For our experiments, for the V/UV classification a simple scheme based on the linear prediction residual energy was used. A threshold was determined, based on the fact that voiced speech tends to be more energetic in the residual signal. The proposed packet loss concealment algorithm for PCM speech thus becomes:

$$S^p(n) = \sum_{i=1}^{P} \left(a_i \times S^p(n-i)\right) + \hat{S}(n) \times G \quad (3)$$

$$S^f(n) = \sum_{i=1}^{P} \left(a_i \times S^f(n+i)\right) + \hat{S}(n) \times G \quad (4)$$

$$S_v(n) = \begin{pmatrix} \left(\alpha_v \times S^p(n) + \beta_v \times \hat{S}(n)\right) \times \text{Hamming win. } 2^{\text{nd}} \text{ half}) \\ + \left(\alpha_v \times S^f(n) + \beta_v \times \hat{S}(n)\right) \times \text{Hamming win. } 1^{\text{st}} \text{ half}) \end{pmatrix} \quad (5)$$

$$S_{uv}(n) = \begin{pmatrix} \left(\alpha_{uv} \times S^p(n) + \beta_{uv} \times \hat{S}(n)\right) \times \text{Hamm. win. } 2^{\text{nd}} \text{ half} \\ + \left(\alpha_{uv} \times S^f(n) + \beta_{uv} \times \hat{S}(n)\right) \times \text{Hamm. win. } 1^{\text{st}} \text{ half} \end{pmatrix} \quad (6),$$

where $S_v(n)$ and $S_{uv}(n)$ are the final concealed signals for voiced and unvoiced frames, respectively, $S^p(n)$ and $S^f(n)$ are the linear prediction result from the past and future samples, respectively, α_v and β_v are summation weights for voiced frames, and α_{uv} and β_{uv} are summation weights for unvoiced frames. The best results were obtained with $\alpha_v = 0.9$ and $\beta_v = 0.1$, $\alpha_{uv} = 0.6$ and $\beta_{uv} = 0.4$.

3. PERFORMANCE OF THE PROPOSED ALGORITHM

The new algorithm was compared to the ITU-T G.711-Appendix A concealment tool and to the packet repetition method. The test was performed on a set of speech files from four speakers (two males and two females) referred to in the results as M1, M2, F1 and F2. For each of those speakers, 10 speech files were used, each containing two sentences in English for a duration of 8 sec. Frames of 80 samples (10 ms) were used. The format of the files was linear PCM, with 8 kHz sampling rate. The files were taken from the ITU-T supplement P.23. The assessment tool used to evaluate the results of the concealment techniques was the Perceptual Estimation of Speech Quality (PESQ) standard P.862 developed by the ITU-T [4]. This tool has shown to give reliable estimation of subjective quality tests. The scores produced by the PESQ are in the range -0.5 (very poor quality) to 4.5 (very good quality), similar to the standard Mean Opinion Score (MOS) scale.

Random loss pattern tests were performed for loss rates of 5%, 10 % and 25%. Table 1 summarizes the average PESQ results for those three loss rates. It can be seen from the above figures that the performance of the new algorithm is superior to both the existing ITU-T G.711A method and the packet repetition method. A significant and almost steady margin appears as a difference between the new proposed algorithm and the ITU-T G.711A method, and both methods perform much better than the packet repetition technique. The margin between the proposed method and the ITU-T G.711A represents the performance gain of incorporating the linear prediction model (and the possible use of future samples) with the pitch-repetition based concealment.

4. CONCLUSION

In this paper, a new concealment algorithm for PCM packetized speech was presented. The model provides very encouraging results for the idea of combining a pitch prediction along with a high-order linear prediction to produce the concealed speech samples. Future samples are taken into account (optionally), as well as the adaptation of the weighting coefficients α and β based on a V/UV classification. The PESQ-MOS scores obtained for the random loss tests have shown that the proposed algorithm exhibits a superior high-quality concealment performance in all cases, when compared to an existing commercial method (packet repetition) or to the ITU-T G.711 A concealment technique.

REFERENCES

[1] ITU-T G.711 Appendix A (2000). A high quality low-complexity algorithm for packet loss concealment with G.711. ITU-T Recommendation.

[2] Montminy, C. and Aboulnasr, T. (2000). Improving the performance of ITU-T G.729A for VoIP. International Conference on Multimedia Exposition 2000 (ICME 2000), vol. 1, 433–436.

[3] Elsabrouty, M., Bouchard, M. and Aboulnasr, T. (2004). Receiver-based packet loss concealment for pulse code modulation (PCM G.711) coder. Signal Processing, vol. 84, 663-667.

[4] ITU-T P.862 (2000). Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrow-band telephone network and speech codecs. ITU-T Recommendation.

Table 1 Average PESQ Results

File	Test	New Algorithm	G.711A	Packet Repetition
M1	5% loss rate	3.98	3.45	3.05
	10% loss rate	3.61	3.09	2.62
	25% loss rate	3.20	2.60	2.28
M2	5% loss rate	3.80	3.41	2.81
	10% loss rate	3.55	3.12	2.61
	25% loss rate	3.02	2.63	2.24
F1	5% loss rate	3.83	3.36	2.88
	10% loss rate	3.53	2.93	2.50
	25% loss rate	3.05	2.58	2.15
F2	5% loss rate	3.76	3.31	2.86
	10% loss rate	3.53	2.87	2.36
	25% loss rate	2.97	2.43	1.83

MODIFIED SPREAD SPECTRUM AUDIO WATERMARKING ALGORITHM

Libo Zhang[1], Sridhar Krishnan[1], and Heping Ding[2]
[1]Department of Electrical and Computer Engineering, Ryerson University, Toronto, Canada, {libo, krishnan}@ryerson.ca
[2]Institute for Microstructural Sciences, National Research Council of Canada, Ottawa, Canada, heping.ding@nrc-cnrc.gc.ca

1. INTRODUCTION

Digital audio watermarking is a technique to embed some information, i.e. the watermark, into the host audio signal in an imperceptible and robust way. The watermark can be used for copyright protection, usage monitoring, etc. The imperceptibility requires that the watermark is embedded without causing perceptual distortion and the robustness requires that the watermark can be extracted even after the watermarked signal is distorted by different signal processing manipulations and/or intentional attacks.

Watermarking can be thought of as a telecommunication means to transmit the watermark signal over the medium of the host signal. The imperceptibility limits the power of the watermark signal to be lower than the hearing threshold and thus the communication is characterized by a low SNR. Many robust audio watermarking algorithms are based on Spread Spectrum (SS) technique due to its robustness to interference and low-energy requirement [1].

In the following, the conventional SS audio watermarking will be reviewed first and then a modification will be presented to increase the robustness significantly.

2. CONVENTIONAL SS WATERMARKING

The projection of one vector $\vec{u}$ onto another vector $\vec{v}$, i.e. *normalized correlation,* is defined as,

$$u_v = v_u = \vec{u} \bullet \vec{v} \equiv \frac{1}{N}\sum_{i=1}^{N} u_i v_i \qquad (1)$$

where N is the length of the vectors $\vec{u}$ and $\vec{v}$.

In the standard SS watermarking scheme, a watermark bit $b = \{\pm 1\}$ is spread by a normalized pseudo-random sequence $\vec{w}$, satisfying $w \bullet w = 1$ and of length N, to generate the spread watermark $b \cdot w$ and this spread watermark is embedded into the host signal vector $\vec{x}$ of the same length. The result is the following watermarked signal,

$$y = x + \alpha \cdot (b \cdot w) \qquad (2)$$

where the perceptual factor $\alpha > 0$ controls the perceptibility of the watermark. The watermark signal power is as follows,

$$W = \| y - x \|^2 = (y - x) \bullet (y - x) = \alpha^2 \qquad (3)$$

At the receiver, the received signal $\vec{r}$, which is the watermarked signal $\vec{y}$ corrupted by an additive transmission noise $\vec{n}$, is projected onto the sequence $\vec{w}$ as

$$c = r \bullet w = (y + n) \bullet w = \alpha \cdot b + (x + n) \bullet w \qquad (4)$$

When the spreading factor N is large enough, the second term decreases so the first term dominates and thus *sign(c)* can be used to determine the embedded watermark bit b.

Assuming $\vec{x} \sim \boldsymbol{N}(0, \sigma_x^2)$ and $\vec{n} \sim \boldsymbol{N}(0, \sigma_n^2)$, it can be easily shown that $c \sim \boldsymbol{N}(\alpha \cdot b, \frac{\sigma_x^2 + \sigma_n^2}{N})$ and thus the corresponding Bit Error Rate (BER) is as follows,

$$BER = Q(\frac{\mu}{\sigma}) = Q\left(\sqrt{\frac{N \cdot \alpha^2}{\sigma_x^2 + \sigma_n^2}}\right) = Q\left(\sqrt{N \cdot WNR \cdot \frac{1}{SNR + 1}}\right) \qquad (5)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^{+\infty} e^{-\frac{u^2}{2}} du$ is normally called the complementary error function, $WNR = \frac{\alpha^2}{\sigma_n^2}$ denotes the power ratio of the watermark to the noise, and $SNR = \frac{\sigma_x^2}{\sigma_n^2}$ is the power ratio of the host signal to the noise.

It can be seen from (5) that both the host signal and the attacks act as noise in watermark extraction. Normally the former impacts the extraction more severely since $\sigma_x^2 >> \sigma_n^2$. For this reason a large spreading factor N is often needed, which results in a very small embedding rate.

3. MODIFIED SS WATERMARKING

Geometrically, the vector $\vec{w}$ and a normalized orthogonal vector $\vec{v}$, satisfying $w \bullet v = 0$ and $v \bullet v = 1$, determine a plane. The projection of the host signal onto this subspace can be represented by the point (x_w, x_v), which maps onto the line $w=v$ at the point $(\frac{x_w + x_v}{2}, \frac{x_w + x_v}{2})$. If the watermark is embedded in the direction of $\vec{w}$ from this point, as shown in Fig. 1, the extraction can be eased. This is shown as follows.

The distortion vector due to this mapping can be expressed as $(\frac{x_v - x_w}{2}, \frac{x_w - x_v}{2})$. The watermarked signal is thus

$$\vec{y} = \vec{x} + \frac{x_v - x_w}{2} \cdot \vec{w} + \frac{x_w - x_v}{2} \cdot \vec{v} + \alpha_{MSS} \cdot (b \cdot \vec{w}) \quad (6)$$

The watermark signal power in this case is

$$W = E[\| \vec{y} - \vec{x} \|^2] = E[(\vec{y} - \vec{x}) \bullet (\vec{y} - \vec{x})] = \alpha_{MSS}^2 + \frac{\sigma_x^2}{N} \quad (7)$$

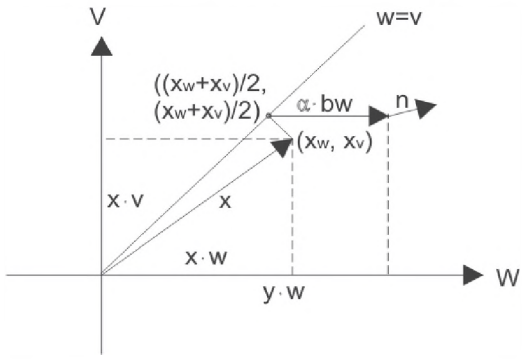


Fig. 1. Proposed watermark embedding algorithm

Based on the above explanation, it can be easily shown that the sign of the following expression can be use to determine the embedded watermark bit,

$$cc = r \bullet w - r \bullet v = \alpha_{MSS} \cdot b + (w - v) \bullet n \quad (8)$$

To illustrate the performance improvement, we compare BERs of the modified and the conventional SS systems with the same watermark signal power, i.e., α_{MSS} is determined as per

$$\alpha_{MSS}^2 + \frac{\sigma_x^2}{N} = \alpha^2 \quad \Rightarrow \quad \alpha_{MSS} = \sqrt{\alpha^2 - \frac{\sigma_x^2}{N}} \quad (9)$$

Thus the BER of the modified system can be expressed as

$$BER = Q(\frac{\mu}{\sigma}) = Q\left(\sqrt{\frac{N \cdot \alpha_{MSS}^2}{2\sigma_n^2}}\right) = Q\left(\sqrt{N \cdot WNR \cdot \frac{N - SWR}{2N}}\right) \quad (10)$$

where $SWR = \frac{\sigma_x^2}{\alpha^2}$ denotes the power ratio of the host signal to the watermark.

This modification improves the conventional scheme under most situations. For example, when SWR=*25dB* and SNR=*5dB*, there is an improvement of about *3dB* at BER of 10^{-5}, as shown in Fig. 2. With the spreading factor *N=4000* (i.e. *N**WNR≈*16* dB), the BER of the conventional SS is 9.675×10^{-4}, while that of the modified SS is 8.864×10^{-6}, an improvement of the order of *2*.

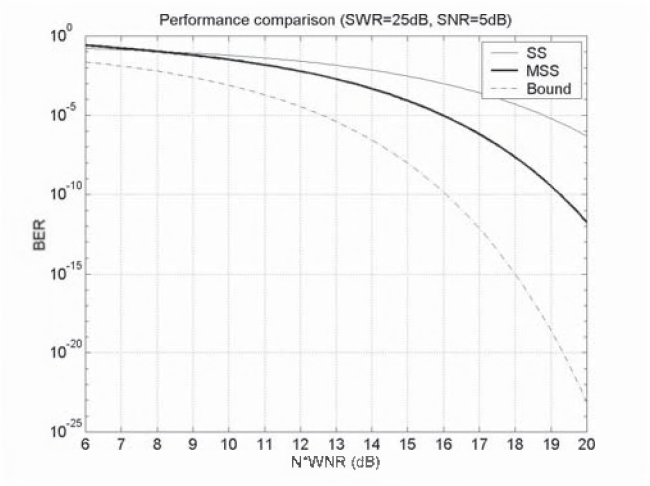


Fig. 2 Bit Error Rate (BER) of the proposed algorithm (Bound: theoretical limit, with host known at receiver)

4. SIMULATION

Audio pieces of different genres, sampled at *44.1* kHz, are tested against the proposed watermarking system. The watermark is spread with the factor *N=4000*, which corresponds to the embedding rate of *11* bps, and the amplitude is controlled by SWR=*32*dB. To evaluate the subjective fidelity, *8* listeners were asked to report dissimilarities between the host and the watermarked using a *5*-point impairment scale (5: imperceptible, 4: perceptible but not annoying, 3: slightly annoying, 2: annoying, 1: very annoying). The average mean opinion score (MOS) was *4.73* with deviation of *0.29*.

To evaluate the robustness, the following attacks are simulated,

- Re-sampling: subsequently down and up sampling.
- Low pass: cutoff at *4 kHz*.
- MP3: compressed at *48kbs*.
- Noise: *-20dB* white Gaussian noise addition.

The mean BER across all the above attacks is *0.075%*. Compared with the conventional SS, the modified SS improves the BER by an averaged order of *2*.

In this paper, we proposed a blind, transparent and simple audio watermarking algorithm. This algorithm can be further improved by introducing more mapping grids.

REFERENCES

1. D. Kirovski and H. S. Malvar, *Robust Spread-Spectrum Audio Watermarking,* IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 1345-1348, 2001.

A SURVEY OF DOUBLE-TALK DETECTION SCHEMES FOR ECHO CANCELLATION APPLICATIONS

Thien-An Vu[1], Heping Ding[2], and Martin Bouchard[1]
[1]School of Information Technology and Engineering, University of Ottawa, Ontario, Canada
[2]Acoustics and Signal Processing, IMS, National Research Council, Ottawa, Ontario, Canada

1. INTRODUCTION

An echo canceller removes undesired echo in full-duplex speech communication. The cancellation is done by modeling the echo path impulse response with an adaptive finite impulse response filter and subtracting the echo estimate from the received signal. A typical diagram of an echo canceller is depicted in **Figure 1**. The signal x(n) and v(n) represent the far-end and near-end speeches respectively. The signal s(n) and y(n) represent the echo signal generated by the actual echo path **h** and the echo estimate produced by the adaptive filter. The signal e(n) denotes the residual error signal, which is transmitted to the far-end side and is used to update the coefficient **w** of the adaptive filter.

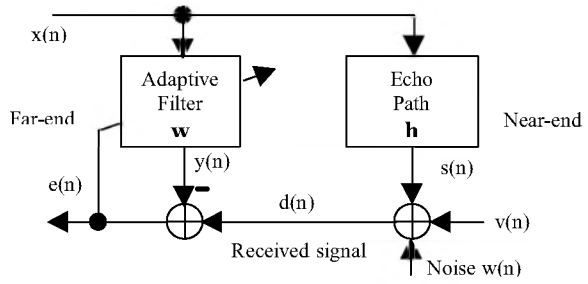


Figure 1: A diagram of an echo canceller.

When the speech signal v(n) is zero and the near-end noise w(n) is assumed to be insignificant, the adaptive filter, **w** can converge to a good estimate of the echo path, **h** and successfully cancel the echo. However, when both v(n) and x(n) are not zero, i.e. double-talk (DT) situation, the near-end speech v(n), which acts as an uncorrelated noise to the adaptive algorithm, may cause the adaptive filter to diverge and allow excessive un-cancelled echo to pass through to the far-end. The common solution to this problem is to slow down or completely stop the filter adaptation when the presence of the near-end speech is detected. This is the role of a double-talk detector.

A special case that some DT detection algorithms seem to have problem with is when there is a change in the echo path, for example in acoustic environment. This can often be falsely detected as a DT condition by those DT detection algorithms. As this is a case the adaptive filter really needs to adapt to the change in the echo path, it is not desired that the adaptation is unnecessarily turned off because DT has falsely been detected. Furthermore, the background noise w(n) at the near-end should not be detected as double-talk.

2. DOUBLE-TALK DETECTION SCHEMES

2.1 Basics

A common basic for most DT detection schemes involves computation of a detection variable from the available data such as the near-end, far-end and/or residue error signal, and comparison of the detection variable with a preset constant threshold. Depending on whether the detection variable is above or below the threshold, a decision is made on whether double-talk condition is present or not. If DT is declared, the filter adaptation is stopped or slowed down for a minimum period of hold time. When the non-DT condition lasts consecutively over the hold time, the adaptation can be resumed until the next DT condition occurs. The hold time is necessary to suppress detection dropouts because of the noisy behavior of the detection variable [1].

There are many different methods existing in the literature on how to form a detection variable for a DT detector. In this paper, attention is paid to some better-known algorithms, which are based on energy comparison and cross-correlation. These algorithms are briefly summarized in the following sub-sections.

2.2 Energy-based algorithms

A simple approach in this category is the well-known Geigel algorithm [2]. The Geigel algorithm compares the magnitude of the near-end received signal d(n) with the maximum magnitude of L most recent samples of the far-end signal x(n), where L is the adaptive filter's length. L past samples are used because of the possible end delay of x(n) through the echo path. The echo path typically dampens the far-end signal x(n), and as a result the magnitude of the received signal d(n) containing only the echo s(n) will be smaller than the received signal d(n) containing both the echo s(n) and the near-end speech signal v(n). The Geigel algorithm computes its detection variable ξ and makes decision as

$$\xi = \frac{|d(n)|}{\max\{|x(n)|,\ldots,|x(n-L+1)|\}} > T$$

If ξ is larger than the threshold T, DT is declared otherwise it is not. The choice of T needs to be made with care, and will strongly affect the performance of the detector. For line echo cancellers, T is set to 0.5 because the hybrid attenuation is assumed to be 6dB. However, for an acoustic echo cancellation environment, not only the background noise level but also the echo path characteristics are time varying. Therefore, it is not easy to decide a proper value for the threshold T. In particular, for the time-varying echo path, the Geigel algorithm can falsely regard a change of the echo path as a DT situation. As a result, the adaptive filter stops updating the coefficients when the coefficient update is actually needed.

Another energy-based DT detection scheme, proposed in [5], exploits the idea of difference in bandwidths of the echo signal s(n) (300-3400 Hz) and the near-end speech v(n) (wider band). Please refer to [5] for more details of the approach.

2.3 Correlation-based algorithms:

A correlation-based DT detector proposed in [3] makes use of the orthogonality principle. When the adaptive filter converges to its optimal solution, the orthogonality principle between the far-end vector $\underline{\mathbf{x}}(n) = [x(n)\ x(n-1)\ \ldots\ x(n-L+1)]$ and the residual error e(n) is satisfied, i.e. $E[e(n)\underline{\mathbf{x}}(n)] = 0$ (E[.] denotes the statistical expectation). In DT situation, the received signal d(n), and therefore the residue echo gets larger abruptly because of the presence of the near-end speech v(n). However, as long as the near end signal v(n) is uncorrelated with the far-end signal x(n), which is usually the case in practice, the orthogonality principle still holds. On the other hand, when the echo path changes, the orthogonality principle cannot be satisfied anymore. The cross-correlation vector between $\underline{\mathbf{x}}(n)$ and e(n) is defined as $\underline{\mathbf{c}}_{xe} = [c_{xe,0}\ \ c_{xe,1}\ \ldots\ \ c_{xe,L-1}]^T$

Where $c_{xe,i} = \dfrac{E[x(n-i)e(n)]}{\sqrt{E[x^2(n-i)][e^2(n)]}}$

The detection variable is defined as $\xi = \dfrac{1}{L}\sum_{i=0}^{L-1} |c_{xe,i}|$. When $\xi \le T$, a properly chosen threshold, the adaptive filter has converged; otherwise, the adaptive has not converged or the echo path has changed. In general, this algorithm does not detect DT condition explicitly. Instead, it decides whether the adaptive filter has converged or not. If the adaptive filter has converged, the adaptation is stopped to protect the filter from being disturbed by DT interference. On the other hand, if the adaptive filter has not converged or the echo path has changed, the adaptive filter will keep adapting.

The algorithm in [3] defines a cross-correlation vector, which is not well normalized. The amount of cross-correlation depends on the statistics of the signals and of the echo path. As a result, the appropriate value for threshold T can vary from one experiment to another [1]. A similar idea that uses the cross correlation vector between $\underline{\mathbf{x}}(n)$ and d(n) is the normalized cross-correlation algorithm, introduced in [4]. The algorithm normalizes the cross-correlation vector in the sense that the detection variable is equal to one when the near-end signal v(n) is zero and less than one when v(n) is not. The normalized cross-correlation vector is defined as $\underline{\mathbf{c}}_{xd} = (\sigma_d^2 \mathbf{R}_x)^{-1/2}\ \underline{\mathbf{r}}_{xd}$ [1], where $\sigma_d^2 = E[d^2(n)]$ is the variance of d(n), $\mathbf{R}_x = E[\underline{\mathbf{x}}(n)\underline{\mathbf{x}}(n)^T]$ is the auto-correlation matrix of $\underline{\mathbf{x}}(n)$, and $\underline{\mathbf{r}}_{xd}$ is the cross-correlation vector between vector $\underline{\mathbf{x}}(n)$ and a scalar d(n). The detection variable is therefore

$$\xi = \sqrt{\underline{\mathbf{r}}_{xd}^T (\sigma_d^2 \mathbf{R}_x)^{-1}\ \underline{\mathbf{r}}_{xd}} = \frac{\sqrt{\underline{\mathbf{h}}^T \mathbf{R}_x \underline{\mathbf{h}}}}{\sqrt{\underline{\mathbf{h}}^T \mathbf{R}_x \underline{\mathbf{h}} + \sigma_v^2(n)}}$$

When $\xi < T$, DT is declared, and when $\xi \ge T$, DT is not present. The threshold T is selected between 0 and 1.

Another DT algorithm, proposed in [5], is based on the orthogonality between e(n) and y(n) and can distinguish between double-talk and echo path change with a low complexity. For more details, please refer to [5].

3. SUMMARY

This paper reviews some typical DT detection schemes existing in the literature, which are based on energy comparison and cross-correlation algorithms. The energy-based algorithms, in general, have the benefit of being computationally simple, needing very little memory, and have been successfully used in line echo cancellation; however, they do not always provide reliable performance in an acoustic echo path environment. On the other hand, the correlation-based algorithms show improved detection performance in such time-varying environment but they would require relatively higher memory storage and computational complexity due to vector or matrix-based operations.

REFERENCES

[1] S. L. Gay, and J. Benesty, *Acoustic Signal Processing for Telecommunication*, Kluwer Academic Publishers, 2000.

[2] D. L. Duttweiller, "A twelve-channel digital echo canceler," IEEE Trans. Comm., vol. 26, pp. 647-653, May 1978.

[3] H. Ye, and B. X. Wu, "A new double-talk detection algorithm based on orthogonality theorem," IEEE Trans. Comm., vol. 39, pp. 1542-1545, November 1991.

[4] J. Benesty, D. R. Morgan, and J. H. Cho, "A new class of doubletalk detectors based on cross-correlation," IEEE Trans. Speech Audio Processing, vol. 8, pp. 168-172, March 2000.

[5] H. Ding, and Frank Lau, "Double-talk detection schemes for echo cancellation," Acoustics Week in Canada - Canadian Acoustical Association, October 2004.

DOUBLE-TALK DETECTION SCHEMES FOR ECHO CANCELLATION

Heping Ding[1] and Frank Lau[2]
[1]Acoustics and Signal Processing, IMS, National Research Council, 1200 Montreal Rd., Ottawa, Ontario K1A 0R6
heping.ding@nrc-cnrc.gc.ca
[2]Nortel Networks, Westwinds Innovation Center, 5050 40th St. NE, Calgary, Alberta T3J 4P8
franklau@nortelnetworks.com

1. Introduction

Widely used in telecommunications, an acoustic or network echo cancellation system, as Fig. 1 shows, subtracts an echo estimate $y(n)$, made by an adaptive filter inputting the far-end signal $x(n)$, from the desired input $d(n)$ to reduce the echo $u(n)$ therein while leaving the near-end signal $s(n)$ intact. In order for $y(n)$ to approximate $u(n)$, the filter coefficients are updated by an adaptation algorithm seeking the minimum of the mean of $e^2(n)$ so that the adaptive filter converges to mimic the echo path. Fundamentals of adaptive filtering and echo cancellation can be found in a text book, such as [1].

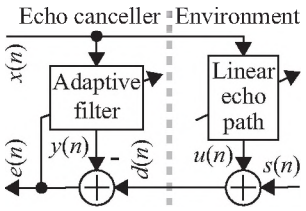


Fig. 1. Echo cancellation

In a practical echo canceller, the adaptation should be a) active when there is little $s(n)$ - so that the filter can converge; and b) halted if $s(n)$ is significant - to prevent the filter from diverging. The key element in controlling this is a so-called double-talk (DT) detector, which detects conditions where $d(n)$ contains a significant near-end signal $s(n)$ mixed with a much stronger $u(n)$ (e.g., 25 dB stronger in a typical speakerphone application), echo of the far-end signal $x(n)$. It is a challenge in academia and industry to find DT detectors that do this job with few misses and false alarms.

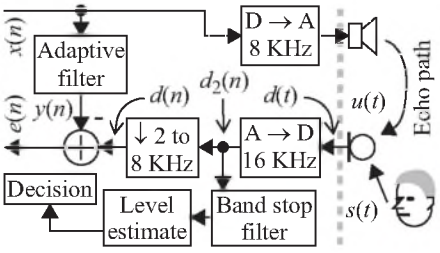


Fig. 2. Proposed Scheme 1 for DT detection

Typical existing DT detection schemes are reviewed in a separate paper [2] and will not be discussed here. This paper presents two new schemes, see for example [3] and [4]†. These schemes have been demonstrated to be pretty reliable, robust, and simple.

2. Scheme 1: Out-of-Telephony-Band Level

In telephony, $u(n)$, linear echo of $x(n)$, has a frequency bandwidth of about 300-3400 Hz, the same as that of $x(n)$. On the other hand, a voice $s(t)$ usually has a wider band, i.e., containing energy outside of this range, and, in a speaker phone, $d(t)$ with a wider band can be readily available since its acoustic source is local. Thus, a DT condition can be claimed if $d(t)$ contains significant energy outside of the telephony band.

Fig. 2 shows the proposed scheme, where $d_2(n)$ is the analog input $d(t)$ sampled at 16 KHz, double the telephony system's sampling rate. In addition to being decimated to form the needed $d(n)$, $d_2(n)$ is band-stop filtered, ridding components within 300-3400 Hz and retaining those within 50-300 Hz and 3400-7000 Hz, for level estimation and decision making.

Simulations with various talkers and speech samples were performed. Fig. 3 shows a typical example, where $d_2(n)$ contains an $s(n)$ 25 dB weaker than $u(n)$ so that it looks almost the same as $u(n)$. This big level difference makes DT detection difficult for some conventional schemes, but the proposed scheme correctly identifies places of significant $s(n)$.

Scheme 1 effectively detects $s(n)$ masked by a much stronger $u(n)$ if a wider band $d(t)$ is made available by the electronics; therefore, it is suitable for speakerphone applications.

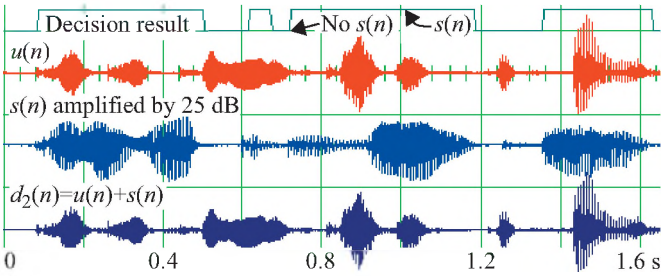


Fig. 3. Simulation result for Scheme 1

† Heping Ding was employed by Nortel Networks during development of the work presented in this article. Nortel Networks owns intellectual property rights associated with the work and this article.

3. Scheme 2: Orthogonality Principle

Comparing levels of various signals in Fig. 1, an energy comparison scheme, discussed in [2], can detect the presence of $s(n)$ *or* a non-convergence echo (uncanceled echo due to non-convergence of the adaptive filter), in $e(n)$, but it cannot distinguish between the two, because either of them results in the same fluctuations in signal levels. Assuming that $s(n)$ is uncorrelated with $x(n)$, the proposed Scheme 2 does the distinction by evaluating the inner product $E[e(n)y(n)]$, where $E[\cdot]$ stands for expectation.

Considering the adaptive filter to be L-tap FIR and with

$$\underline{X}(n) \equiv \left[x(n)\ x(n-1)\ \ldots\ x(n-L+1)\right]^T,\ \boldsymbol{R}(n) \equiv E[\underline{X}(n)\underline{X}^T(n)],$$
$$\underline{W} \equiv \left[w_0\ w_1\ \ldots\ w_{L-1}\right]^T,\ \text{and}\ y(n) = \underline{X}^T(n)\underline{W}, \tag{1}$$

the expectation of interest can be found as

$$E[e(n)y(n)] = (\underline{W}^{\text{opt}} - \underline{W})^T \boldsymbol{R}(n)\underline{W}, \tag{2}$$

where $\underline{W}^{\text{opt}}$ is the optimal value of the filter coefficient vector $\underline{W}$. Once the adaptive filter has converged, i.e., $\underline{W} = \underline{W}^{\text{opt}}$, Eq. (2) vanishes - whether there is an uncorrelated $s(n)$ or not. If the adaptive filter has not converged, Eq. (2) will in general be non-zero - although not guaranteed, it is so in practice. This can also be understood intuitively in an infinite-dimensional Hilbert space, in which $x(n)$, ..., $x(-\infty)$, $y(n)$, $d(n)$, and $e(n)$ are vectors. $d(n)$ consists of: $u(n)$ being a linear combination of $\{x(n), \ldots, x(-\infty)\}$, and $s(n)$ being not. A linear combination of $\{x(n), \ldots, x(n\text{-}L\text{+}1)\}$, the optimum $y(n)$, $y^{\text{opt}}(n) = \underline{X}^T(n)\underline{W}^{\text{opt}}$, can resemble part of $u(n)$, but not the uncorrelated $s(n)$ or the residual echo due to under modeling. $y^{\text{opt}}(n)$ is the projection of $d(n)$ onto the sub-space occupied by $\{x(n), \ldots, x(n\text{-}L\text{+}1)\}$ and is orthogonal to the projection error $e(n)$.

To summarize, $E[e(n)y(n)]$ becomes small with a converged adaptive filter, regardless of $s(n)$, and is large in magnitude if the filter has not converged.

With ensemble means replaced by time averages, the scheme is implemented as

$$\begin{aligned} R_{ey}(n) &= \beta R_{ey}(n-1) + (1-\beta)e(n)y(n) \\ A_{ey}(n) &= \beta A_{ey}(n-1) + (1-\beta)|e(n)y(n)| \end{aligned}, \tag{3}$$

where $0 < \beta < 1$ is a forgetting factor to provide a low-pass smoothing effect, and $A_{ey}(n)$ gives a reference for evaluating the magnitude of $R_{ey}(n)$. When the energy comparison part claims a candidate for the presence of either $s(n)$ or a non-convergence echo in $e(n)$, a distinction is made as per

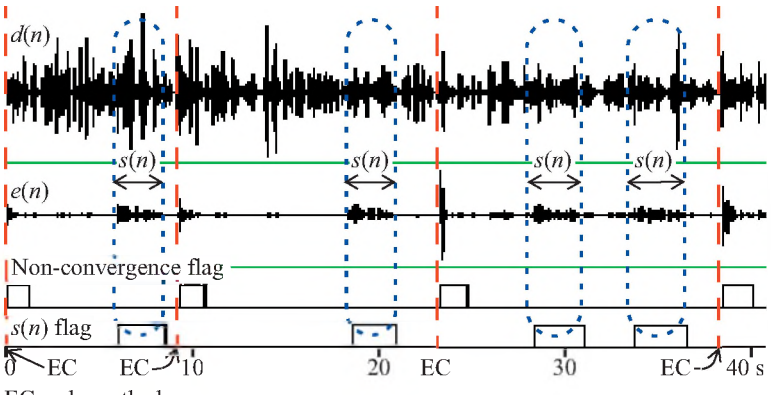
$$|R_{ey}(n)| > T_{\text{echo}} A_{ey}(n) \quad \Rightarrow \text{Candidate is non-conv. echo} \tag{4}$$

$$|R_{ey}(n)| < T_{s(n)} A_{ey}(n) \quad \Rightarrow \text{Candidate is } s(n), \tag{5}$$

where the thresholds $0 < T_{s(n)} \le T_{\text{echo}} < 1$ are experimentally determined.

Unlike Scheme 1, this scheme does not need a wider band $d(t)$, and therefore is suitable for general echo cancellation.

Simulations with similar conditions as that with Scheme 1 were performed, with an example shown in Fig. 4. It is seen that $e(n)$ contains both $s(n)$ and non-convergence echoes, as results of a non-converged adaptive filter after echo path changes. An energy comparison based DT detection scheme is only able to spot these $e(n)$ level increases as candidates for either of the two events, while the proposed Scheme 2 clearly distinguishes between them by raising the non-convergence flag if Eq. (4) is satisfied, or the $s(n)$ flag if Eq. (5) is met.



EC: echo path change

Fig. 4. Simulation result for Scheme 2

4. Summary

This paper presents two DT detection schemes which, compared to other existing ones reviewed in [2], are reliable, robust, and relatively simple in terms of implementation. Both schemes have been verified in simulations and Scheme 2 has further been incorporated into acoustic and network echo cancellation products.

References

[1] Simon Haykin, *Adaptive Filter Theory*, 4th Edition, Prentice Hall, Sept. 2001.

[2] Thien-An Vu, Heping Ding, and Maritn Bouchard, "A survey of double-talk detection schemes for echo cancellation applications," *Acoustics Week in Canada* - Canadian Acoustical Association, Oct. 6-8, 2004, Ottawa, Canada.

[3] Heping Ding and Frank Lau, "Circuit and Method of Double Talk Detection for Use in Handsfree Telephony Terminals," U.S. Patent 6,049,606, issued April 11, 2000, assigned to Nortel Networks.

[4] Heping Ding, "Method of Distinguishing between Echo Path Change and Double Talk Conditions in an Echo Canceller," U.S. Patent 6,226,380, issued May 1, 2001, assigned to Nortel Networks.

PARALLEL RECOGNIZER ALGORITHM FOR AUTOMATIC SPEECH RECOGNITION

Akakpo AGBAGO, Caroline BARRIÈRE
Institute for Information Technology, National Research Council of Canada,
Pavillon Lucien Brault, 101 rue St-Jean-Bosco, Gatineau, Québec, J8Y 3G4, Canada
{Akakpo.Agbago, Caroline.Barriere}@nrc-cnrc.gc.ca

1. INTRODUCTION

Research in Automatic Speech Recognition (ASR) has been very intense in recent years with focus given to accuracy and speed issues. To achieve good accuracy, the employed techniques usually rely on heavy computations. Agbago and Barrière [2] earlier defined a Three-Stage Architecture (TSA) framework for ASR composed of (1) pre-processing stage, (2) phoneme recognition stage, and (3) natural language post-processor stage. Within that TSA framework, our present focus is to improve the speed of Stage 2 which looks specifically at the comparison of low level speech units. It is different from several systems that include HMM processes in this Stage (e.g. Shawn's [5]). We present a new algorithm called Parallel Recognizer that is 320 times faster than a standard Two-Level Dynamic Programming (TLDP) [3]. In comparison, working on speed at low-level, Nkagawa [4] got a reduction factor of 4 to 6 the time needed to compute local distances in the improved DP algorithm of Sakoe [6].

2. PARALLEL RECOGNIZER

A Knowledge Base of Reference Phonemes (KBRP) provides English speakers' phoneme models. To identify the phonemes which are part of a speech segment, our Parallel Recognizer algorithm uses a principle of best-fit in an open competition for all phonemes in the KBRP. It can also embed heuristics that might reduce the search space such as clustering KBRP into a lesser number of models.

2.1 Principle

Parallel Recognizer must take an input speech segment (hereafter T) and recognize it as a sequence of phonemes. These phonemes, already encoded into cepstrums, are selected from the reference base KBRP. In an attempt to reduce the search space during comparison task, heuristics are applied to pre-select some of the units (disqualify unvoiced facing voiced, etc.) that could match parts of segment T. The Parallel Recognizer's originality is that, every reference unit competes and scores a distance value with respect to any possible part of T. This sounds like processing a DP algorithm but it differs by the number of combinations to process that is normally far less than the binomial combinations employed by TLDP.

The algorithm opens a matching competition to fill *segments* of T starting from every frame of it (see section 2.2 for segment definition). It fixes the starting frame but gives the freedom to every competitor reference unit to find the optimal length of T it can best match from that frame (position). Consequently, the number of combination in the comparison is reduced from n^2 to almost n (plus the optimal length computation overhead) as shown in Figure 2.1 where the computation for every node of a matrix space is reduced to a vector space (every row).

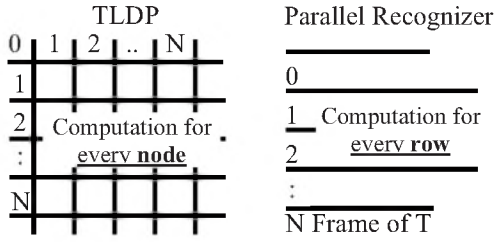


Figure 2.1: Computation combination comparison over the frames of the input utterance T.

The matching process generates a sorted pile for every frame position to the end of T. Using the piles as a lattice of segments we can form numerous chains of segments (hypotheses) by concatenating the best of any pile that comes in a consecutive order. The final result is the best hypothesis of these, based on conditions of our choice: best total chain distortion, fewer/larger number of composing units, etc. The hypotheses can be converted into strings of the ARPABET symbols of the reference speech units.

2.2 Segment Object definition and ordering Principle

Parallel Recognizer relies on an entity we call segment. It has a beginning and an ending position or a length but it encapsulates a third distortion parameter that represents the distance scored from the matching of this segment X in its entire length to a part of another segment Y starting from frame STARTPOS for a length LENGTH. So, we define a segment object (entity) as:
SegO=SegmentObject(X, Y, StartPos, Length, Distortion);

During our comparison process, reference units are successively replacing X and Y is the input segment T. As the resulting pile of segment objects needs sorting, a comparator rule should be defined upon them. The rule could be anything reasonable as how two segments should be ranked. For our experiments, we used the following:

- Order 1 (default): based on starting frame position of segments; used for forward chaining of hypotheses.
- Order 2: based on ending frame of segments to backtrack position to the beginning of a hypothesis.
- Order 3: based on distance to find the closest hypothesis.

The choice of a rule may influence accuracy but not the speed and since our focus is on the latter, we did not evaluate the impact of these different options on the former.

2.2 Strengths and weaknesses

The possibility to choose parameters on which hypotheses are sorted makes Parallel Recognizer very flexible. The segment pile technique *generates plenty and full length strings of phonemes* (no guess of length) as compared to TLDP that *outputs length from 1 unit to a given pre-guessed length L*. This is an advantage for an NLP post processing stage (Stage 3 of TSA) to complete the task. The use of heuristics to reduce search spaces can become weaknesses if they get too heavy in computations just like if naïve methods were used to perform the sorting process.

3. EXPERIMENTS AND RESULTS

In Agbago and Barrière [2], we have presented a Fast Two-Level Dynamic Programming (F-TLDP) approach which already was 5 to 20 times faster than standard TLDP. So, we tested Parallel Recognizer against F-TLDP using an input speech T ("she had your dark suit") and KBRP speech units (a teen phoneme models for every ARPABET symbol) that we derived from TIMIT. Only the execution time of the algorithms was considered excluding the overhead time to compute cepstrums features. Figure 3.1&2 show their contrast with respect to clustering KBRP and the length of the input T. Using the analysis of Figure 3.1, the test for Figure 3.2 is done in the case of no clustering (N=1) and the worst performance of Parallel Recognizer (N=80) compared to F-TLDP. Parallel Recognizer is 4 to 16 times faster than F-TLDP or an overall of 320 times faster than TLDP. The peak in the performance of Parallel Recognizer corresponds to the valley in the performance of F-TLDP [2] which meant better performance of F-TLDP.

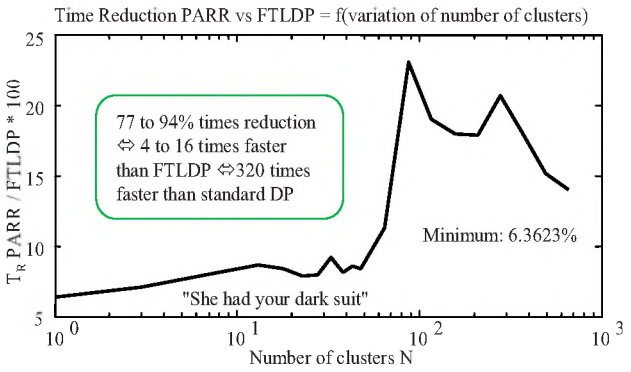


Figure 3.1: Parallel Recognizer vs. F-TLDP as a function of N the number of phoneme clusters in KBRP. PARR = Parallel Recognizer and T_R = time reduction

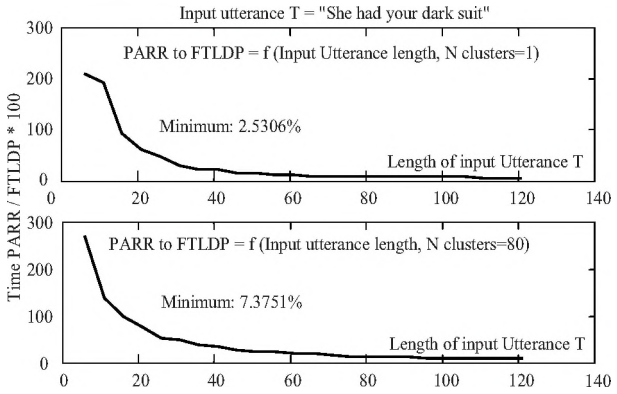


Figure 3.2: Comparison graph between FTLDP and Parallel Recognizer as a function of the length L of the input speech utterance T. PARR = Parallel Recognizer

4. CONCLUSION

With the concept of concurrency (parallel) matching of phonemes in KBRP to an input T using our defined segment entities, we succeeded to convert the matrix computation space of TLDP to a vector space (Figure 2.1). It corresponds to a significant compressing from a computation order of n^2 towards n. The result is 16 times speed increase of Parallel Recognizer over F-TLDP with the advantage of numerous hypotheses output for Stage 3 of the TSA. The speed result is independent of the segment ranking chosen but the choice may have an impact on accuracy.

Agbago has detailed elsewhere [1] that within the Stage 2 (presented here) of the overall Three-Stage Architecture framework, the accuracy of the results depends on an autonomous unit that evaluates the distortion between utterances (segments in this case). In this paper, focus has not been accuracy but speed improvement and it shows quite interesting results.

REFERENCES

[1] Akakpo Agbago, (May 2004) "Investigating Speed Issues In Acoustic-Phonetic Models For Continuous Speech Recognition", Master thesis at the University of Ottawa, SITE.

[2] Akakpo Agbago, Caroline Barriere, (2004) "Fast Two-Level-Dynamic-Programming Algorithm For Speech Recognition", IEEE 2004, May 17-21, Montreal, Canada.

[3] Lawrence R. and Biing-Hwang J., (1993) "Fundamentals of Speech Recognition", Prentice hall signal processing series, Englewood Cliffs New Jersey 07632, pp. 395

[4] Nakagawa S., (1982) "Continuous speech recognition methods by Constant Time Delay DP matching and O(n) DP matching", Trans. Committee Speech Research, ASJ, S82-17.

[5] Rafid A. S., Shawn M.H., Anand R.S. and Carl D.M., (2000), "Reducing computational complexity and response latency through the detection of contentless frames", IEEE-ASSP, Vol.6.

[6] Sakoe H., (1979), "Two-level DP-matching--A dynamic programming-based pattern matching algorithm for connected word recognition", IEEE-ASSP, vol. 27 no. 6, pp. 588 -595.

ON THE STOCHASTIC PROPERTIES OF THE NEURAL ENCODING MECHANISM OF SOUND INTENSITY

Liz C. Chang, Willy Wong
Edward S. Rogers Department of Electrical and Computer Engineering, University of Toronto, Ontario, Canada, M5S 3GS
Institute of Biomaterial and Biomedical Engineering, University of Toronto

1. INTRODUCTION

One of the simplest ways to model the activity in a neural spike train is to use a Poisson process. However, the conventional homogeneous Poisson process (HPP) is not compatible with the results collected from past studies. For example, the pulse-number distribution (PND) extracted from an HPP will follow a Poisson distribution with a mean-to-variance ratio of one. On the other hand, several studies have indicated that the ratio is not unity but is approximately equal to 2 across the dynamic range (e.g. Teich and Khanna, 1985). Additionally many features in the behaviour of real sensory neurons such as rate adaptation, rate-intensity dependence and a dead time in spike activity require that modifications be made to this model.

