ISSN 0711-6659

# canadian acoustics
# acoustique canadienne

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

SEPTEMBER 2003 | SEPTEMBRE 2003
Volume 31 -- Number 3 | Volume 31 -- Numéro 3



OTHER FEATURES / AUTRES RUBRIQUES



## PROCEEDINGS

ACOUSTICS WEEK IN CANADA
SEMAINE CANADIENNE D'ACOUSTIQUE

## 2003

## CAHIERS DES ACTES

# canadian acoustics

# acoustique canadienne

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

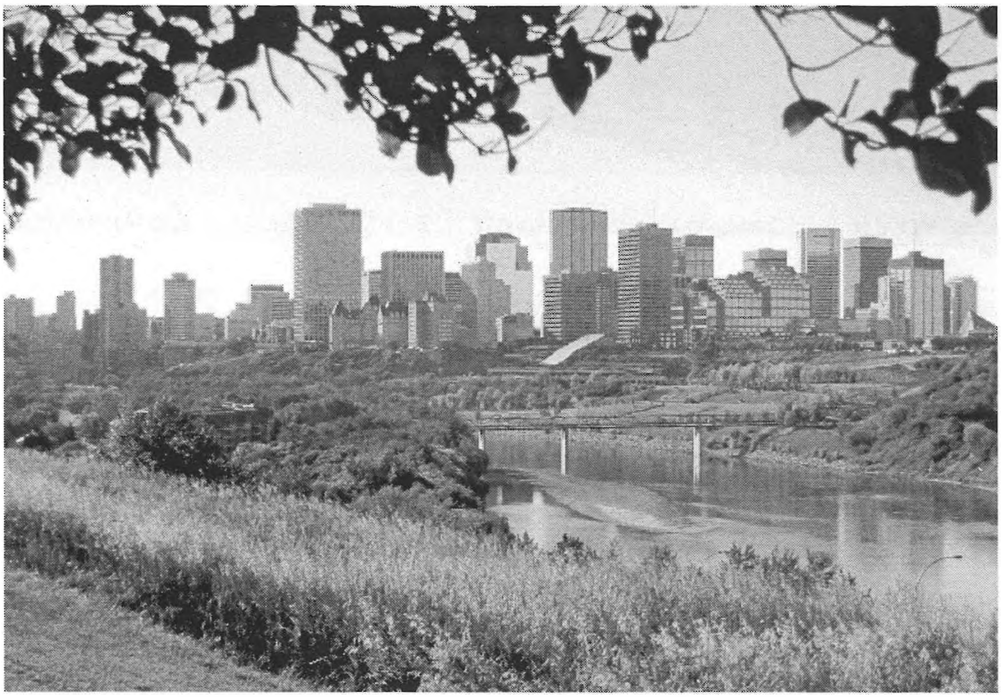
**L'ASSOCIATION CANADIENNE D'ACOUSTIQUE**
**C.P. 1351, SUCCURSALE "F"**
**TORONTO, ONTARIO M4Y 2V9**


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.

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.

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: $15 (student); $55 (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 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: $15 (étudiant); $55 (individuel, société); $250 (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 OF THE ANNUAL MEETING OF THE CANADIAN ACOUSTICAL ASSOCIATION

## ACOUSTICS: THE FUTURE OF QUIET?

## ACOUSTIQUE: L'AVENIR DU SILENCE?

# LES ACTES DE LA RÉUNION ANNUELLE DE L'ASSOCIATION CANADIENNE D'ACOUSTIQUE

ISSN -711-6659

ACOUSTICS WEEK IN CANADA OCTOBER 15-17, 2003 EDMONTON, ALBERTA

# *" The Future of Quiet ? "*

**ORGANIZING COMMITTEE**

Corjan Buma, Convenor
aci Acoustical Consultants Inc.

Steven Bilawchuk, Technical Support, Registrations & Web-Master
aci Acoustical Consultants Inc.

Megan Hodge, Secretary
University of Alberta, Department of Speech Pathology and Audiology

Gary Faulkner, Technical Program Coordinator
University of Alberta, Department of Mechanical Engineering

Izzy Gliener, Exhibition Coordinator, Marketing & Publicity
Western Noise Control, Edmonton, AB

Eugene Bolstad, Treasurer & Hospitality Coordinator
aci Acoustical Consultants Inc.

Ken Fyfe, Registrations, Presentations Coordinator
aci Acoustical Consultants Inc.

Kelly Kruger, Venue/Facilities Coordinator
Alberta Infrastructure, Technical Services, Acoustics Branch

With assistance from:
Melanie Campbell, University of Alberta, Dept. of Speech Pathology and Audiology
Rudy de Bruin, Noise Control Systems, Edmonton
Terry Nearey, University of Alberta, Department of Linguistics
Don Scheirer, COMTEC Associates, Edmonton
Michel Zielinski, Fabra-Wall Ltd., Edmonton

**EXHIBITORS**

H.L. Blachford Ltd. / Bruel & Kjaer / Dalimar Instruments / Eckel Industries / Fabra-wall Ltd. / GWP Wallworks Group Inc. / Norseman Inc. / Scantek Inc. / Acoutherm Insulation Ltd. / Western Noise Control

**TRANSLATIONS**

The 2003 Edmonton Steering Committee also gratefully acknowledges assistance by the following in providing translation into French of the various Announcements and our Guest Editorial:
Annie Ross, University of Moncton
Jean Francois Plourde, University of Moncton
Christian Giguere, University of Ottawa
Véronique Vaillancourt, University of Ottawa