Based on these concerns, the HPP was modified and extended to give a more realistic stochastic model that expressed quantitatively the properties of auditory neurons. The predictions of the model with respect to the mean-to-variance ratio will be taken as an indication of whether the new model outperforms the conventional HPP-based model.

2. METHOD

In an HPP, the inter-event intervals $\mathbf{t}_{1,2,\ldots}$ which specify (in our case) the interval been spikes are governed by independent exponential random variables with a probability density function $f(t_n = x) = \lambda T e^{-\lambda T x}$ (Leon-Garcia, 1994). λ is a constant representing the spike count within a fixed time window **T**. The exponential distribution was used as a basis from which the new model was developed.

2.1 Firing Rate Modifications

We discarded the fixed spike rate λ and used in its place a rate function $\lambda(L, t)$, where **L** is sound intensity level and **t** is stimulus duration. This function was constructed based on the measurements of rate adaptation (Litvak, *et al.*, 2003) and intensity dependence (Yates, *et al.*, 2000; Smith, 1988). Please see Figure 1. With a non-constant firing rate, the process we have described is known as a non-homogeneous Poisson process (NHPP).

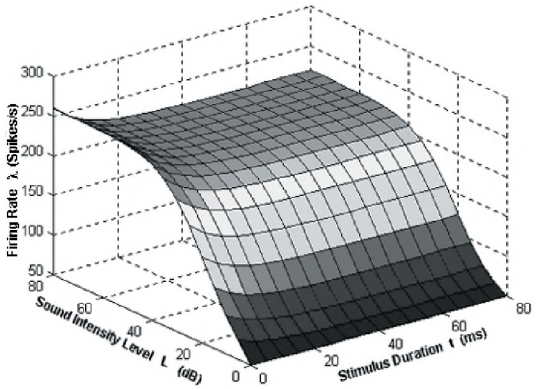


Figure 1. The idealised firing rate behaviour with respect to stimulus duration and sound intensity level for a peripheral neuron.

2.2 Dead Time Modifications

A fixed value τ was used to denote the dead time during which the neuron cannot be activated further. In our model, the inter-spike interval was set equal to the sum of the time value generated by the firing probability function and the dead time τ. This process is known as a dead-time-modified Poisson Process (DTMPP).

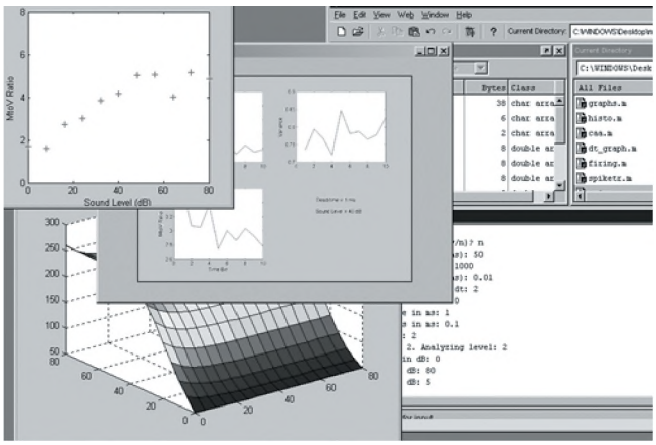


Figure 2. A screen shot of the program in MATLAB.

We implemented our stochastic model within MATLAB. To study the effects of the different components on the mean-variance ratio, a simple command program was written in MATLAB to control the different parameter values (Figure 2). A flowchart illustrating the difference between the various simulations is shown in Figure 3. Each

trial spans over 50 ms and inter-spike intervals were generated with the appropriate firing probability equation. The same trial was repeated 1000 times to collect sufficient data for parameter estimation and PND generation.

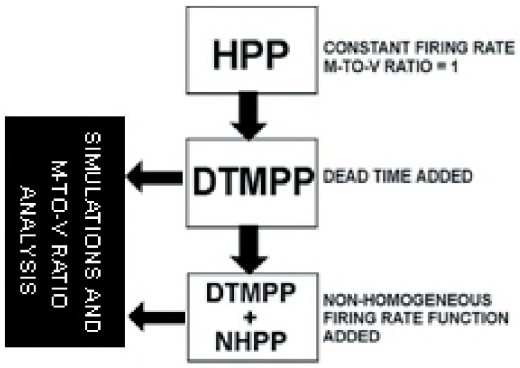


Figure 3. A diagram illustrating the process of model construction and the subsequent simulations and statistical analysis.

3. RESULTS

The results are presented in the form of plots of mean-to-variance ratio against different parameters.

3.1 Dead Time Effects under Constant Firing Rate

We considered the effect of dead time with firing rate as a parameter (Figure 4). Under a fixed firing rate, the mean-variance ratio was found to be a monotonic increasing function of dead time. Furthermore, the mean-to-variance ratio grows with increasing values of firing rate. This dependence becomes stronger as the dead time is increased.

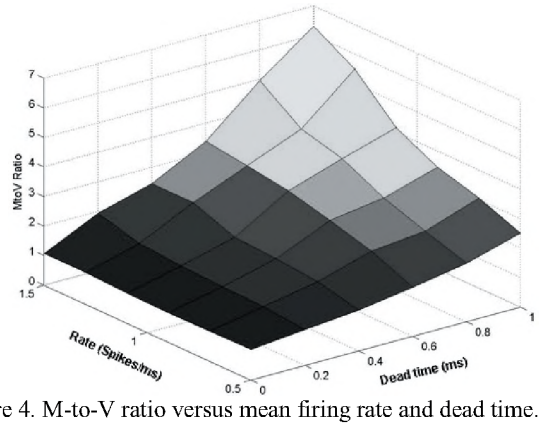


Figure 4. M-to-V ratio versus mean firing rate and dead time.

3.2 Dead Time with Time-varying Firing Rate

The full rate function was used here. Figure 5 shows that the mean-to-variance ratio behaves in a similar manner as an HPP model with dead time. The ratio remains around one without dead time, and is driven above one when a dead time is present. However, the growth is not as rapid as when plotted against the average firing rate. In fact, under a moderate dead time (0.5 ms), the ratio demonstrates little dependence on sound intensity level and appears to be around 2 in agreement with results from Teich and Khanna (1985) .

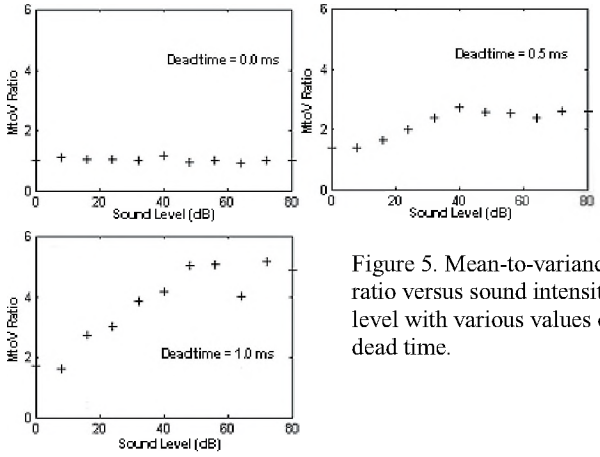


Figure 5. Mean-to-variance ratio versus sound intensity level with various values of dead time.

4. DISCUSSION

Clearly dead time is the dominant factor that drives the mean-to-variance ratio above unity. In fact, in a DTMPP the mean-variance ratio can be shown to equal approximately $(1-\lambda\tau)^{-2}$ (Cantor and Teich, 1975). With a non-zero dead time, the M-V ratio is positively correlated with spike rate. However using a time-varying rate function, the ratio does not grow significantly when plotted against sound intensity. This is most certainly due to the fact that since the growth rate in mean activity is diminished with adaptation, the effects of sound level on the M-V ratio are minimized.

Using a moderate dead time value, we have found that the values obtained for the mean-to-variance ratio are compatible with existing experimental results. However, further investigation will be required to thoroughly evaluate this model with other intensity-coding properties of auditory neurons.

REFERENCES

Cantor, B. I., and Teich, M. C. (**1975**). "Dead-time-corrected photocounting distributions for laser radiation," J. Opt. Soc. Am. **65**, 786-791.

Leon-Garcia, A. (**1994**). *Probability and Random Processes for Electrical Engineering* (Addison-Wesley, New York).

Litvak, L. M., Smith, Z. M., Delgutte, B., and Eddington, D. K. (**2003**). "Desynchronization of electrically evoked auditory-nerve activity by high-frequency pulse trains of long duration," J. Acoust. Soc. Am. **114**, 2066-2078.

Smith, R. L. (**1988**). "Encoding of Sound Intensity by Auditory Neurons," in *Auditory Function: Neurobiological Bases of Hearing*, edited by G. M. Edelman, et al. (Wiley, New York), pp. 243-274

Teich, M. C., and Khanna, S. M. (**1985**). "Pulse-number distribution for the neural spike train in the cat's auditory nerve," J. Acoust. Soc. Am. **77**, 1110-1128.

Yates, G. K., Manley, G. A., and Köppl, C. (**2000**). "Rate-intensity functions in the emu auditory nerve," J. Acoust. Soc. Am. **107**, 2143-2154.

EVALUATION OF OBJECTIVE MEASURES OF LOUDNESS

Gilbert A. Soulodre
Communications Research Centre, Ottawa, Ontario, Canada, K2H 8S2 gilbert.soulodre@crc.ca

1. INTRODUCTION

In many applications it is desirable to measure and control the subjective loudness of typical program material. Examples of this include television and broadcast applications where the nature and content of the audio material changes frequently. In these applications the audio content can continually switch between music and speech, or some combination of the two. The program material can also include sound effects and environmental sounds. These changes in the content of the program material can result in dramatic changes in subjective loudness. Moreover, various forms of dynamic range processing are frequently applied to the signals, which can have a significant effect on the perceived loudness of the signal.

There is currently an effort within broadcast standards organizations (ITU-R, NABA) to identify or develop an objective loudness measure (a loudness meter) that can be used by broadcasters to equalize the perceived loudness of their content. The ultimate goal is to have more consistent broadcast levels across program materials and broadcast stations. The matter is also of great significance to the music industry where dynamic range processing is commonly used to maximize the perceived loudness of a recording. In the present study the performance of various objective loudness measures is evaluated.

2. OVERVIEW OF TESTS

In the first part of the study a series of subjective tests (loudness-matching experiments) were conducted at 5 test sites around the world in order to create a database for evaluating the objective loudness measures. A total of 97 subjects listened to a broad range of typical program material and adjusted the level (in 0.25 dB steps) of each test item until its loudness matched that of a reference signal. The reference signal consisted of English female speech with no background sounds, and was reproduced at a level of 60 dBA.

The program material used in the tests was taken from actual television and radio broadcasts from various locations around the world. The 98 sequences included music, television and movie dramas, sporting events, news broadcasts, sound effects, and advertisements. Included in the sequences were speech segments in several languages.

The test setup employed a single loudspeaker placed directly in front of the listener. The subject could switch instantly between the reference signal and the test items while matching their levels.

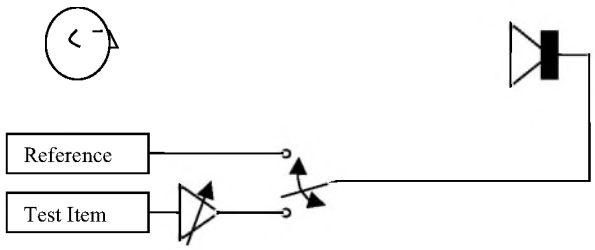


Figure 1: Subjective test setup.

Ten commercially developed loudness meters (labeled A-K in order to hide their identities) were submitted by proponents to be evaluated in their ability to predict the results of the subjective database. In addition, the author contributed two additional basic loudness measures to serve as a performance baseline. One was a simple *Leq* measure, while the other was a frequency-weighted *Leq* using a "*Revised Low-frequency B-weighting*" (referred to as *Leq*(RLB)) [1]. The individual audio sequences of the subjective database were processed through each of the loudness meters and the measured loudness estimates were recorded. These objective readings were then compared against the subjective loudness ratings using a variety of metrics to assess each meter's performance.

In order to assess the performance of the various loudness meters objectively, it was necessary to establish a set of suitable performance metrics that would effectively reflect the requirements of a practical loudness meter. In general, we want the meter to match the relative levels of the database as closely as possible. However, small errors in the meter's predictions are probably acceptable since listeners are unlikely to detect (or be annoyed by) small changes in loudness. Based on previous findings loudness errors of less than 1.25 dB are expected to go largely unnoticed [1]. Therefore, a meter could be considered to be ideal if all of its errors were less than 1.25 dB. Conversely, even a single error beyond some limit (say 10 dB?) could be considered entirely unacceptable, thus disqualifying a given meter from further consideration.

The metrics included, correlation (R), Spearman's rho, the root-mean squared error (RMSE), the maximum absolute

error (MAE), and the average absolute error (AAE).

3. RESULTS AND DISCUSSION

The results of the evaluation are summarized in Table 1 where the performance of the 12 meters is shown in terms of the various performance metrics. The numbers shown in square brackets indicate the relative ranking of the loudness meters for each metric.

Table 1: Performance of the loudness meters for the nine performance metrics. Values in brackets [ ] indicate the relative ranking for each metric.

	R	*Spearman rho*	*RMSE* (dB)	*MAE* (dB)	*AAE* (dB)
A	0.944[10]	0.916 [10]	2.37[11]	6.37[10]	1.88[11]
B [*Leq*(A)]	0.929[11]	0.889 [11]	2.19[10]	6.39[11]	1.77[10]
C	0.955 [9]	0.952 [8]	1.75 [9]	5.76 [9]	1.35 [9]
D	0.976 [3]	0.958 [5]	1.31 [3]	4.70 [5]	0.99 [3]
F	0.965 [8]	0.951 [9]	1.55 [8]	3.61 [1]	1.28 [8]
G [*Leq*(B)]	0.972 [5]	0.952 [7]	1.37 [6]	4.19 [4]	1.07 [5]
H	0.848[12]	0.841[12]	3.33[12]	6.89[12]	2.90[12]
I	0.972 [5]	0.960 [3]	1.36 [5]	4.80 [6]	1.09 [6]
J	0.968 [7]	0.955 [6]	1.51 [7]	4.97 [7]	1.17 [7]
K	0.975 [4]	0.958 [4]	1.33 [4]	5.13 [8]	0.99 [3]
Leq	0.979 [2]	0.971 [1]	1.26 [2]	4.33 [3]	0.93 [2]
Leq(RLB)	0.982 [1]	0.971 [1]	1.15 [1]	3.62 [2]	0.87 [1]

It can be seen that the basic loudness meter *Leq*(RLB) is ranked as the best meter for all of the metrics except the maximum absolute error (MAE). For this metric it is ranked second. However, it can be considered to be effectively equivalent to the first ranked meter for this measure, since its error is only 0.01dB larger. According to the various performance metrics the second best meter is a simple *Leq* measure. Therefore, for the present study, none of the commercially developed loudness meters submitted by the proponents performed as well as *Leq* or *Leq*(RLB).

This finding is quite remarkable given that most of the loudness meters included some form of complex perceptual model. The accuracy of the simple measures is quite impressive. Note that the worst case error for *Leq*(RLB), corresponding to the MAE metric, is only 3.62 dB. It was revealed that Meters B and G were *Leq*(A) and *Leq*(B) respectively.

It is of interest to examine plots of the worst and best performing loudness meters (Meter H and *Leq*(RLB) respectively). The data are plotted in terms of the gain that needs to be applied to a given audio signal in order to match its level to the reference signal. The open circles represent speech-based audio sequences, while the stars are non-speech-based sequences. It should be noted that a perfect objective meter would result in all data points falling on the diagonal line having a slope of 1 and passing through the origin (as shown in the figures). Any data point falling above the diagonal line indicates that the meter overestimated the gain required to match the loudness of that audio sequence to the reference signal. That is, the meter underestimated the perceived loudness of that particular audio sequence. The plots of Figure 2 and 3 clearly demonstrate the difference in performance of these two loudness meters.

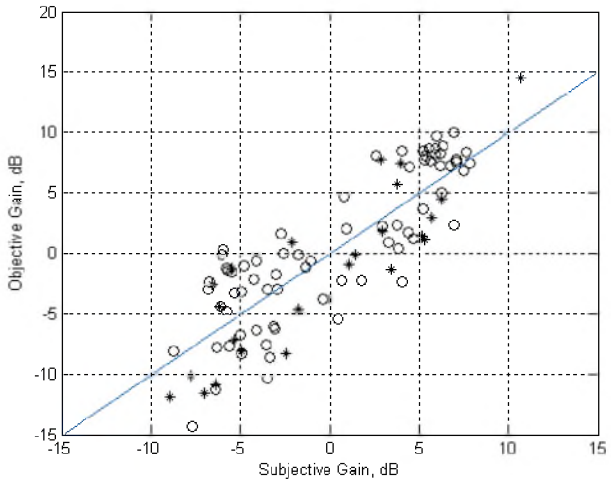


Figure 2: Meter H.

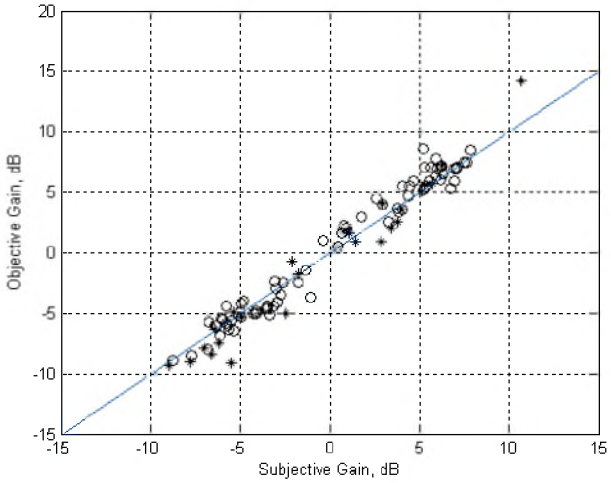


Figure 3: *Leq*(RLB).

The results of the present study indicate that for typical broadcast material, a simple energy-based loudness measure is more robust than more complex measures that may include detailed perceptual models. This finding is supported by the fact that one European broadcaster (TV2/Denmark) has been successfully using high-pass filtered RMS for many years as a measure of loudness.

REFERENCES

[1] Soulodre, G. A. (2004) Evaluation of Objective Loudness Meters, AES 116th Convention, Berlin (preprint 6161).

Audio Processing in Police Investigations

David Luknowsky[1] and Jeff Boyczuk[1]
[1]Audio and Video Analysis Section, Royal Canadian Mounted Police, Ottawa, Ontario

1. INTRODUCTION

The Audio and Video Analysis Section of the Royal Canadian Mounted Police (RCMP) tries to reduce interferences and improve the intelligibility of audio recordings. The primary goal is to be able to clearly understand what was being said. The final result often still contains considerable noise because further noise reduction would also reduce signal information that is important for voice intelligibility. Common sources of the audio recordings are from hidden devices, interview rooms, hand held recorders, wire taps and 911 calls. Many different types of interference are encountered and each case is unique.

Some of the more common types of interference include tones, hum, motor noises, and music. These interferences may be reduced with notches which may be set manually or automatically, comb filters, spectral inverse filters, and adaptive filters. Common problems with voice which have to be corrected are low volume of one or more voices, reverberation, and clipping.

2. METHOD

One of the most common types of interference, and one of the simplest to reduce, is tones and hum. Hum typically comes from power lines, fluorescent lights or other electrical sources. It is seen in the spectrum as spikes at 60 Hz and harmonics thereof (Figure 1).

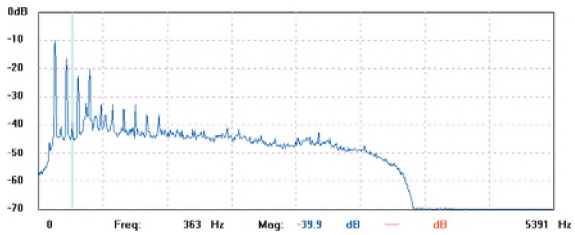


Figure 1. Average voice spectrum with strong power line hum.

60 Hz hum can easily be removed with a comb filter which applies notches every 60 Hz up to a maximum amplitude (Figure 2). The amplitude of the harmonics often falls below the average signal beyond 360 Hz so the maximum of the comb filter may not need to go much beyond this.

A series of individual tones at various, unequally spaced frequencies or tones whose frequency varies over time are also common tonal interference

Figure 2. Spectrum from figure 1 with hum reduced using a comb filter.

problems. For this type of interference, we would probably use an adaptive filter. There are two algorithms which are likely to be used for cases like this: 1) an algorithm which analyzes the amplitude over time of all frequencies in small, discrete bins and identifies frequencies whose amplitude is not changing over time, within certain defined parameters. 2) an adaptive predictive de-convolution filter which de-correlates the input signal and removes long term correlated components [1]. Essentially the filter delays the input signal, attempts to predict the next part of the signal using a least-mean-square estimator, and then subtracts the predicted signal thereby removing predictable components such as hum which remain consistent over time. Adaptive filters such as these can also be useful for reducing other types of predictable interferences such as motor noise from engines, fans, or air conditioners, and with the appropriate parameters can be set to reduce music since it is generally much more predictable than voice.

If tones are strong yet stationary and unevenly spaced (so that a comb filter would not work), a spectral inverse filter may be used. The average spectrum is analyzed and divided up into nearly a thousand equally spaced frequency bands. Amplitude adjustment can then be applied to each of these bands to produce a flat or shaped average spectrum as if one were automatically setting a very high resolution graphic equalizer. While a spectral inverse filter can be used to reduce tonal noises, it is more likely to be used to shape the average spectrum to increase the clarity of the voices. This is especially useful for analogue recordings with poor high frequency response such as those made at a very low record

speed. Such recordings generally have too much bass and sound muffled. A spectral inverse filter with shaping will provide significant amplitude boost to higher voice frequencies and then shape the spectrum to a typical averaged voice spectrum[2]. While this may boost high frequency noise, it can also greatly reduce the muffled sound and increase the crispness, clarity, and intelligibility of the recorded voices.

In addition to problems with interferences, there are often problems with the level of one or more voices. On recordings of interviews, it is common to have the interviewee's voice at a very low level even when the investigator's voice is loud and clear. Similar problems are seen with body-worn recorders where the voice of the person wearing the recorder is clear but the other voices are at a very low level. Sometimes playing the audio through an automatic gain control, which attempts to bring all sounds to about the same record level, is sufficient. When the level adjustment needs to be dramatic or sudden, a digital audio editor is used. This allows very accurate selection of the section that needs to have its level changed and very quick and precise changes in level.

Audio levels which are too high and which result in distortion or clipping are also seen. In some cases where the clipping is not too bad, the waveform can be reconstructed to eliminate the clipping at that point. For example, Figure 3 shows a stereo audio waveform with clipping on the top and just below shows the same waveform reconstructed to remove the clipping effect by interpolating the waveform based on the surrounding data points.

The quality of the recordings we deal with is often quite poor. Sometimes the quality of the media itself is marginal. Before processing, we want to get the best possible signal from the media. The first step in processing forensic audio is to make sure we are dealing with the original recording. We do not want to work on a degraded copy. It is then necessary to ensure the quality of the media. For tape, this may mean cleaning the tape and/or placing it in a new cassette if there is damage. Good playback equipment with clean heads that can be azimuth adjusted for the best signal on playback is important for tape recordings. Such steps to get the best available signal from the media can sometimes make a big difference in the end product when dealing with forensic audio.

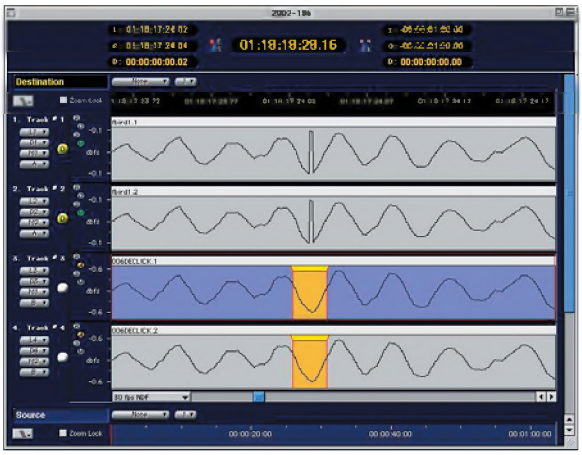


Figure 3. Audio waveform with clipping (top) and corrected (bottom). Total length of audio shown is 9.6 ms.

3. DISCUSSION

The exact combination of filters and filter parameters which provides the best intelligibility may be different for each recording. The analyst is always tasked with determining which processing to use in order to achieve the highest level of intelligibility. This is not the same as producing a nice sounding or noise free recording. Higher frequencies which often make voice recordings sound noisy need to be included (and sometimes even amplified) because voice information crucial to intelligibility would otherwise be lost. Being able to understand what was said is always the most important thing.

REFERENCES

[1] Paul, J.E., "Adaptive Digital Techniques For Audio Noise Cancellation," *IEEE Signal Processing Magazine,* Vol. 1, No. 4, pp. 2-7, 1979.

[2] J. L. Flanagan, Speech Analysis Synthesis and Perception, New York: Springer-Verlag, 1972, p. 163.

ASPECTS OF INVERSE FILTERING FOR LOUDSPEAKERS

Scott G. Norcross
Advanced Audio Systems, Communications Research Centre, 3701 Carling Ave., P.O. Box 11490 Station H, Ottawa, Ontario, K2H 8S2 e-mail: scott.norcross@crc.ca

1. INTRODUCTION

Inverse filtering has been proposed for numerous applications in audio and telecommunications such as loudspeaker equalization and room deconvolution. Its attraction is that it can potentially "undo" a system and correct both the magnitude and phase responses [1][2]. Two methods are compared in this paper: a time-domain least-squares approach and a frequency-domain deconvolution method. Formal subjective tests conducted in previous studies [5][6][7] have shown that the inverse filter can actually degrade the audio quality by producing artifacts or distortions. These artifacts, which include pre-echo and timbre changes, can in some cases result in subjectively poorer performance than if no inverse filter was used. The severity of the degradation depends on both the inversion method and the system that is being corrected. This paper will review the results of the formal subjective tests evaluating the two methods as well as some strategies, such as regularization and smoothing, to remove or control the severity of audible artifacts.

2. INVERSE FILTERING

The inverse filtering process is based around the concept of linear filtering. Assuming that a flat frequency response is desired, we have the following equation to solve

$$\delta(n) = c(n) \otimes h(n) \qquad (1)$$

where $\delta(n)$ is the Kronecker delta function or unit impulse function and *c(n)* is the impulse response (IR) and *h(n)* is the inverse filter.

A least-squares time-domain approach can be expressed in matrix form, as in [2] and a solution for the inverse filter *h(n)* is given by,

$$h(n) = (\mathbf{C}^T\mathbf{C})^{-1} \bullet \mathbf{C}^T\mathbf{a}_m \qquad (2)$$

where $\mathbf{C}$ is the convolution matrix of *c(n)* and $\mathbf{a}_m$ is a modeling delay vector.

Although fast algorithms exist for calculating the time-domain solution given by Equation (2), a fast frequency-domain deconvolution method can be used to speed up this computation. This approach is based on the fact that a time-domain convolution becomes a multiplication in the frequency domain and can be written as

$$H(k) = \frac{D(k)}{C(k)} \qquad (3)$$

where *H(k)* and *C(k)* are the discrete Fourier transform (DFT) of *h(n)* and *c(n)* respectively, and *D(k)* is the DFT of the desired signal, usually a delayed pulse.

A potential problem with inverse filtering is when the denominator in equation (3) is zero or very small. This will result in *H(k)* having an excessive boost. Kirkeby et al. [2] and Craven et al.[3] have used regularization to limit this effect. Kirkeby's implementation can be expressed as,

$$H(k) = \frac{D(k)C^*(k)}{C(k)C^*(k) + \beta B(k)B^*(k)} \qquad (4)$$

where *B(k)* is the regularization term and β is a scaling factor controlling the amount of regularization.

Another approach to overcome the problem of small values of *C(k)* is the use of smoothing. Hatziantonious and Mourjopoulos [4], have suggested complex smoothing given as,

$$C_{cs}(k) = \sum_{i=0}^{N-1} C((k-i) \bmod N) \bullet W_{sm}(m,i) \qquad (5)$$

where W_{sm} is the smoothing window.

3. MEASUREMENTS AND SUBJECTIVE TESTS

The IRs from two different loudspeakers were measured on- and off-axis (45°) in an anechoic environment. The IRs were sampled at 44.1kHz and were 1024 samples in length. Test files were created from a mono recording of a castanets passage. The test strategy was to filter the audio source signal with one of the measured IRs and then process these filtered audio signals with an inverse filter to correct for the loudspeaker's response. Ideally, if the inverse filter

were perfect, the results of this process should yield an audio file identical to the original audio source file.

To evaluate the subjective performance of the different inverse filtering methods (correction methods), double-blind subjective tests were conducted using the MUSHRA method (Recommendation ITU-R 1534) test over headphones.

4. RESULTS

Figure 1 shows mean subjective grades for the five loudspeaker conditions versus correction method. Six frequency-domain inverse filters of increasing length (2k to 64k) are shown along with a 2k time-domain inverse filter. Also included is the case where no correction (labeled 'Filter') was made to the filtered audio signal.

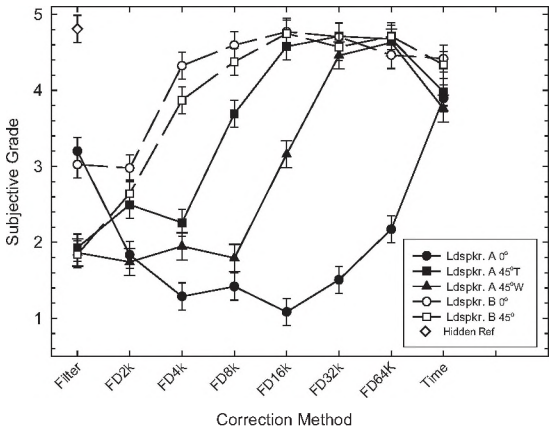


Figure 1. Test results for the five IRs showing the mean subjective grade versus correction method. 'FD' refers to the frequency-domain method and *X*k refers to the length of the inverse filter. Also included are the uncorrected condition ('Filter') and the time-domain inverse (2k in length) labeled with 'Time'.

It can be seen that the time-domain inverse filters scored higher than the frequency-domain filters of similar lengths. The time-domain filters also scorred very similarly for all five loudspeakers whereas with the frequency-domain method there was a larger spread between them. Most of the frequency-domain filters subjective performance increased when the length was increased, except for one. This might indicate the presence of time-aliasing with the frequency-domain deconvolution method.

Three forms of regularization, one frequency-independent and two frequency-dependent, were used to try to improve the subjective performance of the frequency-domain method. One frequency-dependent regularization method consisted of regularizing the extreme low and high frequencies with a value of 1 (0 elsewhere), and the other method was based on the 1/3 octave spectrum of the impulse response. All the regularization methods did improve the performance over the case with no regularization, but the amount required to achieve the performance increase was dependent of the IR.

Third-octave complex smoothing of the IR prior to calculating the inverse was also explored and showed some improvement in the subjective performance. The smoothing did not tend to broaden the corrected IRs as was the case when regularization was added.

5. SUMMARY

Two methods for calculating the inverse filter of a loudspeaker's IR were examined. Formal subjective tests have shown that the inverse filter can produce audible artifacts. The frequency-domain method is not as robust as the time-domain method. Regularization and complex smoothing can help improve the subjective performance of the inverse filter. The amount of regularization required is dependent on the IR that is being inverted, thus requiring hand-tuning to optimize the subjective performance. Complex smoothing of the impulse response with a third-octave smoother can also be used to improve the subjective performance of the inverse filter. Complex smoothing did not broaden the main pulse of the corrected IR as the regularization tended to do.

6. REFERENCES

[1] Mourjopoulos, John, N. (1994). "Digital Equalization of Room Acoustics", *J. Audio Eng.*, vol. 42, pp. 884-900.

[2] Kirkeby, Ole and Nelson, Philip A., "Digital Filter Design for Inverseion Problems in Sound Reproduction", *J. Audio Eng.*, vol. 447, pp. 583-595.

[3] Craven, Peter G. and Gerzon, Michael A., "Practical Adaptive Room and Loudspeaker Equaliser for Hi-Fi Use", AES UK DSP Conference (Sept 1992).

[4] Hatziantoniou, Pantagiotis D., Mourjopoulos, John N.,"Generalized Fractional-Octave Smoothing of Audio and Acoustic Responses", *J. Audio Eng. Soc.*, vol. 48, pp. 259-280 (April 2000).

[5] S. G. Norcross, G. A. Soulodre and M. C. Lavoie, "Evaluation of Inverse Filtering Techniques for Room/Speaker Equalization", 113th AES Convention, LA, 2002, Preprint 5662

[6] S. G. Norcross, G. A. Soulodre and M. C. Lavoie, "Subjective Effects of Regularization on Inverse Filtering", 114th AES Convention, Amsterdam, 2003, Preprint

[7] S. G. Norcross, G. A. Soulodre and M. C. Lavoie, "Further Investigations of Inverse Filtering", 115th AES Convention, New York, 2003, Preprint

[8] Louis D. Fielder, "Analysis of Traditional and Reverberation-Reducing Methods of Room Equalization", *J. Audio Eng. Soc.*, vol. 51, pp.3-26 (January/February 2003).

EXTENSION OF THE AGGREGATE BEAMFORMER TO FILTER-AND-SUM BEAMFORMING

David I. Havelock
National Research Council, 1200 Montreal Rd., Ottawa, ON, K1A 0R6, david.havelock@nrc-cnrc.gc.ca

1. INTRODUCTION

Beamforming combines signals from an array of sensors to obtain a directional response. In basic delay and sum (D&S) beamforming, the signal from each sensor is delayed to contribute constructively to the summed output. Delays are determined by the propagation time across the array. Sources that are off-beam (not in the desired direction) are partially cancelled by incoherent summation, whereas on-beam sources (in the 'steering' direction) are enhanced by coherent summation. The directional response can be adjusted by applying weights to sensor inputs. Moreover, in a more general approach referred to as filter-and-sum (F&S) beamforming, sensor inputs can be filtered to apply a frequency dependent weight and delay (phase), allowing more general optimization.

The aggregate beamformer differs from conventional beamforming in that it converts off-beam signals to noise that can then be filtered. The array is sampled, one sensor at a time, at a relatively high sampling rate. For each sample, the sensor is selected randomly. The sampled data are time-aligned (as in conventional beamforming), low-pass filtered, and then decimated to the Nyquist frequency of the desired bandwidth. The time-aligned data faithfully reconstructs on-beam signal components but reduces off-beam components to random noise.

Aspects of the aggregate beamformer are described in articles by the author [1-4] in the context of delay-and-sum beamforming. In this article, the technique is extended to filter-and-sum beamforming (as suggested in [3]) through the introduction of a 'stochastic filter'. The stochastic filter, similar to the aggregate beamformer, reduces out-of-band signals to random noise while preserving in-band signals.

2. D&S AGGREGATE BEAMFORMING

Conventional D&S beamforming for an array of M sensors, each with a sampled input signal $x_m(n)$, has output

$$B_{\mathrm{C}}(n) = \sum_{m=1}^{M} w_m x_m(n - \tau'_m), \tag{1}$$

where w_m are the 'shading' weights and τ'_m are the beamformer sample delays. The corresponding formulation of the aggregate beamformer output is

$$B_{\mathrm{A}}(n) = \sum h^{\mathrm{d}}(m) x_{s(m')}(m' - \tau_{\sigma(m')}), \tag{2}$$

where we have defined the implicit summation index $m' = (nK_{\mathrm{os}}\text{-}m)$ for brevity. For the aggregate beamformer, $h^{\mathrm{d}}(m)$ are decimation filter coefficients, $K_{\mathrm{os}} \geq 1$is a decimation factor, $t_m = t'_m/K_{\mathrm{os}}$ are sample delays and σ(n) is a random sequence of sensor indices that are independently and identically distributed as

$$\Pr\{\sigma(n) = m\} = w_m. \tag{3}$$

($\Pr\{\cdot\}$ indicates probability.)

Array weights w_m are applied by adjusting the sampling probabilities; no scaling of data is necessary. The residual noise level is controlled by adjusting the over-sampling factor K_{os} and can also be reduced by noise shaping [4].

3. STOCHASTIC FILTERING

The output $y_{\mathrm{c}}(n)$ of a conventional FIR filter with impulse response $\{h(m)\}$ $(0 \leq m \leq M)$ is

$$y_{\mathrm{c}}(n) = \sum_{m} h(m) x(n - m). \tag{4}$$

We define the output y_{s}(n) of the corresponding stochastic filter as

$$y_{\mathrm{s}}(n) = \varepsilon_{\rho(n)} x(n - \rho(n)), \tag{5}$$

where ρ is a random sequence of sample indices that are independently and identically distributed with

$$\Pr\{\rho(n) = m\} = |h(m)| \quad \text{and} \quad \varepsilon_m = \frac{h(m)}{|h(m)|}. \tag{6}$$

It is assumed that the FIR coefficients satisfy $\sum |h(m)| = 1$. The conventional and stochastic FIR filters are equivalent in the mean (both in magnitude and phase);

$$\langle y_s(n) \rangle = y(n). \tag{7}$$

where $\langle\cdot\rangle$ indicates expected value. Both the conventional and stochastic filters are linear but the stochastic filter is only time invariant in the mean. We define the residual noise of the stochastic filter as

$$\eta(n) = y_s(n) - y_c(n). \tag{8}$$

It can be shown that η has zero mean and is spectrally white irregardless of the input signal. If the input signal is stationary then the total output power is equal to the input signal power. The SNR at the output depends upon how much of the input signal is in-band but is bounded above by

$$\frac{\sum h_m^{\,2}}{1-\sum h_m^{\,2}}. \tag{9}$$

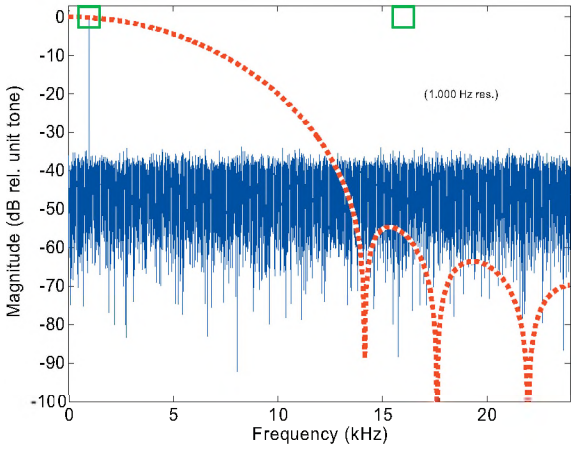


Fig. 1 *Stochastic filter output spectrum. The input tones (large squares) are 1 kHz and 16 kHz. The dashed curve is the filter magnitude response.*

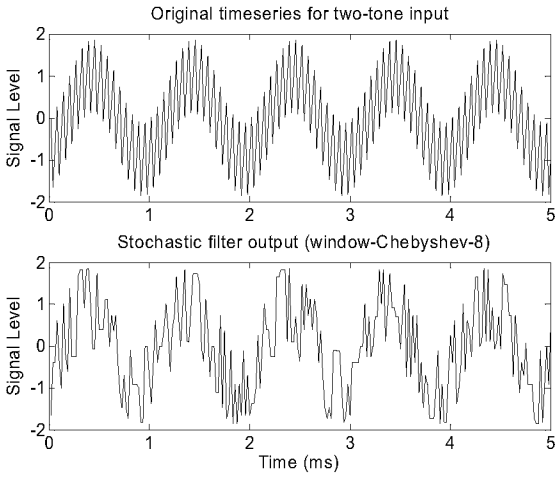


Fig. 2 *Stochastic filter input and output time series.*

Figure 1 shows the output magnitude spectrum of a stochastic filter for an input consisting of two equal-amplitude tones. The filter is an 8th order Chebyshev window filter, which has no negative coefficients. The dotted curve is the frequency response of the conventional FIR filter. The large squares indicate the components of the input. The lower frequency component is passed by the stochastic filter while the higher frequency component is converted to white noise.

The input and output time series of the example stochastic filter are shown in Fig. 2.

4. F&S Aggregate Beamformer

The filter-and-sum aggregate beamformer output B_{AFS} is derived from the stochastic filter by combining the random indices defined in Eqs. (3) and (6),

$$B_{\mathrm{AFS}}(n) = \sum_m h^{\mathrm{d}}(m)\, x_{\sigma(m')}\left(m' - \tau_{\sigma(m')} - \rho(m')\right). \tag{10}$$

We have again used the implicit index $m' = (nK_{\mathrm{os}}\text{-}m)$ in the summation. Recall that $h^{\mathrm{d}}(m)$ are decimation filter coefficients and that the stochastic filter coefficients are incorporated through $\rho(m')$.

Unlike conventional beamforming, extending D&S aggregate beamforming to F&S does not incur additional computation. The residual noise level, however, may be significantly increased.

5. CONCLUSION

The delay and sum aggregate beamformer can be extended to filter and sum beamforming. The conventional and the aggregate F&S beamformers are equivalent in the mean. The residual noise in the F&S beamformer may be significantly greater than the D&S beamformer.

REFERENCES

[1] D.I Havelock, "The aggregate beamformer," Canadian Acoustics 30(3)104, Sept. 2002.

[2] D.I Havelock, "Sensor array beamforming using random channel sampling: The aggregate beamformer," JASA 114(4)1997-2006, October 2003.

[3] D.I Havelock, "Beamforming arrays with many sensors: the aggregate beamformer," Paper Th4.D.4, 18th Int'l Congress on Acous., Kyoto Japan, April 4-9, 2004.

[4] D.I Havelock, "Residual noise in the aggregate beamformer," 2003 IEEE WASPAA, New Paltz NY, Oct. 19-22, 2003.

ACOUSTIC PULSE PROPAGATION THROUGH STABLY STRATIFIED LOWER ATMOSPHERE

Igor Chunchuzov[1], and Sergey Kulichkov[2]

[1]Obukhov Institute of Atmospheric Physics, 3 Pyzhevsky Per., 119017 Moscow, Russia
Corresponding Address: 2 Putman Ave, Ottawa, Ontario, K1M 1Y9, Canada, oksana@achilles.net
[2]Obukhov Institute of Atmospheric Physics, 3 Pyzhevsky Per., 119017 Moscow, Russia, snk@ifaran.ru

1. INTRODUCTION

There are several problems in atmospheric acoustics, which require taking into account the influence of both the mean stratification and fluctuations of wind speed and temperature on acoustic signal propagation through atmosphere. These problems include a location of various acoustic sources in the atmosphere, a prediction of sound levels from pulse and noise sources, and acoustic remote sensing of the atmosphere. To solve them one needs to parameterize the statistical characteristics of wind speed and temperature fluctuations for calculating a statistics of the parameters of acoustic signals such as travel time, duration and arrival angles.

It is necessary to note that the parameterization of turbulence statistics in stably stratified atmospheric boundary layer (ABL) is still a problem, since these characteristics at high Richardson numbers can not be described by means of the existing similarity theories, most of which are based on the assumption about the turbulence as being locally homogeneous and stationary. It is recognized now that different types of instabilities of internal gravity waves (IGWs) may be a source of a small-scale turbulence and meso-scale eddy structures (such as "cat eyes", "filaments", "banks" and others) in stable ABL (Gossard and Hooke, 1975). Such turbulence coexists and continuously interacts with the IGWs, so the statistical properties of the IGWs affect the statistics of turbulence in stably stratified ABL.

The IGWs itself induce meso-scale wind speed and temperature fluctuations in stable ABL with the periods from a few minutes to a few hours. Their horizontal scales are ranging from hundred metres to a few km. These fluctuations cause fluctuations of the parameters of acoustic signals propagating through ABL, such as travel time, amplitude, duration, and the angles of arrival of the acoustic signals. To obtain statistics of the parameters of acoustic signals one needs to know a form of the frequency-wavenumber spectrum of the wind speed and temperature fluctuations induced by IGWs. Such a need motivated our study of the effects of meso-scale wind speed fluctuations associated with IGWs on the statistical characteristics of the parameters of the acoustic pulses in stable ABL. Below we describe experimental and theoretical results of this study.

2. MEASUREMENT OF ACOUSTIC TRAVEL TIME FLUCTUATIONS.

To measure meso-scale wind speed and temperature fluctuations we used an acoustic pulse sounding method. This method is some kind of acoustic tomography of stable ABL based on the existence of an acoustic wave guide near ground surface due to formation of the temperature inversion and vertical wind shear.

The field experiment was conducted in August 2002 at the base of the Institute of Atmospheric Physics near Zvenigorod. For sounding of the ABL we used a special acoustic pulse generator and four receivers placed on the ground at different distances from the generator. The stratification of the ABL was controlled by Doppler Sodar, Temperature Profiler and meteorological mast with the acoustic anemometers mounted at the heights of 6 m and 56 m above ground surface.

Acoustic pulses of stable form were generated due to detonation of the air-propane mixture with the repetition period of 30sec. One of the acoustic pulses at a distance of 20m from the generator is shown in the upper part of Fig.1. When propagating in the acoustic wave guide the initial signal "splits" at a distance of 2.6 km into a set of arrivals A, B, C, and D, as seen in Fig.1. Such evolution of the pulse wave form with distance was earlier explained theoretically.

For each signal radiated at the moment t we measured the time interval $\Delta T(t)$ between the arrivals (A) and (D) with the accuracy of 2 msec . By measuring temporal variations of this interval $\delta T(t) = \Delta T(t) - \Delta T_0$ relative to its initial value ΔT_0 we were able to calculate the temporal variations $(\delta c_{eff})(t)$ of the effective sound speed $c_{eff} = c + V_e$ averaged over the selected ray trajectory connecting the source and each receiver. Here c is the sound speed and V_e is the projection of the wind velocity along the radius-vector directed from the source to the receiver. The main contribution to the travel time fluctuations comes from some portion of the ray trajectory near its turning point.

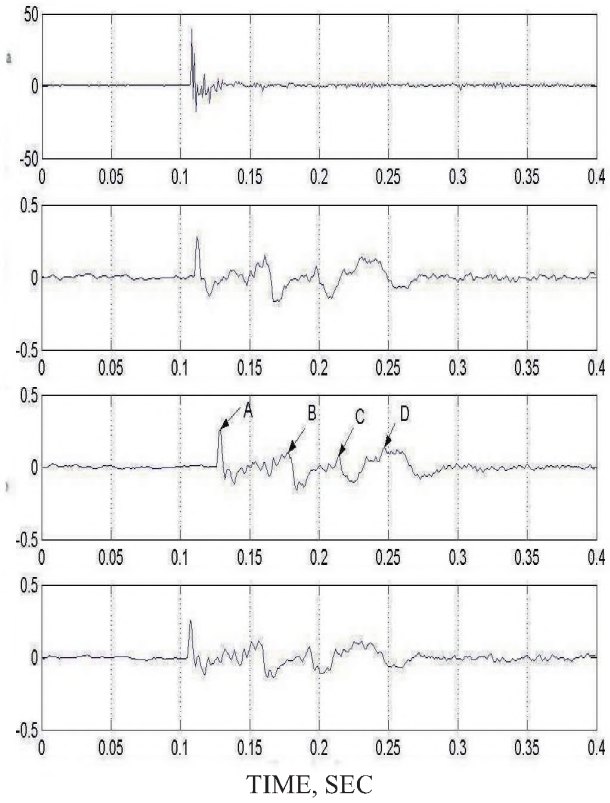


Fig.1. One of the acoustic signals recorded near the source at a distance r=20m (top), and far from the source at r=2550m (Vertical axis is an acoustic pressure in Pa). The signal was received by a 5-m triangle array of receivers. A,B,C, and D are different arrivals of the signal.

3. RESULTS AND DISCUSSIONS.

The temporal fluctuations of the pulse travel time measured in the night hours of August 13, 2002 by a 200-m triangle array are shown in Fig.2. For these time series we have calculated the coherences K_{ij} and the corresponding phase spectra φ_i- φ_j between the pairs of receivers i and j of a triangle array, where i,j=1,2,3. The calculated coherences showed a low-frequency peak ~ (0.7÷0.98) within a low frequency interval (0.6÷1.7) 10^{-3} Hz (the corresponding periods are between 10 min and 28 min). Within the same frequency interval the sum of phase differences was close to zero, and this fact indicated a wave like character of the observed low-frequency fluctuations. Only those frequencies were selected, for which the condition $\Sigma\varphi\approx 0$ was met. Beside a low-frequency peak this condition was also met for the frequencies 6.7 10^{-3} Hz (period 2.5 min) and 1 10^{-2} Hz (period 1.5 min). Thus, there existed a number of dominant frequencies, for which all the coherences reached maxima.

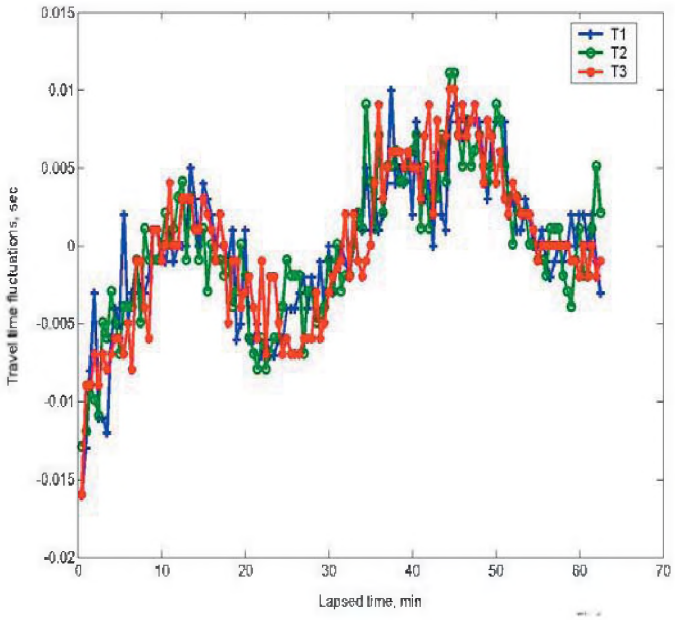


Fig.2. Temporal fluctuations $\delta T(t)$ (vertical axis) of the pulse travel time measured on August 13, 2002 (22:40-22:50) by a 200-m triangle array R1-R3.

For the obtained dominant frequencies we have estimated (by using phase spectra at a triangle array) the horizontal phase velocities V_p of the observed fluctuations and their direction of propagation. The value of V_p as found decreases from 5.5 m/s at a period of 28min to about 2m/s at a period of 8min. The obtained periods and phase velocities are typical for the internal waves often observed in stably stratified ABL. The corresponding horizontal wave lengths of these fluctuations $\lambda = V_p / f$ decrease from 8.6 km at a period of 28 min to 960m at a period of 8min. At shorter period (2.5 min) the wavelength $\lambda\approx 270$m.

Within the range of periods (2.5 ÷10)min the observed fluctuations are likely induced by the trapped IGWs in the wave guide formed near ground surface by the mean stratification of the Brunt-Väisäla frequency and wind speed in stable ABL. This was clearly seen from the vertical profiles (not shown here) of BV-frequency, and wind speed in the lower 400-m atmospheric layer. The same dominant periods were also found in the frequency auto-spectra of travel time fluctuations obtained during the period of measurements (August 9-August 13, 2002). Despite a temporal variability of the obtained experimental spectra their average power law decay was close to the predicted theoretical power law $\varepsilon\ \omega^{-2}$, where ε is the mean rate of wave energy input from the random internal wave sources.

REFERENCES

GOSSARD, E.E., AND W.H. HOOKE (1975) "WAVES IN THE ATMOSPHERE," *ELSEVIER, AMSTERDAM*, 456 PP.

WAVE PROPAGATION IN CURVED LAMINATE COMPOSITE STRUCTURES

SEBASTIAN GHINET [1], NOUREDDINE ATALLA [1], HAISAM OSMAN [2]

[1] Department of Mechanical Engineering, Université de Sherbrooke, 2500 Boulevard Université Sherbrooke, QC, J1K 2R1, Canada

[2] The Boeing Company, 5301 Bolsa Avenue, Huntington Beach, CA 92647, USA

1. INTRODUCTION

The principal aim of this work is to present a model to compute the transmission loss of sandwich composite cylindrical shells. The effects of membrane, bending, transverse shearing as well as rotational inertia are considered in all of the layers composing the structure. The elastic constants of any layer are related to the orthotropic angle-ply defined as the angle of the principal directions of the layer's material to the global axis of the shell. Fundamental relations are expressed using the dynamic equilibrium relations of the unit forces in the structure. A general eigenvalue approach to compute the dispersion curves of such structures is presented. Using these curves, the radiation efficiency, the modal density, the group velocity, the resonant and non-resonant transmission coefficients are computed and used within SEA framework to predict the sound transmission loss of these structures. The described model is shown to handle accurately, both laminate and sandwich composite shells. Comparisons with existing models and experimental data are also discussed.

2. THEORY

This paper describes the SEA modeling of the transmission loss through finite laminate and sandwich composite curved panels. Both laminate composite and sandwich composite are modeled using a discrete thick laminate composite theory. The studied transmission problem has three primary resonant systems such as: two reverberant rooms separated by the composite curved panel. The dispersion curves of the structure are used to compute the modal density and the radiation efficiency. Several models to compute the radiation efficiency were tested[1,2,3]. Identical results were obtained but the model of Leppington[3] was preferred for its accuracy and fast convergence. These indicators allow the calculation of the radiation loss factor and also the resonant contribution of the transmission loss. The standard flat panel theory[15] is used to compute the non-resonant transmission but it is adapted here to the particular vibration behaviours of the curved panels. In particular, a sub-coincident modes selection method is used to compute the non-resonant transmission contribution. Moreover, the classical wave approach non-resonant contribution is corrected using the spatial windowing method presented in reference [2]. Finally, a transmission loss experimental result of a curved sandwich composite panel is successfully compared with numerical estimation.

2.1. Geometry

Figure 1 represents the global geometrical configuration of the composite shell, where R is the curvature radius and h is the total thickness. The layered construction is considered, asymmetrical.

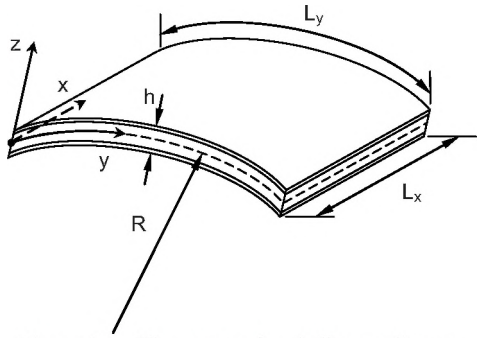


Figure 1. The composite shell co-ordinates

2.2. Dispersion relation

The displacement field of any discrete layer "i" of the panel is of Mindlin's type. For any layer of the shell, Flüge's theory is used to describe the strain-displacement relations. The resulting dynamic equilibrium system has 5N+3(N-1) variables regrouped in two vectors; a displacement–rotation vector {U}, and an interlayer forces vector {F}. The associated 5N+3(N-1) equations are composed of 5 equations of dynamic equilibrium for each of the N layers plus 3 equations of interlayer continuity of displacements for each of the N-1 interlayer surfaces.

To solve for the dispersion relations, the system of dynamic equilibrium equations is expressed in terms of a hybrid displacement-force vector $\langle e \rangle^T = \{U\ F\}$.

Assuming $\langle e \rangle = \{e\}\exp(jk_x x + jk_y y - j\omega t)$, a harmonic solution, the system is expressed in the form of a generalized polynomial complex eigenvalue problem:

$$k_c^2[A_2]\{e\} - ik_c[A_1]\{e\} - [A_0]\{e\} = 0 \qquad (1)$$

where, $k_c = \sqrt{k_x^2 + k_y^2}$, $i = \sqrt{-1}$ and $[A_0]$, $[A_1]$, $[A_2]$ are real square matrices of dimension 5N+3(N-1). Relation (1) has 2(5N+3(N-1)) complex conjugate eigenvalues and represents the dispersion relations of the laminated composite shell.

At any heading direction the curved panel has two propagating solutions below the ring frequency. At the ring frequency a third solution becomes propagating thus, in the dispersion field context the ring frequency is mathematically perceived as a cut-off or transition frequency. Two other cut-off frequencies appear at high

frequencies where two additional solutions become propagating.

2.3. Transmission problem

Below the ring frequency, the non-resonant transmission is dependent on the direction of the incident acoustic waves. For a given excitation frequency band (with ω_{cen} the center band frequency and ω_1, ω_2 the frequency limits of the band), and an incidence direction (θ, φ) the structural and the forced wave numbers are calculated from the dispersion relation (1) and the following condition is checked to ensure that the forced modes are non-resonant:

$$k_0(\omega_{cen})\sin\theta < k_s(\omega_1) \quad \text{OR} \quad k_0(\omega_{cen})\sin\theta > k_s(\omega_2)\cdot \qquad (2)$$

This accounts for both mass and stiffened controlled non-resonant modes. Usually, stiffened controlled modes contribution is neglected and the mass-controlled non resonant transmission coefficient is used. The allowable heading directions are obtained using the dispersion equation (1) and the first condition in (2).

In order to improve the low frequency predictions of the non-resonant transmission coefficient, a geometrical windowing correction method is also used. The correction method used here, is detailed in reference [2] and examples of its validation are given in references [2] and [4].

The resonant transmission coefficient is calculated from the radiation efficiency of the panel and its modal density.

A simple SEA acoustic transmission scheme is used here and consists of two reverberation rooms separated by the studied curved panel. One of the rooms is excited by a diffuse field and the acoustic transmission problem is assumed to encompass two transmission contributions: resonant and non resonant transmission.

3. EXPERIMENTAL VALIDATION

Transmission loss tests were performed on the singly curved sandwich composite panel described in the previous section. The tests were performed at the Canadian National Research Center transmission loss facility located in Ottawa. Measured transmission loss and predictions with both the wave approach and the modal approach are given in Figure 2. It should be noted at this stage that the mounted panel damping was not measured, and that a nominal modal damping ratio of 2.5% was assumed in the analysis. The wave approach with the geometrical correction leads to a very good agreement throughout the frequency range of the test apart from the ring frequency region. On the other hand, the modal approach shows a higher transmission loss than the test below the panel ring frequency. Above the panel critical frequency, both the wave and modal approaches yield almost identical results. Note that the number of resonant modes is not sufficient in the first 1/3 octave bands for the modal method to be reliable at low frequencies (less than 1 mode at 100 Hz).

4. DISCUSSIONS AND CONCLUSIONS

An efficient model to compute the transmission loss of sandwich and laminate composite curved panels was developed. The physical behaviour of the panel is represented using a discrete lamina description. Each lamina is represented by membrane, bending, transversal shearing and rotational inertia behaviours. The model is developed in the context of a wave approach. Using the dispersion relation's solutions, the modal density, the radiation efficiency as well as the resonant and non-resonant transmission loss are calculated. The acoustic transmission problem is represented within statistical energy analysis context using two different schemes for the non-resonant path. It is observed that for the presented problem, the modal energy exchange is dominated by the first wave solution. It is concluded that for classical acoustic transmission problems, the SEA scheme used here is accurate. The results were compared successfully to the transmission loss test of a singly curved sandwich panel and to two other models. In particular, the presented model is applicable to both sandwich panels and composite laminate panels with thin and/or thick layers.

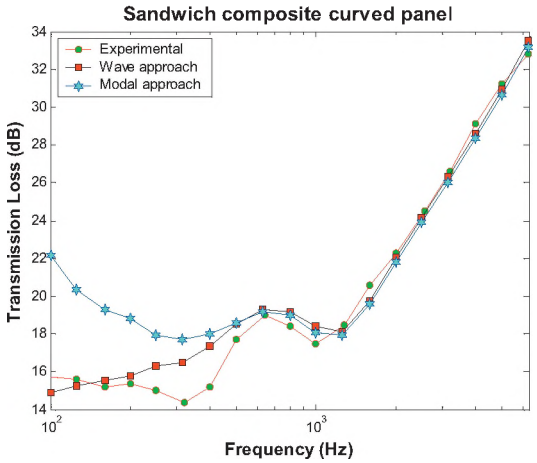


Figure 2. Total transmission loss of a sandwich composite curved panel.

REFERENCES

1. Villot M, Guigou C. and Gagliardini L, "Predicting the acoustical radiation of finite size multi-layered structures …", JSV, 245(3), 433-455, 2001.
2. Ghinet S and N. Atalla « Sound transmission loss of insulating complex structures », Canadian Acoustics, Vol. 29, 3, 26-27, 2001.
3. Leppington F. G, Broadbent E. G. and Heron K. H. "The acoustic radiation from rectangular panels with constrained edges", Proc. R. Soc. Lond. A382, pp 245-271, 1982
4. Osman H., Atalla N., Atalla Y., Panneton R. "Effects of acoustic blankets on the insertion loss …", Tenth international congress on sound and vibration, ICSV 10, Stockholm, Sweden, July 2003

ON TRANSMISSION OF STRUCTURE BORNE POWER FROM WOOD STUDS TO GYPSUM BOARD MOUNTED ON RESILIENT METAL CHANNELS – PART 1: FORCE AND MOMENT TRANSMISSION

Andreas Mayr[†] and T.R.T. Nightingale[§]

[†]Department of Bauphysik, University of Applied Science Stuttgart, Schellingstrasse 24, D-70174 Stuttgart, Germany
[§] Institute for Research in Construction, National Research Council, Ottawa, Ontario, Canada K1A OR6

1. INTRODUCTION

A systematic study [1], recently conducted at the NRC has verified most of the fundamental assumptions made when using the mobility approach for direct attached gypsum board. These assumptions need to be verified for the more complex and practical situation of gypsum board mounted on resilient channels (RC's). This paper, the first of two, examines assumptions relating to power flow as a function of position and number of fasteners between the gypsum board and RC's. Also considered is if the power transmitted by a rotational component of the stud is negligible compared the translational component. The second paper [2] examines further assumptions and gives estimates of the change in radiation efficiency and acoustic power due to adding RC's.

2. SPECIMEN AND EVALUATING TRANSMISSION

Figure 1 shows the wall evaluated consists of a single sheet of 16 mm Type X gypsum board attached to one, or more, rows of resilient metal channels, RC's. These are attached to 35 x 85 mm clear western red cedar studs, spaced nominally 406 mm on centre. A single point force was applied to one of the studs.

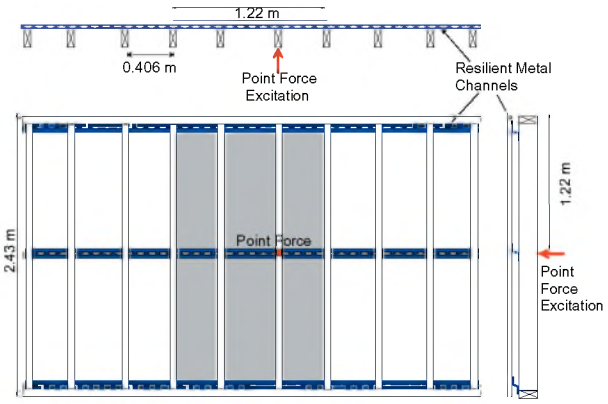


Figure 1: Sketch showing wall framing and the single sheet of gypsum board mounted on RC's.

It is not practical to measure power transmission through a junction directly – indirect evaluation is necessary. Statistical energy analysis (SEA) may be used if both connected elements satisfy the conditions of a subsystem – modes are spaced equally in each frequency band and create a uniform energy density proportional to the damping. This allows one to write,

$$\eta_{12} = \left(\frac{W_{12}}{E_1 \omega}\right) = \left(\frac{m_2}{m_1}\eta_2 \left(\frac{\langle v_1^2 \rangle}{\langle v_2^2 \rangle}\right)^{-1}\right) \qquad (1)$$

where subscript 1 indicates source, 2 receiver, W is transmitted power, E is energy, and ω is angular frequency, m is mass, η_2 is the total loss factor (TLF) of the gypsum board and $\langle V^2 \rangle$is space averaged RMS velocity. Because the mass of the stud and gypsum board are constant, the measured velocity ratio $\langle V_1^2 \rangle / \langle V_2^2 \rangle$ is proportional to the ratio of E_1/W_{12} or inversely proportional to the ratio of gypsum board TLF and the coupling loss factor (CLF), η_{12}, between the stud and gypsum board.

This paper uses a CLF to describe the power flow from the stud to the gypsum board, which will be expressed in dB and can be thought of as being the fraction of the stud energy transmitted to the gypsum board in one cycle. Measurements of the velocity ratio are obtained from differences in the space average stud level (sampled using 14 points) and the gypsum board level (sampled using 98 points).

3. TRANSLATION AND ROTATION DISPLACEMENT

In theory, a point force applied to the neutral axis of a homogenous isotropic beam will cause only pure translation (displacement in the direction of the applied force). Wood studs are not homogeneous and isotropic so there will be both translation parallel to the force and rotation about an effective centre of mass. Structure borne transmission for both types of motion are shown in Figure 2 and must be evaluated because both will transmit power.

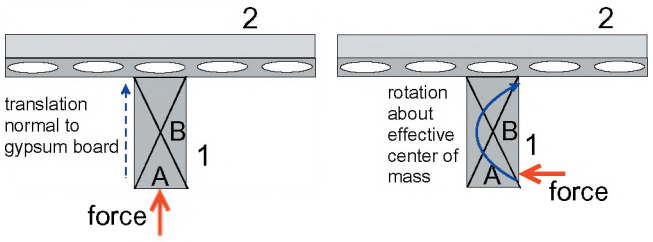


Figure 2: Locations of the point force applied to the stud to induce translation (left) and rotation (right). The stud is labelled 1 with its faces A and B. The gypsum board is labelled 2.

For each source location the velocity was measured at 14 positions on faces A and B along a line that passed through the applied force. Also, the resulting gypsum board velocity was measured. The measurements gave a set of two simultaneous equations from which the CLF for the two types of motion can be determined.

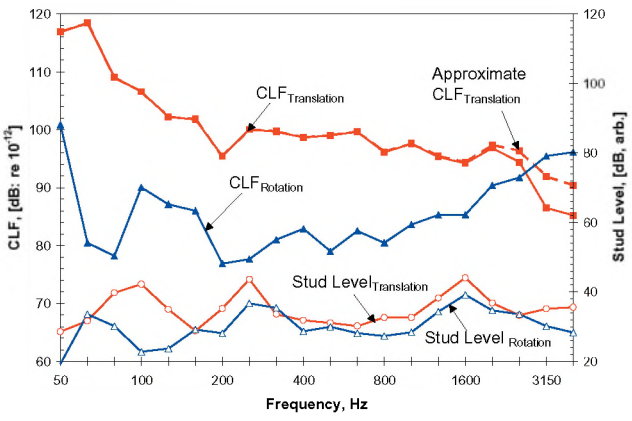


Figure 3: Measured $CLF_{Translation}$ and $CLF_{Rotation}$. Also given is the Approximate-$CLF_{Translation}$ obtained assuming rotation transmission is negligible. Stud velocity levels due to a point force applied to face A of the stud are shown on a separate axis.

Figure 3 shows that below 2500 Hz, the CLF is considerably greater (i.e., more power is transmitted) for stud translation than for rotation. Above this frequency more power is transmitted by rotation. Figure 3 also shows the average stud levels on faces A and B when a point force on face A excites the stud. Since the velocity due to rotation is less than translation it is possible to state that for frequencies less than 2500 Hz power transmission due to rotation will be less than translation when the stud of Figure 2 (left) is excited on face A by a force normal to the gypsum board. The figure also shows that the Approximate $CLF_{Translation}$ (obtained by exciting the stud on face A and collecting the resulting space average stud velocity on face A and gypsum board velocity) is an excellent approximation for frequencies below 3150 Hz. Henceforth, in this paper and in Part 2 we will use the Approximate $CLF_{Translation}$ and simply call it CLF, recognising that above 2500 Hz the value will be an overestimation for pure translation.

4. FASTENERS BETWEEN GYPSUM BOARD AND RC'S

To determine if location and number of fasteners attaching the gypsum board to the RC's are factors in determining power transmission, fasteners were systematically installed in all five columns, one column at time, as shown by the insert in Figure 4. Three resilient channels were installed.

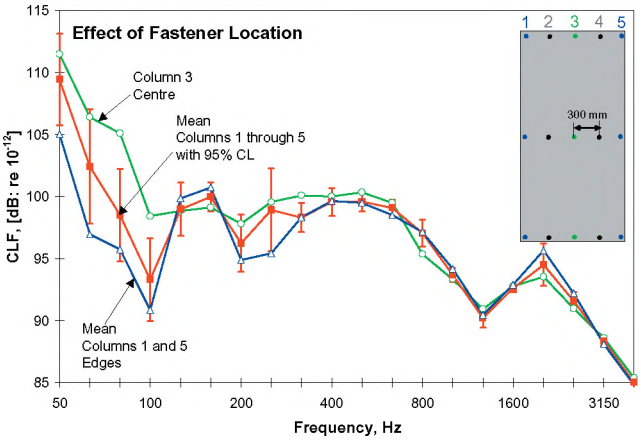


Figure 4: Sensitivity of CLF to location of fasteners attaching the gypsum board to RC's. Insert shows location of the five columns.

Comparing the CLF's it is clear that below 125 Hz the strongest transmission occurs at column 3 (near the gypsum board centre and excited stud). While, the weakest coupling is at columns 1 and 5 (near the edges of the gypsum board). Above 250 Hz there is very little effect associated with a change in fastener location. Thus, fastener location would not have to be modelled explicitly for this frequency range. However, below 250 Hz, assuming that all positions could be represented by a mean CLF could sometimes underestimate or overestimate by about 8 dB. Although, on average the estimate would be correct.

A measurement series investigated increasing the number of fasteners attaching the gypsum board to the RC's, from 1 to 5. Figure 5 shows the results for two different channel configurations. For both, the change in CLF above 315 Hz is considerably less than 7 dB (10log(5)) expected from simple theory. Below 315 Hz, simple theory based on incoherent motion at the fastener reasonably approximates the change.

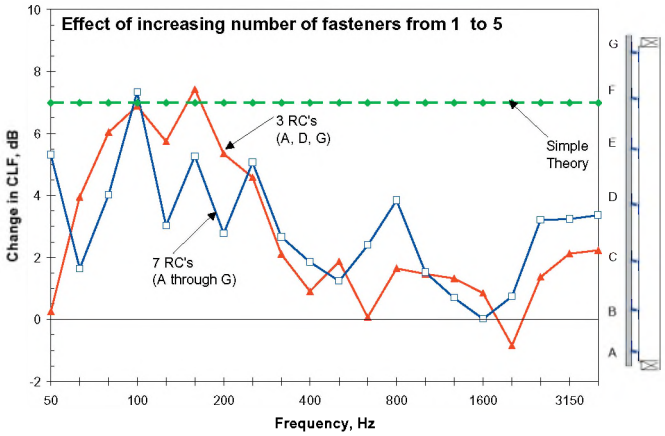


Figure 5: Change in CLF due to increasing the number of fasteners from 1 (located at column 3) to 5 in each RC with three and seven RC's installed.

5. DISCUSSION AND CONCLUSIONS

This paper has shown that typically studs will have both a rotational and translation motion, but that with RC's the translation displacement (normal to the gypsum board) couples more strongly than rotation. For practical purposes if the rotation velocity is less than or equal to the translation velocity, then structure borne transmission due to the rotation component can be ignored for frequencies below about 2500 Hz.

Results showed that above about 315 Hz there is little effect associated with either the location or the number of fasteners attaching the gypsum board to the RC's. However, below this frequency there is a rather significant effect associated with both the number and location. This is somewhat surprising as it is the opposite of what was observed for direct attached gypsum board [1]. The cause was not examined in detail because this is a preliminary study. However, it is speculated that in the high frequencies the RC's are sufficiently stiff that vibration levels are reasonably uniform with distance along the RC and that a series of ill-defined contact points exist so that adding fasteners or changing their location does not have a significant effect. But in the low and mid frequencies there is a complex modal interaction between the studs and the RC's that span the width of the wall, such that the vibration level of the RC varies significantly with distance away from the excited stud. As shown in Part 2 of this paper [2] the system becomes more complicated and the CLF is more sensitive to physical changes in the low frequencies because of a very low mode count.

A detailed examination of the sensitivity of the number and location of fasteners between the gypsum board and RC's is suggested for future work.

6. REFERENCES

1 T.R.T. Nightingale, Katrin Kohler, Jens Rohlfing (2004), "On predicting structure borne sound transmission from wood stud to direct-attached gypsum board," Proceedings of ICA 2004, Kyoto Japan, April 5-9.

2 T.R.T. Nightingale, Andreas Mayr, (2004), "On transmission of structure borne power from wood studs to gypsum board mounted on resilient metal channels – Part 2: Some simplifications for modelling", Canadian Acoustics, Vol. 32(3).

ON TRANSMISSION OF STRUCTURE BORNE POWER FROM WOOD STUDS TO GYPSUM BOARD MOUNTED ON RESILIENT METAL CHANNELS – PART 2: SOME SIMPLIFICATIONS FOR MODELLING

T.R.T. Nightingale[§] and Andreas Mayr[†]

[§] Institute for Research in Construction, National Research Council, Ottawa, Ontario, Canada K1A OR6

[†]Department of Bauphysik, University of Applied Science Stuttgart, Schellingstrasse 24, D-70174 Stuttgart, Germany

1. INTRODUCTION

This is the second of two papers that consider the effect of resilient metal channels (RCs) on structure borne transmission from wood studs to gypsum board. Readers are referred to the first paper [1] for wall details, definitions, and measurement procedures. This paper examines several simplifications that might be used when creating a model for structure borne transmission from a wood stud to gypsum board that is mounted on RCs. Possible simplifications include assuming transmission is reasonably independent of the location of the RCs, but is dependent on the number, and that the total power transmission is simply the sum of transmission through individual RCs. The paper concludes with the reduction in radiated sound power velocity that might be expected if direct attached gypsum board is mounted on RCs.

2. EFFECT OF CHANNEL LOCATION

To estimate the variation in structure borne transmission due to a change in the location of the RC, a single RC was installed at each of the seven locations A-G shown in the insert to Figure 1 and the coupling loss factor (CLF) measured. Resilient pads supported the gypsum board at the bottom. The confidence intervals indicate that above about 400 Hz CLF does not change much with RC location. However, below 400 Hz the large confidence intervals indicate a significant variation in CLF with channel location.

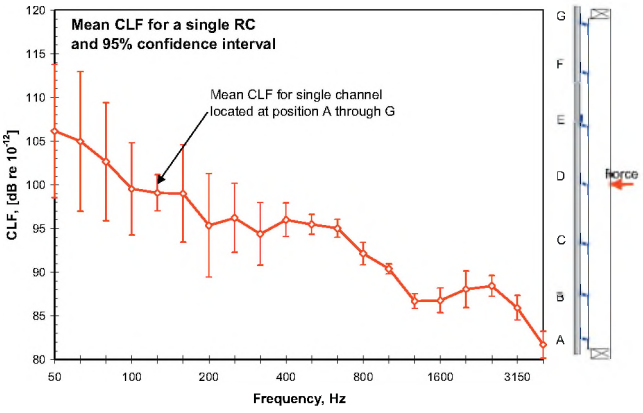


Figure 1: Mean CLF for a single RC installed at positions A through G, when the stud is excited by a single point force at D.

To gain some insight into why the CLF below 400 Hz is sensitive to RC location, Figure 2 shows the RMS velocity of the stud for 125 and 200 Hz when excited at position D. The figure shows the stud can only support low order modes in these frequencies and that the RMS velocity varies significantly with location. The horizontal lines indicate the location where the resilient channels would be attached and that some RCs will be located near antinodes while others will be located near nodes. More power will be transmitted by an RC located at an antinode than the same RC located at a node despite constant stud energy. This means statistical approaches (like SEA) cannot accurately predict the power flow through individual RCs where there are few modes because estimates of the mean energy in the stud might not be highly correlated with the velocity of the stud at the RC. With several resilient channels spaced a significant distance apart and considerably more modes in the frequency band of interest, statistical approaches will have greatly improved accuracy, because the mean stud energy will be highly correlated with the mean of the stud velocities at the RCs. Finite Element or Modal Methods could have given more accurate estimates of low frequency transmission, if the boundary conditions of the studs and gypsum board are known. But, in reality the boundary conditions are not well known and will vary from one stud to the next because of workmanship and variations in material properties.

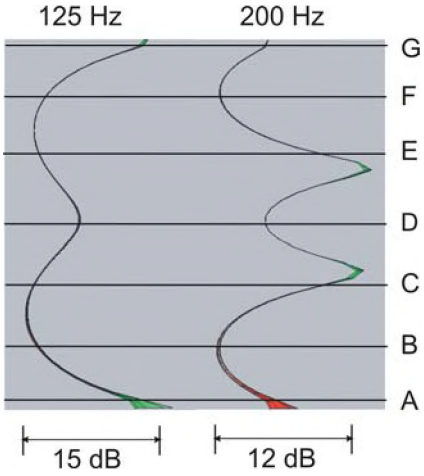


Figure 2: Stud vibration response at 125 and 200 Hz expressed as RMS velocity when excited by a point force applied at Position D.

3. APPROXIMATING TOTAL POWER FLOW

It is often assumed that the total power flow between systems coupled by a series of well-defined point contacts can be approximated by the sum of the power flow at each point. This implies that the motion of each fastener is incoherent such that each fastener appears to be an incoherent source and the powers simply add. This will not be the case when the spacing between the fasteners is such that the points are coherent which as some have suggested occurs when the spacing is less than half a wavelength.

To verify this assumption, the individual CLF's obtained from the work described in Figure 1 were used to estimate the total CLF, and hence total power flow, that would be exist for an arbitrary number and distribution of RCs. If the points act incoherently then there will be good agreement between measured and summed CLF's. Figure 3 shows results for the case where there are 7 RCs (A through G) and 3 RCs (A, D, and G). There is good agreement between the measured total CLF with 3 RCs and the estimate obtained from the sum of CLF's. The agreement for 7 RCs is not as good. The sum of individual CLF's slightly overestimates the total CLF. With 7 RCs, the spacing between the RCs is about 400 mm and is small compared to the bending wavelength in the studs at low frequencies. Thus, the motion of the fastening points will not be fully incoherent and the total power flow will be less than that predicted from a simple sum. The cause for the 1-2 dB

overestimation for frequencies above 315 Hz is not known. The effectiveness of RCs to attenuate structure borne transmission can be seen from Figure 3 by the significantly greater CLF for directly attached gypsum board.

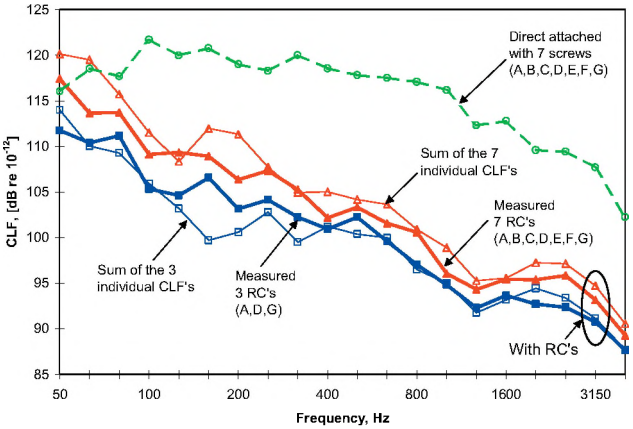


Figure 3: Comparison of measured CLF's and those computed by summing the CLF's for the individual RCs.

4. RADIATION EFFICIENCY

It is possible that the radiation efficiency of gypsum board will be different when directly attached to the studs and when mounted on RCs. To evaluate this, the radiation efficiency [2] of the gypsum board for resonant motion was measured when a single stud was excited at position D by a point force. The gypsum board radiation efficiency did not have a systematic dependence on the number of fasteners when directly attached to the studs or when mounted on RCs. Figure 4 shows the means for 3, 5 and 7 fasteners or RCs. The 95% confidence interval can be used to gauge the range of values. Below about 1250 Hz the radiation efficiency is between 5 and 10 dB lower when the gypsum board is mounted on RCs.

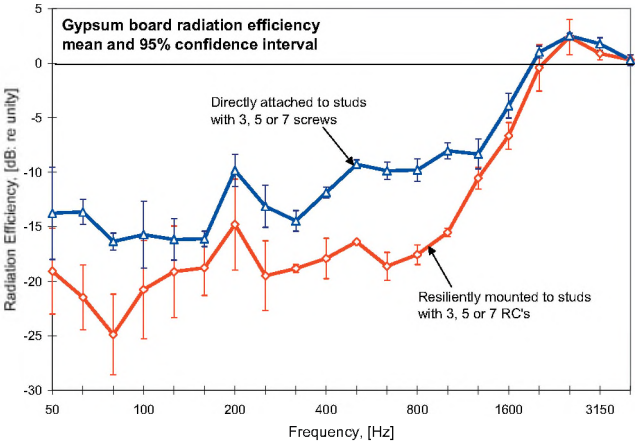


Figure 4: Estimates of radiation efficiency for gypsum board that is directly attached and mounted on RCs.

This can be explained by recognising that the radiation efficiency of 16mm gypsum board below the critical frequency (~2500 Hz) is strongly affected by the aspect ratio. Because the studs are effectively line connected to the gypsum board below about 1250 Hz due to ill-defined contact points, the gypsum board is effectively divided into a series of high aspect ratio panels each 0.4x2.4m. However, when the same panel is resiliently mounted it acts like a single low aspect ratio 1.2x2.4m sheet.

5. SUPPRESSION OF RADIATED POWER

The ability of RCs to attenuate the sound power radiated by gypsum board when one or more studs are structurally excited is the combined effect of the change in velocity level difference, VLD (between stud and gypsum board) and the change in radiation efficiency. A rough estimate of VLD change can be obtained from Figure 3; by taking the difference in the CLFs. The change in radiation efficiency is given by the difference in the curves of Figure 4. The combination of these two differences is the change in radiated sound power for a structurally excited single leaf wall as shown in Figure 5. The figure indicates that resiliently mounting gypsum board can be a very effective method to control flanking transmission involving gypsum board surfaces because the path is attenuated by 10 to 20 dB throughout the building acoustics range (100-4000 Hz).

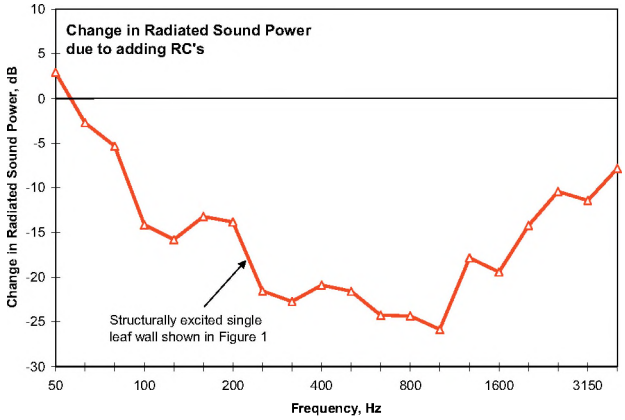


Figure 5: Change in radiated sound power when direct attached gypsum board (seven screws) is mounted on five resilient channels.

6. DISCUSSION AND CONCLUSIONS

This paper showed that RCs attenuate both structure borne transmission and acoustic radiation. The result is a reduction in the radiated sound power of at least 10 dB for frequencies greater than 100 Hz for structural excitation.

Parts 1 and 2 showed that above 400 Hz power transmission is almost entirely determined by stud translation (except at very high frequencies), is proportional to the number of RCs, and is independent of number and location of fasteners between gypsum board and RCs. These properties suggest in this frequency range it might be possible to use the mobility approach if RCs can be approximated by a series of idealised springs located at points of intersection between stud and RCs. The dynamic stiffness of each spring would be equal to that of the RC at a stud. Because of the simplicity of this approach it is suggested for future work. Below 400 Hz there are few modes and RC location becomes an important factor so a mobility approach would have significant errors.

7. REFERENCES

1 Andreas Mayr, T.R.T. Nightingale, (2004), "Part 1: Force and moment transmission", Canadian Acoustics, Vol. 32(3)

2 L. Cremer, M. Heckl, "Structure borne sound", Springer Verlag, New York 1988.

FORWARD AND REVERSE TRANSMISSION LOSS MEASUREMENTS

A.C.C. Warnock
Institute for Research in Construction, National Research Council Canada, 1200 Montreal Road, Ottawa, Ontario K1A 0R6, Canada. Alf.Warnock@nrc-cnrc.gc.ca

1. INTRODUCTION

The two main standard methods for measuring sound transmission loss using the two-room method – ASTM E90[1] and ISO 140-3[2] – only require that measurements be made in one direction. One of the rooms is designated as the source room, the other automatically is the receiving room. Neither standard gives any guidance on how to select the role for each room in most cases. Users may choose to test in two directions but it is not mandatory.

At low frequencies the room modal response influences the measured transmission loss values, and so is a factor in determining the reproducibility of the two-room test methods. Measurements made in both directions[3] at the National Research Council showed differences in transmission loss at low frequencies that were large enough to give different STC[4], or R_w[5] ratings. This observation caused us to routinely measure transmission loss in both directions in our test facilities. This paper presents the findings from the data collected.

2. TEST ROOMS

In the wall test suite originally built at NRC around 1955, one room had a volume of 65 m^3 and the other a volume of 250 m^3. Both rooms were essentially box-shaped and the smaller was used as the source room for many years. In 1998 the smaller room was replaced[6] with one having a volume of 145 m^3. This new room has a pentagonal floor plan. The NRC floor test suite[3] was commissioned in 1992 and has room volumes that are approximately equal at 175 m^3. Since 1997 enough data have been collected in all three test suites to allow a closer examination of differences in transmission loss when the direction of the test is changed.

3. EXAMPLE OF DIFFERENCES.

When transmission loss plots for the two measurement directions are compared on a single chart, it is not always obvious that there are significant differences. Occasionally, however, differences are enough to cause the STC or R_w ratings to be different. Figure 1 shows one such result measured in the floor facility. In this case, "forward" and "reverse" mean the upper room and the lower room respectively act as the source room.

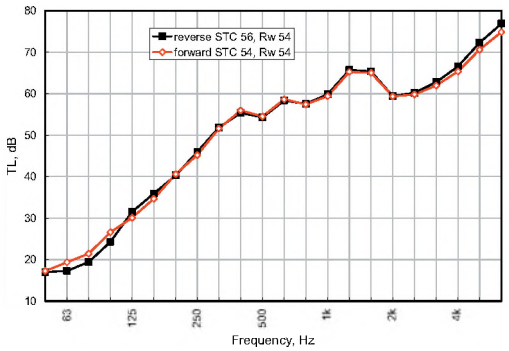


Figure 1: Sound transmission loss measured in two directions in the floor facility.

The graph shows differences in transmission loss at low frequencies as might be expected but there are also differences at high frequencies. The STC rating differs by two points for this floor while, in this case, the R_w rating is the same. When only differences are plotted, as in Figure 2, it is clear that the differences are much greater than the repeatability limits for the facility. These differences in transmission loss due to direction are not random. Once a specimen has been installed, the differences do not change significantly when the measurements are repeated.

4. MEAN DIFFERENCES

Inspection of many of these difference plots suggests that the average difference for the frequency range 200 to 800 Hz is zero. In a few cases there were small but obvious biases that were attributed to the uncertainty associated with calibration. To minimize such effects, each difference spectra was normalized so the mean difference in the frequency range 200 to 800 Hz was zero.

The normalized differences for the floor test suite are shown in Figure 3. The graph shows the mean difference, the minimum and maximum differences observed and the standard deviation of the differences. It is disturbing that the average difference is not zero at all frequencies. It is even more disturbing that the maximum and minimum differences observed are so different from zero. Such differences can be expected to lead to differences in single number ratings such as STC. While the details differ, each test suite shows the same kind of general behavior.

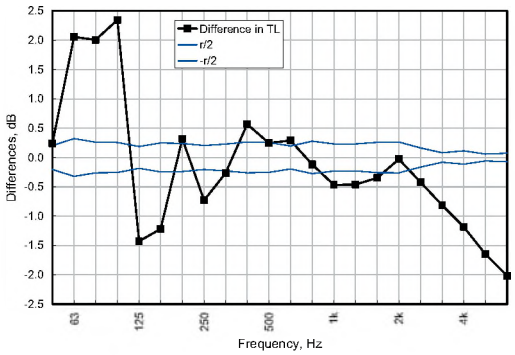


Figure 2: Difference in transmission loss for two measurement directions compared with repeatability limits for the floor facility.

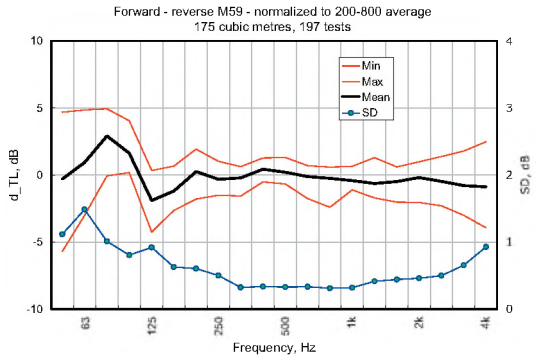


Figure 3: Normalized forward-reverse differences for the NRC floor test facility and the standard deviation of the differences.

The conclusion that derives from these graphs is not really new. The expression used to calculate transmission loss, *TL = L(source) – L(receive) + 10 log S/A*, is only an approximation to the reality of reverberation room test suites. Dependence on test direction is evident throughout the frequency range. It is commonly assumed that when the measurement frequency is high enough – above the Schroeder frequency for example – that the rooms provide a good approximation to the ideal diffuse field. If this were so, transmission loss would not depend on measurement direction, but it does.

5. VARIATIONS IN RATINGS

These differences cause differences in the summary ratings generated by the test methods. The table below shows the distribution of STC and OITC[7] differences for 507 tests run in the three NRC facilities. Simply by reversing the test direction, ratings can be obtained for a specimen that are different enough to make the difference between meeting or not meeting building code requirements in North America.

For many lightweight stud walls and joist floors, the STC rating is determined by application of the 8 dB rule to the transmission loss values below 250 Hz; quite often the STC is determined by the transmission loss in one band. Because of the 8 dB rule, the STC rating is very sensitive to changes in transmission loss caused by changing test direction. The 8 dB rule is sometimes applied at high frequencies. In one measurement, the STC changed by 3 points when the test direction was changed because of the application of the 8 dB rule at 2500 Hz.

Table 1: Distribution of forward-reverse differences for one ISO and two ASTM ratings.

Difference	STC	OITC	R_w
-4	1	0	0
-3	12	0	0
-2	38	6	1
-1	186	103	48
0	222	202	433
1	46	141	22
2	2	45	1
3	0	9	1
% different	56%	57%	14%

6. SUMMARY

Changing the direction of a transmission loss test can change the STC or other ratings generated. A laboratory operator who chooses to run a standard test in both directions has no way to decide which of the two sets of results obtained is correct. It would be preferable for standard test methods[1,2] to require measurements in both directions as a means of improving reproducibility for these test methods. The number of microphone positions required for a test in one direction could be reduced somewhat to avoid doubling the time for testing if this is thought necessary. With automated systems, however, the measurement time is negligible compared to time spent on construction, administration and report preparation.

REFERENCES

[1] ASTM E90. Standard Test Method for Laboratory Measurement of Airborne Sound Transmission Loss of Building Partitions.
[2] ISO 140-3. Acoustics, Measurement of sound insulation in buildings and of building elements – laboratory measurements of airborne sound insulation of building elements.
[3] R.E. Halliwell, J.D. Quirt, and A.C.C. Warnock. Design And Commissioning of a New Floor Sound Transmission Facility. Proceedings of INCE93.
[4] ASTM E413. Classification for Rating Sound Insulation.
[5] ISO 717-1. Acoustics, Rating of sound insulation in buildings and of building elements. Airborne sound insulation in buildings and of interior building elements.
[6] R.E. Halliwell. Renovations of the IRC/NRC Acoustical Transmission Loss Facility for Walls, and their Effects, Internal Report 826, Institute for Research in Construction, National Research Council Canada, 2001
[7] ASTM E1332. Standard Classification for Determination of Outdoor-Indoor Transmission Class.

IMPACT SOUND RATINGS: ASTM VERSUS ISO

A.C.C. Warnock
Institute for Research in Construction, National Research Council Canada, 1200 Montreal Road, Ottawa, Ontario K1A 0R6, Canada. Alf.Warnock@nrc-cnrc.gc.ca

1. INTRODUCTION

ASTM[1] and ISO[2] impact sound tests use the same standardized tapping machine and are essentially the same. Four possible metrics, described below, are available for rating the effectiveness of the floor systems[3,4]. These metrics are or can be used to evaluate floor coverings or toppings that may be placed on concrete slabs or on lightweight joist floors. Correlations among the metrics are poor and numerical differences can be very large. Data collected in the NRC laboratory over several years are used to illustrate the problems facing the standards-writing committees.

2. METRICS

The IIC[4] fitting procedure differs in only two respects from the ISO[3] $L_{n,w}$ procedure: data are rounded to the nearest decibel instead of to the nearest 0.1 dB and the maximum deficiency allowed during the fitting procedure is 8 dB – the "8 dB rule". When fitting is complete, IIC is obtained by subtracting the value of the reference contour at 500 Hz from 110. This has the effect that higher IIC numbers mean greater protection against impact sound. In contrast, the $L_{n,w}$ rating decreases as the floor impact sound attenuation increases. When the 8 dB rule is not invoked, the two ratings are related by IIC = 110 – $L_{n,w}$.

In an informative Annex to ISO 717, spectrum adaptation terms are proposed that are meant to deal with low frequency impact sound. The sum of the adaptation term and the $L_{n,w}$ rating is equal to the unweighted energy sum of the tapping machine levels minus 15 dB. The suggested frequency range is 100 to 2500 Hz but it can be extended down to 50 Hz. In this paper C50 and C100 denote the energy sums from 50 or 100 to 2500 Hz minus 15 dB.

3. COMPLETE FLOOR SYSTEMS

IIC, $L_{n,w}$, C100 and C50 have been calculated for 407 floor tests carried out over several years at NRC. As might be expected, the 8 dB rule means IIC does not always agree with $L_{n,w}$. 152 of the floors tested were given poorer ratings by the ASTM metric[4]. The differences are not all small. The origins of the 8 dB rule are shrouded in historical mist and it is not the point of this paper to examine its validity. The effect on IIC of rounding to the nearest 0.1 dB instead of 1 dB is negligible; only 41 results changed when the rounding rule was changed and the changes were all by only by ±1 point. Neither IIC nor $L_{n,w}$ correlate very well with the C100 rating as expected. The distribution of differences is shown for 110-IIC and $L_{n,w}$ in Figure 1. Where the differences are large and positive, the floor tested usually had a concrete or other very hard surface.

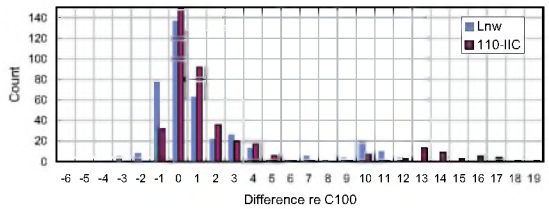


Figure 1: Distribution of differences re C100.

C50 ought to deal more effectively with the low frequency sounds since it includes data down to 50 Hz. The histogram of differences between C50 and the other ratings is shown in Figure 2.

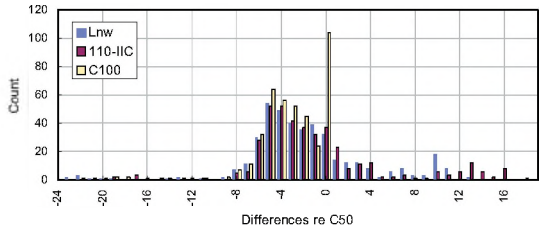


Figure 2: Distribution of differences re C50

The range of differences is much greater as might be expected since many of the floors in the database were joist floors which usually have most of the transmitted impact energy at low frequencies. Even C100 does not compare very well with the C50 rating. The situation becomes worse when ratings for floor toppings are compared.

4. TOPPINGS ON CONCRETE SLABS

The ASTM[5] and ISO[6] test procedures for evaluating floor coverings or toppings are almost identical. The rating procedures are based on E989 and ISO 717-2 and are denoted as ΔIIC and $\Delta L_{n,w}$. Because of the 8 dB rule in E989, the reference contour for the bare reference slab is 4 dB lower when it is fitted than in the case of the ISO 717

fit. This bias can be further modified by a second application of the 8 dB rule to the reference slab plus the topping.

Toppings might also be evaluated using the differences in the two quantities C50 and C100. Figure 3 compares the two standard ratings with ΔC100. Figure 4 uses ΔC50 as the common reference. These variations are large.

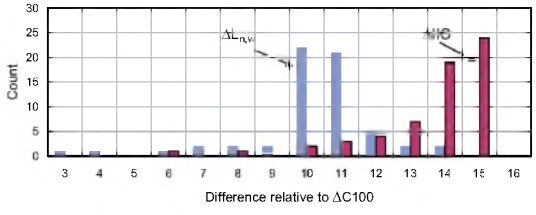


Figure 3: Differences between ΔIIC and $\Delta L_{n,w}$ and ΔC100

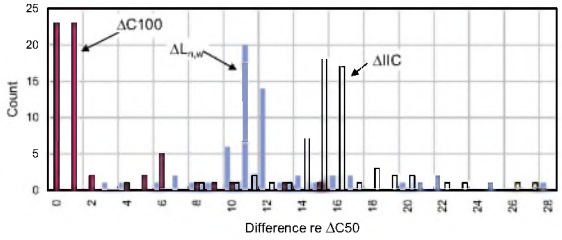


Figure 4: Differences between ΔIIC, $\Delta L_{n,w}$, ΔC100 and ΔC50 for toppings on a concrete slab

5. TOPPINGS ON JOIST FLOORS

An ISO working group has almost finished creating a test method for evaluating toppings placed on a standard joist floor[7]. Toppings placed on joist floors with wood subfloor do not give the same reduction in impact sound pressure level as they do when placed on a concrete slab. To estimate the magnitude of the differences among ratings, improvement spectra from 41 measurements of toppings on joist floors were applied to a reference spectrum and rated using the four rating schemes. Four new symbols are introduced. These are Δ_{joist}IIC , $\Delta_{joist}L_{n,w}$, Δ_{joist}C100, and Δ_{joist}C50.

Less than half of the differences between $\Delta_{joist}L_{n,w}$ and Δ_{joist}IIC are in the range ±1. The ISO 717 based rating tends to give higher ratings, which was not the case for toppings on concrete.

Figure 5 compares the contour-based ratings with .Δ_{joist}C100. The $\Delta_{joist}L_{n,w}$ rating tends to be higher than Δ_{joist}C100; Δ_{joist}IIC tends to be lower.

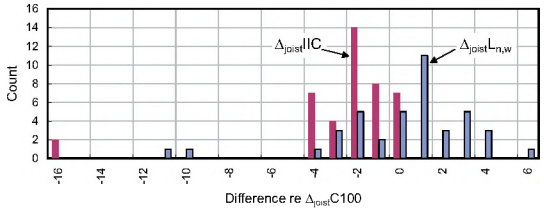


Figure 5: Difference in ratings for toppings on joist floors relative to Δ_{joist}C100 , the difference in the flat level from 100 to 2500 Hz.

Δ_{joist}C50 ought to be the best of these ratings since it includes low frequencies. It is compared with the other three in Figure 6. Including frequencies below 100 Hz increases disagreement among the ratings.

Some of the differences seem unreasonably large but their cause can be understood when the individual spectra are examined.

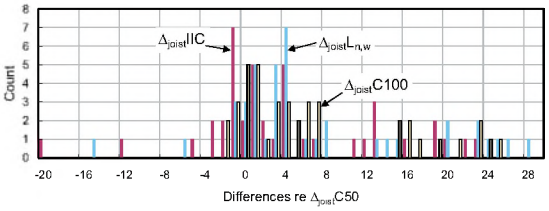


Figure 6: Difference in ratings for toppings on joist floors relative to Δ_{joist}C50.

6. SUMMARY

The poor correlation among the metrics means that complete floor systems are not ranked consistently. When differences in these metrics are used to rank floor toppings, not only are the rankings inconsistent but there is not always agreement that a topping actually improves the floor system.

The C type ratings are not mandatory in ISO 717. It would be unwise to make them so without simultaneously abandoning the use of the contour fitting procedures in E989 and ISO 717-2. To have four contradictory metrics in place would lead to intolerable confusion. It would be ideal if both ISO and ASTM could change simultaneously to an improved rating. The C50 rating, because of the low frequencies included, might lead to unacceptably large reproducibility ranges for tests on complete floors but should be acceptable for toppings.

REFERENCES

[1] ASTM E492. Standard Test Method for Laboratory Measurement of Impact Sound Transmission through Floor-ceiling Assemblies using the Tapping Machine.

[2] ISO 140-6. Laboratory measurements of impact sound insulation of floors

[3] ISO 717. Rating of sound insulation in buildings and of building elements, Impact sound insulation.

[4] ASTM E989. Standard Classification for Determination of Impact Insulation Class (IIC).

[5] ASTM E2179. Standard Test Method For Laboratory Measurement Of The Effectiveness Of Floor Coverings In Reducing Impact Sound Transmission Through Concrete Floors.

[6] ISO 140-8. Measurement of sound insulation in buildings and of building elements, Laboratory measurement of the reduction of transmitted impact noise by floor coverings on a standard floor

[7] ISO 140-11. Laboratory measurements of the reduction of transmitted impact noise by floor coverings on lightweight framed standard floors.

PRACTICAL GUIDANCE FOR SEISMICALLY RESTRAINING PIPING AND DUCTWORK

Paul Meisel, PE
Kinetics Noise Control, 6300 Irelan Place, Dublin, Ohio 43017 USA pmeisel@kineticsnoise.com

1. INTRODUCTION

While seemingly simple, the wide range of factors that comes into play when adding restraints to hanging pipe and ductwork makes this the most difficult and misunderstood seismic task that mechanical contractors have to face. The requirements vary significantly with size, application, location in the structure, mounting arrangement and building use.

In practice, piping and ductwork is commonly re-routed from the initial design path. Other components compete for space, significant structure to which to attach it is unavailable and access is difficult. The net result is that the individual doing the installation does not have a clear "road map" to apply. He needs to understand fully what he is trying to accomplish so as to be able to devise workable solutions to each problem that comes up.

The focus here is on problem issues most frequently encountered in the field. Offered are suggestions as to how to address them as well as strategies to avoid the need for restraint, selecting, sizing and placing them as well as dealing with tight spaces.

2. WHERE ARE RESTRAINTS REQUIRED?

The 1995 NBC is currently the master code driving seismic issues in Canada. Although it is due to be replaced in 2005 with an "Objective" based code, this has not yet been finalized and for the purpose of this discussion, the 1995 code will be referenced.

Within the Code, Chapter 4 addresses Structural issues and section 4.1.9 offers guidance with respect to seismic design. In addition, the current addendum to the 95 NBC references NFPA 13 1999 as the appropriate standard for Fire Piping. To the best of my knowledge, there are no exclusions listed in the NBC code for small pieces if equipment, pipes, ducts or special conditions. The SMACNA design standard has however, been historically accepted in both the United States and Canada as a sufficient design standard. It does carry with it some exceptions that have been universally applied. Drawing from SMACNA, the following piping and ductwork does not require restraint as long as the likely swinging motion of these components will not damage the components themselves or through impact, will not result in damage to other components.

SMACNA Exemptions

Piping: All non-hazardous piping under 2-1/2 in. dia
All non-hazardous piping in mech rooms 1 in. dia and smaller.
All piping runs suspended on hanger rods 12 in. or less in length

Ducts: All ducts under 6 sq ft in. area
All duct runs suspended on hanger rods 12 in. or less in length

(Note: Multiple trapeze mounted pipes or ducts where the summed weight equals the above are also exempted.)

NFPA Piping Exemptions

Lateral Braces for all branch lines under 2-1/2 in. dia
Lateral Braces for all piping runs suspended on hanger rods 6 in. or less in length

3. ADHERENCE TO THE 12 IN. OR 6 IN. RULES

If installing a system tight to the ceiling to take advantage of the 6 or 12 in. hanger exclusion rule, the dimension is measured from the top of the pipe or duct if it is individually supported on a hanger. If supported by a trapeze bar or at its centerline, the dimension to the support point is used. The top of the hanger rod is at its connection to the structure. All supports for a given run of pipe or duct must comply to use the 12 in. rule.

An additional requirement of the 6 or 12 in. rule is that the hanger rod must include a non-moment generating (free-swinging) connection to the structure. This is to allow the pipe or duct to swing without stressing the hanger rod. A swivel, a cable or an isolation hanger can accomplish this function. If using an isolation hanger, a vertical limit stop must be positioned on the hanger rod just below the isolator housing so that when subjected to an uplift load, the limit stop will come into contact with the isolator housing and prevent significant upward motion of the rod. For a strap-supported duct, a half twist in the strap will usually allow the duct to swing freely in all directions.

4. PROS AND CONS OF STRUTS VS CABLES

In general, piping restrained by struts will require only 1 strut per restraint location while piping restrained with cables requires that 2 cables be fitted forming an "X" or a "V". As a trade-off, the number of restraint points

needed on a run of piping or duct will typically be considerably higher for a strut-restrained system than for the cable-restrained system. As a rule, strut-restrained systems will be more costly to install.

The obvious advantage to struts is that, when space is at a premium, cables angling up to the ceiling on each side of a run may take more space than is available. Struts can be fitted to one side only, allowing a more narrow packaging arrangement. As struts can be loaded in compression however, they impact the tensile forces in hanger rods. This often requires the hanger rod size and its anchorage to be upsized.

The advantages of cables are numerous. First, they can usually be spaced less frequently along a pipe than can struts. Second, they cannot increase the tensile forces in the hanger rod, so rod and rod anchorage capacities are not impacted. Third, they are easily set to the proper length. And fourth, they are well suited to isolated piping applications.

5. MIXING CABLE AND STRUTS

Within a run, cables and struts cannot be mixed. This is because the relative stiffness of the two restraint methods are significantly different and in mixed systems, the stiffer strut components will absorb an excessive amount of the load and fail prematurely.

6. RESTRAINT SPACING AND COMPONENTRY SELECTION

There are many factors that can impact the restraint spacing. The primary factor is the buckling capacity of the pipe or duct and this varies with the relationship between its weight and its stiffness. The spacing also varies with the magnitude of the force applied. For added safety, this spacing must be cut in half for hazardous or life safety systems.

The spacing is also controlled by the capacity of the restraint device or anchorage. In many cases, especially those involving struts, practical limits for the restraint system components reduce the allowable spacing between restraints.

Tables are available from SMACNA or from the various providers of these types of systems that list the allowed axial and lateral restraint spacing by pipe or duct size, components used and ground acceleration. Because of the inter-relationship between the tabulations and the components, extreme caution must be used to ensure that components and tabulations are not mismatched.

7. DEALING WITH TIGHT SPACES

Most projects involve areas where space is at a premium and access is poor. An awareness that there are many equivalents to the classic "_/" cable or "_/" strut arrangement can often offer a solution to these issues. While there are far too many to include here, below are illustrated several equivalent restraint arrangements for ducts.

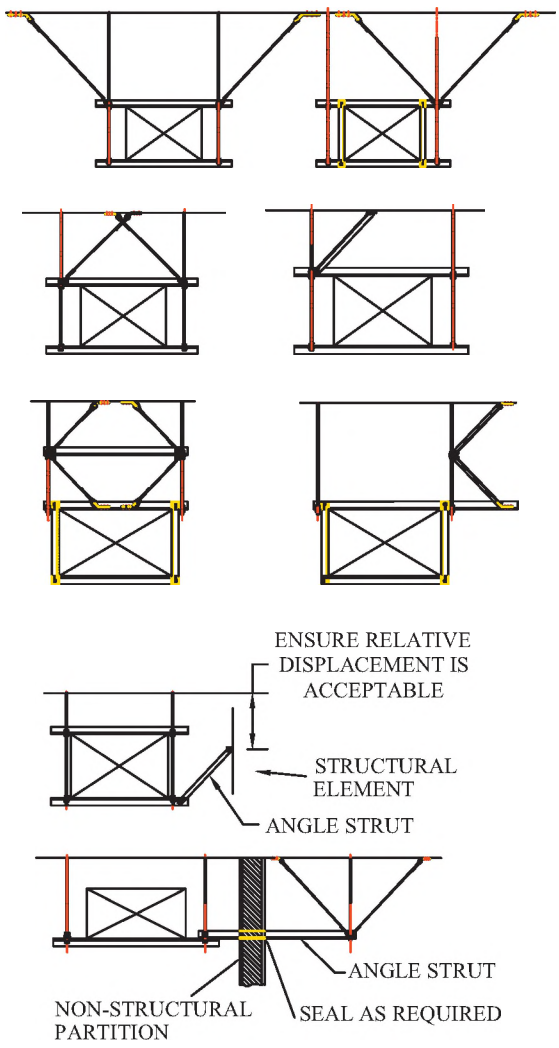


Figure 1

REFERENCES

Kinetics Noise Control Seismic Design Manual (2004)

NFPA*(National Fire Protection Association)* Standard 13 Installation of Sprinkler Systems (1999 Edition)

NRC-CNRC National Building Code (1995 + Addenda)

SMACNA*(Sheet Metal and Air Conditioning Contractors National Association)* Seismic Restraint Manual (1998)

USING A SPHERICAL MICROPHONE ARRAY FOR IDENTIFICATION OF AIRBORNE SOUND TRANSMISSION PATHS

Bradford N. Gover
Institute for Research in Construction, National Research Council, Ottawa, Ontario K1A 0R6 brad.gover@nrc-cnrc.gc.ca

1. INTRODUCTION

Assessment of the sound insulation properties of building partitions (e.g., walls, floors) is usually conducted according to standard procedures. ASTM E90 and E336 and ISO 140-3 and 140-4 provide means of rating the sound insulation performance of partitions in the laboratory or in the field. The results of these tests do not necessarily indicate whether an otherwise highly-insulating partition contains weak spots or leaks, other than a lowered overall rating. This paper presents the results of some initial efforts using a highly directional microphone array measurement system for identification of such weak spots.

Major weak spots such as holes may be obvious, but even subtle ones may be of importance. Recent work has shown that, particularly near the threshold of intelligibility, a slight increase in speech signal level can lead to a large increase in intelligibility [1]. This may have important implications on rating the "speech security" level of a meeting room, for instance. In these cases, finding even minor weak spots in the room boundaries may be of interest.

2. MICROPHONE ARRAY SYSTEM

A spherical array measurement system has been previously developed for assessment of directional variation of sound within rooms [2,3]. By measuring the impulse response at each array microphone, then beamforming, a directional impulse response due to sound arriving from within the array aperture is obtained. This is simultaneously computed for steering directions distributed over all 3D angles. Analysis of this set of directional impulse responses allows determination of the temporal and angular variations of arriving sound at the array position.

Two free-field arrays of 32 omnidirectional electret microphones, differing only in diameter, were constructed. The layout of the array microphones is shown in Fig. 1(a). Filter-and-sum beamforming is used with each array to generate a narrow beam over a different frequency range. The beampattern of the 48 cm-diameter array in the 800 Hz 1/3-octave band, and that of the 16 cm-diameter array in the 2500 Hz 1/3-octave band are shown in Fig 1(b). The 3-dB beamwidth of each pattern is only 28 degrees; the directivity index is over 14 dB.

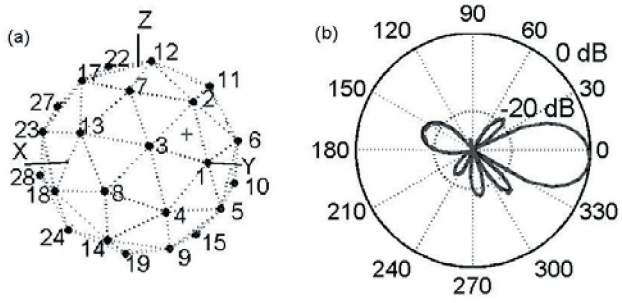


Figure 1 (a) Layout of microphones in 32-element spherical array, (b) beampattern in 1/3-octave bands: 800 Hz (48 cm array) and 2500 Hz (16 cm array). The are two curves, nearly coincident.

3. MEASUREMENTS

A 2.44-by-3.66 m test wall was constructed between two reverberation chambers at NRC. The wall had a single row of 90 mm wood studs (610 mm on centre), with a single layer of 16 mm drywall directly attached to the studs on one side, and a single layer of 16 mm drywall mounted on 13 mm resilient channels (610 mm on centre) on the other side. The cavity was completely filled with glass fibre batts. A standard ASTM E90 test was performed and the wall was found to have an STC rating of 52. Array measurements were subsequently made with the array 0.8 m from the centre of the wall in one chamber, and an omnidirectional loudspeaker 2.3 m from the wall in the other chamber. The room containing the array had a 1/3-octave band reverberation time of 4.3 s in the 800 Hz band, and 3.0 s in the 2500 Hz band. To more accurately represent typical office and residential rooms, all measurements were also made with absorptive material in the chamber containing the array. With the absorption present, the 1/3-octave band reverberation time was lowered at 800 Hz to 0.65 s, and at 2500 Hz to 0.60 s.

3.1 Partial Slit in Test Wall

A seam between two sheets of drywall on the side facing the array (the layer mounted on resilient channels) was intentionally widened to 5 mm. Such a "slit" could arise from improper cutting or installation of drywall, for instance. After this modification, the STC of the wall was measured to be 50. A directional measurement made with the absorption in the room with the array is presented in

Fig. 2. Panel (a) shows the geometry of the setup, indicating the wall and the slit with heavy lines. Panels (b) and (c) show a plot of the sound energy arriving during the first 35 ms of the response, viewed from above. The arrow indicates the expected direction of sound incidence from the slit. The shape of the surface indicates that the sound transmission through the slit is detected in both the 800 Hz and 2500 Hz 1/3-octave bands.

3.2 Partial Holes in Test Wall

After the slit was caulked and taped, two 8-by-10 cm holes were cut in the drywall, one on either side of the wall. The holes did not line up, and were cut through one layer of drywall only, the glass fibre batts were left intact. The STC of the wall in this state was still 52. A directional measurement made without the absorption in the room with the array is presented in Fig. 3. Panel (a) again shows the geometry, indicating the holes in the wall. Panels (b) and (c) show a plot of the sound energy arriving during the first 35 ms of the response, viewed from above. The arrows indicate the expected directions of sound incidence from the holes. The shape of the surface indicates that the sound transmission through the holes is detected in the 800 Hz 1/3-octave band, but not in the 2500 Hz 1/3-octave band. This is presumably due to the fact that less sound power is transmitted at 2500 Hz. The difference in level arriving from the two holes is likely due to the difference in distance of each from the array.

4. CONCLUSIONS

A spherical array measurement system has been employed to locate weak spots in a wall. In conditions where the "defects" in the wall were minor enough to cause no, or only a slight, reduction in STC, they were localized accurately. The technique was effective even under conditions of high reverberation, and showed that the "detectability" of the defects varied with frequency. These initial investigations indicate that the present approach seems capable of identifying weak spots in a reliable and effective manner.

REFERENCES

[1] B.N. Gover and J.S. Bradley, "Measures for assessing architectural speech security (privacy) of closed offices and meeting rooms," J. Acoust. Soc. Am. (accepted August 2004).

[2] B.N. Gover, J.G. Ryan, and M.R. Stinson, "Microphone array measurement system for analysis of directional and spatial variations of sound fields," J. Acoust. Soc. Am., **112**, 1980–1991 (2002).

[3] B.N. Gover, J.G. Ryan, and M.R. Stinson, "Measurements of directional properties of reverberant sound fields in rooms using a spherical microphone array," J. Acoust. Soc. Am. (in press).

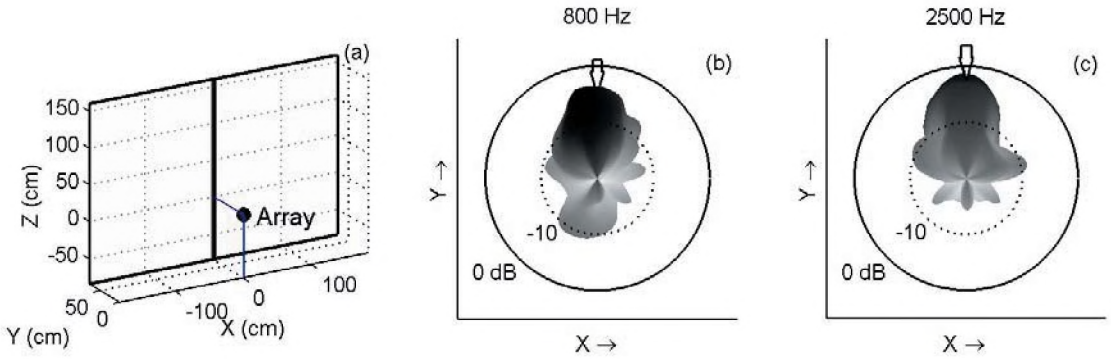


Figure 2 Measurement with slit in one side of test wall: (a) geometry of setup. Energy arriving at array during first 35 ms of impulse response is shown in (b) 800 Hz and (c) 2500 Hz 1/3-octave band.

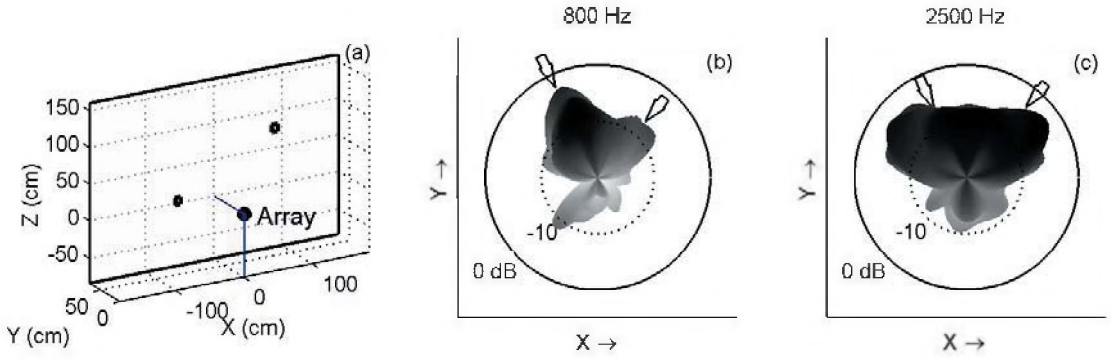


Figure 3 Measurement with two partial holes in test wall: (a) geometry of setup. Energy arriving at array during first 35 ms of impulse response is shown in (b) 800 Hz and (c) 2500 Hz 1/3-octave band

LES LIEUX DE LA MUSIQUE AU QUÉBEC ; ACOUSTIQUE DE SIX SALLES DE CONCERT

Jean-François Hardy et Jean-Gabriel Migneron
Laboratoire d'acoustique de l'Université Laval, École d'architecture, Université Laval, 1 Côte de la Fabrique, Québec, (Québec), Canada, G1K 7P4

1. INTRODUCTION

Les salles de concert fournissent une terre fertile à l'exploitation des connaissances acquises en acoustique architecturale. Dans la province de Québec, il existe aujourd'hui un grand nombre de salles qui accueillent les différentes activités musicales classiques. Avec l'institutionnalisation de la musique classique au Québec, au milieu du XIX^{e} siècle, un réseau de salles utilisées en partie pour les concerts commença à se tisser dans les centres urbains de Montréal, Québec, Trois-Rivières et Sherbrooke. Ces lieux possédaient une acoustique souvent inadéquate pour la présentation de concerts. Ils furent remplacés par des salles publiques construites avec la participation des gouvernements, suite à la démocratisation de la musique classique et à l'accroissement des auditoires. Une meilleure formation pour les musiciens québécois, la multiplication des ensembles musicaux professionnels et le soutien des médias pour la musique classique créèrent un dynamisme qui sera à l'origine de la construction de différentes salles de concert.

La période musicale contemporaine au Québec débute avec de grands changements dans la société québécoise. La Révolution Tranquille, la perte du pouvoir politique et culturel de l'église catholique, le mouvement du Refus Global et l'ouverture de la province sur le monde, qui culmine avec Expo '67, sont autant de facteurs qui favorisent de grands projets de société. Depuis lors, de nombreuses salles à l'acoustique plus conforme à la musique classique verront le jour. Les salles décrites dans cette étude sont issues de projets culturels divers de cette période moderne. Il s'agit de salles de concert construites à l'intérieur de grands complexes culturels, de salles conçues pour des institutions d'enseignement ou de salles bâties pour des festivals de musique estivaux. Leur analyse acoustique permettra d'évaluer leur comportement et les différentes solutions de design, qui apportent un caractère unique à chacun de ces lieux musicaux.

2. MÉTHODE

La méthode choisie s'inspire, pour le choix de la procédure globale et des indices de mesure, de quelques-uns des nombreux travaux réalisés à ce jour sur des groupes de salles de concert. Les publications de L.L. Beranek, M. Barron et J.S. Bradley citées dans la bibliographie furent consultées pour dresser une méthodologie cohérente.

2.1 Échantillon

Cette étude offre une description de l'acoustique de quelques lieux musicaux québécois dont la majorité n'ont jamais bénéficié de l'apport de relevés acoustiques. Il s'agit d'un groupe de salles qui se distinguent les unes des autres par leurs formes et leurs dimensions, mais elles sont toutes utilisées fréquemment pour la présentation de concerts classiques. Le tableau 1 en décrit les principales caractéristiques.

Tableau 1. Salles de l'échantillon et typologies.

Salle	Forme et année d'inauguration	Nombre de places	Volume (m^3) approximatif
Amphithéâtre de Lanaudière (AL)	Amphithéâtre extérieur, 1989	2000	20000
Wilfrid-Pelletier (WP)	Éventail, 1963	2982	26500
Louis-Fréchette (LF)	Rectangulaire, 1971	1875	19800
Claude-Champagne (CC)	Fer à cheval, 1964	1010	10000
Pollack (POL)	Rectangulaire, 1975	600	7000
Françoys-Bernier (FB)	Rectangulaire, 1996	604	13600

2.2 Relevés quantitatifs

Les indices mesurés dressent un portrait sommaire des caractéristiques acoustiques des salles de concert de l'échantillon. Il s'agit des temps de réverbération globaux et initiaux, des décroissances des niveaux de pression, des réponses impulsionnelles, des indices de clarté, des indices d'intelligibilité et des niveaux de bruit de fond. Différents appareils, tels que les analyseurs B&K 2133 et TEF-20, ont été utilisés.

2.3 Écoute

Des commentaires sur l'acoustique de ces salles sont notés suite à l'audition de nombreux concerts aux caractéristiques musicales différentes. Diverses positions d'écoute servent à évaluer l'uniformité des ambiances acoustiques dans les salles.

3. RÉSULTATS DES RELEVÉS

Les résultats moyens de l'ensemble des points mesurés sont exprimés dans le tableau 2. Comme la salle Françoys-Bernier possède un grand nombre de rideaux rétractables, les mesures de certains indices sont disponibles dans les deux configurations extrêmes.

Tableau 2. Résultats moyens des mesures quantitatives.

Salles	TR (sec.)	EDT (sec.)	Niv. press. (dB)	C_{80} (dB)	STI (%)	BF (NC)
AL	2,51	1,16	16,6	5,83	67	33
WP	1,88	1,73	17,6	1,13	55	21
LF	1,34	1,22	19,4	3,94	65	17
CC	1,47	1,01	18,0	7,30	68	33
POL	1,95	1,51	12,6	3,65	61	17
FB fermés	1,49	0,88	16,8	6,85	71	<10
FB ouverts	2,38	0,72	n/d	n/d	n/d	<10

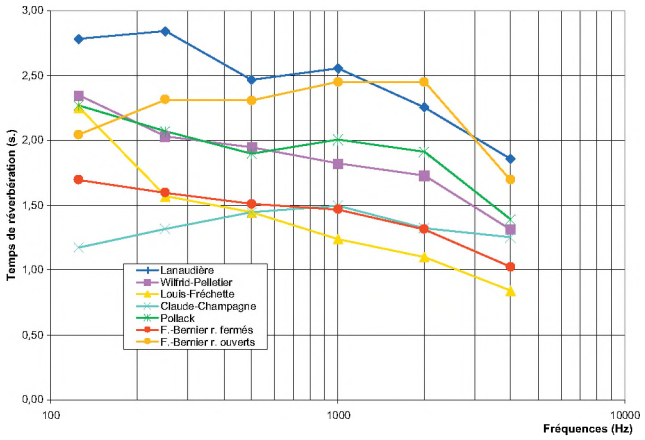


Fig. 1. Temps de réverbération en fonction de la fréquence.

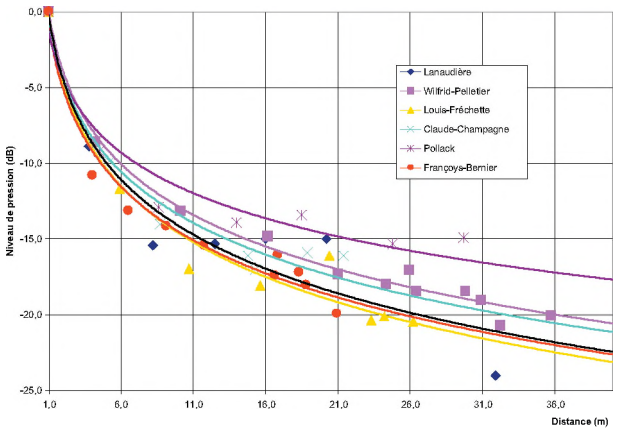


Fig. 2. Atténuation des niveaux de pression à 1 kHz dans l'axe central.

Les temps de réverbération mesurés à l'octave (fig. 1) démontrent que la plupart des salles ont peu de support de la réverbération dans les basses fréquences. Ils permettent aussi de constater la différence entre les deux configurations de la salle Françoys-Bernier. Dans le graphique de l'atténuation des niveaux en fonction de la distance (fig. 2) on constate que les plus petites salles ne sont pas nécessairement celles qui conservent les niveaux sonores les plus élevés. Cela dépend aussi de la diffusion des surfaces et de la quantité de réverbération présente.

4. RÉSULTATS DE L'ÉCOUTE

L'ensemble des salles de l'échantillon possède des caractéristiques acoustiques favorables à la présentation de concerts classiques. Les mieux adaptées sont tout de même les salles conçues spécifiquement dans la perspective d'un tel usage, ce qui n'est pas forcément le cas des grandes salles multifonctionnelles comme Wilfrid-Pelletier et Louis-Fréchette. Par contre, ces dernières sont dotées d'équipements scéniques qui permettent de pallier les carences causées par un manque de surfaces réfléchissantes.

Ainsi, chacun de ces lieux musicaux québécois est en mesure d'offrir un espace favorable aux performances artistiques des musiciens classiques et ils bénéficient d'un caractère acoustique spécifique qui fait leur âme propre : Lanaudière pour un son riche au niveau intéressant, Wilfrid-Pelletier et Louis-Fréchette pour un son direct franc et puissant, Claude-Champagne pour la qualité des timbres des instruments et des voix, Pollack pour la grande présence sonore et l'enveloppement agréable et Françoys-Bernier pour le caractère unique de sa réverbération variable.

RÉFÉRENCES

Barron, M. (1988), « Subjective study of british symphony concert halls », Acustica, 66, 1, 1-14.

Barron, M. (1993), *Auditorium Acoustics and Architectural Design*, Londres : Spon, 443p.

Beranek, L. L. (1992), « Concert hall acoustics – 1992 », Journal of Acoustical Society of America , 92, 1, 1-39.

Beranek, L. L. (2003), 2è éd., *Concert and Opera Halls : Music, acoustics and architecture*, New York : Springer, 661p.

Bradley, J. S. (1991), « A comparison of three classical concert halls », Journal of Acoustical Society of America , 89, 3, 1176-1192.

Migneron, J. G., Woodcock, R. et Asselineau, M. (1986), « Analyse acoustique du Grand Théâtre de Québec avec utilisation de l'intensimétrie acoustique pour l'évaluation des réflecteurs », International Congress on Acoustics, 12, 99-103.

Vincent, O. (2000), *La vie musicale au Québec – Art lyrique, musique classique et contemporaine*, Québec : Éditions de L'IQRC, 157p.

REMERCIEMENTS

Les auteurs tiennent à remercier les responsables des salles visées par cette étude et les collègues qui les ont assistés lors de la cueillette des données pour leur aimable participation à ce projet.

EFFET DE L'ECLAIREMENT ET DU TEINT DU LOCUTEUR SUR L'INTELLIGIBILITE VISUELLE DE LA PAROLE

Ariane Laplante-Lévesque et Jean-Pierre Gagné
École d'orthophonie et d'audiologie, Faculté de médecine, Université de Montréal
C.P. 6128, Succursale Centre-Ville Montréal (Québec) H3C 3J7

1. INTRODUCTION

L'ajout d'indices visuels aux indices auditifs améliore de façon significative les performances recueillies lors de tâches de perception de la parole (Sumby et Pollack, 1954 et autres). Les indices visuels de la parole permettent, entre autres, la lecture labiale. En 1997, une compagnie québécoise, *Audisoft Technologies,* a développé un système MF audio-visuel qui permet aux personnes ayant une déficience auditive d'avoir accès aux indices visuels de la parole en tout temps (voir figure 1). *AudiSee* consiste en un casque muni d'une mini-caméra et d'un microphone, ces derniers étant fixés sur une tige placée devant le visage du locuteur. L'image (le visage du locuteur) parvient à un moniteur par ondes MF. Le signal sonore est aussi transmis. Cependant, seuls les signaux visuels furent à l'étude lors du projet ici présenté.

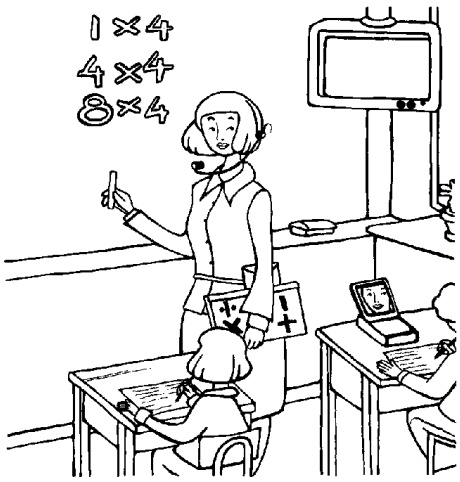


Fig. 1: Le système MF audio-visuel *AudiSee* (voir http://www.audisoft.net pour de plus amples informations)

Lors d'une observation du système *AudiSee* sur le terrain, des étudiants malentendants ont rapporté qu'il était parfois difficile d'extraire les indices visuels de la parole lorsque la quantité de lumière dans la classe était diminuée afin de permettre l'utilisation d'un rétroprojecteur ou d'un vidéoprojecteur (Gagné, Le Monday, Boisclair, Gagnon, Lapalme et Ducas, 2000). Des essais informels effectués en laboratoire ont suggéré que la quantité de lumière (ex.: éclairage important ou éclairage réduit) et le teint de la peau du locuteur (ex.: teint de type caucasien ou teint de type africain) peuvent compromettre la perception visuelle de la parole.

Différentes mesures photométriques sont utilisées. L'éclairement constitue la quantité de lumière en un point spécifique et se mesure en lux. La luminance constitue la quantité de lumière réfléchie par un objet et se mesure en candela par mètre carré. Dans le cadre d'expérimentations reliées à la perception visuelle de la parole, l'éclairement et la luminance sont mesurés au visage du locuteur.

À ce jour, peu d'études ont évalué l'effet de la quantité de lumière sur la perception visuelle de la parole. Un biais méthodologique a été introduit dans une des études visant à évaluer l'effet de l'éclairement (Thomas, 1962, cité dans Berger, 1972). Erber (1974) a mesuré l'effet de la luminance sur les performances en perception visuelle de la parole de cinq adolescents ayant une déficience auditive profonde. Les résultats de cette étude ont démontré une diminution significative des performances des sujets lorsque que la luminance a été réduite de 0,1 à 0,03 cd/m^2, ce qui constitue une très faible quantité lumineuse. Récemment, McCotter et Jordan (2003) ont comparé l'effet de certains paramètres de luminance sur les performances en perception visuelle de la parole. Leurs résultats suggèrent que la distribution des contrastes de luminance sur le visage du locuteur pourrait être critique pour la perception visuelle de la parole.

Une seule étude portant sur l'effet du teint du locuteur sur les performances en perception visuelle a été répertoriée (Berger, Perry, Hoffman et Smith, cité dans Berger, 1972). La description limitée de l'étude rend difficile l'émission de conclusions.

Les objectifs de la présente étude sont: 1) de mesurer l'effet de l'éclairement au visage du locuteur sur les performances en perception visuelle de la parole, 2) de mesurer l'effet du teint du locuteur sur les performances en perception visuelle de la parole et 3) d'explorer la possibilité d'une interaction statistique entre l'éclairement au visage du locuteur et le teint de ce dernier sur les performances en perception visuelle de la parole.

2. MÉTHODOLOGIE

Quinze sujets devant compléter une épreuve de perception visuelle de la parole ont été recrutés. Les sujets étaient tous francophones, âgés de 18 à 50 ans et avaient une audition périphérique normale et une vision périphérique normale ou normale lorsque corrigée. Un ensemble de phrases comportant 3 mots clés chacune (sujet, verbe et adjectif) a été utilisé. Puisque 7 sujets, 7 verbes et 7 adjectifs différents étaient disponibles, 343 phrases pouvaient être générées. Les 14 conditions expérimentales incluaient 7 niveaux du facteur Éclairement (22, 32, 43, 172, 646, 2755 et 6456 lux) et 2 niveaux du facteur Teint du locuteur (teint de type caucasien et teint de type africain). Les enregistrements ont été réalisés à l'aide d'un système AudiSee et 280 phrases ont été enregistrées (20 phrases pour chacune des 14 conditions expérimentales). Une seule locutrice a énoncé l'ensemble des phrases afin de minimiser la variabilité inter-individuelle d'intelligibilité de la parole. Pour modifier le facteur Teint du locuteur, la locutrice a été enregistrée soit à l'état naturel (teint de type caucasien) ou soit maquillée par une professionnelle afin de reproduire un teint de type africain. Pour modifier le facteur Éclairement, la quantité de lumière fut variée et mesurée au visage de la locutrice grâce à un photomètre. Les stimuli ont été visionnés en ordre aléatoire par les 15 sujets. Ces derniers étaient invités à choisir les mots clés qu'ils croyaient avoir vus à partir d'un choix fermé. Il s'agissait donc d'une tâche d'identification visuelle de la parole.

3. RÉSULTATS

Les performances moyennes pour chacune des 14 conditions expérimentales sont présentées à la figure 2. Une ANOVA à mesure répétée impliquant deux facteurs a révélé une interaction significative entre les facteurs Éclairement et Teint du locuteur. Des comparaisons pairées (tests t) en effectuant un ajustement de Bonferroni ont révélé des différences significatives entre les deux conditions du facteur Teint du locuteur à quatre divers niveaux d'éclairement: 32, 43, 172 et 2755 lux.

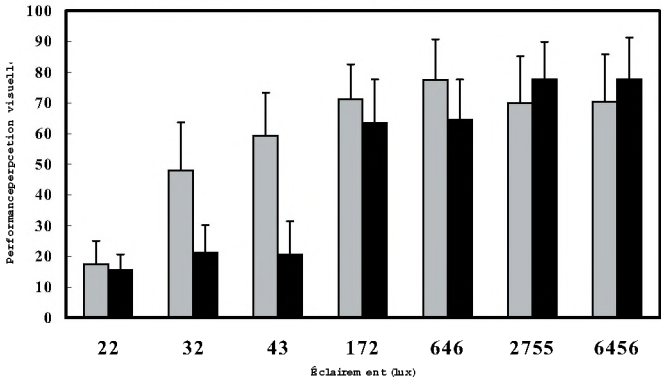


Fig. 2: Performancesmoyennes en perception visuelle de la parole en fonction de l'éclairement. Les barres grises représentent la condition teint de type caucasien et les barres noires, la condition teint de type africain.

4. DISCUSSION ET CONCLUSION

De façon générale, les performances moyennes augmentent en fonction de l'éclairement. Lorsque l'éclairement est moindre, les performances chutent, et ce pour les deux conditions du facteur Teint du locuteur. Lorsque l'éclairement est très faible (22 lux), les performances moyennes pour chacune des conditions du facteur Teint du locuteur sont similaires. Les différences observées lorsque l'éclairement est important (2755 lux) ne sont probablement pas significatives d'un point de vue clinique. L'effet du teint du locuteur sur les performances est surtout apparent à des niveaux d'éclairement intermédiaires, soit entre 32 et 172 lux. À ces niveaux d'éclairement, les performances en perception visuelle sont plus élevées pour la condition teint de type caucasien que pour la condition teint de type africain. Ces résultats peuvent potentiellement être expliqués par un plus important contraste de luminance dans ces conditions d'éclairement lorsque la locutrice a un teint de type caucasien. Ces contrastes de luminance pourraient faciliter l'identification des articulateurs.

Cette recherche démontre que l'éclairement et le teint du locuteur influencent les performances en perception visuelle de la parole. D'autres études sont nécessaires afin de mieux cerner l'effet des contrastes de luminance sur la perception visuelle de la parole.

REFERENCES

Berger, K.W. (1972). *Speechreading: Principles and methods.* Baltimore: National Educational Press.

Erber, N.P. (1974). Effects of angle, distance, and illumination on visual reception of speech by profoundly deaf children. *Journal of Speech and Hearing Research, 17*, 99-112.

Gagné, J.-P., Le Monday, K., Boisclair, A., Gagnon, L., Lapalme, M. et Ducas, L. (2000). *Projet d'expérimentation du système MF-visuel : Rapport final.* Déposé à l'Office des personnes handicapées du Québec.

McCotter, M.V., and Jordan, T.R. (2003). The role of facial color and luminance in visual and audiovisual speech perception. *Perception, 32*, 921-936.

Sumby, W.H. et POLLACK, I. (1954). Visual contributions to speech intelligibility in noise. *Journal of the Acoustical Society of America, 26*, 212-215.

REMERCIEMENTS

Norman P. Erber, Katrine Doucet, Maude Labelle, Marie-Christine Potvin et *Audisoft Technologies.* Financement provenant du CRSNG et de la Faculté de Médecine de l'Université de Montréal.

DOES A CONTINUOUS MASKER MAKES SPEECH COMPREHENSION IN NOISE EFFORTFUL?

Antje Heinrich and Bruce A. Schneider
Department of Psychology, University of Toronto at Mississauga, 3359 Mississauga Rd., Mississauga, ON, Canada, L5L 1C6
aheinrich@utm.utoronto.ca

1. INTRODUCTION

A common complaint among older adults is that it is difficult for them to understand spoken language, especially in a noisy environment (CHABA, 1988). To comprehend speech in noise listeners not only have to attend to the target speech while ignoring distractors, they also have to extract the meaning from the individual words and phrases and store the result in memory for future use. It is possible that task of segregating the speech signal from the background, because it might require the redeployment of cognitive resources to perceptual processes, could adversely affect memory of the heard material, thereby reducing comprehension on a cognitive level. To see whether the effort of extracting the speech signal from noise adversely affected memory, Murphy, Li, Craik and Schneider (2000) investigated recall performance in a paired-associate memory paradigm when the word pairs were presented in quiet or embedded in a background babble presented at a signal-to-noise ratio (SNR) of –7 dB. Murphy et al. demonstrated that memory performance was especially affected for word pairs presented early on the list (early serial positions) when the word were presented in a background babble compared to the quiet baseline condition. In fact, presenting the words in noise decreased young adults' memory performance to a level comparable to old adults' when they listened to the word pairs in quiet.

Two explanations are conceivable for this result. First, the babble may have degraded the perceptual representation of the words, and an instable or incomplete representation compromised their subsequent processing. Second, the young adults may have had to engage top-down processes when listening to the words in background noise to recover the incomplete or distorted words. These top-down processes may have taken up cognitive resources that otherwise would have been available for the encoding of the words.

We tested both hypotheses. To test the first hypothesis we temporally distorted the word pairs. This manipulation decreased perceptual accuracy of the words to the same extent as the babble, yet, no masker was present from which to extract the words. The results are published elsewhere (Heinrich and Schneider, submitted) and, in short, show that temporal distortion did not adversely affect recall as long as the words were presented on a moderate intensity level. Based on these results we concluded that it was not the degraded perceptual representation of the words themselves that led to the decrement in memory performance in the Murphy et al. study. Rather, it was the effort in extracting the words from the background babble that caused the drop in memory performance. Evidence that a noise masker may involuntarily engage top-down processes to improve accuracy of stimulus perception stems from research on stimulus detection in gated noise. A number of studies show that detection performance is improved when a noise masker is gated on prior to stimulus onset compared to simultaneous onset of stimulus and masker (e.g., Zwicker, 1965). This result suggests that listeners track the noise in order to minimize its effect on stimulus perception. We speculated that a similar mechanism operated when words were presented in babble– the auditory system tracked to background babble between words to enable it to be better able to extract the target words. We reasoned that if this tracking process was effortful (requiring considerable cognitive processing) that fewer resources would be available to store these words in memory.Hence we would expect a drop in memory performance under this condition.

In order to test our hypotheses, we presented the word pairs in discontinuous babble that was present only during word presentation but not in the interval between words. This preserved the masking character of the babble at word presentation but again did prevent the auditory system from effectively being to monitor the babble over longer periods of time in order to minimize its influence on word perception. Subsequently, we compared recall performance from this condition with memory in conditions in which word pairs were presented in quiet and in a continuous babble.

2. METHOD

2.1 Participants

Sixteen young adults (mean age: 19.69 years; s.d.: 1.35, 11 females) were tested in this experiment. All individuals were undergraduate students at the University of Toronto, had completed an average of 15.06 years of education (s.d.: 1.88), and scored 13.38 (s.d.: 1.62) on the Mill-Hill vocabulary test. In exchange for their participation they received $10 per hour. All participants spoke English before the age of five. The demographic data are

comparable to the young adults tested by Murphy et al. in the other two conditions.

2.2 Material

The same word pairs as in Murphy et al. (2000) were used. The material consisted of a total of four hundred two-syllable common nouns arranged in 40 lists of 5 word pairs each. The words were digitized at a sampling rate of 20 kHz. For stimulus presentation, words as well as babble, stimuli were delivered through a 16-bit digital-to-analog converter (TDT DD1) followed by a 10-kHz low-pass filter (TDT FT6-2, 60 dB attenuation at 11.5 kHz), a programmable attenuator (TDT PA4), and a weighted signal mixer (TDT SM3). All testing took place in a double walled sound attenuating chamber.

2.3 Procedures

For each listener pure-tone air-conduction thresholds were determined for nine frequencies between 250 and 8000 Hz in the left and right ear. All listeners had pure tone audiometric hearing thresholds within the normal range (< 25 dB HL) for frequencies between 250 and 8000 Hz.. Thresholds for the detection of speech babble were also determined for each individual. A recording of a twelve-talker babble, taken from the modified Speech Perception in Noise (SPIN) test (Bilger, Nuetzel, Rabinowitz, & Rzeczkowski, 1984) was used. The speech threshold was obtained for the right ear first and the left ear second. For the presentation of the words in babble, the presentation level of the words was individually adjusted for each participant so that each word was presented 50 dB above the individual's babble threshold. The babble level was also individually adjusted based on each individual's low-context SPIN threshold to reflect a SNR=-7 dB. For more details on the adjustment procedure please refer to Murphy et al. (2000). For the memory experiment, all participants listened to forty lists containing five word pairs each. In the encoding phase, word pairs were played with a presentation rate of 4 seconds per pair. The two words in a pair were separated by a silence period of 100 ms. The words were randomly paired, any obvious association between two words was avoided. In the retrieval phase, participants were cued with the first word from one of the five previously presented word pairs and were asked to recall the second word of the pair. Only one pair from each list was cued. There was no time limit placed on the recall. Participants were encouraged to guess.

3. RESULTS

The recall performance under discontinuous noise conditions displayed a primacy as well as a recency effect. Correct recall was close to 45% for the first two word pairs in a list, dropped to 38% for the third word pair, and went up to 54% and 79% for serial positions 4 and 5. This current result together with the results from the quiet baseline and continuous noise conditions from the Murphy et al. study was submitted to a mixed measures ANOVA with three background conditions (quiet, continuous noise, discontinuous noise) as between and serial position as within-subject variables. The results show a main effect of serial position ($F(4, 172)= 67.9$, $p<0.001$), condition ($F(2, 43)=4.5$, $p<0.05$), and a significant interaction between the two factors ($F(8, 172)=2.6$, $p<0.05$). Post-hoc test reveal that the main effect of background condition is caused by significant differences in memory recall between quiet and discontinuous noise on the one hand and continuous noise on the other hand. There are no significant differences in memory between the former two conditions. Moreover, the interaction effect is due to the fact that memory performance under continuous noise is considerably decreased for early serial positions (word pairs 1 to 2) compared to the other two conditions but not for serial positions 3 to 5.

4. DISCUSSION

The results demonstrate that it is the continuity of the noise that caused the drop in memory. When the noise was only present during word presentation but not in the interval between words, recall was not affected compared to the baseline condition. Under both these conditions the memory performance is considerably higher than when the words are masked with a continuous masker. Why does a continuous masker cause a drop in recall when a discontinuous masker does not? We assume that the auditory system uses cues available in the continuous babble in order to improve word perception. To do this, it monitors the babble in intervals where no words are presented so that it can use abrupt changes in the auditory stream to identify when words are being presented. However, this process is effortful and resource demanding and resources invested in minimizing the influence of the background noise are not available for memory processing. This lack of resources at encoding leads to the decrement in memory performance

REFERENCES

Bilger, R.B., Nuetzel, M.J., Rabinowitz, W.M., & Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. Journal of Speech and Hearing Research, 27, 32-48.

CHABA (Committee on Hearing, Bioacoustics, and Biomechanics) Working Group on Speech Understanding and Aging, National Research Council (1988). Speech understanding and aging. Journal of the Acoustical Society of America, 83, 859-895.

Murphy, D.R., Craik, F.I.M., Li, K.Z.H. & Schneider, B.A. (2000). Comparing the effects of aging and background noise on short-term memory performance. Psychology and Aging, 15 (2), 323-334.

Zwicker, E. (1965). Temporal effects in simultaneous masking by white-noise bursts. Journal of the Acoustical Society of America, 37(4), 653-663.

THE INFLUENCE OF SPECTRAL AND TEMPORAL ACUITIES IN HEARING ON SPEECH INTELLIGIBILITY

Gaston Hilkhuysen, Tammo Houtgast, and Joost Festen
Dept. of Otolaryngology, VU University Medical Center, de Boelelaan 1117, 1081 HV Amsterdam, the Netherlands
G.Hilkhuysen@vumc.nl

1. INTRODUCTION

Hearing impaired listeners with cochlear hearing loss (HI) often complain about their abilities to understand speech, especially in the presence of concurrent background noise. Objective measurements underline their subjective findings: to obtain equal speech intelligibility HI need higher speech to noise ratios (S/N) than normal hearing persons (NH). In a model Plomp (1978) distinguishes two factors that account for HI's deteriorated speech intelligibility: attenuation and distortion.

Elevated hearing thresholds illustrate attenuation. If cochleae are less sensible, speech becomes inaudible. Mere amplification, as provided by hearing aids, restores audibility. But in general HI still perform less even when speech is presented well above their hearing thresholds, giving rise to the idea that impaired cochleae distort speech. Two possible sources of distortion are impairments of spectral resolution and temporal neural coding.

Models of speech intelligibility, like the Speech Intelligibility Index (ANSI, 1997) and the Speech Transmission Index (Steeneken & Houtgast, 1980), stress the transfer of speech modulations in different frequency bands. Reduced spectral or temporal acuities can impair this transfer. For HI with reduced spectral acuity, modulations mix over frequency bands, contaminating frequency-specific modulations. Reduced temporal acuity might diminish the modulation depth especially of higher modulation frequencies within a band.

Noordhoek et al. (2001) measured spectral and temporal acuities as well as speech intelligibility for a small group of HI. She found that in this group 1) spectral and temporal acuity scores are independent and 2) impaired spectral or temporal acuity is associated with reduced speech intelligibility. The current study tries to confirm these findings with a large group of naïve HI.

2. METHOD

Speech intelligibility is measured in two ways. 1) The Speech Reception Threshold test in speech fluctuating noise (SRT_{fluc}) estimates the S/N that gives 50 % intelligibility in speech-like modulated masking noise (Festen & Plomp 1990). We expect that the SRT_{fluc} test puts demands on listeners' temporal acuities in particular. 2) The Speech Reception *bandwidth* Threshold (SR*b*T) test measures the bandwidth needed for 50 % intelligibility employing band-pass filtered speech centered around 1 kHz in complementary band-stop filtered masking noise (Noordhoek et al. 2001). We anticipate poor SR*b*T scores to coincide especially with impaired spectral acuity.

A recently developed clinically applicable test measures spectral and temporal acuities of listeners at a specified frequency (Hilkhuysen et al., submitted). The current study focuses on acuities around 1 kHz. Listeners detect sweeps in spectrally or temporally modulated noise grids. Detection thresholds reveal the widths of auditory filters and time windows.

The long-term average spectrum of masking noises and signals at 0 dB S/N is placed halfway between the listener's hearing threshold and uncomfortable loudness level (UCL), that is in the middle of the listeners dynamic range. This eliminates in both SRT tests audibility effects on speech intelligibility. Spectral and temporal acuities vary with intensity. In order to acquire acuity estimates at levels present during the SRT tests, the masking noise in the spectral and temporal acuity test is also positioned in the middle of the listener's dynamic range.

Only HI participated who have at least 30 dB dynamic ranges and whose speech intelligibilities in quiet are 80 % or more. Local standard clinical audiometry provides this information. The study focuses on the ear with the highest intelligibility in quiet. A preliminary 46 HI, who visited the hearing center for hearing-aid adjustment, subsequently carried out all tests twice.

3. RESULTS

Final scores are means over double measurements: metric means for SRT_{fluc} scores and geometric means for the other tests. Statistical analyses employ logarithmically transformed SR*b*T and spectral and temporal acuity scores. Table 1 shows Pearson correlations, the corresponding scatterplot matrix (not included here) shows homogeneous clouds of points. Entries on the diagonal of Table 1 are

Table 1. Correlation between test scores (n=46). Diagonal entries represent reliability estimates. Test reliability influences ordinary off-diagonal correlations. Estimates of correlations that would be obtained with 100 % reliable tests are parenthesized. Numbers between brackets represent 95% confidence intervals for uncorrected correlations.

	SR*b*T	SRT_{fluc}	spectral	temporal
SR*b*T	.80 [.66-.89] (1.00)	.63 [.42:.80] (.73)	.59 [.36:.75] (.68)	.23 [-.06:.49] (.27)
SRT_{fluc}		0.92 [.85-.95] (1.00)	.42 [.25:.63] (.46)	.24 [-.05:.50] (.25)
spectral			.93 [.88-.96] (1.00)	.24 [-.06:.49] (.25)
temporal				.96 [.92-.98] (1.00)

reliability estimates based on correlations between test and retest scores using Spearman-Brown prophecy formula. Reduced reliabilities attenuate the off-diagonal correlations, corrected estimates are expressed by numbers in parentheses.

The two SRT tests correlate highest, both measure partly the same phenomenon. Spectral acuity correlates well with SR*b*T and moderately with SRT_{fluc}, but in neither case enough to account for all differences in speech intelligibility among HI. Correlations of temporal acuity with other tests are low.

4. DISCUSSION

The present study finds reliable differences in temporal acuities among HI, but these differences have no influence on speech intelligibility. Looking at speech intelligibility models one might theorize that reduced temporal acuity possibly causes distortion, but the current data find no evidence for this hypothesis. HI with reduced temporal acuities seem to have comparable access to those speech modulations that are essential for speech intelligibility as HI with normal temporal acuities. However, it is remarkable that temporal acuity varies independently from spectral acuity, implying impairments of different auditory mechanisms.

The current results confirm the findings of Noordhoek et al (2001) that reduction of spectral acuity distorts speech intelligibility. It deteriorates in particular SR*b*T scores, simply explained by spectral masking. Noise from lower frequencies masks the modulations in the enclosed band-passed filtered speech. At the same time energy from low-frequencies within the band of speech can mask speech modulations in the higher part of the band. Consequently HI with reduced spectral acuities need broader bands of speech to acquire equal intelligibility, that is to obtain an equal amount of speech modulations.

Reduction in spectral acuity appears to be one source of distortion but explains only part of the problems that HI experience while listening to speech in concurrent noise. Among the speech intelligibility tests SRT_{fluc} scores have the lowest proportion of explained variance, hence SRT_{fluc} scores seem more sensitive to this unknown distortion source.

REFERENCES

ANSI (1997). ANSI S3.5-1997, American national standard methods for calculation of the speech intelligibility index (American National Standards Institute, New York).

Festen, J.M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. J Acoust Soc Am, 88, 1725-36.

Hilkhuysen, G.L.M., Houtgast, T., & Festen, J.M. (submitted). Fast and reliable measurements of spectral and temporal acuities in noise for naïve listeners. Submitted to J Acoust Soc Am.

Noordhoek, I.M., Houtgast, T., & Festen, J.M. (2001). Relations between intelligibility of narrow-band speech and auditory functions, both in the 1-kHz frequency region. J Acoust Soc Am, 109, 1197-212.

Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. J Acoust Soc Am, 63, 533-49.

Steeneken, H.J., & Houtgast, T. (1980). A physical method for measuring speech-transmission quality. J Acoust Soc Am, 67, 318-26.

ACKNOWLEDGEMENTS

The Foundation "Heinsius-Houbolt Fonds" in the Netherlands supported this research. We thank Lieselot van Deun for testing part of the 46 HI, as well as Hans van Beek who implemented the SR*b*T and SRT_{fluc} tests in a user-friendly Delphi program.

HEARING ONE OR TWO VOICES : F0 AND VOWEL SEGREGATION IN YOUNGER AND OLDER ADULTS

Tara Vongpaisal and Kathy Pichora-Fuller
Dept. of Psychology, University of Toronto at Mississauga, 3359 Mississauga Rd. N., Ontario, Canada, L5L 1C6.
tarav@psych.utoronto.ca; kpfuller@utm.utoronto.ca

In the present study, we investigate the effect of age on the ability to detect differences in the fundamental frequency (F0) of vowels and the use of this cue to identify simultaneous vowels.

Older listeners with good hearing in the speech range experience relatively little difficulty understanding one talker in a quiet listening environment. However, they report difficulty understanding speech in multi-talker situations. Age-related changes in auditory temporal processing are believed to contribute to these difficulties (Schneider & Pichora-Fuller, 2001; Pichora-Fuller & Souza, 2003). Periodicity coding is one aspect of auditory temporal processing. It enables a listener to use information conveyed by the F0 and harmonic structure of speech. This information is important for the perception of voice pitch and quality (Assmann & Summerfield, 1990). Voice pitch and quality may help listeners to segregate voices in multi-talker situations.

Compared to normal hearing listeners, hearing impaired listeners are less able to detect differences in F0—a problem that is associated with the reduced ability to identify concurrently spoken vowels (e.g., Summers & Leek, 1998). Although Summers and Leek (1998) set out to examine the effects of hearing loss on F0 detection and concurrent vowel identification, they noticed that regardless of audiometric status, older listeners performed more poorly than younger listeners. In the present study, the effect of age on F0 detection and its relationship to vowel identification is directly investigated in listeners with good audiograms.

Experiment 1

Method

Participants. Fifteen younger adults (M = 26 years of age, SD = 3.2) and 15 older adults (M = 74 years of age, SD = 5.6) were recruited from the Mississauga area. All had good hearing with pure-tone, air-conduction thresholds ≤ 25 dB HL between .25 and 3 kHz in their better ear. These participants also completed Experiment 2.

Stimuli and Apparatus. Tokens of the vowel [a] were synthesized using five fixed formant frequencies (Assmann and Summerfield, 1994) and with F0 varying from 120 to 130Hz in increments of .1 Hz. The synthesized vowel stimuli were presented through a TDT System II and presented monaurally through TDH49 headphones at 80 dB SPL. Testing took place in a double-wall sound-attenuating booth (Industrial Acoustics Corporation).

Procedure. A practice block of 100 trials was administered before the test phase. Each trial of the practice consisted of three successively presented tokens (duration = 260 ms) of the vowel [a]: the standard token (F0 = 120 Hz) followed by two comparisons tokens. The interstimulus interval was 150 ms. One of the comparison tokens matched the standard while the other differed from the standard (F0 = 145 Hz). Listeners indicated which of the two comparison tokens was different from the standard by pressing the corresponding button on a button box. Feedback was given after each trial. In the test phase, the F0 difference limen threshold for each participant was determined using an adaptive task. The initial step size was 30 Hz and the step size on subsequent trials was halved following a correct response or doubled following three incorrect responses. After five reversals, the increments were reduced from 2 to 1.25 and decrements were increased from .5 to .8. F0 was determined from the mean of the last 10 reversals. Each participant completed the adaptive task three times. The final ΔF0 threshold was the average of the three runs.

Results and Discussion

In the practice block, both groups achieved a high degree of accuracy in detecting the contrast from the standard token: younger adults (M = 98.1%, SD = 2.3), older adults (M = 94.3%, SD = 7.8). Younger adults had reliably lower ΔF0 thresholds (M = .6 Hz, SD = .4, range = .1 to 1.0 Hz) than older adults (M = 1.8, SD = .8, range = .3 to 3.1), $t(28)$ = 5.063, $p < .001$. Not surprisingly, the F0 thresholds for the younger adults were better than those found previously for middle-aged adults, while the F0 thresholds for older adults were similar to those found previously for older adults with normal hearing and better than those of older adults with hearing loss (Summers & Leek, 1998). The reduced ability of older adults to detect differences in F0 could reduce their ability to use this cue in segregating voices.

Experiment 2

Method

Apparatus and Stimuli. Five synthesized vowels [a, i, æ, er, u] were created using formant frequencies corresponding to Assmann and Summerfield (1994). There were six tokens of each vowel differing in F0 (F0 = 120, 122, 124, 127, 135, 151 Hz, respectively corresponding to increases of 0, .25, .5,

1, 2, and 4 semitones from an F0 of 120 Hz). All possible paired combinations of vowels were formed. Thus, a total of 150 vowel pairs were created (5 x 5 pairs x 6 F0 difference levels). All stimuli were 260 ms in duration and delivered monaurally to the better ear via headphones at an overall level of 80 dB SPL.

Procedure. Listeners were first familiarized with all 150 vowel tokens in a single vowel identification task. They identified each vowel from a closed set of five vowel alternatives. Feedback was provided after each response and this procedure was repeated as necessary until listeners achieved an accuracy of at least 80%. Next, listeners completed the concurrent vowel labeling task. They heard pairs of simultaneously presented vowels and identified each vowel by pressing the corresponding vowel button on the button box. If listeners were uncertain on the identity of the vowel, they were instructed to select their best guess. No feedback was provided.

Results and Discussion

All participants identified single vowels with a high degree of accuracy, with younger adults obtaining a higher proportion of correct responses (M = .98, SD = .02) than older adults (M = .93, SD = .06), $t(28) = 3.209$, $p = .003$.

Figure 1 shows the mean proportion of correct identification of both vowels by the younger and older groups as a function of F0 difference in the concurrent vowel labeling task. Overall, a mixed design analysis of variance (2 age groups x 6 F0 difference levels) revealed a significant main effect for age group, with younger adults performing with greater accuracy (M = .58 , SD = .09) than older adults (M = .31, SD = .04) across all F0 difference levels, $F(1, 28) = 21.06$, $p < .001$. There is also a significant main effect for F0 difference, $F(5, 140) = 30.56$, $p < .01$, as well as a significant two-way interaction, $F(5, 140) = 4.42$, $p = .001$. To examine the interaction, Tukey HSD tests were conducted to determine whether each group's performance differed across F0 difference levels. For young adults, there was a significant increase ($p < .001$) in accuracy when F0 separation increased from 0 to .25 semitones. However, for older adults, the improvement in accuracy when F0 separation increased from 0 to .25 semitones approached, but did not reach, conventional levels of significance (p = .068). Beyond an F0 difference of .25 semitones, there was no significant increase in accuracy for younger and older adults ($ps > .05$). Subsequent analyses on the individual data of older adults reveal that three of these participants performed within 1 SD of the mean of younger adults across all F0 difference levels. All others performed significantly below younger adults, but above chance performance.

Simple correlations were conducted to examine associations between participants' performance on the smallest F0 separation values (i.e., 0, and .25 semitones) with age and ΔF0 threshold. Both variables correlated negatively with performance at zero semitone separation: age, $r(29) = -.669$, $p < .001$; and ΔF0 threshold, $r(29) = -.563$, $p = .001$. Similarly, performance at .25 semitone separation is negatively correlated with age, $r(29) = -.651$, $p < .001$, and with ΔF0 threshold, $r(29) = -.594$, $p = .001$.

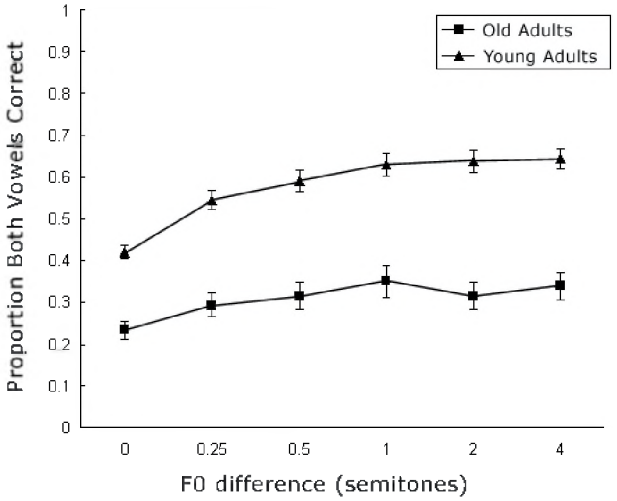


Fig. 1. Proportion of correct identification of both vowels as a function of F0 difference. Error bars indicate standard error.

Analyses of partial correlations revealed that both age and ΔF0 thresholds retained reliable associations with performance at both of these F0 separation values. In sum, these correlations suggest that both age and ΔF0 threshold have a role in the ability to segregate and identify simultaneously presented vowel sounds.

In summary, there are age differences in the ability to detect ΔF0 and the ability to use this cue to identify concurrent vowels. These findings support the hypothesis of loss of periodicity coding as a characteristic of auditory aging (Schneider & Pichora-Fuller, 2001). It sheds new light on older listeners' difficulties in using voice pitch and quality to follow speech in multi-talker environments.

References

Assmann, P. F., & Summerfield, Q. (1990). Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies. *Journal of the Acoustical Society of America, 88,* 680-697.

Assmann, P. F., & Summerfield, Q. (1994). The contribution of waveform interactions to the perception of concurrent vowels. *Journal of the Acoustical Society of America, 95,* 471-484.

Pichora-Fuller, M. K., & Souza, P. E. (2003). Effects of aging on auditory processing of speech. *International Journal of Audiology, 42,* 2S11-2S32.

Schneider, B., & Pichora-Fuller, M. K. (2001). Age-related changes in temporal processing: Implications for speech perception. *Seminars in hearing, 22,* 227-239.

Summers, V, & Leek, M. R. (1998). F0 processing and the separation of competing speech signals by listeners with normal hearing and with hearing loss. *Journal of Speech, Language, and Hearing Research, 41,* 1294-1306.

Perceived Spatial Separation Induced by the Precedence Effect Releases Chinese Speech From Informational Masking

Jing Chen[1], Chun Wang[1], Hongwei Qu[1], Wenrui Li[1], Yanhong Wu[1], Xihong Wu[1], Bruce A. Schneider[2], Liang Li[1,2]

[1]National Key Laboratory on Machine Perception, Speech and Hearing Research Center, Center for Brain and Cognitive Sciences, Department of Psychology, Peking University, Beijing, China, 100871. chenj@cis.pku.edu.cn
[2]Centre for Research on Biological Communication Systems, Department of Psychology, University of Toronto at Mississauga, Mississauga, Ontario, Canada L5L 1C6

INTRODUCTION

Spatial separation of the source of signal sound from the source of interference sound improves the recognition of the signal (Zurek et al., 1993). In reverberant environments, listeners usually "fuse" the direct sound wave from a source with its reflections and perceive a single event as originating from the direction of the source. This phenomenon is called the precedence effect (Litovsky, 1999). Using the perceptual consequence of the precedence effect (Freyman et al., 1999) induced a perceived separation of images of target and masking stimuli, and found a much larger advantage of the perceived spatial separation in English speech recognition when masking stimuli were informational than energetic.

Compared to English, Chinese syllables have more voiceless consonants and less voiced consonants. Chinese syllable initials would be more vulnerable to energetic masking. It has been reported that intelligibility of Chinese speech is considerably worse than that of English speech under conditions with noise masking (Kang, 1998). On the other hand, standard Chinese (also called Mandarin) is a type of tonal language and the pitch pattern of a Chinese single syllable is lexically meaningful. Fox and Unkefer (1985) reported that the perception of tones in Chinese is affected by the lexical status of the speech token. These features in Chinese syllables would induce a distinct pattern of informational masking that does not exit in English, when masking stimuli are Chinese speeches.

In the present study, we used the precedence effect to induce perceived spatial separation of target Chinese speech from either informational or energetic maskers. Instead of a single-sized separation within one semifield (within perceptual channel) as used in the study by Freyman et al., we also investigated the advantages of perceived spatial separation with two different sizes (45° and 90°).

MATERIALS AND METHODS

Target speech stimuli were Chinese "nonsense" sentences spoken by a young female talker. These sentences are syntactically correct but not meaningful. In each of the target sentences there are 3 key words that are scored during speech recognition testing. Target sentences were presented by both the right and the left loudspeakers with the right speaker leading the left speaker by 3 ms. Listeners perceived the target sentence images as coming from the right side. The single-source presentation level of the target sentences was fixed at 54 dB SPL.

There were two types of masking stimuli: noise and speech. The noise masker was steady speech-spectrum noise made from Chinese speech. The speech masker was a recording of continuous Chinese nonsense sentences spoken by other two young female talkers and the contexts were different from target speech. There were 3 perceived locations for the masking stimuli: right, central and right. There were 4 single-source intensity levels for both the speech masking and the noise masking stimuli: 66, 62, 58, and 54 dB SPL, which corresponded to the 4 signal/noise ratios: -12, -8, -4, and 0 dB, respectively.

RESULTS

The percent-correct speech rrecognition across 12 listeners for 2 masking conditions and 3 perceived locations of each masking condition are presented in Figure 1. Under noise masking conditions, the location effect on threshold was not quite significant, $F(2, 22) = 3.430$, $MSE = 1.898$, $p = 0.051$. Under speech masking conditions, the location effect on threshold was significant, $F(2, 22) = 15.896$, $MSE = 2.697$, $p = 0.00$. Pairwise comparisons indicated that the perceived left and central locations of speech maskers did not differ from one another ($p = 1.000$) but both locations differ significantly from the right location ($p = 0.000$, $p = 0.003$, respectively).

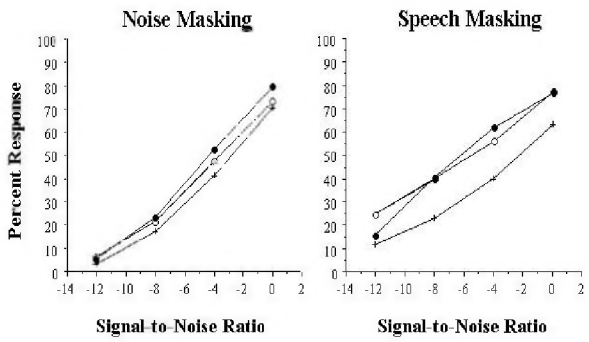


Figure 1. Symbols: crosses, perceived right location; filled circles, perceived center location; open circles, perceived left location.

DISCUSSION

The present study used Chinese nonsense sentences as speech signals and obtained results that are comparable to those reported from Freyman et al's study. When the masker was noise, the improvement of recognition of nonsense Chinese speech was minor (1dB), event though a large perceived spatial separation (45° or 90°) was induced by the precedence effect. When the masker was nonsense speech, the perceived spatial separation of target speeches from masking speeches markedly improved recognition of the target speeches. The analyses of psychometrical function reveal a 3.3-dB improvement. This improvement is somehow smaller than those (4-9 dB) reported by Freyman et al. For Chinese speeches, recognition of syllable initials is critical to recognition of the associated words. Since initials have broad spectrums like those of noises, Chinese words would be more vulnerable to energetic masking than English words (Kang, 1998). Also, perception of tones of syllable finals in Chinese is closely linked to lexical meaning, which may provide listeners with additional cues to connect syllables in target speeches across time. In spite of these characteristics of Chinese speeches, the results indicate that advantage of perceived separation in releasing speeches from masking can extend to tonal Chinese. Interestingly, the perceived separation across hemifields (perceived 90° separation) produced equivalent unmasking effect, compared to the perceived separation within hemifield (perceived 45° separation). These results are not consistent with the prediction of the two-channel model (Boehnke et al. 1999) since the perceived across-channel masking (perceived 90° separation) did not cause less masking than the perceived within-channel masking (perceived 45° separation). A prospective explanation of this disagreement is that the two spatial channels are more associated with processing for locations of sound sources, while perceived locations of "fused" image induced by the precedence effect are more related to higher-order processing.

At this moment it is not clear why under speech masking conditions the perceived-spatial-separation advantage obtained here is smaller than those reported by Freyman et al. In the future, at least two cross-language issues should be addressed: (1) whether Chinese and English speeches have different or similar vulnerabilities to noise masking, and (2) whether Chinese and English masking speeches have different or similar interference effects on recognition of either Chinese or English speeches.

REFERENCES

Zurek, P. M., 1993. Binaural advantages and directional effects in speech intelligibility. In: Studebaker, G.A.,

Litovsky, R. Y., Colburn, H. S., Yost, W. A., Guzman, S. J., 1999. The precedence effect. J. Acoust. Soc. Am. 106, 1633-1654.

Freyman, R.L., Helfer, K.S., McCall, D.D., Clifton, R.K.,1999. The role of perceived spatial separation in the unmasking of speech. J. Acoust. Soc. Am. 106, 3578-3588.

Kang, J., 1998. Comparison of speech intelligibility between English and Chinese. J. Acoust. Soc. Am. 103,1213-1216.

Fox, R.A, Unkefer, J., 1985. The effect of lexical status on the perception of tone + Chinese phonology. J. Chinese Linguist.13, 69-90.

Boehnke, S.E., Phillips, D.P., 1999. Azimuthal tuning of human perceptual channels for sound location. J. Acous. Soc. Am.106, 1948-1955.

PERIPHERAL VERSUS CENTRAL PROCESSING OF A GAP BETWEEN TWO COMPLEX TONES IN YOUNG AND OLD ADULTS

Stephan de la Rosa, Antje Heinrich, and Bruce A. Schneider
Department of Psychology, University of Toronto at Mississauga, 3359 Mississauga Rd, ON, Canada, L5L 1C6

1. INTRODUCTION

Older adults often report difficulties in understanding conversations in noisy environments, especially in the presence of multiple talkers. These difficulties are likely due to age-related losses in both spectral and temporal resolution (e.g., Pichora-Fuller, 1997; Schneider, 1997; Stuart & Phillips, 1996). This study explores the nature of age-related losses in the ability to detect a gap between two complex tones differing in spectral content and perceived fundamental frequency.

1.1 Gap detection with simple tones

In a gap detection task, the listener is asked to detect a period of silence between two sounds, the leading and lagging marker. When each marker is composed of a single frequency, the frequencies can be identical or dissimilar. If these frequencies are identical, the markers stimulate the same region on the cochlea and the listener can perform the gap detection task by detecting a gap or discontinuity in a single auditory channel (within-channel). Several studies suggest that gap thresholds in this condition are generally small for all listeners, yet age differences exist between young and old adults (e.g. Schneider et al. 1994). When the spectral content of the two markers does not overlap, the task becomes one of detecting a gap between different auditory channels (between-channel). Studies that investigated gap detection thresholds in between-channel tasks found that gap detection thresholds are larger for between-channel gap detection tasks than for within-channel gap detection tasks (e.g. Formby et al., 1998). Nothing is known about the extent of age differences in between-channel tasks. In contrast to within-channel tasks, it has been suggested that the gap information from different auditory channels is recovered centrally (Formby et al., 1998).

1.2 Gap detection and complex tones

Vowels consist of a fundamental frequency, f_0, and its overtones or harmonics. Hence, in order to understand how age differences in temporal resolution might affect speech, it might be useful to employ complex tones as markers in gap detection task. Complex tones possess two properties that can potentially influence gap detection: 1) The degree of overlap of the spectral content between leading and lagging marker and 2) the harmonic structure of the markers. Regarding the first characteristic, if spectral content is not overlapping, the task is essentially a between-channel one and high thresholds should be expected. With respect to the second property, Oxenham (2000) used harmonic tone complexes to investigate the effect of a change in the fundamental frequency (f_0) between leading and lagging marker on gap detection thresholds and found elevated gap detection thresholds when f_0 was changing.

In the present study we investigated age differences in gap detection performance when the two markers differed with respect to whether or not 1) they had a frequency in common (both markers had energy at 1 kHz), and 2) whether the two markers shared a common f_0 (e.g., marker 1 had energy at 250, 500, 750, and 1000 Hz, and marker 2 at 1000, 1250, 1500, and 1750 Hz). We hypothesized that the presence of a common frequency should improve gap detection if the listener could make use of the discontinuity information in a more central auditory channel that processes tones having the same perceived pitch. Note that because one of the markers does not have energy at the common fundamental frequency, this pitch channel must be central rather than peripheral.

2. METHOD

2.1 Subjects

Ten young adults (mean age: 21.2 years; SD: 1.81) and ten older adults (mean age: 74.5; SD: 4.35) participated. All participants had good hearing (hearing thresholds below 30 dB up to 3000 Hz).

2.1 Apparatus and stimuli

The stimuli were generated digitally with a sampling rate of 20 kHz, and played using TDT System II. In each condition, the leading marker had energy at 250, 500, 750, and 1000 Hz (f_0.= 250 Hz). Four different lagging markers were employed in a 2 (presence versus absence of a common fundamental) by 2 (presence versus absence of a common frequency) design. The lagging markers defining these four conditions were : 1.) Common f_0 and overlap at 1000 Hz (lagging marker had energy at 1000, 1250, 1500, and 1750 Hz); 2.) Common f_0 and no overlapping frequency (lagging marker had energy on 1250, 1500, 1750,

and 2000 Hz); 3.) Different f_0's but energy overlap at 1000 Hz (lagging marker had energy on 1000, 2000, 3000, and 4000 Hz); 4.) Different f_0's and no overlapping frequency (lagging marker had energy on 2000, 3000, 4000, and 5000 Hz). All participants were tested in each of these four conditions. In addition, we also tested them in a condition in which there was no frequency overlap and the lagging marker did not have a recognizable f_0 (lagging marker had energy on 1300, 1900, 2100, and 4100 Hz). The latter condition served as a check to see whether having a perceptually identifiable f_0 in the first marker but not the second marker had the same effect on gap detection threshold as having two perceptually different f_0's in the two markers defining a gap. Both leading and lagging markers were 20 ms long. The no-gap stimulus was created by filling in the gap between the two 20 ms markers, with the gap and the no-gap stimulus having the same energy throughout the experiment. The presentation level was 90 dB SPL.

2.1 Procedure

Gap detection thresholds were determined in a 2IFC paradigm and a staircase procedure was used to determine the 79.4% point on the psychometric function. The beginning gap size was 300 ms. The inter-stimulus interval was 100 ms. Each participant was tested in all conditions. The testing order of the conditions was counterbalanced across participants.

3. RESULTS AND DISCUSSION

We first checked to see whether having a perceptually identifiable f_0 in the first but not the second marker had the same effect on gap detection thresholds as having perceptually-identifiable f_0's in both markers. (This analysis only included the 3 conditions in which no frequencies were common to the two markers.) A 2 (Age) by 3 second-marker f_0 condition (second marker f_0 absent, the same, different from first marker f_0) found a significant main effect for second marker f_0 ($F(2,34)=7.71$, $p=0.002$) but no main effect for age, and no significant age by second marker condition interaction. Post hoc analysis revealed that the absence of an identifiable second marker f_0 had the same effect on gap detection thresholds as when the second marker f_0 was the same as that of the first marker. Hence, there does not appear to be an appreciable advantage in having a common fundamental frequency between the two markers when there is no overlapping frequency content. However, thresholds for the condition where the lagging marker had a different f_0 than the first marker were signifcantly higher than in the other two conditions, indicating that when there is no overlapping frequency content, having identifiably different f_0's interferes with gap detection. To evaluate the relative contribution of having overlapping frequencies versus same or different f_0's, we conducted a 2 age (young vs old) by 2 frequency overlap (presence vs absence of shared energy at 1 kHz), by 2 fundamental frequency (same or different f_0 in leading and lagging markers) ANOVA. We found significant main effects for frequency overlap ($F(1,52)=20.94$; $p=0.000$) and f_0 ($F(1,52)=5.16$; $p=0.027$), and a significant interaction between f_0 and frequency overlap ($F(1,52)=5.38$; $p=0.024$). There was no significant main effect of age, nor did age interact with any other factor. Post hoc analyses indicated that gap detection thresholds were elevated only when leading and lagging markers differed in perceived fundamental frequency and did not share a frequency in common. Hence, the pattern of results indicates that gap detection thresholds are only elevated when the two markers do not share a common frequency, and have identifiably different fundamental frequencies. One possible reason for this might be that the presence of a different f_0 in the second marker might have led the participant to focus on the frequencies regions associated with the two f_0's, thereby ignoring information that might be available at other frequencies. However, when the two markers contain at least one frequency in common, the presence of the common frequency prevents this from happening.

In all conditions there was no indication of age differences in gap detection thresholds. This was somewhat surprising given the evidence in the literature for age differences when leading and lagging markers are identical. This suggests that age differences might be limited to situations in which there is substantial spectral overlap (more than a single frequency) between leading and lagging markers. .

4. REFERENCES

Formby, C., Gerber, M.J., Sherlock, L.P., & Magder, L.S. (1998). Evidence for an across-frequency, between channel process in asymptotic monaural temporal gap detection. *Journal of the Acoustical Society of America*, 103, 3554-3560.

Oxenham, A.J. (2000). Influence of spatial and temporal coding on auditory gap detection. *The Journal of the Acoustical Society of America*, 107, 2215-2223.

Pichora-Fuller, M.K. (1997). Language comprehension in older listeners. *Journal of Speech-Language Pathology & Audiology*, 21, 125-142.

Schneider, B. (1997). Psychoacoustics and aging: Implications for everday listening. *Journal of Speech-Language Pathology and Audiology,* 21, 111-124.

Schneider, B. A., Pichora-Fuller, M.K., Kowalchuk, D., & Lamb, M. (1994). Gao detection and the precedence effect in young and old adults. *The Journal of the Acoustical Society of America*, 95, 980-991.

Stuart, A.M. & Phillips, D.P. (1996). Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal hearing, and presbycusic listeners. *Ear and Hearing*, 17, 478-489.

ACKNOWLEDGEMENTS

This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada and Canadian Institutes of Health Research. The first and second author was supported, in part, by a Canadian Institutes of Health Research Training grant.

BEAMFORMING A BENT ARRAY

Brian H. Maranda and Nicole E. Collison
DRDC Atlantic, P.O. Box 1012, Dartmouth, NS, Canada B2Y 3Z7
Brian.Maranda@drdc-rddc.gc.ca and Nicole.Collison@drdc-rddc.gc.ca

1. INTRODUCTION

The application of interest is the processing of a towed array that continuously changes shape as the tow ship maneuvers. It is assumed that the deviation of the array shape from a linear geometry can be so large that it is desirable to compensate for the array shape within the beamformer. In such a system, the signal-processing algorithms and display layouts must be designed to handle the changing array shape without human intervention. In this paper, we treat such issues as defining the number of beams and their steering angles, as well as the reference line from which these angles are measured.

2. BEAMFORMING [1]

It is assumed that the towed array is confined to the horizontal plane. The positions of the hydrophones will be denoted by (x_n, y_n) for $n = 0,\ldots,N-1$. Figure 1 illustrates the geometry in plan view. Here the x-axis is shown aligned with the axis of the tow ship, but the exact positioning of the coordinate system is not important for the following theory.

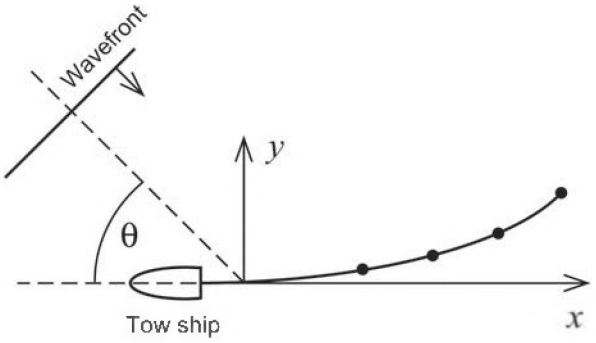


Fig. 1. Plan view of the towed array and coordinate system.

We consider the beam-pattern of a narrowband beamformer processing a signal of spatial dependence $\exp[-i\omega\tau_n(\theta)]$, where ω is the frequency and $\tau_n(\theta)$ the propagation delay between the coordinate origin and the n-th hydrophone for an arrival angle θ. Angles are measured with respect to the negative x-axis, and range between $\pm 180^\circ$. For this plane-wave signal, $\tau_n(\theta) = (x_n \cos\theta - y_n \sin\theta)/c$, where c is the speed of sound. The beam-pattern for the m-th beam ($m = 0,\ldots,M-1$) is given by $|b(\theta;\theta_m)|^2$, where

$$b(\theta;\theta_m) = \frac{1}{W}\sum_{n=0}^{N-1} w_n \exp[i\omega\{\tau_n(\theta_m) - \tau_n(\theta)\}] \qquad (1)$$

and $\tau_n(\theta_m)$ is the delay inserted into the beamformer to steer the m-th beam at angle θ_m. Also, the w_n are hydrophone weights, with sum W.

For an operational sonar system, there is strong motivation to minimize the number of beams while avoiding gaps in detection coverage, such as occur if the beams are spaced too far apart ("scalloping" loss, or "picket fence" effect). Most of the time the tow ship is on a steady course, for which the array is approximately linear, and hence the number of beams and their steering angles would be optimized for the linear geometry. Our purpose is to see how well the coverage is maintained during a tow-ship maneuver, when the array shape is distorted. As will be seen, in order to reduce scalloping loss it is important to choose an appropriate reference axis for the steering angles.

3. REFERENCE AXIS

Given the hydrophone positions, an obvious approach to defining a reference baseline is to apply linear regression. However, the orientation of the line in physical space should not depend on the coordinate system in which it is computed. The standard method of linear regression [2] does not satisfy this property, because deviations from the line are measured in the y-direction. Hence a rotation of the coordinate system, changing the y-direction, also changes the regression solution. An approach that is independent of the coordinate system is to find the line that minimizes the sum of squares of the *perpendicular* deviations from the line. This method of regression, which we call "p-directed", is quite old [3]; see also [4–5].

Although the hydrophone positions are assumed to be known at any instant of time, and are not random in any sense, it will nevertheless be useful to introduce the language and notation of statistics. We define a weighted mean $\overline{x} = {}^1\!/\!_W \sum_n w_n x_n$, with similar definitions for $\overline{y}$, $\overline{xy}$, etc. The variance and covariance are then defined as $\mathrm{var}(x) = \overline{x^2} - (\overline{x})^2$ and $\mathrm{cov}(x,y) = \overline{xy} - \overline{x}\cdot\overline{y}$. It can be shown [3–5] that the slope of the p-directed regression line relative to the x-axis is $\tan\psi$, where the angle ψ is given by

$$\psi = \frac{1}{2}\arctan\left[\frac{2\,\mathrm{cov}(x,y)}{\mathrm{var}(x)-\mathrm{var}(y)}\right]. \qquad (2)$$

In many cases p-directed regression gives a numerical value of the slope differing little from the conventional regression slope, which is $\tan\psi = \mathrm{cov}(x,y)/\mathrm{var}(x)$. However, p-directed regression has properties convenient for theoretical analysis; for example, if the x-y coordinate system is rotated by the angle ψ in Eq. (2) to define new hydrophone positions (x'_n, y'_n), then $\mathrm{cov}(x', y') = 0$.

4. EXAMPLE

In this example, the towed array comprises 100 hydrophones spaced 2 m apart. The beam patterns shown below are for a narrowband signal at frequency 350 Hz, slightly below the array design frequency of 375 Hz (for c = 1500 m/s). The weights w_n used in the beamformer are Kaiser-Bessel weights with –30 dB sidelobe suppression.

When the array is linear, good coverage is provided by M = 99 beams spaced equally in cosine of angle over 0° to 180°. Mathematically, the angles relative to the array axis are $\theta_m = \arccos(1-2m/98)$ for $m = 0,\ldots,98$. In the linear case, the cross-over point of any two adjacent beams has the same numerical value, –1.8 dB (the maximum scalloping loss). Now suppose that the towed array lies on an arc of a circle of diameter 1 km. The head of the array is assumed tangent to the x-axis, as illustrated in Fig. 2. To allow for the bent array, where 360° coverage is necessary, we steer at angles $\pm\theta_m$ relative to a given reference axis (using the same θ_m as for the linear array).

When the x-axis is used as the reference axis for the bent array, the beam coverage is that of Fig. 3(a), which shows the *envelope* of the patterns of all the beams. The maximum scalloping loss is seen to vary with angle, being smaller than necessary in some regions and larger than desired in others. Figure 3(b) shows the beam coverage when the steering angles are taken relative to the regression line, which is at an angle ψ = 11.3° to the x-axis. The coverage is now much more uniform: the maximum scalloping loss is almost the same as for the linear array over the entire circle, although it dips slightly below –2 dB near the endfire directions. (The regression line is used here to define the endfire directions of the curved array.)

5. SUMMARY

A good choice of reference axis has been described for beamforming a bent towed array when the number of beams and their steering angles are optimized for the linear geometry. In particular, the use of an appropriate reference axis can reduce the loss in detection coverage that results from scalloping.

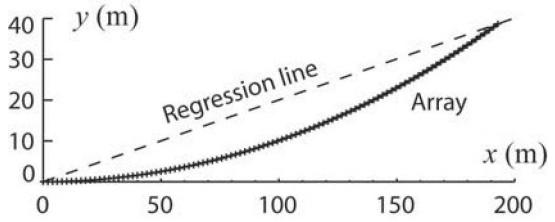


Fig. 2. The curved array and its regression line. The line has been translated so that it goes through the origin.

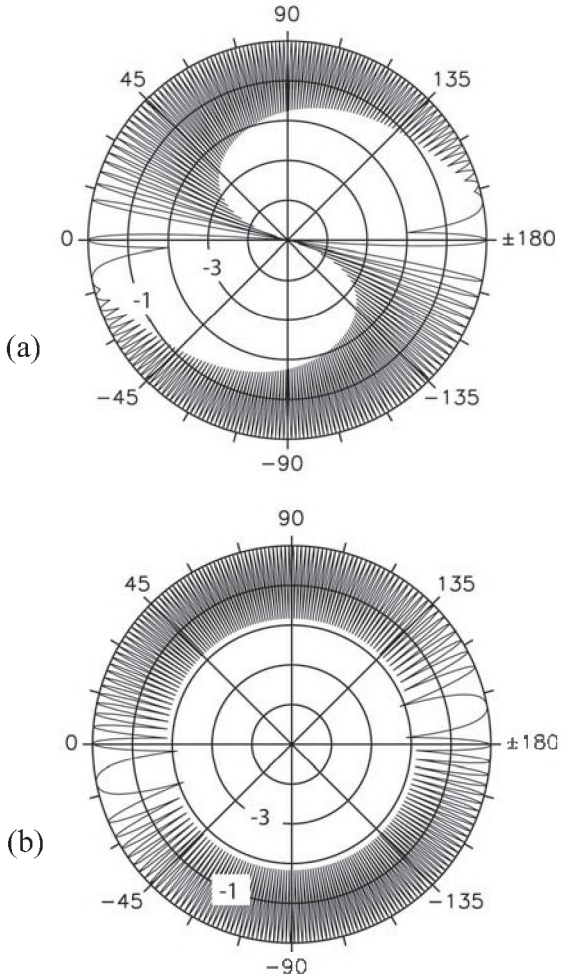


Fig. 3. Polar plots of the envelope of the beam-patterns for the bent array. Steering angles are relative to (a) the x-axis and (b) the regression axis. Beam peaks are at 0 dB and rings at –1-dB intervals.

REFERENCES

[1] Johnson, D.H. and Dudgeon, D.E. (1993). *Array Signal Processing*. Upper Saddle River, NJ: PTR Prentice-Hall.

[2] Draper, N.R. and Smith, H. (1966). *Applied Regression Analysis*. New York: John Wiley & Sons.

[3] Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. Phil. Mag., ser. 6, vol. 2, 559–572.

[4] Cramer, H. (1946) *Mathematical Methods of Statistics*. Princeton: Princeton Univ. Press. (See pp. 275–276.)

[5] Zucker, L.M. (1947). Evaluation of slopes and intercepts of straight lines. Human Biology, vol. 19, 231–259.

DATA ERROR ESTIMATION IN MATCHED-FIELD GEOACOUSTIC INVERSION

Stan E. Dosso, Michael J. Wilmut, and Jan Dettmer
School of Earth and Ocean Sciences, University of Victoria, Victoria BC Canada V8W 3P6

1. INTRODUCTION

The problem of estimating seabed geoacoustic parameters by inverting measured ocean acoustic fields has received considerable attention in recent years. Matched-field inversion (MFI) is based on searching for the set of geoacoustic model parameters $\mathbf{m}$ that minimizes an objective function quantifying the misfit between measured and modelled acoustic fields. A number of approaches have been applied to this challenging nonlinear optimization problem. In particular, adaptive simplex simulated annealing (ASSA) [1], a hybrid optimization algorithm that combines local (gradient-based) downhill simplex moves within a fast simulated annealing global search, has proved highly effective for MFI.

In a Bayesian formulation of MFI, the objective function to be minimized is derived from the likelihood function corresponding to the assumed data uncertainty distribution. The likelihood depends on parameters describing data uncertainties (e.g., standard deviations) which are nuisance parameters in terms of recovering seabed properties, but must be accounted for in a rigorous inversion. Data uncertainties include both measurement errors (e.g., due to instrumentation and ambient noise) and theory errors (due to the simplified model parameterization and approximate acoustic propagation model). Theory errors in particular are generally not well known, and tend to increase with frequency due to the effects of scattering, 3-D environmental variability, sensor location errors, etc. [2]. This paper derives several likelihood-based objective functions and examines their performance in MFI of acoustic data with unknown, frequency-dependent uncertainties.

2. THEORY

For acoustic data $\mathbf{d}_f$ measured at an N-sensor array at $f=1,F$ frequencies contaminated by independent, complex Gaussian-distributed errors with standard deviations σ_f, it can be shown that the likelihood function when source amplitude and phase are unknown is given by [2]

$$L(\mathbf{m},\boldsymbol{\sigma}) = \prod_{f=1}^{F} \frac{1}{(\pi\sigma_f^2)^N} \exp[-(1-B_f(\mathbf{m}))\,|\,\mathbf{d}_f\,|^2 / \sigma_f^2],$$

where

$$B_f(\mathbf{m}) = |\,\mathbf{d}_f(\mathbf{m})^{\mathrm{T}}\mathbf{d}_f\,|^2 / |\,\mathbf{d}_f(\mathbf{m})\,|^2 |\,\mathbf{d}_f\,|^2$$

is the normalized Bartlett (linear) correlator, $\mathbf{d}_f(\mathbf{m})$ are the data predicted for model $\mathbf{m}$, and $^{\mathrm{T}}$ denotes conjugate transpose. Maximum-likelihood parameter estimates are obtained by maximizing the likelihood over $\mathbf{m}$. If the standard deviations σ_f are *known*, this is equivalent to minimizing the objective function

$$E_1(\mathbf{m}) = \sum_{f=1}^{F} (1-B_f(\mathbf{m}))\,|\,\mathbf{d}_f\,|^2 / \sigma_f^2 .$$

However, as mentioned above, data uncertainties are rarely well known due to theory errors. The standard approach is to assume that the uncertainty weighting factor $|\mathbf{d}_f|^2/\sigma_f^2$ is *uniform* over frequency, and minimize an objective function

$$E_2(\mathbf{m}) = \sum_{f=1}^{F} \left[1-B_f(\mathbf{m})\right].$$

However, this is often a poor assumption in practice [2]. A straightforward approach for unknown uncertainties is to *explicitly* estimate the standard deviations as part of the inversion by minimizing the objective function

$$E_3(\mathbf{m},\boldsymbol{\sigma}) = \sum_{f=1}^{F} \left[(1-B_f(\mathbf{m}))\,|\,\mathbf{d}_f\,|^2 / \sigma_f^2 + N \ln \sigma_f^2\right]$$

over $\mathbf{m}$ and $\boldsymbol{\sigma}$. The disadvantage to this approach is that it introduces F new unknown parameters σ_f, resulting in a more difficult inverse problem. An alternative approach is to maximize the likelihood over σ_f, yielding the analytic solution

$$\sigma_f = (1-B_f(\mathbf{m}))\,|\,\mathbf{d}_f\,|^2 / N .$$

Substituting this back into the likelihood function leads (after some algebra) to an objective function

$$E_4(\mathbf{m}) = \prod_{f=1}^{F} (1-B_f(\mathbf{m})).$$

Minimizing this objective function treats the data standard deviations as *implicit* unknowns without increasing the number of parameters in the inversion.

3. RESULTS

This section considers a synthetic study of inversion performance for the various objective functions based on a shallow-water geoacoustic experiment carried out in the

Mediterranean Sea southeast of Elba Island [2]. The seabed model consists of a sediment layer over a semi-infinite basement. The unknown geoacoustic parameters are the sediment thickness, h, and sound speed, attenuation and density of the sediment and of the basement, c_1, α_1, ρ_1 and c_2, α_2, ρ_2, respectively. In addition, small corrections to the water depth, D, and source range and depth, r and z, are also included in the inversion as these geometric parameters are generally not know to sufficient accuracy. Synthetic acoustic fields were generated at 50-Hz intervals from 300-500 Hz, with Gaussian noise added so that the signal-to-noise ratio decreased uniformly from 12 to 0 dB across the band, reflecting the observed increase in theory error with frequency [2].

ASSA inversions were carried out for 200 different realizations of random noise on the data using each of the four objective functions. In all cases the inversion found a set of model parameters with an objective function value less than that for the true parameters, indicating an excellent solution. The explicit formulation (objective function E_4) required about three times more forward model calculations for convergence than the other cases due to the increased number of unknowns. The inversion results are quantified in terms of standard deviations about the true model parameters. The model-parameter standard deviations for the case where the data standard deviations are known exactly (objective function E_1) are given in Table 1. These results indicate that, relative to the parameter search bounds, the basement sound speed is the most accurately estimated parameter, followed by the sediment thickness, sediment sound speed, and basement attenuation. The other parameters have relatively large standard deviations and are poorly undetermined in the inversion

To compare inversion performance for the various objective functions, the model-parameter standard deviations obtained in all cases were divided by those obtained for objective function E_1 (known data standard deviations) to obtain normalized errors, which are shown in Fig. 1. This figure shows that objective function E_2, based on the assumption of uniform data uncertainty weighting, leads to normalized errors greater than unity for all parameters, with particularly large errors for the well-determined parameters. By contrast, objective functions E_3 and E_4 (which explicitly and implicitly include data standard deviations in the inversion, respectively) produce much smaller normalized errors which are close to unity for all parameters.

These results suggest that the acoustic data contain sufficient information content to estimate data standard deviations as well as (some) geoacoustic parameters, and that including standard deviations in the inversion (explicitly or implicitly) is preferable to the standard assumption of uniform data uncertainty weighting over frequency. Further, the implicit formulation requires no greater computational burden than the standard approach, but produced significantly better results.

Table 1. True parameter values, search bounds, and model-parameter standard deviations obtained using objective function E_1 (known data standard deviations).

Parameter	True Value & Search Bounds		Stnd Dev
h (m)	7	[0- 30]	1.5
c_1 (m/s)	1495	[1460 -1550]	15
c_2 (m/s)	1530	[1500-1600]	4.0
α_1 (dB/λ)	0.1	[0-0.5]	0.2
α_2 (dB/λ)	0.2	[0-0.5]	0.1
ρ_1 (g/cm^3)	1.4	[1.0-1.8]	0.25
ρ_2 (g/cm^3)	1.6	[1.2-2.2]	0.30
D (m)	130	[128-132]	1.5
r (km)	3.9	[3.8-4.0]	0.05
z (m)	10	[8-14]	2.0

REFERENCES

[1] S.E. Dosso, M.J. Wilmut & A.-L.S. Lapinski, 2001. An adaptive hybrid algorithm for geoacoustic inversion, *IEEE J. Oceanic Eng.*, **28**, 24-336.

[2] S.E. Dosso & P.L. Nielsen, 2002. Quantifying uncertainty in geoacoustic inversion, II. Application to broadband, shallow-water data, *J. Acoust. Soc. Am.*, **111**, 143-159.

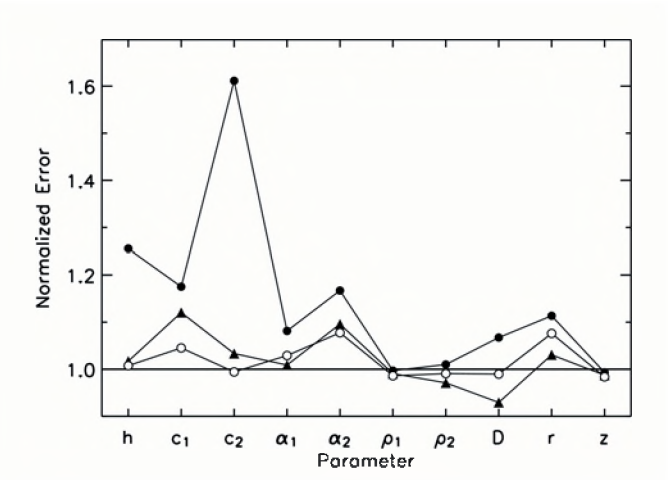


Fig. 1. Normalized parameter standard deviations for uniform, explicit, and implicit processors (filled circles, triangles, and open circles, respectively).

GEOACOUSTIC INVERSION WITH STRONGLY CORRELATED DATA ERRORS

Jan Dettmer[1], Stan E. Dosso[1], and Charles W. Holland[2]
[1]School of Earth and Ocean Sciences, University of Victoria, BC, Canada jand@uvic.ca
[2]Applied Research Laboratory, The Pennsylvania State University, State College, PA, USA

1. INTRODUCTION

Geoacoustic inversion infers seabed properties from acoustic measurements in the water column. A typical assumption in geoacoustic inversion is that the data errors are uncorrelated and obey certain statistics, e.g., are Gaussian distributed. However, real data often show strong error correlations. In the past, this has been dealt with by sub-sampling the data to a point where correlations disappear. However, sub-sampling the data is always a trade off between reducing correlations and loosing information.

This paper considers pre-processing of synthetic single bounce reflection-loss data with strongly correlated data errors to improve application of a nonlinear Bayesian inversion to recover geoacoustic parameters from a viscoelastic model. Correlated Gaussian errors are generated using a realistic synthetic covariance matrix derived from experimental measurements [1]. The inverse problem is solved with fast Gibbs sampling [2], which provides parameter estimates and credibility intervals by sampling the posterior probability density. The error correlations are taken into account by estimating a covariance matrix from the data residuals obtained from a solution to the inverse problem. This covariance matrix is then used in the cost function of the fast Gibbs sampler. The recovered covariance matrix is compared to the original and stringent statistical tests are applied to the data residuals to illustrate the benefits of this rigorous error treatment. The geoacoustic parameters of the viscoelastic model are clearly resolvable and show reasonable error bounds for correlated errors.

2. Method

The pre-processing of fast Gibbs sampling [2] to apply it to synthetic data with correlated data errors involves three basic steps. First, zero mean Gaussian noise n_i' of standard deviation 1 is generated for each frequency f_i. In a second step, the noise is then correlated using a covariance matrix for each frequency $C^{(d_i)}$. This can be done by forming the Cholesky decomposition of the matrix

$$C^{(d_i)} = LL^T, \quad (1)$$

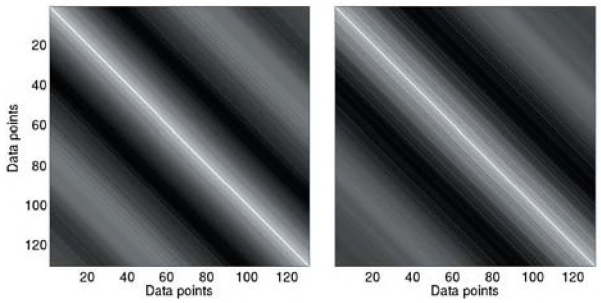


Fig. 1. Original (left) and recovered (right) covariance matrices for 1600 Hz.

where L is the lower triangular matrix. Then, calculating

$$n_i = Ln_i' \quad (2)$$

gives noise that is drawn from a Gaussian distribution with respect to the data covariance matrices $C^{(d_i)}$ at each frequency. The noise is then added to the synthetic data that were computed with a forward model that calculates the plane wave response of the sub-bottom as a function of frequency and angle.

To recover $C^{(d_i)}$, residuals $\tilde{n}_i$ are calculated from predicted and real data. The predicted data are calculated from a maximum likelihood model that is found by applying global optimization to determine the geoacoustic model parameters which best fit the noisy data. The autocovariance can then be calculated as a function of $\tilde{n}_i$. Theoretically, the covariance should be calculated as an ensemble average. In reality, however, we have access to only a single finite subset of the random process. By assuming an ergodic process, we replace the ensemble average by an average over angles, and the j[th] element of the autocovariance function for the finite subset $\tilde{n}_i$ can be estimated as

$$c_j^{(d_i)} = \frac{1}{N_i} \sum_{k=0}^{N_i-j-1} n_i(k+j)n_i(k), \quad (3)$$

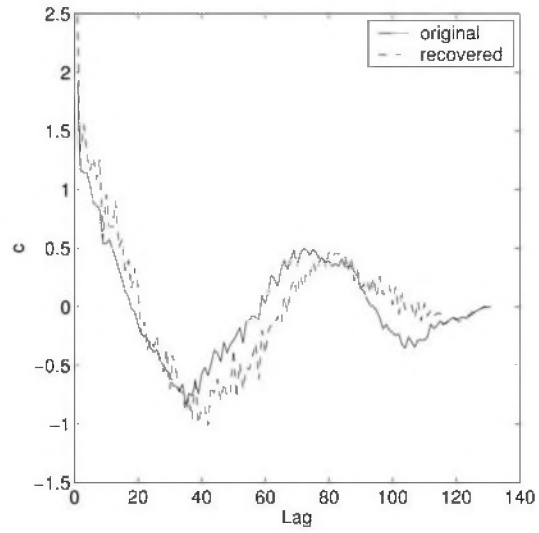


Fig. 2. First row of original covariance matrix and recovered covariance matrix after one iteration for a frequency of 1600 Hz.

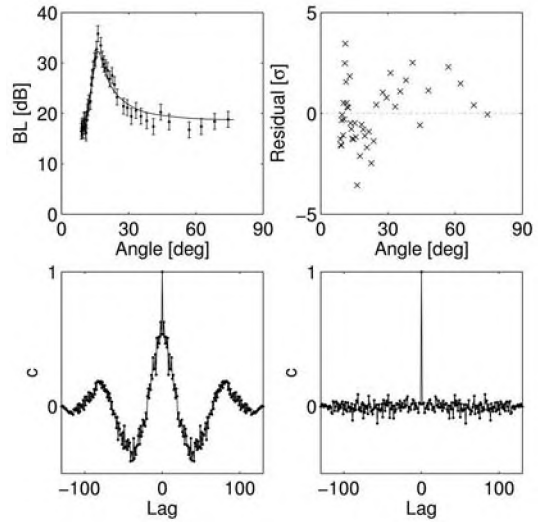


Fig. 3. (a) Fit of predicted data (solid line) and synthetic data with correlated errors and (b) the residuals that clearly show the correlations for 1600 Hz. (c) and (d) show the normalised autocovariance of the residuals with and without the recovered covariance matrix was taken into account, respectively.

where N_i is the number of angles at frequency f_i. Every term $c_j^{(d_i)}$ then builds one diagonal in the symmetric covariance matrix. The normalisation factor N_i does not strictly result in an average but rather damps terms in eq. (3) that do not have many samples to average over and hence have higher uncertainty. The covariance matrix estimate can be iteratively improved. by using the estimates in subsequent global optimizations. The recovered covariance matrix can then be used in the Gibbs sampler by implementing the likelihood function L for multiple frequencies

$$(4) \qquad L(m) \propto \prod_{i=1}^{F} \exp(-n_i^T (C^{(d_i)})^{-1} n_i),$$

where F is the number of frequencies and m is the model.

3. RESULTS

The data set used here contained 8 frequencies in a band from 300 to 1600 Hz with different numbers of angles at each frequency (between 54 and 131 data points at each frequency). For simplicity, the results are illustrated with only one frequency, 1600 Hz. The covariance matrices used here are taken from an analysis of real data collected in the Strait of Sicily [1] that showed strongly correlated errors.

Fig. 2 compares the original and the recovered covariance matrix after one iteration for 1600 Hz. It can be seen that there is a close match between the two. Inversion results (not shown here) with the two matrices show that both the recovered models and the parameter uncertainty bounds are very close as well.

Fig. 3 (a) and (b) show the fit of the predicted data to the synthetic data, and the residuals. It is obvious that the data have strong correlations. Fig. 3 also shows the auto covariance function for the data residuals and for residuals to which the Cholesky decomposition of the recovered covariance matrix was applied by

$$(5) \qquad \tilde{n}_i = L^{-1} n_i .$$

These two plots indicate that the recovered covariance matrix is capable to take all important correlations into account. To statistically quantify the effect of using $C^{(d_i)}$, a runs test was performed. The data residuals n_i strongly fail the runs test ($p < 0.001$), whereas the data residuals $\tilde{n}_i$ easily pass the tests ($p = 0.39$). In the inversion, the geoacoustic parameters show more realistic uncertainty bounds if $C^{(d_i)}$ is used in the energy function. It is thus important to correctly treat error correlations if reasonable uncertainties are to be estimated from the data.

REFERENCES

[1] C. W. Holland and J. Osler. High-resolution geoacoustic inversion in shallow water: A joint time- and frequency-domain technique. J. Acoust. Soc. Am., 107, pp. 1263-1279. 2000.

[2] S. E. Dosso. Quantifying uncertainty in Geoacoustic Inversion. I. A fast Gibbs sampler approach. J. Acoust Soc. Am., 111, pp. 129-142. 2002.

SWAMI AND ASSA FOR GEOACOUSTIC INVERSION

James A. Theriault [1] and Colin Calnan [2]
[1]DDRDC Atlantic, P.O. Box 1012, Dartmouth, NS, B2Y 3Z7-e-mail jim.theriault@drdc-rddc.gc.ca
[2] xwave solutions, 36 Solutions Drive, Halifax, N.S., B3S 1N2

1. INTRODUCTION

The SWAMI (Shallow Water Active-sonar Modelling Initiative) (Theriault and Ellis 1997) toolset in use at DRDC Atlantic contains modules to produce predictions of transmission loss, reverberation, signal excess, and probability of detection. The toolset includes a capability to consider many source and receiver configurations; from omni directional to line-arrays (horizontal and vertical), and to volumetric arrays. The toolset allows the environment to vary both azimuthally and radially (Nx2D).

Simulated reverberation data can be compared with measured reverberation data and a quantitative measurement can be made that essentially tests the goodness of fit of the modelled data to the "measured" values. This paper briefly presents an overview of the SWAMI toolset, the type of active sonar of interest, the inversion technique, and results. For the purposes of this paper, a reverberation model has generated the "measured" input data.

1.1 SWAMI

SWAMI is an Nx2D active sonar performance prediction toolset based on adiabatic normal-mode theory (Bucker and Morris 1968). The current reverberation model, MONOGO is based on Ellis' OGOPOGO (Ellis 1992 and Ellis 1995) model. The key differences between the models are that OGOPOGO computes reverberation predictions for range-independent environments with bistatic geometries whereas MONOGO computes reverberation predictions for Nx2D environments for the monostatic geometry case. MONOGO, as do all of the SWAMI components, computes its predictions for multiple receiver steering angles in parallel.

1.2 System

The DRDC Atlantic TIAPS (Towed Integrated Active-Passive Sonar) system consists of a two-element vertical source array and two high-dynamic-range towed arrays. One of the arrays (MANTArray) consists of a large set of omni directional hydrophones while the second array (DASM) consists of a set of directional sensors (Theriault and Hood 2004). These towed arrays allow the reverberation environment to be sampled in both range (time) and azimuth (steered beams). It is this type of system that is of interest for the inversion presented in this paper.

The goal of the effort is be able to produce sonar operator decision aids. By inverting reverberation measurements, estimates of geoacoustic parameters, and therefore predicted target echo strength can be obtained. A comparison of predicted target echo strength with the original reverberation data yields a measure of performance.

2. METHOD

The reverberation inversion is performed using an Adaptive Simplex Simulated Annealing (ASSA) technique, where the implementation is essentially the same as that described by Dosso (2001). The ASSA-derived geoacoustic parameter values are used as input to the model MONOGO, which uses them to produce reverberation time series. An energy value E related to a set of model parameters is obtained by calculating the differences between adjacent model time series values, subtracting them from the corresponding differences produced from the measured data, and summing the absolute values of the differences. Both model and measured time series must have the same time origin and increment so that E can be produced by comparing the slopes of the two time series. Slope comparison, as opposed to direct comparison, is used based on an earlier observation by Ellis (1994). Ellis observed that the bottom-loss related parameters seemed to be more sensitive to reverberation decay than overall level. Scattering strength is more sensitive to the overall level. Predicting echo levels is not dependant on scattering strength so the alternative energy value is used in order to generate a potentially faster inversion methodology.

The geoacoustic parameters are found at a number of evenly spaced points along a user-selectable number of fixed-length radials. A feature of the analysis program is that besides specifying the number of points per radial (N_R) to solve for, the user is able to give the program a range of N_R values to use. When this option is used the program produces minimum E results for each N_R value and indicates which N_R produces the overall minimum E. The user can examine the results for all values of N_R before choosing which geoacoustic values to use in later processing. This is done since it may turn out that the results from certain N_R values may produce low E values but conflict with reality in the form of known parameter values, topography, etc.

3. RESULTS

For the purposes of this paper, a three radial environment was generated, with each radial being divided into six segments. For each radial the environment segments started at 0, 12.04, 24.08, 36.12, 48.12 and 60.20 km. The three radials were unequally spaced with bearings of 40°, 90°, and 250° respectively. The center environment (range 0 km) was the same for all radials.

Table 1 shows the values used to represent the seabed characteristics for the simulated environment. MONOGO was used to generate a reverberation time series. The sound speed in the water column was held constant (1500 m/s). The columns of Table 1 show the input radial number, point on the radial, density, compressional sound speed, compressional attenuation, Lambert scattering coefficient, and water depth.

Table 1. Description of Radials

#	Pt	Density g/cm^3	Comp. Sound Speed m/s	Comp. Atten. dB/km	Lambert Scattering Coef. dB/m^2	Depth m
	1	2.07	1782	0.218	-29.3	83
1	2	2.01	1786	0.212	-28.7	78
1	3	1.94	1792	0.207	-27.6	75
1	4	1.86	1799	0.203	-26.8	72
1	5	1.78	1810	0.198	-26.2	68
1	6	1.73	1818	0.193	-25.3	63
2	2	2.05	1788	0.216	-29.3	78
2	3	1.99	1794	0.214	-28.9	67
2	4	1.95	1801	0.213	-28.2	71
2	5	1.91	1807	0.212	-27.4	75
2	6	1.85	1812	0.211	-26.6	68
3	2	2.13	1776	0.221	-29.3	90
3	3	2.18	1771	0.222	-29.0	96
3	4	2.22	1765	0.225	-28.4	85
3	5	2.25	1758	0.227	-27.6	80
3	6	2.28	1753	0.229	-26.9	74

The system assumed for the simulation consisted of omni-directional source and receiver with a depth of 50m. The transmitter projected a 1s CW waveform with a source level of 210 dB re 1μPa @ 1m.

After generating the simulated time series, the ASSA method for the geoacoustic parameters was used to invert the given reverberation. For the purposes of this paper, averages of 10 results from 10 ASSA runs were computed. Each ASSA run required an average of 1683 MONOGO runs. The simulated reverberation data and the results from the inversion are shown in Figure 1. Figure 2 shows the difference between the simulated reverberation time series and the inversion results.

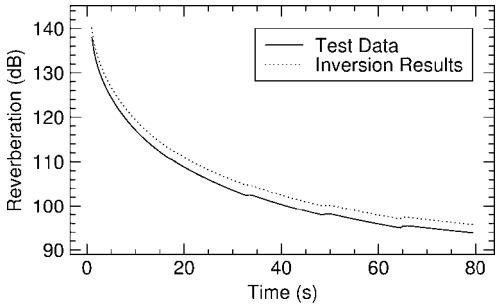


Fig. 1. Reverberation Predictions Using Input and Simulated Geoacoustic Parameters.

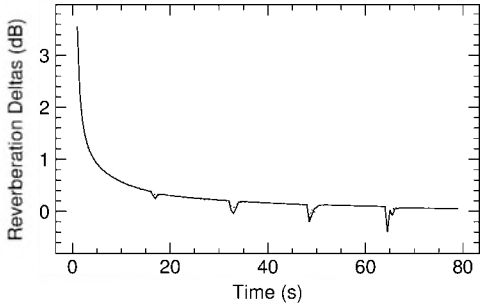


Fig. 2. Difference between Original Simulated Reverberation Levels and Results of Inversion.

4. DISCUSSION

As shown in Figures 1 and 2 the inversion approach is capable of generating geoacoustic parameters that will closely replicate the input time series. The sharp differences occurring at ranges that correspond to radial segments are likely a shortcoming in the MONOGO model rather than the inversion technique.

The approach presented above shows some promise for producing relevant geoacoustic parameters. However, experience with simulated noisy input time series and measured data is required.

REFERENCES

Bucker HP and Morris HE (1968) Normal-mode reverberation in channels or ducts, *J. Acoust. Soc. Am.* **44**, 827-828, 1968.

Dosso SE, Wilmut MJ, and Lapinski A-L.S. (2001) An Adaptive-Hybrid Algorithm for Geoacoustic Inversion, *IEEE J. Ocean Eng.*, **26**, pp. 324-336, July 2001.

Ellis DD (1992) A Shallow-Water Normal-mode Reverberation Model, SACLANTCEN Report 92-196, September 1992.

Ellis DD (1994) Personal Communication.

Ellis DD (1995) A Shallow-Water Normal-Mode Reverberation Model, *J. Acoust. Soc. Am* **97**, 2804-2814, 1995.

Theriault JA and Ellis DD (1997), Shallow-water low-frequency active-sonar modelling issues, in *Oceans '97 MTS/IEEE Conference Proceedings* (IEEE, Piscataway, NY, USA) Vol **1**, pp. 672-678, 1997.

Theriault JA and Hood J (2004) Using Directional Sensors to Enhance the Performance of Towed Receiver Array of Active Sonar, in *Undersea Defence Technology Proceedings*, October 2004.

FIELD TRIALS OF GEOPHONES AS ARCTIC ACOUSTIC SENSORS

Stan E. Dosso[1], Garry J. Heard[2], Michael Vinnins[3] and Slobodan Jovic[3]
[1]School of Earth and Ocean Sciences, University of Victoria, Victoria BC Canada, sdosso@uvic.ca
[2]Defence Research and Development Canada–Atlantic, Halifax NS Canada
[3]Defence Research and Development Canada–Ottawa, Ottawa ON Canada

1. INTRODUCTION

This paper considers two practical issues concerning the use of ice-mounted geophones as Arctic acoustic sensors: the ability to resolve the relative bearing to an acoustic source in the water column, and the ability to determine absolute sensor bearing via short baseline GPS. Two approaches to bearing estimation are compared, including beamforming seismo-acoustic arrivals at an array of geophones and resolving the incident power versus arrival angle at a tri-axial geophone [1, 2]. The latter approach is preferable logistically, as it requires only a single sensor. However, due to the complexities of seismo-acoustic propagation in Arctic pack ice, it is difficult to predict *a priori* the effectiveness of these approaches, and source-bearing estimation must be studied *in situ* via Arctic field trials.

2. FIELD TRIALS

A linear array of five geophones, spaced at 20-m intervals, was deployed by hand on the surface of a 4.5-m thick multi-year ice floe in the Lincoln Sea. The three sensors at the centre of the array were tri-axial geophones; the two outer sensors were vertical-component geophones. The geophone signals were transmitted via over-ice cables to a heated tent and recorded on a 12-channel digital seismograph at a sampling rate of 2000 Hz (results shown here are low-pass filtered 0–200 Hz since the geophone instrument response is unreliable at higher frequencies). Impulsive acoustic sources were deployed at 18-m depth in the water column at ranges of 2, 5, 10, and 25 km and bearings of 0°, 30°, 60°, and 90° (with respect to the geophone array axis) and at 50-km range and 60° and 90° bearing. Source deployments involved drilling through the ice with a power auger, and moving between sites using a helicopter with GPS positioning. The ice cover was continuous in the study area.

3. RESULTS

Figure 1 shows the source-bearing estimates obtained by beamforming the seismograms recorded at the five vertical-component geophones. An analysis of the optimal wave-propagation speed for the beamforming yielded a speed of 1439 m/s, consistent with the average water-column sound speed, indicating that the dominant propagation path consists of water-borne acoustic waves coupling locally into ice seismic waves near each geophone. Symmetry produces

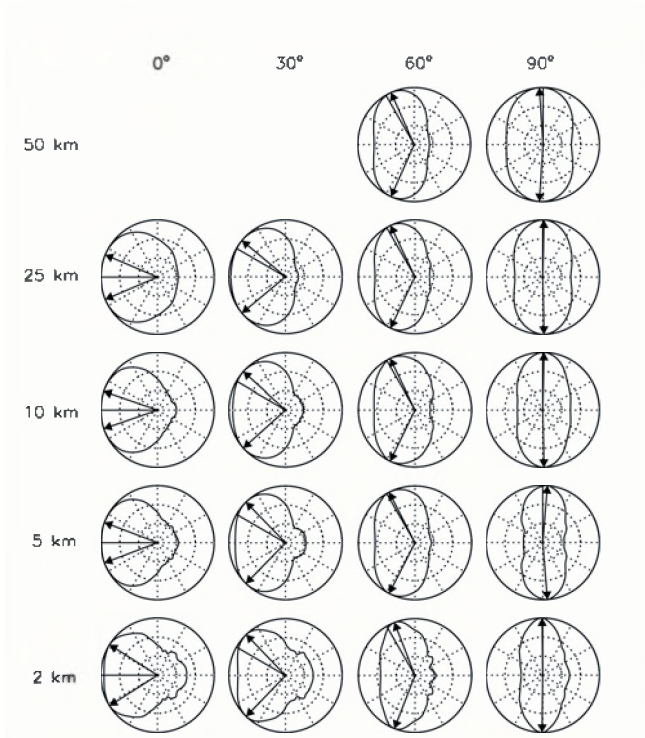


Fig. 1. Beamforming bearing estimates (source bearings and ranges indicated on figure). Solid curves represent beamformer power as a function of look angle from 0–360°; dotted circles indicate 5 dB intervals. True and (ambiguous) estimated source bearings are indicated by solid lines and arrows, respectively.

an ambiguity in the beamforming results about the array axis, represented in Fig. 1 by two arrows on each panel indicating the ambiguous optimal bearing estimates. Figure 1 shows that the beamforming results are relatively poor for bearings of 0° (endfire to the array), but improve substantially towards 90° bearing (broadside). The bearing estimates and the beamformer angular responses do not appear to degenerate significantly as the source range increases from 2 to 50 km.

Figure 2 shows the bearing-estimate results obtained from rotational analysis of particle motion at a single tri-axial geophone (results are shown for the centre geophone, although similar results were obtained for all three tri-axial sensors). The rotational analysis consists of geometrically combining the horizontal seismograms to compute the wave

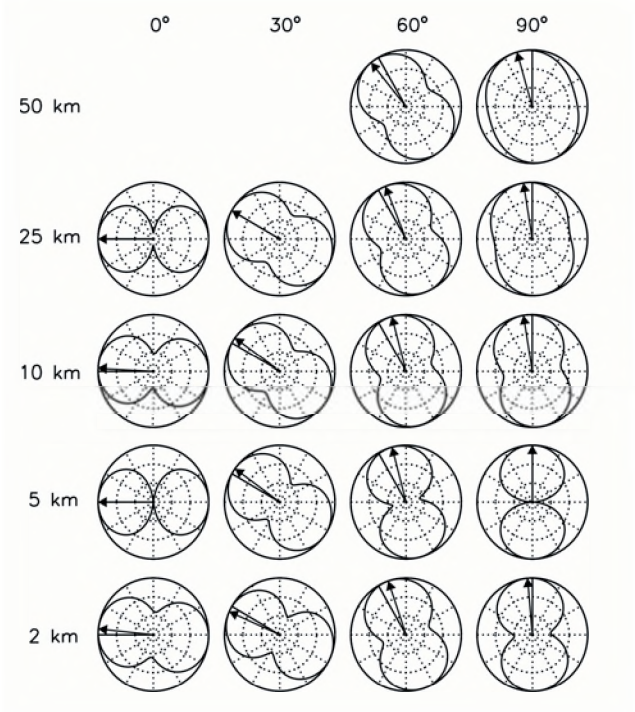


Fig. 2. Bearing estimates from rotational analysis. Solid curves represent signal power as a function of bearing angle; true and estimated bearings given by solid lines and arrows.

power in radial look angles from 0–360° (plotted as solid curves). This analysis is enhanced by applying seismic polarization filters, originally developed for earthquake seismology [3]. These filters make use of theoretical phase relationships between the vertical and horizontal components of the various wave types to selectively suppress waves with transverse particle motion (e.g., shear waves) which degrade the rotational analysis, while passing waves with radial particle motion. In addition, with three-dimensional measurements the 180° ambiguity inherent in the rotational analysis can be resolved by requiring outgoing (prograde) particle motion in the vertical-radial plane, providing a unique bearing estimate (indicated by arrows in Fig. 2). Figure 2 shows that bearing estimates of similar quality are obtained for all source bearings, and that bearing estimates do not degrade significantly with range.

The results for the two approaches to bearing estimation are summarized in Fig. 3, neglecting the ambiguity in the beamforming results. At source bearings of 0° and 30° (at and near endfire), the single-sensor rotational analysis provides substantially better results than array beamforming, with average bearing-estimate errors for rotational and beamforming analyses of 3° and 24° at 0° bearing, respectively. However, beamforming generally produces better bearing estimates for sources at 60° and 90° (broadside), with rotational and beamforming analyses yielding average errors of 8° and 2° at 90°, respectively. Considering all recordings, the rotational and beamforming analyses produced average source-bearing estimation errors of 6.8° and 10.2°, respectively. Given the additional ability

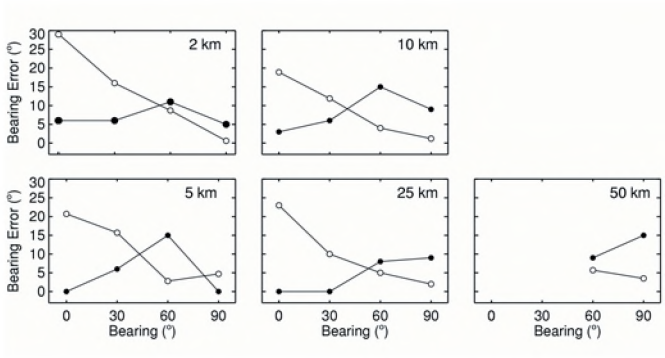


Fig. 3. Bearing estimate errors for rotational analysis (solid circles) and beamforming (open circles).

to provide a unique bearing estimate and the logistical advantages to using a single sensor, rotational analysis at a tri-axial geophone appears to be the superior approach for source-bearing estimation in the Arctic.

Given the ability to estimate the bearing to an acoustic source relative to the horizontal components of a tri-axial geophone, the absolute orientation of the sensor is required to obtain the source bearing in geographical coordinates. In the case of the absolute bearing determination for geophone sensors, the practical limitations dictate that the GPS antennas be mounted directly on the top of the sensor, which has a diameter of approximately 10 cm. This limits the bearing accuracy that can be achieved due to carrier phase noise, multi-path, and antenna phase-centre errors. A real-time heading determination algorithm using a double-difference carrier phase approach and integer ambiguity resolution on-the-fly was developed to determine the three-dimensional vector between two GPS antennas [4]. Arctic field trials indicate that this approach can resolve absolute bearing at high latitudes (83° N) to uncertainties of less than 3° employing only low-cost commercial antennas with separations of as little as 4 cm.

REFERENCES

[1] S. E. Dosso, G. J. Heard & M. Vinnins, 2002. Source bearing estimation in the Arctic Ocean using ice-mounted geophones, *J. Acoust. Soc. Am.*, **112**, 2721–2734.

[2] S. E. Dosso, M. Vinnins & G. J. Heard, 2003. Arctic field trials of source bearing estimation using ice-mounted geophones, *J. Acoust. Soc. Am.*, **113**, 2980–2983

[3] J. E. White, Motion product seismograms, 1964. *Geophysics*, **29**, 288–298.

[4] M. Vinnins, G. Lachapelle, E. Cannon, S. E. Dosso & G. J. Heard, 2003. High latitude attitude: Military testing in the Arctic, *GPS World*, **14**, 16–27.

AURAL ANALYSIS OF THE HARMONIC STRUCTURE OF SONAR ECHOES

G. Robert Arrabito[1], Nancy Allen[2], Taresh D. Mistry[1], and Kyle Leming[1]
[1] DRDC Toronto, P.O. Box 2000, Toronto, ON M3M 3B9 Canada
[2] DRDC Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7 Canada

1. INTRODUCTION

The return echoes of many types of sonar systems from targets of interest (e.g., man-made objects such as submarines and sea mines) are often indistinguishable from non-targets. This is especially evident for low-frequency active sonar systems operating under conditions of substantial geological clutter. Geological clutter, or geoclutter, refers to strong coherent sonar returns from geological features on or beneath the seafloor, such as buried channels, found in a number of coastal regions. The failure to distinguish between targets and geoclutter results in substantially increased false alarm rates [1]. One approach towards addressing this problem is to use aural cues for distinguishing between the two classes of signals. Although not new, it is generally agreed in the sonar community that targets and geoclutter are often aurally distinguishable [2]. However, this is based primarily on anecdotal evidence and has yet to be formally analyzed.

An initial investigation was recently undertaken to identify aural properties of targets and non-targets that could explain their dissimilarities when judged by listeners. Using examples from real sonar data, auditory recordings of sonar echoes were decomposed into individual complex tones; each complex tone had a fundamental frequency (F0) and harmonics (i.e., a harmonic is an integral multiple of F0). Based on the strongest amplitude of the complex tones for each sonar echo, preliminary observations revealed that both classes of sounds differ little in F0, and targets have fewer higher order harmonics than non-targets. These observations suggest that timbre may be one attribute of auditory sensation that could account for the aural differences in the discrimination of these classes of sounds. Timbre is defined as that attribute of a sound that a listener can judge as different when two or more sounds have the same loudness and pitch [3]. For example, listeners can distinguish between a C4 note played on a trumpet from the same note played on a piano. We report on a pilot study that investigated the perceptual interactions between timbre and pitch; pitch is defined as the ordering of sounds on a musical scale [3]. Subjects were asked to rate the timbre of complex tones for different levels of pitch.

2. METHOD

2.1. Subjects

Two subjects with musical training and two with no musical training participated. The subjects with musical training had formal instruction in playing a musical instrument. All subjects self-reported normal hearing.

2.2. Stimuli and apparatus

Three pairs of complex tones were synthesized using the Cmusic program. The sounds were sampled at 48 kHz, and had a duration of 500 ms with a 10 ms onset/decay cosine envelope. The F0s for these pairs of sounds were 30 Hz, 40 Hz, and 50 Hz. Low frequencies were employed in order to vary the pitch, thereby allowing the F0s and their corresponding harmonics to span the frequency range of human audibility. The sounds in each pair had the same F0; for each pair of sounds one sound had harmonics 1-5, and the other had harmonics 6-10. Varying the spectrum of the sound via its harmonics provided a means of assessing timbre based on previous findings which showed that complex tones with strong harmonics below the 6th sound mellow, whereas complex tones with strong harmonics beyond the 6th or 7th sound sharp and penetrating [3]. All the F0s and harmonics had the same phase and amplitude. Testing took place in a quiet room that contained a personal computer, headphones, and chair.

2.3. Experimental design

There were 12 experimental conditions arranged as a between- and within-subjects repeated measures design. The between-subject factor was musical training. The within-subject factors were 2 harmonic patterns for each of the 3 pairs of sounds, and 6 variations of pitch. Pitch was varied by successively doubling the frequency of the F0s of the 6 sounds (i.e., combinations of 3 F0s and 2 harmonic patterns). Each frequency doubling raised the pitch by an octave on a musical scale. A block comprised the 6 sounds with pitch held constant.

2.4. Procedure

Subjects were individually tested. On each trial a sound was dichoticly presented over the subject's headphones. Subsequently, a rating scale representing a continuum from 0-100% appeared on the computer monitor. A rating of 0 denoted the sound was extremely mellow, and a rating of 100 denoted that the sound was extremely sharp. The subject's task was to rate the perceived timbre of the sound by moving the cursor (positioned by default at 50%), using the computer mouse, to the desired position on the rating scale followed by pressing the "enter" key. Each block

contained 30 trials. The first 6 trials comprised 1 repetition of a random ordering of the 6 sounds. These were designated as practice trials and were intended as a means for the subject to become familiarized with the sounds. The remaining 24 trials comprised 4 repetitions of a random ordering of each sound. The presentation of blocks was in ascending order of pitch because a software bug prevented the randomization of the ordering of blocks. All the subjects participated in each experimental condition. The duration of the session was approximately 25 minutes.

3. RESULTS AND DISCUSSION

The practice trials were excluded from the analysis. The data were not statistically analyzed given that only 4 subjects participated. For both subject groups, the mean of the subjects' median ratings of perceived timbre for each sound in each block was calculated. The means were subsequently transformed to ranks; a rank of 1 indicates the sound having the highest sharpness. The judgement of timbre differed slightly between subject groups. For 5 of the 6 blocks, the sounds with harmonics 6-10 were ranked higher than the sounds with harmonics 1-5 which is in agreement with earlier findings [3].

These limited results suggest that subjects had no difficulty judging timbre. Musical training may thus not necessarily be advantageous in classifying targets and non-targets. The implications of the results obtained from the study on the perceptual interactions of timbre and pitch should help to further the understanding of the aural qualities that discern classes of sonar sounds, with the aim of leading to a reduction of false alarms. We emphasize that the present results are based on a pilot study and acknowledge that more work in this investigation is required.

4. REFERENCES

1. Makris NC, Ratilal P. A unified model for reverberation and submerged object scattering in a stratified ocean waveguide. J. Acoust. Soc. Am. 2001;109(3):909-941.
2. Coon AC, Ross CA, Chalmers RW, Gallati PC. The extended echo ranging aural and visual support trainer. Johns Hopkins APL Technical Digest 1997;18:113-123.
3. Moore BCJ. An introduction to the psychology of hearing, 3rd ed. San Diego, Academic Press, 1989.

RECENT PROJECTS IN BUILDING ACOUSTICS AT NRC

J.S. Bradley, B.N. Gover, R.E. Halliwell, T.R.T. Nightingale, J.D. Quirt, and A.C.C. Warnock
Indoor Environment Program, Institute for Research in Construction,
National Research Council, Ottawa, K1A 0R6

1. INTRODUCTION

This series of papers presents results from recent major projects on building acoustics at the National Research Council. In each case, major reports providing more detail on the projects are listed at the end of the brief abstracts. The reports are available from the NRC/IRC website at http://irc.nrc-cnrc.gc.ca/ircpubs/, using the code (e.g.- RR-170) in Report# field. For readers with detailed questions, the e-mail address for the most suitable contact is listed at the end of each item.

2. ACOUSTICAL DESIGN OF OPEN-PLAN OFFICES

J. S. Bradley

The recently completed COPE (Cost-effective Open Plan Environments) project examined all aspects of the design of open-plan offices: including lighting and air quality issues, as well as acoustical concerns. The research included both laboratory and field studies including measurements in 700 workstations in a number of buildings.

This presentation will describe how the acoustical aspects of this work and related studies at NRC have be used to develop a better approach to the acoustical design of open-plan offices. Important design parameters will be discussed and new software to simplify design calculations will be demonstrated.

COPE-Calc acoustical design software can be downloaded from http://irc.nrc-cnrc.gc.ca/ie/cope/07.html along with software related to the other aspects of open-plan design.

(Contact: John.Bradley@nrc-cnrc.gc.ca)

Reports available at http://irc.nrc-cnrc.gc.ca/ircpubs/:

1. Describing Levels of Speech Privacy in Open-Plan Offices (RR-138),
2. Measurements of Sound Propagation between Mock-Up Workstations (RR-145),
3. J.S. Bradley, The acoustical design of conventional open plan offices," Canadian Acoustics, 31, (2), (NRCC-46274)

3. METHODS FOR ASSESSMENT OF ARCHITECTURAL SPEECH SECURITY OF CLOSED OFFICES AND MEETING ROOMS

B. N. Gover and J. S. Bradley

For speech security, conversations taking place within closed offices and meeting rooms should be difficult or impossible to understand outside those rooms, in adjoining spaces. The degree to which words are audible or intelligible to a listener outside such a room depends on the transmitted speech level and background noise level at the position of the listener, and of course on the listener's hearing abilities.

The transmitted speech level depends on how loudly people are speaking inside the room, and on the amount of sound attenuation provided by the building. For a given voice level, a given background noise level, and a listener with typical hearing abilities, one can rate the degree of architectural speech security of the room. This enables specification, at the design stage, of criteria or constructions that will ensure an adequate level of speech security. Additionally, this allows existing rooms to be measured and rated in terms of the degree of speech security they provide.

This presentation will describe recent and ongoing work at NRC-IRC in these areas, including: subjective testing to determine a suitable signal-to-noise measure that reliably indicates audibility and intelligibility of speech in noise; measurements during actual meetings to determine speech levels within meeting rooms and background noise levels in adjoining spaces; and development of a field measurement protocol for assessment of existing rooms.

(Contact: Brad.Gover@nrc-cnrc.gc.ca)

Reports available at http://irc.nrc-cnrc.gc.ca/ircpubs/:

1. Speech and Noise Levels Associated with Meeting Rooms (RR-170)
2. Measures for assessing architectural speech security (RR-171 to be published late 2004)

4. STUDIES OF SOUND TRANSMISSION THROUGH FLOORS AND WALLS

A.C.C. Warnock

The paper will present an overview of practical information derived from projects focusing on airborne and impact sound transmission through walls and floors. Several projects have been completed in the last ten years or so. Factors such as the number of layers, type of studs, type of joists, and type and thickness of sound-absorbing material have been varied.

An accurate model that predicts transmission loss or impact sound pressure levels has not been developed. Instead, as a practical measure, regression equations have been developed that can be used to estimate single number ratings. The precision of the estimates is thought to be adequate for most practical purposes. A program that implements the regression equations is available and will be demonstrated.

Databases containing the one-third octave band data for the test results used to develop the regression equations are available. The paper will also describe some of the work that has been done on lightweight floating floors used to reduce impact sound transmission.

(Contact:Alf.Warnock@nrc-cnrc.gc.ca)

Reports at http://irc.nrc-cnrc.gc.ca/ircpubs/:

1. Airborne sound transmission through gypsum board walls (IR-693, IR-761, IR-832) and concrete block walls (BRN-217 and IR-586)
2. Airborne and Impact sound transmission through floors (IR-766, IR-811, and RR-169 to be published late 2004)
3. Change in impact sound transmission due to floor coverings (IR-802)
4. Other aspects (BRN-172 on windows, IR-772 on effect of electrical outlets)

5. DESIGN PRINCIPLES AND DETAILS FOR CONTROL OF FLANKING TRANSMISSION IN WOOD FRAMED MULTIFAMILY CONSTRUCTION

T.R.T. Nightingale, R.E. Halliwell, J.D. Quirt

This summary of two recent IRC projects provides a systematic investigation of design details that affect sound insulation, and evaluates several options for treating important flanking paths for both horizontally and vertically separated spaces.

The paper begins by showing the importance of wall type (single, double stud, or shear wall), joist type (wood-I or dimensional lumber), joist orientation (parallel or perpendicular to the wall/floor junction), joist and subfloor continuity (across the wall/floor junction). In general, these details change the magnitude of flanking for particular paths, but typically do not change their ranking.

- For vertically separated rooms, the dominant flanking path (if the floor does not have a topping) is from the subfloor to the lower wall. The path from the upper wall to the lower wall can be ignored for most practical constructions.
- For horizontally separated rooms, the dominant flanking path involves the floor if either the subfloor or the joists provide structural continuity at the wall/floor junction.

Typically, the most effective measure to control flanking will be in the form of a topping to the subfloor, since the dominant horizontal and vertical flanking paths involve this surface. The effectiveness is strongly dependent on several factors relating to the topping (i.e., mass, damping, interlayer, etc.) but also on the details of the floor to which it is applied, namely joist orientation. The paper concludes with estimates of the apparent sound insulation that might be expected for several complete building assemblies.

(Contact: Trevor.Nightingale@nrc-cnrc.gc.ca)

Reports at http://irc.nrc-cnrc.gc.ca/ircpubs/:

1. Sound Isolation and Fire Resistance of Assemblies with Fire Stops (IR-754).
2. Flanking Transmission in Multifamily Dwellings: Effects of Continuous Structural Elements at the Wall/Floor Junction (RR-103 and RR-168 to be published late 2004)

ABSTRACTS FOR PRESENTATIONS WITHOUT SUMMARY PAPERS

Acoustical Materials

Designing layered sound absorbing materials
Youssef Atalla
Groupe d'Acoustique et de Vibrations de l'Université de Sherbrooke (GAUS), 2500 Boul. de l'Université
Sherbrooke (Québec) J1K 2R1 Canada

Over the years, noise control in a vehicle is becoming a challenging part of the automobiles industry since more importance are paid to acoustic features. Sound packages particularly play a significant part to achieve the desired interior acoustics in a car within bounds to meet customer requirements. Floor covering as well as the headliner represent the two largest pieces of acoustic treatment in the interior of a vehicle in which a composite construction is considered. The construction involve usually multiple layers of different sound absorbing materials. However the overall acoustic performance of multiple layered system is dependent on the type of the materials selected and how they are arranged in the construction. For a well acoustically designed construction, some guidelines on how to select the porous layers involved in the construction are necessary, which in turn require the knowledge of the material properties. The porous elastic modeling of sound absorbing materials and the knowledge of the material properties of each porous layer allow to perform by computer a systematic evaluation of potential layer combinations rather than do it experimentally which is time consuming and always not efficient. In this paper, a guidelines are proposed to design efficiently such systems. The design method proposed takes into account some physical aspect of the absorption mechanism at low and high frequencies and the porous material properties. As a first attempt, a two-layer configuration is retained. It is shown that the proposed final two-layer designs provide a good acoustic performance while reducing weight.

Acoustics of Educational Facilities

Questionnaire study of the effect of renovations on student perception of the listening environments in university classrooms
Murray Hodgson and Hugh Davies
School of Occupational and Environmental Hygiene, University of British Columbia, Vancouver, BC V6T 1Z3

The research reported here investigated the effect of renovations on student perception of the listening environment in university classrooms. It involved four large classrooms at the University of British Columbia which were studied before and after acoustical renovation. Details of the renovations were identified. Measurements of reverberation time, background-noise level and sound propagation were made in each case. Room-average Speech Transmission Indices were calculated as a summary physically-based quality measure, and their changes on renovation determined. An existing questionnaire, designed to determine student perception of the listening environment, as well as the personal, academic and environmental factors that might affect it, was administered to a large number of students taking classes in each classroom before and after renovation. A summary Perceived Listening Ease (PLE) score, measuring student perception of the quality of the listening environment, was calculated for each questionnaire from the responses. Room-average PLE scores, and the changes on renovation, were then calculated. These results were related to the results of the physical measures. Physical, personal, academic and environmental factors affecting the changes in PLE score were investigated. The implications of the results for classroom design were considered.

Speech recognition in classroom noise by school-age children
Michel Picard, Tony Leroux, Benoit Jutras
Ecole d'orthophonie et d'audiologie, Université de Montréal, Montréal, Qc, Canada, H3C 3J7

Classroom noise and interacting poor acoustical design are serious obstacles to children's understanding of instructions because typically interfering with speech recognition. Given the importance of these issues, we conducted two experiments with normal-hearing students attending regular school. Participants were 1st graders, 3rd graders, and 6th graders based on the assumption of a developmental speech recognition function in noise between ages of approximately 6 and 10-12 years old. The purpose of experiment 1 was to determine how speech recognition varies with signal-to-noise ratio (SNR) in groups of young speakers of English using a closed-set response format test, the WIPI. Experiment 2 was intended to determine the similarity of findings in groups of young native speakers of French compared to English ones at a constant SNR of +5 dB. Participants were tested individually in a sound-proof booth. They were required to listen to the speech material presented by a loudspeaker at 80 cm in front of them over a background of noise. These findings were interpreted as adding generality to the particular phenomenon of developmental speech recognition function in noise in between ages of 6 and 12 as may be representative of speech understanding in elementary school.

Impact of noise on academic skills in elementary school-aged children.
Prudence Allen, Nashlea Brogan, Chris Allan and Michelle Baker

National Center for Audiology, University of Western Ontario, London ON

The level of noise in classrooms is often much higher than is recommended for good speech intelligibility and, in retrospective studies, it is often suggested that higher noise levels correlate significantly with poorer academic performance [e.g., Shield., et al. (2002)]. There has been, however, little data examining the effect of noise on the academic performance of individual children obtained from well controlled laboratory studies. In a recent series of studies the effect of noise on the academic abilities (vocabulary, oral and silent reading, general academic achievement including word reading, spelling, and arithmetic) of children was examined using parallel forms of standardized tests. Noise consisted of pre-recorded general classroom sounds presented at 60 dB SPL. Signals, when required, were presented at a +10 dB Signal-to-Noise ratio. Results showed fairly small effects of the noise. Data will be discussed in terms of age-related, task specific, and individual differences. [Work supported by CLLRNet.]

Hearing and the Workplace

Risk factors for noise-induced hearing loss in Canadian Forces personnel.

Sharon M. Abel and Stephanie Jewell
Defence Research and Development Canada – Toronto, 1133 Sheppard Ave., W., Toronto, Ontario, Canada M3M 3B9

The cost of claims for noise-induced hearing loss in the Canadian Forces supports the need to review and upgrade current hearing conservation practices. A prospective study was conducted to assess risk factors for the development of hearing loss in a wide range of military trades. A total of 1,057 individuals working in land, sea and air environments contributed their most recent audiogram and first audiogram on record and completed a 56-item questionnaire relating to demographics, occupational and non occupational noise exposure history, training in and utilization of personal hearing protectors, and factors other than noise which might affect hearing (e.g., head injury, solvent exposure and leisure noise). Medical personnel at five participating military bases recruited the subjects and assessed hearing. Apparatus and protocols for the latter conformed to current clinical practice. The prevalence of moderate to severe hearing loss progressed with noise exposure, with hearing thresholds in those over 46 years ranging broadly from normal to over 70 dB HL. Solvent exposure appeared to be a significant determinate of adverse outcome, while the effect of head injury, history of ear disease, and the use of medications was small. The survey suggested that training on the hazards of noise exposure and hearing protector selection and utilization was poor. Hearing protectors were sometimes reported to be incompatible with other gear, uncomfortable and an impediment to communication. [Work supported by Veterans Affairs Canada]

Instrumentation and Measurements

Motion of microobject induced by the resonant vibration via acoustic power

Yujie Han, Furuichi Takuya and Yuji Furukawa
Department of Mechanical System, Tokyo University of Agriculture and Technology, 2-24-16 Naka-cho, Koganei, Tokyo, 184-8588, Japan

Compact and high-efficiency micro-power sources are a key technology in future integrated micro-systems that enable multifunctional, high-performance sensing, computing, actuation, control, and communication on a single chip for remote and autonomous operation. Fuel cell, solar cell, Coulomb electrostatic force, and optical, magnetic, acoustic or thermal effects have been used to actuate microobjects and small parts with some advantages and drawbacks.

Sound wave allows generating and sensing vibrations on a large dynamic range and a wide frequency band so that a simple system or part is possible. Sound energy can be efficiently transmitted through solid, liquid and gas. Vibration control can be realized by adopting different shapes and structures for modal control of its natural frequencies. Models to move small parts have been made and observed using a CCD camera system. ANSYS is used to analyze the vibration modes of the prepared models. Both translational and rotational movements have been observed.

– Modes occur at certain frequencies
– At these frequencies, input amplitude is much smaller than the output amplitude of the object.
– The frequency depends on geometry, strain, Modulus of elasticity, density, and constraints.
– The first mode has the largest amplitude which makes knowledge of the frequency of the first mode important.

Noise Emission and Noise Control

Availability of noise emission data for machinery sold in Canada: a pilot survey

Stephen Bly1 and Bhawani Pathak2
1Health Canada, Consumer and Clinical Radiation Protection Bureau, 775 Brookfield Road, Ottawa, Ontario K1A 1C1
2Canadian Centre for Occupational Health and Safety, 250 Main St. E., Hamilton, Ontario L8N 1H6

Noise emission declarations are intended to help reduce noise from workplace machinery and outdoor equipment. In December, 2003, the CSA standard, Z107.58–02 "Noise Emission Declarations for Machinery" became a National Standard of Canada. This paper presents the results of a pilot survey completed in June 2003, to obtain baseline data on the availability of noise emission data to Canadian purchasers. One hundred manufacturers and distributors of machinery used in Canadian workplaces were surveyed. Manufacturers or distributors of specific types of machines, such as compressors and fans were included as well as a variety of equipment

categorized by use in a number of industries. These categories were: (i)auto repair, (ii)construction, (iii)food processing and packaging, (iv)furniture making, (v)logging, (vi)metal manufacturing, (vii)mining, (viii)pulp and paper, (ix)saw mills and (x)shingle mills. The survey was done by means of telephone interviews using a questionnaire. Information was obtained on whether noise emission data was provided and, if so, the method used. The survey data revealed that 54 of the manufacturers /suppliers responded that they provided noise emission data either voluntarily or on request by the buyer. Twenty-nine of these 54 responses indicated that the data was provided in conformance with a standard.

The Influence of Area and Aspect Ratio on Sound Power Splitting at Duct Branches

Werner Richarz1 and Harrison Richarz2
1Aercoustics Engineering Ltd., 50 Ronson Dr., Toronto, ON M9W 1B3
2Dept. of Music, McGill University, 555 Sherbrooke St. W, Montreal, QC H3A 1E3

In HVAC noise prediction one normally splits sound power entering branches according to the area ratio. The practise is based on the results of the so-called 'lumped impedance' model. This split assumes that the termination impedances are identical. For frequencies where the wave nature of the sound field becomes important, the simple analysis breaks down. A series of tests has been conducted to document the effect of the duct geometry on the sound power splitting. To this end a 90o take-off has been configured in a variety of area ratios as well as duct aspect ratios. The sound power radiated as well as the effective impedance offered by the branch have been measured. The results are summarized in a format suitable for design and analysis.

Effectiveness of noise barriers tilted towards the traffic way over bridges

M.A. Ibrahim1, M.A Kenawy2, M. ABD EL Gawad Saif3, and A.A Mahmoud1
1Prof. Dr. National Institute for Standards Giza 12211 – POB 136 (Egypt)
2Prof. Dr. Ain-Shams University (Egypt)
3Dr. Eng. Housing & Building Research Center (HBRC) (Egypt)

Traffic noise over bridges and its impacts can be substantially reduced with road side barriers. A single barrier erected on one side of a bridge way can often provide the most effective solution. A single barrier may not always be sufficient, particularly in urban sittings. In these cases, barriers may be necessary on both sides of the bridge. This paper presents a study of the effectiveness of barriers tilted towards the traffic way with different angles and shapes to reduce the effect of the reflected sound waves from the second barrier. The insertion loss was determined at many points distributed horizontally and vertically behind the barrier.

Outdoor Sound

Validation of ISO 9613 Part 2

Ramani Ramakrishnan1, Colin Novak2, Helen Ule2, and Robert Gaspar2
1Department of Architectural Science, Ryerson University, 350 Victoria Street, Toronto
2University of Windsor, Mechanical, Automotive and Materials Engineering, 401 Sunset Ave. Windsor

ISO 9613 Standard provides a framework for the evaluation of noise levels radiating from sources. It generates simple schemes to account for the topographical as well as atmospheric complexities that may exist between the source region and the receptor locations. However, there are a number of unanswered concerns that remain to be addressed as far the accuracy of ISO-9613 procedures. An attempt will be made in this paper to validate the ISO-9613 procedures. Actual industrial sources will be used for this investigation. A number of modeling programs, to predict outdoor noise levels from industrial sites, with various levels of complexity are available in the market. These programs use the modeling algorithm developed in the ISO 9613 Standard. The current version of an outdoor noise propagation modeling software, Cadna, will be used for the predictive calculation of noise levels radiating from the industrial site. The sound power levels, known a priori, of the noise sources at the industrial sites will be used in the modeling software. The sound pressure levels, both near-field and far-field, will be measured over an extended duration. The sound levels will be expressed both as octave band sound pressure levels and averaged A-weighted overall levels. In addition to the noise measurements, the weather conditions during the monitoring time will also be monitored. The measurements will be compared to the predicted sound pressure levels. The accuracy of the ISO-9613 procedures can thus be verified.

Signal Processing

Practical applications of conformal microphone arrays

Stéphane Dedieu
Acoustic Design, Mitel Networks, 350 Legget Dr., Kanata (Ottawa), Ontario

Microphone arrays in research applications have generally been free field where the object is to make the array as unobtrusive as possible. In trying to incorporate microphone arrays in consumer product generally this has also been attempted. Until Mitel undertook collaborative research with NRCC it seems that exploiting the diffractive effects of the mounting object has not really been recognised. In that application the object was to sample the pressure at the surface of the object and to use it in a beamformer.
Mitel has since continued research on the modelling and use of such arrays and our current results will be presented. Exploiting the object and carefully choosing the shape of the

diffractor permit us to reduce the computational complexity of beamformers in high frequencies and also permit more accurate source localisation algorithms.

Sound Transmission and Room Acoustics

Case Study: Retrofit Acoustic Treatments in a Heritage Apartment Building

Hugh Williamson

Hugh Williamson Associates Inc., 12 Maple Lane, Ottawa, Ontario, K1M 1G7, Canada

When renovating older buildings there is frequently a desire to improve sound isolation between rooms or apartments. However, the renovator is faced with a dilemma: little acoustical guidance or data exists to assist in the selection of retrofit techniques for older types of construction such as lath and plaster or balloon framing. In this case study, sponsored by Canada Mortgage and Housing Corporation, a series of retrofit treatments have been applied to floor/ceilings in a heritage apartment building. The study provides field data on the improvements in sound and impact isolation achieved through the retrofit treatments. Critical issues such as flanking transmission are discussed in detail. Comparative costs for retrofit treatments are also presented.

Comparison and validation of prediction models based on Modal and Image-Phase methods

Galen Wong

1328 East 27th Avenue, Vancouver, B.C. V5V 2L8

A prediction model for calculating the sound pressure field in a room based on its modal response was developed and compared with a model based on the image-phase method. The objective for the use of these prediction models was to be able to predict with good reliability the low frequency response of rooms with respect to pure tone sources. Both methods employ a pure tone source, are currently limited to predictions in rectangular rooms, and are suitable for low-frequency predictions. Results show good agreement in some limited cases. Differences may be somewhat explained with the complications which arose in the specification of wall impedance and absorption.

Speech Understanding and Psychoacoustics

Vision and Speech Perception

Julie N. Buchan and Kevin G. Munhall

Department of Psychology, Queen's University, Kingston, Ontario, K7L 3N6

Vision and audition are separate modalities with different processing stages, but in our everyday experience we usually integrate the information from both senses to form a single perceptual event. This unity of sensory information is also true in communication. Visual information plays an important role in how we perceive auditory speech. Watching someone speak improves speech intelligibility when the auditory signal is degraded. Visual information can also change the perception of perfectly audible speech signals. For example, when conflicting visual and auditory information are presented simultaneously to a participant, the visual information can influence what the participant "hears". This is known as the McGurk effect and can be used as a measure of audiovisual speech integration. In this presentation we will summarize the current literature on audiovisual perception of speech-in-noise and the McGurk effect. We will report recent findings from our laboratory on eye tracking and spatial frequency filtering of the visual stimuli that show that only relatively crude visual information is required for audiovisual integration. Our presentation will include demonstrations of the McGurk effect and other audiovisual speech phenomena.

Predicting speech perception in noisy environments from the clinical HINT

C. Laroche, C. Giguère, S. Soli, V. Vaillancourt

Programme d'audiologie et d'orthophonie, Université d'Ottawa, Ottawa, Ontario K1H 8M5

Many of the jobs in Canadian Coast Guard (CCG) are hearing-critical. These jobs have several features in common: they are often performed in noisy environments and involve a number of auditory skills and abilities. These skills and abilities include speech communication, sound localization, and sound detection. If an individual lacks these skills in a sufficient amount, it may constitute a safety risk for this individual, as well as for fellow workers and the general public. A number of scientific models have been developed that allow one to predict performance on these auditory skills based on diagnostic measures of hearing such as pure-tone audiograms. While these models have significant scientific and research value, they are unable to make sufficiently accurate predictions of real life performance on auditory skills to be used for individual personnel actions regarding hearing-critical jobs. An alternative and more accurate approach has been developed in this research project. Direct measures of functional hearing ability have been identified and validated for use in screening applicants for hearing-critical jobs in the CCG. These tests have adequate psychometric properties, i.e., reliability, sensitivity, and validity, so that test results accurately predict the individual's ability to use critical auditory skills in the sound environment of the hearing-critical jobs. The detailed approach will be presented.

Location-Based Auditory Inhibition of Return Is Perceptually Driven

Chun Wang1, Yang Wang1, Hao Wu1, Qiang Huang1, Tianshu Qu1, Xihong Wu1, Xiaolin Zhou1, Liang Li1,2

1Department of Psychology, Speech and Hearing Research Center, National Key Laboratory on Machine Perception, Center for Brain and Cognitive Sciences, Peking University, Beijing, China, 100871
2Centre for Research on Biological Communication Systems, Department of Psychology, University of Toronto at Mississauga, Mississauga, Ontario, Canada L5L 1C6

Location-based auditory inhibition of return (IOR) refers to the delay in detecting an auditory target presented at the same spatial location as a preceding cue. Previous studies have suggested that IOR reflects a tendency to inhibit a return of attention to recently inspected locations. The present study attempted to distinguish whether location-based auditory IOR is a binaural-channel driven process or a higher-order perceptually driven process. When the delay between two similar and spatially separated sounds is sufficiently short, only a "fused" image is perceived as originating from the leading sound region, a phenomenon called precedence effect (PE). When the stimulus-onset asynchrony (SOA) between the cue and target was 60 or 100 ms, the reaction time (RT) to the target was shorter when the PE-manipulated perceived location of the target was the same as the cue location than when different from the cue location. However, when the SOA was 260 or 300 ms, the RT was longer when target and cue were perceived at the same location than when they were perceived to originate from different locations. These results indicate that location-based auditory IOR still operates even when the perceived target location is determined by PE and the activated binaural channels are unchanged. Thus location-based auditory IOR is mainly a high-order perceptually driven process.

Detection of Uncorrelated Sound Fragment Embedded in Correlated Noise Sounds: A Measurement of the Interaural Memory and Auditory Aging

Juan Huang1, Xihong Wu1, James G. Qi1, Bruce A. Schneider2, Liang Li1,2

1Department of Psychology, Speech and Hearing Research Center, National Key Laboratory on Machine Perception, Peking University, Beijing, China, 100871
2Department of Psychology, Center for Research on Biological Communication Systems, University of Toronto at Mississauga, Mississauga, Ontario, Canada L5L 1C6

In reverberant environments, when delays between the direct sound wave and its reflections are sufficiently short, only a single "fused" image is perceived (the precedence effect). Such sound fusion occurs only when there is a strong correlation between the original sound and its reflections. The present study shows that listeners' detection of a break in the correlation (a correlation gap) between a sound and its hypothetical reflection dramatically improves as the delay between the two sounds decreases. Younger listeners could detect a correlation gap even when the sound delay was larger than the threshold for fusing the two sounds (so called echo threshold). Older listeners could detect the correlation gap only when the sound delays were shorter than those for younger listeners. These results indicate that the processing of dynamic correlations between two noise sounds is strongly affected by the delay between the two sounds and involves high-order central processes. That older listeners cannot detect a change in correlation at delays that are detectable to younger listeners suggests that there are age-related declines in auditory memory, because the detection of interaural correlations when delays are involved, requires the comparison of the lagging signal with a memory representation of the leading signal [work supported by the Natural Sciences and Engineering Research Council of Canada, the Canadian Institutes of Health Research, and the Ministry of Science and Technology of China].

Underwater Sound

Estimating geoacoustic properties of marine sediments by matched field inversion using ship noise

M.G. Morley, N.R. Chapman, and S.E. Dosso
School of Earth and Ocean Sciences, University of Victoria, P.O. Box 3055, Victoria, B.C.

Geoacoustic inversion for seabed properties using ambient noise sources in the ocean has received a great deal of interest recently. A project is currently underway to establish a remote gas hydrate monitoring station in the Gulf of Mexico. The goal of the project is to detect changes that occur in the acoustic properties of gas hydrate bearing sediments due to the formation and dissocation of hydrates. Temporal monitoring of these properties will be achieved by matched field inversion using noise from passing ships of opportunity as a sound source. In this study, the method is investigated using ship noise data collected on a prototype vertical line array deployed in the Atwater Valley region of the Gulf of Mexico. A Bayesian inversion approach is applied to estimate the posterior probability distributions (PPD) of the geoacoustic model parameters. The moments of the PPD yield estimates of the parameter uncertainties and correlations. A sensitivity study is performed to determine the minimum change in the model parameters that can be resolved by the data.

The story behind the June 2004 issue of Canadian Acoustics

Francine Desharnais
Defence R&D Canada - Atlantic, P.O. Box 1012, Dartmouth, Nova Scotia, B2Y 3Z7, Canada

The concern that acoustic signals can affect marine mammals has increased over the past few years, mainly within the con-

text of military active sonars or seismic operations. Whether it is in support of mitigation measures, or in the larger context of marine mammal studies, recent years have seen a significant increase in research on marine mammal detection and localization techniques using passive acoustics. As more groups are looking at automating some of these algorithms in real-time or near real-time systems, there was a need for a forum in which the efficiency of state-of-the-art algorithms could be compared. In November 2003, DRDC Atlantic and Dalhousie University jointly hosted a first Workshop on Detection and Localization of Marine Mammals Using Passive Acoustics. In June 2004, Canadian Acoustics took on a new initiative by printing a Special Issue with the Proceedings of the workshop. This special issue significantly increased the distribution of Canadian Acoustics, and enhanced its international visibility. This presentation will discuss the primary results from the workshop, and the benefits to the CAA organization.

NEWS / INFORMATIONS

CONFERENCES

If you have any news to share with us, send them by mail or fax to the News Editor (see address on the inside cover), or via electronic mail to stevenb@aciacoustical.com

2004

13-16 September: Subjective and Objective Assessment of Sound, Poznan, Poland. Contact: Institute of Acoustics, Adan Mankiewicz University, Poznan, Poland. Web: www.ia.amu.edu.pl/index.html

13-17 September: 4th Iberoamerican Congress on Acoustics, 4th Iberian Congress on Acoustics, 35th Spanish Congress on Acoustics, Guimarães, Portugal. Contact: Sociedade Portuguesa de Acústica, Laboratório Nacional de Engenharia Civil, Avenida do Brasil 101, 1700-066 Lisboa, Portugal; Fax: +351 21 844 3028; E-mail: dsilva@lnec.pt

14 - 16 September: International Conference on Sonar Signal Processing and Symposium on Bio-Sonar Systems and Bioacoustics. Loughboro, UK. Contact: D. Gordon, Fax: +44 1509 22 7053, Web: http://ioa2004.lboro.ac.uk

15-17 September: 26th European Conference on Acoustic Emission Testing. Berlin, Germany. Contact: DGZIP, Max-Planck-Str. 25, 12489 Berlin, Germany. Web: www.ewgae2004.de

20-22 September: International Conference on Noise and Vibration Engineering. Leuven, Belgium. Contact: Fax +32 16 32 29 87. Web: www.isma-isaac.be/fut_conf/default_en.phtml

20-22 September: ACTIVE 2004, International Symposium on Active Control of Sound and Vibration. Williamsburg, Virginia, USA. Contact: Richard J. Silcox, Mail Stop 463, NASA Langley Research Center, Hampton VA 23681, Ph: 757-864-3590, email: r.j.Silcox@larc.nasa.gov

20-22 September: 9th International Workshop "Speech and Computer". St. Petersburg, Russia. Web: www.spiiras.nw.ru/speech

27 - 29 September: Hellenic National Conference on Acoustics. Thessaloniki, Greece. Web: www.wlc.ee.upatras.gr/helina/ac2004.htm

28-30 September: Autumn Meeting of the Acoustical Society of Japan. Naha, Japan. Contact: Fax +81 3 5256 1022. Web: www.soc.nii.ac.jp/asj/index-e.html

4-9 October 8th Conference on Spoken Language Processing (INTERSPEECH). Jeju Island, Korea. Web: www.icslp2004.org

5-8 October: 11th Mexican International Congress on Acoustics. Morelia, Mexico. Email: sberista@hotmail.com

CONFÉRENCES

Si vous avez des nouvelles à nous communiquer, envoyez-les par courrier ou fax (coordonnées incluses à l'envers de la page couverture), ou par courriel à stevenb@aciacoustical.com

2004

13-16 septembre: Subjective and Objective Assessment of Sound, Poznan, Poland. Contact: Institute of Acoustics, Adan Mankiewicz University, Poznan, Poland. Web: www.ia.amu.edu.pl/index.html

13-17 septembre: 4e Congrès ibéro-américain d'acoustique, 4e Congrès ibérien d'acoustique, 35e Congrès espagnol d'acoustique, Guimarães, Portugal. Info: Sociedade Portuguesa de Acústica, Laboratório Nacional de Engenharia Civil, Avenida do Brasil 101, 1700-066 Lisboa, Portugal; Fax: +351 21 844 3028; Courriel: dsilva@lnec.pt

14 - 16 septembre: Conférence internationale sur Sonar Signal Processing et Symposium sur Bio-Sonar Systems et Bioacoustics. Loughboro, UK. Contact: D. Gordon, Fax: +44 1509 22 7053, Web: http://ioa2004.lboro.ac.uk

15-17 septembre: 26e Conférence European sur Acoustic Emission Testing. Berlin, Germany. Contact: DGZIP, Max-Planck-Str. 25, 12489 Berlin, Germany. Web: www.ewgae2004.de

20-22 septembre: Conférence internationale sur Noise and Vibration Engineering. Leuven, Belgium. Contact: Fax +32 16 32 29 87. Web: www.isma-isaac.be/fut_conf/default_en.phtml

20-22 septembre: ACTIVE 2004, International Symposium on Active Control of Sound and Vibration. Williamsburg, Virginia, USA. Contact: Richard J. Silcox, Mail Stop 463, NASA Langley Research Center, Hampton VA 23681, Ph: 757-864-3590, email: r.j.Silcox@larc.nasa.gov

20-22 septembre: 9th Internationale Workshop "Speech and Computer". St. Petersburg, Russia. Web: www.spiiras.nw.ru/speech

27 - 29 septembre: Hellenic National Conference on Acoustics. Thessaloniki, Greece. Web: www.wlc.ee.upatras.gr/helina/ac2004.htm

28-30 septembre: Rencontre d'automne de l'Acoustical Society of Japan. Naha, Japan. Contact: Fax +81 3 5256 1022. Web: www.soc.nii.ac.jp/asj/index-e.html

4-9 octobre 8e Conférence sur Spoken Language Processing (INTERSPEECH). Jeju Island, Korea. Web: www.icslp2004.org

5-8 octobre: 11th Mexican Congress Internationale sur Acoustics. Morelia, Mexico. Email: sberista@hotmail.com

6-7 October: Institute of Acoustics (UK) Autumn Conference. Oxford, UK. Web: www.ioa.org.uk/

6-8 October: Acoustics Week in Canada, Conference of the Canadian Acoustical Association. Ottawa, Ontario. Contact: John Bradley, National Research Council of Canada, Ottawa Ontario. Phone (613) 993-9747, Fax (613-954-1495, web: www.caa-aca.ca

13-16 October: 7th annual Canadian Academy of Audiology Conference, Quebec City. Contact: 250 Consumers Road, Suite 301, Toronto, Ontario M2J 4V6. Phone 1-800-264-5106, Fax (416) 495-8723. web: www.canadianaudiology.ca

27-29 October: 25th Symposium on Ultrasonic Electronics. Sapporo, Japan. Web: http://use-jp.org/USEframepage_E.html

3-5 November: Australian Acoustical Society Conference - Transportation Noise and Vibration. Surfers Paradise/Gold Coast, Queensland, Australia. Web: www.acoustics.asn.au

4-5 November: Autumn Meeting of the Swiss Acoustical Society, Rapperswil, Switzerland. Contact: SGA-SSA, c/o Akustik, Suva, P.O. Box 4358, 6002 Luzern, Switzerland; Fax: +41 419 62 13; Web: www.sga-ssa.ch

8-9 November: 17th Biennial Conference of the New Zealand Acoustical Society. Wellington, New Zealand. Web: www.acoustics.org.nz

15-18 November: 15th Session of the Russian Acoustical Society. Nizhny Novgorod, Russia. Contact: N.N. Andreyev Acoustics Institute, 4 Shvernika ul, Moscow 117036, Russia. Fax: +7 95 1260100, web: www.akin.ru

15-19 November: 148th Meeting of the Acoustical Society of America, San Diego, CA. Contact: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; Web: asa.aip.org

17-19 October: 7th Congress of the Turkish Acoustical Society. Nevsehir (Cappadocia), Turkey. Web: www.tak.der.org

8-10 December: 10th Austrailian International Conference on Speech Science and Technology. Sydney, Australia. Web: www.assta.org

2005

14-17 March: 31st Annual Meeting of the German Acoustical Society (DAGA). Munich, Germany. Web: www.daga2005.de

19-23 March: International Conference on Acoustics, Speech, and Signal Processing. Philadelphia, PA, USA. Web: www.icassp2005.org

18-21 April: International Conference on Emerging Technologies of Noise and Vibration Analysis and Control. Saint Raphael, France. Fax: +33 4 72 73 87 12, email: goran.pavic@insa-lyon.fr

6-7 octobre: Institute of Acoustics (UK) Autumn Conference. Oxford, UK. Web: www.ioa.org.uk/

6-8 octobre: Congrès annual de l'Association canadienne d'acoustique. Ottawa, Ontario. Contact: John Bradley, National Research Council of Canada, Ottawa Ontario. Phone (613) 993-9747, Fax (613-954-1495, web: www.caa-aca.ca

13-16 octobre: 7e Congrès annual de l'Academy canadian de audiology, Quebec City. Contact: 250 Consumers Road, Suite 301, Toronto, Ontario M2J 4V6. Phone 1-800-264-5106, Fax (416) 495-8723. web: www.canadianaudiology.ca

27-29 octobre: 25th Symposium sur Ultrasonic Electronics. Sapporo, Japan. Web: http://use-jp.org/USEframepage_E.html

3-5 novembre: Australian Acoustical Society Conference - Transportation Noise and Vibration. Surfers Paradise/Gold Coast, Queensland, Australia. Web: www.acoustics.asn.au

4-5 novembre: Rencontre d'automne de la Société suisse d'acoustique, Rapperswil, Suisse. Info: SGA-SSA, c/o Akustik, Suva, P.O. Box 4358, 6002 Luzern, Switzerland; Fax: +41 419 62 13; Web: www.sga-ssa.ch

8-9 novembre: 17th Biennial Conference of the New Zealand Acoustical Society. Wellington, New Zealand. Web: www.acoustics.org.nz

15-18 novembre: 15th Session de l'Acoustical Society du Russia. Nizhny Novgorod, Russia. Contact: N.N. Andreyev Acoustics Institute, 4 Shvernika ul, Moscow 117036, Russia. Fax: +7 95 1260100, web: www.akin.ru

15-19 novembre: 148e rencontre de l'Acoustical Society of America, San Diego, CA. Info: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tél.: 516-576-2360; Fax: 516-576-2377; Courriel: asa@aip.org; Web: asa.aip.org

17-19 octobre: 7th Congress de l'Acoustical Society du Turkey. Nevsehir (Cappadocia), Turkey. Web: www.tak.der.org

8-10 decembre: 10th Austrailian Internationale Conference sur Speech Science et Technology. Sydney, Australia. Web: www.assta.org

2005

14-17 mars: 31st Annual Meeting de l'Acoustical Society du Germany (DAGA). Munich, Germany. Web: www.daga2005.de

19-23 mars: Conference Internationale sur Acoustics, Speech, et Signal Processing. Philadelphia, PA, USA. Web: www.icassp2005.org

18-21 avril: Conference International sur Emerging Technologies de Noise et Vibration Analysis et Control. Saint Raphael, France. Fax: +33 4 72 73 87 12, email: goran.pavic@insa-lyon.fr

16-19 May: SAE Noise and Vibration Conference, Grand Traverse Resort, Traverse City Michigan. Contact: Mrs. Patti Kreh, SAE International, 755 W Big Beaver Rd, Ste 1600, Troy, Michigan, 48084. Tel: (248) 273-2474, E-mail: pkreh@sae.org

16-20 May: 149th Meeting of the Acoustical Society of America, Vancouver, BC, Canada. Contact: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; Web: asa.aip.org

1-3 June: 1st International Symposium on Advanced Technology of Vibration and Sound. Hiroshima, Japan. Web: http://dezima.ike.tottori-u.ac.jp/vstech2005

20-23 June: IEEE Oceans05 Europe. Brest, France. Web: www.oceans05europe.org

23-24 June: 2nd Congress of the Alps-Adria Acoustics Association (AAAA2005). Opatija, Croatia. Web: http://had.zea.fer.hr

28 June - 1 July: International Conference on Underwater Acoustic Measurements: Technologies and Results. Heraklion, Crete, Greece. Web: http://UAmeasurements2005.iacm.forth.gr

11-14 July: 12th International Congress on Sound and Vibration (ICSV12). Lisbon, Portugal. Web: www.icsv12.ist.utl.pt

18-21 July: 17th International Symposium on Nonlinear Acoustics (ISNA 17). Pennsylvania State University, PA, USA. Web: http://outreach.psu.edu/c&i/isna17

6-10 August: Inter-Noise, Rio de Janeiro, Brazil. Details to be announced later.

28 August – 2 September: Forum Acusticum Budapest 2005, Budapest, Hungary. Fax: +36 1 202 0452; Web: www.fa2005.org; E-mail: sea@fresno.csic.es

4-8 September: 9th Eurospeech Conference, Lisbon, Portugal. Contact: Fax: +351 213145843. Web: www.interspeech2005.org

4-8 September: 9th Eurospeech Conference (EUROSPEECH2005). Lisbon, Portugal. Web: www.interspeech2005.org

5-9 September: Boundary Influences in High Frequency, Shallow Water Acoustics. Bath, UK (Details to be announced later)

11-15 September: 6th World Congress on Ultrasonics (WCU 2005). Beijing, China. Web: www.ioa.ac.cn/wcu2005

14-16 September: Autumn Meeting of the Acoustical Society of Japan. Sendai, Japan. Web: www.asj.gr.jp/index-en.html

18-21 September: IEEE International Ultrasonics Symposium. Rotterdam, The Netherlands. Web: www.ieee-uffc.org

16-19 mai: SAE Noise and Vibration Conference, Grand Traverse Resort, Traverse City Michigan. Contact: Mrs. Patti Kreh, SAE International, 755 W Big Beaver Rd, Ste 1600, Troy, Michigan, 48084. Tel: (248) 273-2474, E-mail: pkreh@sae.org

16-20 mai: 149e rencontre de l'Acoustical Society of America, Vancouver BC, Canada. Info: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tél.: 516-576-2360; Fax: 516-576-2377; Courriel: asa@aip.org; Web: asa.aip.org

1-3 juin: 1st Symposium International sur Advanced Technology de Vibration et Sound. Hiroshima, Japan. Web: http://dezima.ike.tottori-u.ac.jp/vstech2005

20-23 juin: IEEE Oceans05 Europe. Brest, France. Web: www.oceans05europe.org

23-24 juin: 2nd Congress de l'Association Acoustique Des Alps-Adria (AAAA2005). Opatija, Croatia. Web: http://had.zea.fer.hr

28 juin - 1 juillet: Conference Internationale sur Underwater Acoustic Measurements: Technologies and Results. Heraklion, Crete, Greece. Web: http://UAmeasurements2005.iacm.forth.gr

11-14 juillet: 12th Congress Internationale sur Sound et Vibration (ICSV12). Lisbon, Portugal. Web: www.icsv12.ist.utl.pt

18-21 juillet: 17th Symposium Internationale sur Nonlinear Acoustics (ISNA 17). Pennsylvania State University, PA, USA. Web: http://outreach.psu.edu/c&i/isna17

6-10 août: Inter-Noise, Rio de Janeiro, Brésil. Information à suivre.

28 août – 2 septembre: Forum Acusticum Budapest 2005, Budapest, Hongrie. Fax: +36 1 202 0452; Web: www.fa2005.org; E-mail: sea@fresno.csic.es

4-8 septembre: 9e Conférence d'Eurospeech, Lisbon, Portugal. Contact: Fax: +351 213145843. Web: www.interspeech2005.org

4-8 septembre: 9th Eurospeech Conference (EUROSPEECH2005). Lisbon, Portugal. Web: www.interspeech2005.org

5-9 septembre: Boundary Influences in High Frequency, Shallow Water Acoustics. Bath, UK (Details to be announced later)

11-15 septembre: 6th World Congress sur Ultrasonics (WCU 2005). Beijing, China. Web: www.ioa.ac.cn/wcu2005

14-16 septembre: Autumn Meeting de la Society Acoustical du Japan. Sendai, Japan. Web: www.asj.gr.jp/index-en.html

18-21 septembre: IEEE International Ultrasonics Symposium. Rotterdam, The Netherlands. Web: www.ieee-uffc.org

17-21 October: 150th Meeting of the Acoustical Society of America JOINT with NOISE-CON 2005, Minneapolis, Minnesota. Contact: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; Web: asa.aip.org

19-21 October: 36th Spanish Congress on Acoustics and 2005 Iberian Meeting on Acoustics. Terrassa-Barcelona, Spain. Web: www.ia.csic.es/sea/index.html

2006

15-19 May: IEEE International Conference on Acoustics, Speech, and Signal Processing (IEEE ICASSP 2006). Toulouse, France. Web: http://icassp2006.org

5-9 June: 151st Meeting of the Acoustical Society of America, Providence, Rhode Island. Contact: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; Web: asa.aip.org

26-28 June: 9th Western Pacific Acoustics Conference. Seoul, Korea. Web: www.wespac8.com/WespacIX.html

3-7 July: 13th International Congress on Sound and Vibration (ICSV13). Vienna, Austria. Http://info.tuwien.ac.at/icsv13

28 November – 2 December: 152nd meeting, 4th Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, Honolulu, Hawaii. Contact: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; Web: asa.aip.org

3 - 6 December: INTER-NOISE 2006, Honolulu HA, USA (Same Hotel at ASA meeting the week preceeding)

2007

17-20 April. IEEE International Congress on Acoustics, Speech, and Signal Processing (IEEE ICASSP 2007). Honolulu, HI, USA

9-12 July: 14th International Congress on Sound and Vibration (ICSV14). Cairns, Australia. Email: n.kessissoglou@unsw.edu.au

2-7 September 19th International Congress on Acoustics (ICA2007), Madrid Spain. (SEA, Serrano 144, 28006 Madrid, Spain; Web: www.ia.csic/sea/index.html)

9-12 September: ICA2007 Satellite Symposium on Musical Acoustics (ISMA2007). Barcelona, Spain. Web: www.ica2007madrid.org

2008

23-27 June: Joint Meeting of European Acoustical Association, Acoustical Society of America, and Acoustical Society of France. Paris, France (Details to be announced later)

17-21 octobre: 150e rencontre de l'Acoustical Society of America AVEC NOISE-CON 2005, Minneapolis, Minnesota. Info: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tél.: 516-576-2360; Fax: 516-576-2377; Courriel: asa@aip.org; Web: asa.aip.org

19-21 octobre: 36th Spanish Congress dur Acoustics et 2005 Iberian Meeting sur Acoustics. Terrassa-Barcelona, Spain. Web: www.ia.csic.es/sea/index.html

2006

15-19 mai: IEEE Conference Internationale sur Acoustics, Speech, et Signal Processing (IEEE ICASSP 2006). Toulouse, France. Web: http://icassp2006.org

5-9 juin: 151e rencontre de l'Acoustical Society of America, Providence, Rhode Island. Info: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tél.: 516-576-2360; Fax: 516-576-2377; Courriel: asa@aip.org; Web: asa.aip.org

26-28 juin: 9e Conférence Western Pacific Acoustics. Seoul, Korea. Web: www.wespac8.com/WespacIX.html

3-7 juillet: 13th Congress Internationale sur Sound et Vibration (ICSV13). Vienna, Austria. Http://info.tuwien.ac.at/icsv13

28 novembre – 2 decembre: 152e rencontre, 4e Rencontre acoustique jointe de l'Acoustical Society of America, et l'Acoustical Society of Japan, Honalulu, Hawaii. Info: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tél.: 516-576-2360; Fax: 516-576-2377; Courriel: asa@aip.org; Web: asa.aip.org

3 - 6 December: INTER-NOISE 2006, Honolulu HA, USA (Same Hotel at ASA meeting the week preceeding)

2007

17-20 avril. IEEE Congress Internationale sur Acoustics, Speech, et Signal Processing (IEEE ICASSP 2007). Honolulu, HI, USA

9-12 juillet: 14th Congress Internationale sur Sound et Vibration (ICSV14). Cairns, Australia. Email: n.kessissoglou@unsw.edu.au

2-7 septembre 19e Congrès international sur l'acoustique (ICA2007), Madrid Spain. (SEA, Serrano 144, 28006 Madrid, Spain; Web: www.ia.csic/sea/index.html)

9-12 septembre: ICA2007 Satellite Symposium sur Musical Acoustics (ISMA2007). Barcelona, Spain. Web: www.ica2007madrid.org

2008

23-27 juin: Rencontre jointe de l'European Acoustical Association, l'Acoustical Society of America, et l'Acoustical Society of France. Paris, France (Details to be announced later)

NEWS / INFORMATIONS

CANADIAN ANNOUNCEMENTS

It is with much regret that we announce the tragic passing of **Mark Derrick, Ph.D., MBA, P.Eng.** of HFP Acoustical Consultants in Calgary, Alberta on June 13, 2004. Mark was a member of the Canadian Acoustical Association and registered as a Professional Engineer in the province of Alberta. As described by Mr. Leslie Frank of HFP:

> "Mark's contributions to noise control engineering consisted of leading-edge work, performed with enthusiasm and dedication, and with an intention to make a difference in our clients' lives. Mark loved the outdoors life, being an avid skier, mountain biker, scuba diver, sail-boater, and kayaker. Mark leaves behind a loving wife, many friends, appreciative clients, and our staff at HFP that will always remember him for his brilliant mind, his superb technical talents, his good nature, and his commitment to our firm. He lived his life to the fullest; energetic in his work activities, his relationships, and his recreational activities. We know that he died pursuing one of his passions; leaving us all far too soon with sorrow in our hearts."

FIRST ANNUAL NOISE POLICY DEVELOPMENT WORKSHOP A SUCCESS

More than 75 attendees participated in the First Annual Workshop on Noise Policy Development held during NOISE-CON 2004 in Baltimore, MD, USA on Tuesday, July 13, 2004. Represented were members of several national and local government agencies, noise control engineers, environmental consultants, and concerned citizens.

The workshop panelists presented a broad view of the need for a revitalized and coordinated United States National Noise Policy in the areas of aircraft/airport noise, industrial noise, occupational noise, consumer product noise, and surface transportation noise. In addition, panel members spoke on local community/citizens' issues and state policies.

For more information contact Dr. William W. Lang, Noise Control Foundation, P.O. Box 3067, Arlington Branch, Poughkeepsie, NY 12603. Telephone: +1 845 471 5493; Fax: +1 845 473 9325; email: langww@alum.mit.edu.

EXCERPTS FROM "*WE HEAR THAT*", IN *ECHOS*, ASA

Robert A. Frosch, senior research fellow at the Belfer Center for Science and International Affairs, John F. Kennedy School of Government, Harvard University, was awarded the 2003 Bueche Award by the National Academy of Engineering. He was cited "for a career of advances in aerospace and automotive technology, and industrial ecology; and for administration of R&D in industry, government, and academia."

Victor W. Zue, Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, has been elected to the National Academy of Engineering. He was cited for advances in the understanding of acoustic phonetics and systems for understanding spoken language.

Hugo Fastl, University of Munich was awarded the 2003 Rayleigh Medal by the Institute of Acoustics in the UK. He was cited for his "outstanding and lasting contributions to acoustics, in particular his leadership in psychoacoustics and his pioneering work on sound quality."

Juergen Meyer was awarded the 2004 Helmholtz medal of the German Acoustical Society (DEGA) for his research and teaching in musical acoustics. The Medal was awarded during the joint French-German acoustics meeting in Strasbourg in March.

Manfred Schroeder was elected an honorary member of the German Society for Audiology in 2003.

Mike and **Wendy Gluyas** were recently awarded the Kelvin medal and prize by the Institute of Physics for their outstanding lecture-demonstrations on the physics of sound and music, delivered to more than 200,000 schoolchildren, university students and members of the public throughout the UK, Eire, and internationally. Both retired from full-time education in 1993, but have continued to give their presentation as a retirement activity ever since.

EXCERPTS FROM "*SCANNING THE JOURNALS*", IN *ECHOS*, ASA

Killer whales adjust their **vocal behavior** when boat traffic reaches a certain level, according to a brief communication in the 29 April issue of *Nature*. Recordings in near-shore waters of Washington state made in 1977-81 and in 1989-92 showed no significant

difference in the duration of primary calls in the presence or absence of boats, but recordings during 2001-03 showed a 15% increase in call duration in the presence of a high boat density.

The March/April issue of *Acta Acustica/Acustica* has a review article on "Sound Sources Reconstruction Techniques: A Review of Their Evolution and New Trends," a technical review of the main approaches to **sound source reconstruction** techniques. Methods include 2-D spatial Fourier transform, acoustical holography, boundary elements method, and equivalent sources method. The applicability of these methods for forward propagation and back propagation problems is discussed. Current research subjects are outlined.

The **acoustic enhancement** system in the new Conference Center at the world headquarters of the Church of Jesus Christ of Latter Day Saints in Salt Lake City is described in the April 19 issue of *Sound & Communication*. With more than 21,000 seats, the Conference Center is the world's largest indoor conference facility. Although the volume is enormous, the natural reverberation time is less than 1.4 seconds when unoccupied due to the enormous amount of absorptive materials in the building. The only way to meet the mandate for speech intelligibility and simultaneously provide optimum acoustical conditions for the Tabernacle Choir and the new pipe organ was to incorporate electronic acoustic enhancement on a grand scale, according to the article. It is claimed that he enhancement system, which includes more than 300 speakers, can make the hall "sound wet like the Tabernacle or dry like the natural conference hall."

The brain's capacity to perceive modest pitch changes may be impaired from birth in persons with what is called **amusia**, according to a paper in the May issue of *Psychological Science*. These same persons, however, may recognize slight changes in timing in music.

An **acoustic radiometer**, which demonstrates acoustic radiation pressure, is described in an apparatus note in the June issue of *American Journal of Physics*. The apparatus consists of two panes attached to opposite ends of a horizontal arm that is pivoted at its center. One side of either pane is acoustically reflective and the other absorptive. The apparatus rotates when placed in an enclosure of high-intensity acoustic noise.

The ear needs a **noise workout** to keep in shape, according to a note in the 29 May issue of *New Scientist*. A test of 10,000 people around the World revealed that people exposed to loud noises at work had poor hearing, as expected. But the hearing of those living in quiet, rural areas was just as bad, an unexpected result. This note is based on paper 1aNS6 at the New York ASA meeting.

Physicists at Brigham Young University have developed an **active noise control** system to reduce the noise from cooling fans such as those in desktop computers, according to a note in the 29 May issue of *New Scientist*. The system, which can reduce computer noise by 20 dB, uses four 20-mm speakers at the corners of an 80-mm fan unit along with 4 microphones to sense the noise. This note is based on paper 2pNSb3 at the New York ASA meeting.

JVC (Victor of Japan) has introduced a new **loudspeaker with a wood cone**, according to a short piece in the April issue of *IEEE Spectrum*. In order to mold the wood, it is soaked in sake, which makes it malleable but allows it to retain its strength. Sound propagates more quickly in wood than in other cone materials such as paper and plastic, and this means the speakers can produce a wider range of frequencies, according to the article.

Bats living side-by-side may, in effect, be in different worlds, according to a paper in the 10 June issue of *Nature*. Their **hearing is differently tuned** in ways likely to affect their mating and hunting. For echolocation, each subspecies emits a different harmonic of the same fundamental frequency, and their ears are tuned to this frequency. They seem to communicate with each other via their own echolocation frequency, so they find mates with similarly tuned hearing, a recipe for species splitting.

An informative article on **hearing loss** appeared in the February issue of *Military Officer*. According to the Centers for Disease Control and Prevention, more than 34 million Americans have some form of hearing loss. Conductive loss occurs when blockage prevents sounds from being carried from the eardrum to the inner ear; sensorineural loss occurs when the inner ear is unable to convert sound waves or the auditory nerve can't transmit them to the brain.

Sound is arguably the most elusive and personal of all the human senses, according to an article in the 4 March issue of *Architectural Record*. Acousticians now have at their disposal 2D and 3D **ray-tracing** techniques that help determine the path that sound waves will take from source to receiver and communicate the acoustic consequences to visually-oriented architects. The impulse response of a room can be regarded as an acoustic signature, as unique to every space as a fingerprint.

Supersonic sound waves have been created in human tissue for the first time by exploiting the **acoustic equivalent of Cerenkov radiation**—the light emitted by charged particles when they travel through a medium faster than the speed of light in that medium, according to a paper in the March 22 issue of *Applied Physics Letters*. The results could have implications for ultrasound imaging in medicine. Most ultrasound imaging is done with longitudinal sound waves, since transverse shear waves are strongly absorbed by human tissue. However by focusing longitudinal sound waves in the tissue shear waves are generated and move in a Mach cone. These shear waves can be analyzed to produce an image.

Wave fronts on the surface of water can be viewed directly, but **acoustic wave fronts** in air require special techniques to make them visible. In the April issue of *The Physics Teacher*, a technique first used by physics teacher Arthur Foley in 1912 is described, and a set of slides of circular wave fronts are shown. Foley's apparatus can be easily reproduced by acoustics teachers and students.

Archaeologists believe they've found Europe's **oldest wooden musical instrument**, a set of yew-wood pipes found on Ireland's coast, according to an note in the 21 May issue of *Science*. The six pipes, found lying in a wood-lined trough during excavations for a new housing development in Dublin, have been dated at more than 4000 years old, 1500 years older than previous finds. The pipes are in graduated lengths from 29 to 59 cm, and they were likely placed vertically and attached to a bellows.

The idea that the deep, rumbling sounds used by **elephants** to "talk" to each other may be true, but not over long distances, according to a note in the 19 June issue of *New Scientist*. Researchers took an African elephant out into the farmland around Salinas, California and used a network of geophones to pick up the vibrations in the ground when the elephant rumbled. The rumbles produced very low-frequency surface waves, similar to those generated by earthquakes, but the vibrations traveled only a few hundred meters before dying out.

EXCERPTS FROM "*ACOUSTICS IN THE NEWS*", IN *ECHOS*, ASA

Researchers at the Dryden Flight Research Center at Edwards Air Force Base drastically reduced the sonic booms from a U.S. Navy jet by giving it a "nose job," according to a *Yahoo! News* story dated April 27. A custom nose glove was attached to a Navy F-5E jet in order to better distribute the air pressure build-up in front of the supersonic plane and thus soften the sonic booms. Instead of spending hundreds of millions of dollars on a completely new vehicle right away, the NASA program will look at other ways to shape a sonic boom. Modifications to a supersonic aircraft's engine inlets and lift surfaces, for example, could also help shape the sonic boom it creates, according to Ed Haering, principal investigator.

The Siemens Foundation held an unusual symposium "Beautiful Minds, Beautiful Music" that explored the relation between artistic and scientific brilliance, according to a note in the 4 June issue of *Science*. The event featured short performances by five winners of the Siemens/Westinghouse student science competition, followed by a panel of experts discussing the phenomenon. "We were amazed to find that nearly 100% of past winners played a musical instrument," said a spokesperson for the foundation.

Using digital enhancements of skull fragments dating more than 350,000 years ago, anthropologists argue that these human ancestors probably had hearing similar to that of people today, according to a news note in the 26 June issue of *Science News*. The researchers used a computerized tomography scanner to measure ear structures on three skulls and cranial pieces. Since the ears of social mammals typically match sounds made by fellow species members, the human-like hearing of these ancient folk probably was accompanied by speech.

Sound recorded on antique wax cylinders and early 78 rpm records can now be recovered without damaging the fragile media, according to a story in the 5 June issue of *New Scientist*, which is based on paper 2pMU4 at the 147th ASA meeting in New York. The source material is scanned using confocal microscopy using scanning and imaging techniques developed for nuclear physics. A noise reduction routine is used to clean up noise from dust and scratches.

About 80 computer manufacturers met in Beijing to consider a "super-high-fidelity" audio standard aimed at better sound quality over the Internet, according to a story in the 24 April issue of *New Scientist*. Intel high-definition audio (IHDA), for example, will capture sound well beyond human hearing up to 100 kHz. The thinking is that frequencies above the range of human hearing can still subtly improve sound perception.

Applying time-reversal acoustics (see *ECHOES*, Winter 2002), it is possible to communicate by tapping on tables or desks, according to a story in the 1 July issue of the *New York Times*. A system developed in France by Mathias Fink, Ros Ing, and colleagues uses one or two inexpensive accelerometers and a computer program to locate the source of the tap. The computer calculates the source of the vibration. Dividing the desk into a grid of 500 locations, for example, is like having a switch with 500 levels. "This is a very clever application of time-reversal acoustics,"ASA President William Kuperman is quoted as saying.

Lie-detector tests more and more use voice stress analysis, according to a story in the 1 July issue of the *New York Times*. Law-enforcement agencies have applied voice-based testing to question thousands of suspects, sometimes by telephone recording. British insurance companies have used it to screen telephone claims in order to detect fraud. The several available applications of the technology work on the same basic principle: that the human voice contains telltale signals that betray a speaker's emotional state. While academic and legal experts debate the merits of the technology, its developers continue to find new uses for it, including airport screening, post-traumatic stress assessment, and even matchmaking.

Sound Solutions from

Novel Dynamics Inc.

Dynamic Test and Analysis Systems

Integrated Solutions from World Leaders

- Precision Measurement Microphones
- Intensity Probes
- Outdoor Microphones
- Hydrophones
- Ear Simulation Devices
- Speech Simulation Devices
- Calibrators
- Array Microphones
- Sound Quality
- Sound Intensity
- Sound Power
- Room Acoustics
- Noise Monitoring
- Dynamic Signal Analyzers
- Multi Channel Dynamic Analyzer/Recorders
- Electro Dynamic Shaker Systems
- Advanced Sound & Vibration Level Meters

New

BEST PRICE AND PERFORMANCE REDEFINED™

LDS DACTRON

G.R.A.S. SOUND & VIBRATION

01dB-Stell MVI technologies group

SVANTEK

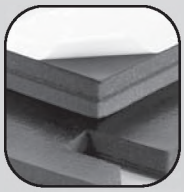
Ottawa
200-440 Laurier Ave. West, K1R 7X6
613-598-0026 fax: 616-598-0019
info@noveldynamics.com

Toronto
RR#2 13652 4th Line. Acton, Ont. L7J 2L8
519-853-4495 fax: 519-853-3366
metelka@aztec-net.com

Better testing... better products.

The Blachford Acoustics Laboratory

Bringing you superior acoustical products from the most advanced testing facilities available.



Our newest resource offers an unprecedented means of better understanding acoustical make-up and the impact of noise sources. The result? Better differentiation and value-added products for our customers.

Blachford Acoustics Laboratory features

- Hemi-anechoic room and dynamometer for testing heavy trucks and large vehicles or machines.
- Reverberation room for the testing of acoustical materials and components in one place.
- Jury room for sound quality development.

Blachford acoustical products

- Design and production of simple and complex laminates in various shapes, thicknesses and weights.
- Provide customers with everything from custom-engineered rolls and diecuts to molded and cast-in-place materials.

Blachford QS 9000 REGISTERED

www.blachford.com | Ontario 905.823.3200 | Illinois 630.231.8300

There´s a *New Kid in Town*

Created for You

Brüel & Kjær presents its innovative 4th generation hand-held instrument for sound and vibration. Experienced users from all over the world assisted us in setting the requirements for the new Type 2250.

- Color Touch Screen is the easiest user interface ever
- Non-slip surfaces with contours designed to fit comfortably in any hand
- Incredible 120 dB measurement range so you can't mess up your measurement by selecting an improper measurement range
- User log-in so the meter is configured the way you want it to be, and templates to make it easy to find user defined setups
- Optional Applications for Frequency Analysis and Logging are seamlessly integrated with the standard integrating SLM application
- High Contrast display and "traffic light" indicator make it easy to determine the measurement status at a distance even in daylight
- SD and CF memory and USB connectivity make Type 2250 the state-of-the-art sound analyzer

Created, built and made for you personally, you'll find Type 2250 will make a wonderful difference to your work and measurements tasks.

For more info please go directly to
www.type2250.com

HEADQUARTERS: DK-2850 Nærum · Denmark · Telephone: +45 45 80 05 00
Fax: +45 45 80 14 05 · www.bksv.com · info@bksv.com

USA: 2815 Colonnades Court, Building A · Norcross, GA 30071
Toll free (800) 332-2040 · www.BKhome.com · bkinfo@bksv.com

Australia (+61) 2 9889-8888 · Austria (+43) 1 865 74 00 · Brazil (+55) 11 5188-8166
Canada (+1) 514 695-8225 · China (+86) 10 680 29906 · Czech Republic (+420) 2 6702 1100
Finland (+358) 9-755 950 · France (+33) 1 69 90 71 00 · Germany (+49) 421 17 87 0
Hong Kong (+852) 2548 7486 · Hungary (+36) 1 215 83 05 · Ireland (+353) 1 807 4083
Italy (+39) 025 768061 · Japan (+81) 3 3779 8671 · Republic of Korea (+82) 2 3473 0605
Netherlands (+31) 318 55 9290 · Norway (+47) 66 77 11 55 · Poland (+48) 22 816 75 56
Portugal (+351) 21 4711 453 · Singapore (+65) 377 4512 · Spain (+34) 91 659 0820
Slovak Republic (+421) 25 443 0701 · Sweden (+46) 8 449 8600
Switzerland (+41) 1 880 7035 · Taiwan (+886) 22 713 9303
United Kingdom (+44) 14 38 739 000
Local representatives and service organizations worldwide

Type *2250*

Brüel & Kjær

The Canadian Acoustical Association
L'Association Canadienne d'Acoustique

PRIZE ANNOUNCEMENT • ANNONCE DE PRIX

A number of prizes and subsidies are offered annually by The Canadian Acoustical Association. Applicants can obtain full eligibility conditions, deadlines, application forms, past recipients, and the names of the individual prize coordinators on the CAA Website (http://www.caa-aca.ca). • Plusieurs prix et subventions sont décernés à chaque année par l'Association Canadienne d'Acoustique. Les candidats peuvent se procurer de plus amples renseignements sur les conditions d'éligibilités, les échéances, les formulaires de demande, les récipiendaires des années passées ainsi que le nom des coordonnateurs des prix en consultant le site Internet de l'ACA (http://www.caa-aca.ca).

CAA conference Student Travel subsidies: **consult www.ottawa2004.ca**
Subventions pour étudiants pour frais de déplacement au congrès annuel de l'ACA : **consulter le www.ottawa2004.ca**

EDGAR AND MILLICENT SHAW POSTDOCTORAL PRIZE IN ACOUSTICS • PRIX POST-DOCTORAL EDGAR AND MILLICENT SHAW EN ACOUSTIQUE

$3,000 for full-time postdoctoral research training in an established setting other than the one in which the Ph.D. was earned. The research topic must be related to some area of acoustics, psychoacoustics, speech communication or noise. • $3,000 pour une formation recherche à temps complet au niveau postdoctoral dans un établissement reconnu autre que celui où le candidat a reçu son doctorat. Le thème de recherche doit être relié à un domaine de l'acoustique, de la psycho-acoustique, de la communication verbale ou du bruit.

ALEXANDER GRAHAM BELL GRADUATE STUDENT PRIZE IN SPEECH COMMUNICATION AND BEHAVIOURAL ACOUSTICS • PRIX ÉTUDIANT ALEXANDRE GRAHAM BELL EN COMMUNICATION VERBALE ET ACOUSTIQUE COMPORTEMENTALE

$800 for a graduate student enrolled at a Canadian academic institution and conducting research in the field of speech communication or behavioural acoustics. • $800 à un(e) étudiant(e) inscrit(e) au 2e ou 3e cycle dans une institution académique canadienne et menant un projet de recherche en communication verbale ou acoustique comportementale.

FESSENDEN GRADUATE STUDENT PRIZE IN UNDERWATER ACOUSTICS • PRIX ÉTUDIANT FESSENDEN EN ACOUSTIQUE SOUS-MARINE

$500 for a graduate student enrolled at a Canadian academic institution and conducting research in underwater acoustics or in a branch of science closely connected to underwater acoustics. • $500 à un(e) étudiant(e) inscrit(e) au 2e ou 3e cycle dans une institution académique canadienne et menant un projet de recherche en acoustique sous-marine ou dans une discipline reliée à l'acoustique sous-marine.

ECKEL GRADUATE STUDENT PRIZE IN NOISE CONTROL • PRIX ÉTUDIANT ECKEL EN CONTRÔLE DU BRUIT

$500 for a graduate student enrolled at a Canadian academic institution and conducting research related to the advancement of the practice of noise control. • $500 à un(e) étudiant(e) inscrit(e) au 2e ou 3e cycle dans une institution académique canadienne et menant un projet de recherche relié à l'avancement de la pratique du contrôle du bruit.

RAYMOND HÉTU UNDERGRADUATE PRIZE IN ACOUSTICS • PRIX ÉTUDIANT RAYMOND HÉTU EN ACOUSTIQUE

One book in acoustics of a maximum value of $100 and a one-year subscription to *Canadian Acoustics* for an undergraduate student enrolled at a Canadian academic institution and having completed, during the year of application, a project in any field of acoustics or vibration. • Un livre sur l'acoustique et un abonnement d'un an à la revue *Acoustique Canadienne* à un(e) étudiant(e) inscrit(e) dans un programme de 1er cycle dans une institution académique canadienne et qui a réalisé, durant l'année de la demande, un projet dans le domaine de l'acoustique ou des vibrations.

CANADA-WIDE SCIENCE FAIR AWARD • PRIX EXPO-SCIENCES PANCANADIENNE

$400 and a one-year subscription to *Canadian Acoustics* for the best project related to acoustics at the Fair by a high-school student • $400 et un abonnement d'un an à la revue *Acoustique Canadienne* pour le meilleur projet relié à l'acoustique à l'Expo-sciences par un(e) étudiant(e) du secondaire.

DIRECTORS' AWARDS • PRIX DES DIRECTEURS

One $500 award for the best refereed research, review or tutorial paper published in *Canadian Acoustics* by a student member and one $500 award for the best paper by an individual member • $500 pour le meilleur article de recherche, de recensement des travaux ou d'exposé didactique arbitré publié dans l'*Acoustique Canadienne* par un membre étudiant et $500 pour le meilleur article par un membre individuel.

STUDENT PRESENTATION AWARDS • PRIX POUR COMMUNICATIONS ÉTUDIANTES

Three $500 awards for the best student oral presentations at the Annual Symposium of The Canadian Acoustical Association. • Trois prix de $500 pour les meilleures communications orales étudiant(e)s au Symposium Annuel de l'Association Canadienne d'Acoustique.

STUDENT TRAVEL SUBSIDIES • SUBVENTIONS POUR FRAIS DE DÉPLACEMENT POUR ÉTUDIANTS

Travel subsidies are available to assist student members who are presenting a paper during the Annual Symposium of The Canadian Acoustical Association if they live at least 150 km from the conference venue. • Des subventions pour frais de déplacement sont disponibles pour aider les membres étudiants à venir présenter leurs travaux lors du Symposium Annuel de l'Association Canadienne d'Acoustique, s'ils demeurent à au moins 150 km du lieu du congrès.

UNDERWATER ACOUSTICS AND SIGNAL PROCESSING STUDENT TRAVEL SUBSIDIES • SUBVENTIONS POUR FRAIS DE DÉPLACEMENT POUR ÉTUDIANTS EN ACOUSTIQUE SOUS-MARINE ET TRAITEMENT DU SIGNAL

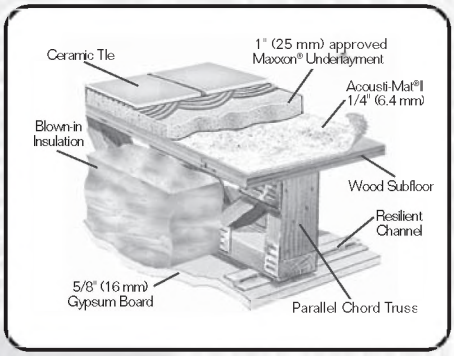
One $500 or two $250 awards to assist students traveling to national or international conferences to give oral or poster presentations on underwater acoustics and/or signal processing. • Une bourse de $500 ou deux de $250 pour aider les étudiant(e)s à se rendre à un congrès national ou international pour y présenter une communication orale ou une affiche dans le domaine de l'acoustique sous-marine ou du traitement du signal.

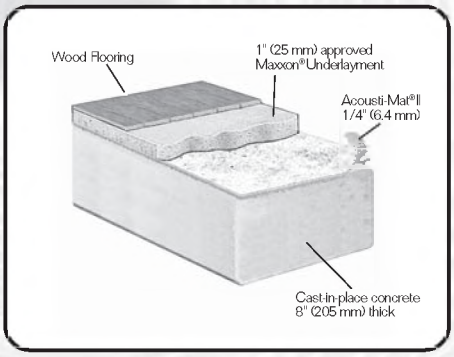
Acousti-Mat

The Best Barrier Between Floors and Noise Complaints.

For Upgraded Sound Control

ACOUSTI-MAT® II





- ***Renovation***
- ***New Construction***
- ***Concrete Construction***
- ***Open Beam Construction***
- ***Hard Surface Floor Goods Areas***

Acousti-Mat® II is the low-profile system that actually isolates sound waves, reducing transmission of impact and airborne sound up to 50 percent over wood frame, and up to 75 percent over concrete.

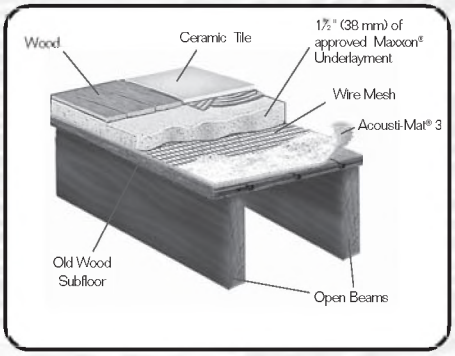
Acousti-Mat® 3 is the answer for "impossible" sound control challenges, such as open beam construction.

A poured Maxxon® Underlayment covers either system option, increasing IIC and STC up to 17 rating points for wood frame, and up to 25 IIC for concrete construction.

Proven on over 23 million square feet.
Specify Acousti-Mat for proven results.

For Super Sound Control

ACOUSTI-MAT® 3



For more information or a FREE GUIDE to acoustical construction, call 800-356-7887 or visit www.maxxon.ca

ACOUSTI-MAT®

Superior Sound Control Systems

INSTRUCTIONS TO AUTHORS FOR THE PREPARATION OF MANUSCRIPTS

Submissions: The original manuscript and two copies should be sent to the Editor-in-Chief.

General Presentation: Papers should be submitted in camera-ready format. Paper size 8.5" x 11". If you have access to a word processor, copy as closely as possible the format of the articles in Canadian Acoustics 18(4) 1990. All text in Times-Roman 10 pt font, with single (12 pt) spacing. Main body of text in two columns separated by 0.25". One line space between paragraphs.

Margins: Top - title page: 1.25"; other pages, 0.75"; bottom, 1" minimum; sides, 0.75".

Title: Bold, 14 pt with 14 pt spacing, upper case, centered.

Authors/addresses: Names and full mailing addresses, 10 pt with single (12 pt) spacing, upper and lower case, centered. Names in bold text.

Abstracts: English and French versions. Headings, 12 pt bold, upper case, centered. Indent text 0.5" on both sides.

Headings: Headings to be in 12 pt bold, Times-Roman font. Number at the left margin and indent text 0.5". Main headings, numbered as 1, 2, 3, ... to be in upper case. Sub-headings numbered as 1.1, 1.2, 1.3, ... in upper and lower case. Sub-sub-headings not numbered, in upper and lower case, underlined.

Equations: Minimize. Place in text if short. Numbered.

Figures/Tables: Keep small. Insert in text at top or bottom of page. Name as "Figure 1, 2, ..." Caption in 9 pt with single (12 pt) spacing. Leave 0.5" between text.

Line Widths: Line widths in techincal drawings, figures and tables should be a minimum of 0.5 pt.

Photographs: Submit original glossy, black and white photograph.

Scans: Should be between 225 dpi and 300 dpi. Scan: Line art as bitmap tiffs; Black and white as grayscale tiffs and colour as CMYK tiffs;

References: Cite in text and list at end in any consistent format, 9 pt with single (12 pt) spacing.

Page numbers: In light pencil at the bottom of each page.
Reprints: Can be ordered at time of acceptance of paper.

DIRECTIVES A L'INTENTION DES AUTEURS PREPARATION DES MANUSCRITS

Soumissions: Le manuscrit original ainsi que deux copies doivent être soumis au rédacteur-en-chef.

Présentation générale: Le manuscript doit comprendre le collage. Dimensions des pages, 8.5" x 11". Si vous avez accès à un système de traitement de texte, dans la mesure du possible, suivre le format des articles dans l'Acoustique Canadienne 18(4) 1990. Tout le texte doit être en caractères Times-Roman, 10 pt et à simple (12 pt) interligne. Le texte principal doit être en deux colonnes séparées d'un espace de 0.25". Les paragraphes sont séparés d'un espace d'une ligne.

Marges: Dans le haut - page titre, 1.25"; autres pages, 0.75"; dans le bas, 1" minimum; latérales, 0.75".

Titre du manuscrit: 14 pt à 14 pt interligne, lettres majuscules, caractères gras. Centré.

Auteurs/adresses: Noms et adresses postales. Lettres majuscules et minuscules, 10 pt à simple (12 pt) interligne. Centré. Les noms doivent être en caractères gras.

Sommaire: En versions anglaise et française. Titre en 12 pt, lettres majuscules, caractères gras, centré. Paragraphe 0.5" en alinéa de la marge, des 2 cotés.

Titres des sections: Tous en caractères gras, 12 pt, Times-Roman. Premiers titres: numéroter 1, 2, 3, ..., en lettres majuscules; sous-titres: numéroter 1.1, 1.2, 1.3, ..., en lettres majuscules et minuscules; sous-sous-titres: ne pas numéroter, en lettres majuscules et minuscules et soulignés.

Equations: Les minimiser. Les insérer dans le texte si elles sont courtes. Les numéroter.

Figures/Tableaux: De petites tailles. Les insérer dans le texte dans le haut ou dans le bas de la page. Les nommer "Figure 1, 2, 3,..." Légende en 9 pt à simple (12 pt) interligne. Laisser un espace de 0.5" entre le texte.

Largeur Des Traits: La largeur des traits sur les schémas technique doivent être au minimum de 0.5 pt pour permettre une bonne reproduction.

Photographies: Soumettre la photographie originale sur papier glacé, noir et blanc.

Figures Scanées: Doivent être au minimum de 225 dpi et au maximum de 300 dpi. Les schémas doivent être scannés en bitmaps tif format. Les photos noir et blanc doivent être scannées en échelle de gris tifs et toutes les phoots couleurs doivent être scannées en CMYK tifs.

Références: Les citer dans le texte et en faire la liste à la fin du document, en format uniforme, 9 pt à simple (12 pt) interligne.

Pagination: Au crayon pâle, au bas de chaque page.
Tirés-à-part: Ils peuvent être commandés au moment de l'acceptation du manuscrit.

The Canadian Acoustical Association
l'Association Canadienne d'Acoustique

Application for Membership

CAA membership is open to all individuals who have an interest in acoustics. Annual dues total $55.00 for individual members and $15.00 for Student members. This includes a subscription to ***Canadian Acoustics***, the Association's journal, which is published 4 times/year. New membership applications received before September 1 will be applied to the current year and include that year's back issues of ***Canadian Acoustics***, if available. New membership applications received after September 1 will be applied to the next year.

Subscriptions to *Canadian Acoustics* *or* Sustaining Subscriptions

Subscriptions to ***Canadian Acoustics*** are available to companies and institutions at the institutional subscription price of $55.00. Many companies and institutions prefer to be a Sustaining Subscriber, paying $250.00 per year, in order to assist CAA financially. A list of Sustaining Subscribers is published in each issue of ***Canadian Acoustics***. Subscriptions for the current calendar year are due by January 31. New subscriptions received before September 1 will be applied to the current year and include that year's back issues of ***Canadian Acoustics***, if available.

Please note that electronic forms can be downloaded from the CAA Website at **caa-aca.ca**

Address for subscription / membership correspondence:

Name / Organization _____

Address _____

City/Province _____ Postal Code _____ Country _____

Phone _____ Fax _____ E-mail _____

Address for mailing Canadian Acoustics, if different from above:

Name / Organization _____

Address _____

City/Province _____ Postal Code _____ Country _____

Areas of Interest: (Please mark 3 maximum)

1. Architectural Acoustics
2. Engineering Acoustics / Noise Control
3. Physical Acoustics / Ultrasound
4. Musical Acoustics / Electro-acoustics
5. Psychological / Physiological Acoustic
6. Shock and Vibration
7. Hearing Sciences
8. Speech Sciences
9. Underwater Acoustics
10. Signal Processing / Numerical Methods
11. Other

For student membership, please also provide:

_____ _____ _____ _____

(University) *(Faculty Member)* *(Signature of Faculty Member)* *(Date)*

I have enclosed the indicated payment for:
[ ] CAA Membership $ 55.00
[ ] CAA Student Membership $ 15.00
[ ] Institutional Subscription $ 55.00
[ ] Sustaining Subscriber $ 250.00
includes subscription (4 issues /year) to ***Canadian Acoustics.***

Payment by: [ ] Cheque
[ ] Money Order
[ ] VISA credit card *(Only VISA accepted)*

For payment by VISA credit card:

Card number _____

Name of cardholder _____

Expiry date _____

_____

(Signature) (Date)

Mail this application and attached payment to:
D. Quirt, Secretary, Canadian Acoustical Association, PO Box 74068, Ottawa, Ontario, K1M 2H9, Canada

l'Association Canadienne d'Acoustique
The Canadian Acoustical Association

Formulaire d'adhésion

L'adhésion à l'ACA est ouverte à tous ceux qui s'intéressent à l'acoustique. La cotisation annuelle est de 55.00$ pour les membres individuels, et de 15.00$ pour les étudiants. Tous les membres reçoivent ***l'Acoustique Canadienne***, la revue de l'association. Les nouveaux abonnements reçus avant le 1 septembre s'appliquent à l'année courante et incluent les anciens numéros (non-épuisés) de ***l'Acoustique Canadienne*** de cette année. Les nouveaux abonnements reçus après le 1 septembre s'appliquent à l'année suivante.

Abonnement pour la revue *Acoustique Canadienne* et abonnement de soutien

Les abonnements pour la revue ***Acoustique Canadienne*** sont disponibles pour les compagnies et autres établissements au coût annuel de 55.00$. Des compagnies et établissements préfèrent souvent la cotisation de membre bienfaiteur, de 250.00$ par année, pour assister financièrement l'ACA. La liste des membres bienfaiteurs est publiée dans chaque issue de la revue ***Acoustique Canadienne***. Les nouveaux abonnements reçus avant le 1 juillet s'appliquent à l'année courante et incluent les anciens numéros (non-épuisés) de ***l'Acoustique Canadienne*** de cette année. Les nouveaux abonnements reçus après le 1 septembre s'appliquent à l'année suivante.

Pour obtenir des formulaires electroniques, visitez le site Web: **caa-aca.ca**

Pour correspondence administative et financière:

Nom / Organisation _____

Adresse _____

Ville/Province _____ Code postal _____ Pays _____

Téléphone _____ Téléc. _____ Courriel _____

Adresse postale pour la revue *Acoustique Canadienne*

Nom / Organisation _____

Adresse _____

Ville/Province _____ Code postal _____ Pays _____

Cocher vos champs d'intérêt: (maximum 3)

1. Acoustique architecturale
2. Génie acoustique / Contrôle du bruit
3. Acoustique physique / Ultrasons
4. Acoustique musicale / Electro-acoustique
5. Physio / Psycho-acoustique
6. Chocs et vibrations
7. Audition
8. Parole
9. Acoustique sous-marine
10. Traitement des signaux /Méthodes numériques
11. Autre

Prière de remplir pour les étudiants et étudiantes:

_____ _____ _____ _____

(Université) (Nom d'un membre du corps professoral) (Signature du membre du corps professoral) (Date)

Cocher la case appropriée:

[ ] Membre individuel $ 55.00
[ ] Membre étudiant(e) $ 15.00
[ ] Abonnement institutionnel $ 55.00
[ ] Abonnement de soutien $ 250.00
(comprend l'abonnement à ***L'acoustique Canadienne)***

Méthode de paiement:

[ ] Chèque au nom de l'Association Canadienne d'Acoustique
[ ] Mandat postal
[ ] VISA — *(Seulement VISA)*

Pour carte VISA: Carte n° _____

Nom _____

Date d'expiration _____

_____

(Signature) (Date)

Prière d'attacher votre paiement au formulaire d'adhésion. Envoyer à:
D. Quirt, Secrétaire exécutif, Association Canadienne d'Acoustique, Casier Postal 74068, Ottawa, K1M 2H9, Canada

The Canadian Acoustical Association
l'Association Canadienne d'Acoustique

SUSTAINING SUBSCRIBERS / ABONNES DE SOUTIEN

The Canadian Acoustical Association gratefully acknowledges the financial assistance of the Sustaining Subscribers listed below. Annual donations (of $250.00 or more) enable the journal to be distributed to all at a reasonable cost. Sustaining Subscribers receive the journal free of charge. Please address donation (made payable to the Canadian Acoustical Association) to the Secretary of the Association.

L'Association Canadienne d'Acoustique tient à témoigner sa reconnaissance à l'égard de ses Abonnés de Soutien en publiant ci-dessous leur nom et leur adresse. En amortissant les coûts de publication et de distribution, les dons annuels (de $250.00 et plus) rendent le journal accessible à tous nos membres. Les Abonnés de Soutien reçoivent le journal gratuitement. Pour devenir un Abonné de Soutien, faites parvenir vos dons (chèque ou mandat-poste fait au nom de l'Association Canadienne d'Acoustique) au secrétaire de l'Association.

ACI Acoustical Consultants Inc.
Mr. Steven Bilawchuk - (780) 414-6373
stevenb@aciacoustical.com - Edmonton, AB

ACO Pacific
Mr. Noland Lewis - (650) 595-8588
acopac@acopacific.com - Belmont, CA

Acoustec Inc.
Dr. J.G. Migneron - (418) 682-2331
courrier@acoustec.qc.ca - Montréal, QC

Acoustik GE Inc.
C/o Gilles Elhadad - (514) 487
7159ge@acoustikge.com - Cote St Luc, QC

Aercoustics Engineering Limited
Mr. John O'Keefe - (416) 249-3361
aercoustics@aercoustics.com - Toronto, ON

Bruel & Kjaer International
Mr. Andrew Khoury - (514) 695-8225
- Pointe-Claire, QC

Dalimar Instruments Inc.
Mr. Daniel Larose - (450) 424-0033
daniel@dalimar.ca - Vaudreuil-Dorion, QC

Dodge-Regupol Inc.
Mr. Paul Downey - 416-440-1094
pcd@regupol.com - Toronto, ON

Earth Tech Canada Inc.
Mr. Christopher Hugh - (905) 886-7022,2625
chris.hugh@earthtech.ca - Markham, ON

Eckel Industries of Canada Ltd.
Mr. Blake Noon - (613) 543-2967
eckel@eckel.ca - Morrisburgh, ON

H. L. Blachford Ltd.
Mr. Dalton Prince - (905) 823-3200
amsales@blachford.ca -Mississauga ON

Hatch Associates Ltd.
Mr. Tim Kelsall - (905) 403-3932
tkelsall@hatch.ca - Mississauga, ON

HGC Engineering
Mr. Bill Gastmeier - (905) 826-4044
info@hgcengineering.com -Mississauga, ON

Hydro-Quebec
Mr. Blaise Gosselin - (514) 840-3000ext5134
gosselin.blaise@hydro.qc.ca - Montréal, QC

Industrial Metal Fabricators Ltd.
Mr. Frank Van Oirschot - (519) 354-4270
frank@indmetalfab.com -Chatham, ON

Integral DX Engineering Ltd.
Mr. Greg Clunis - (613) 761-1565
au741@ncf.ca - Ottawa, ON

J. E. Coulter Associates Ltd.
Mr. John Coulter - (416) 502-8598
jcoulter@on.aibn.com - Toronto, ON

J. L. Richards & Assoc. Ltd.
Mr. Fernando Ribas - (613) 728-3571
mail@jlrichards.ca -Ottawa, ON

Jade Acoustics Inc.
Ms. Dalila Giusti - (905) 660-2444
dalila@jadeacoustics.com - Concord, ON

John Swallow Associates
Mr. John C. Swallow - 905-271-7888
jswallow@jsal.ca - Mississauga, ON

MJM Conseillers en Acoustique
M. Michel Morin - (514) 737-9811
mmorin@mjm.qc.ca - Montréal, QC

Novel Dynamics Test Inc.
Mr. Andy Metelka - (519) 853-4495
metalka@aztec-net.com - Acton, ON

Owens Corning Canada Inc.
Mr. Keith Wilson - (705) 428-6774
keith.wilson@owenscorning.com-Stayner,ON

OZA Inspections Ltd.
Mr. Gord Shaw - (905) 945-5471
oza@ozagroup.com - Grimsby, ON

Peutz & Associés
Mr. Marc Asselineau - +33 1 45 23 05 00
marc.asselineau@club-internet.fr - Paris, France

Michel Picard
(514) 343-7617; FAX: (514) 343-2115
michel.picard@umontreal.ca - Brossard, QC

Pyrok Inc.
Mr. Howard Podolsky - (914) 777-7770
info@pyrokinc.com - Mamaroneck, NY US

Scantek Inc.
Mr. Peppin - (410) 290-7726
info@scantekinc.com - Columbia, MD

Silex Inc.
Mr. Mehmoud Ahmed - (905) 612-4000
mehmouda@silex.com - Mississauga, ON

SNC/Lavalin Environment Inc.
Mr. J-L. Allard - (514) 651-6710
jeanluc.allard@snclavalin.com-Longueuil,QC

Soft dB Inc.
M. André L'Espérance - (418) 686-0993
contact@softdb.com - Sillery, Québec

Spaarg Engineering Limited
Dr. Robert Gaspar - (519) 972-0677
gasparr@kelcom.igs.net - Windsor, ON

State of the Art Acoustik Inc.
Dr. C. Fortier - 613-745-2003
sota@sota.ca - Ottawa, ON

Tacet Engineering Ltd.
Dr. M.P. Sacks - (416) 782-0298
mal.sacks@tacet.ca - Toronto, ON

Valcoustics Canada Ltd.
Dr. Al Lightstone - (905) 764-5223
solutions@valcoustics.com - Richmond Hil

Vibro-Acoustics Ltd.
Mr. Tim Charlton - (800) 565-8401
tcharlton@vibro-acoustics.com - Scarborough, ON

West Caldwell Calibration Labs
Mr. Stanley Christopher - (905) 624-3919
info@wccl.com - Mississauga, On

Wilrep Ltd.
Mr. Don Wilkinson - (905) 625-8944
www.wilrep.com - Mississauga, ON