## GUEST EDITORIAL

# ACOUSTICS: THE FUTURE OF QUIET

### ANNUAL MEETING OF THE CANADIAN ACOUSTICAL ASSOCIATION, EDMONTON, OCTOBER 14 -17, 2003

**Corjan Buma, Convener**
**aci** Acoustical Consultants Inc., Suite 107, 9920-63 Ave, Edmonton, Alberta, Canada, T6E 0G9

With this year's Conference theme "The Future of Quiet" we don't just mean "the-absence-of-noise". Both intentional privacy and communication depend in large degree on appropriate use or control of sound (and vibration) in its various forms. The underlying matter is a good, or ever-improving, "quality-of-life". An overview of the papers being contributed at the 2003 Conference of the Canadian Acoustical Association indicates clearly that all our work relates to this in some way. Whether characterizing acoustical environment in concert halls (Bradley et al.), seeing how drivers' abilities are influenced by sound-based signals (Slawinski et al.), improving modeling capability (Hodgson et al; and others), or considering directly the effects of sound on human listeners (various presenters), these are but a sampling of where, in 2003, research and problem-solving are at in achieving a somewhat "quieter" condition for us all. A Conference such as CAA-03 allows us to see again where we've collectively gotten to and what areas to look at next.

Thank you ! to all those who have chosen to help make **Acoustics Week in Canada 2003** the unique event that it is by either contributing a paper, setting up an Exhibitor booth, financially sponsoring an aspect of the Conference or helping on the organizing Committee. A special thank you to Senator Banks, Mr. De Gagne and Professor Eggermont for serving as Plenary speakers. A hearty Edmonton welcome to all delegates, especially our international guests; we trust the (n)ever-too-modest Canadian way will encourage your future/continued involvement. A special welcome too, to all first-time student delegates; we trust your experience here will help integrate your talent and potential and thereby translate into solid growth for the CAA.

In many ways Edmonton-2003 is maintaining a well-established tradition of Canadian Acoustical conferences. As a relatively "small" Conference, having just two parallel sessions reduces the common problem of having to chose between too many sessions, in all of which the papers are compellingly interesting. There is opportunity to gain insights into areas other than our own specialization. This year we hope to facilitate an extra bit of cross-fertilization by means of our closing combination luncheon/panel-discussion.

Again, thank you to all for your attendance and participation. We trust the initiatives making up "Edmonton-2003" will maintain and promote the vitality of all types of acoustical work in Canada and beyond.

For the Edmonton-2003 Steering Committee,

*Corjan Buma*, Chair

## WHAT'S NEW ??

| | |
|---|---|
| Promotions | Retirements |
| Deaths | Degrees awarded |
| New jobs | Distinctions |
| Moves | 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 | Retraites |
| Décès | Obtention de diplômes |
| Offre d'emploi | Distinctions |
| Déménagements | Autres nouvelles |

Avez-vous des nouvelles que vous aimeriez partager avec les lecteurs de l'Acoustique Canadienne? Si oui, écrivez-les et envoyer à:

Francine Desharnais, DREA Ocean Acoustics, P.O. Box 1012, Dartmouth NS, Email: desharnais@drea.dnd.ca

## ÉDITORIAL

# ACOUSTIQUE: L'AVENIR DU SILENCE?

## LES ACTES DE LA RÉUNION ANNUELLE DE L'ASSOCIATION CANADIENNE D'ACOUSTIQUE

**Corjan Buma, Convener**

**aci** Acoustical Consultants Inc., Suite 107, 9920-63 Ave, Edmonton, Alberta, Canada, T6E 0G9

Le thème du congrès de cette année, "L'avenir du silence", ne signifie pas uniquement l'absence de bruit. Des activités désirées comme l'intimité personnelle ou la communication verbale dépendent en grande partie de l'usage et du contrôle approprié du son (et des vibrations) sous ses diverses formes. Le principe sous-jacent est une qualité de vie favorable ou en constante évolution. Un aperçu général des articles soumis pour le congrès annuel 2003 de l'Association acoustique canadienne indique clairement que tous nos travaux se rattachent, d'une façon ou d'une autre, à ce concept. Que ce soit par la caractérisation des environnements acoustiques des salles de concert (Bradley et al.), la détermination de l'influence des signaux sonores sur les capacités des conducteurs d'automobile (Slawinski et al.), l'amélioration des capacités de modélisation (Hodgson et al; et autres auteurs) ou même par la considération des effets du son sur les humains (plusieurs auteurs), il est possible de démontrer à quel point la recherche et la résolution de problème en sont rendus en 2003 dans la réalisation d'une condition plus « silencieuse » pour nous tous. Le congrès 2003 de l'ACA nous permet à nouveau de faire le point sur où nous en sommes collectivement et de déterminer quels domaines devront ensuite être explorés.

Merci à tous ceux qui ont contribué, soit par l'entremise d'une soumission d'article, d'une participation au salon d'exposition, d'une commandite financière ou par une participation dans le comité organisateur, à transformer la Semaine canadienne de l'Acoustique en un événement unique. Un remerciement spécial au sénateur Banks, à M. De Gagne et au professeur Eggermont, nos invités en conférencières plénières. Un mot chaleureux de bienvenue à Edmonton à tous les délégués, particulièrement nos invités internationaux. Nous avons confiance que la façon de faire canadienne, jamais trop modeste, incitera votre participation future et continue. Également, une bienvenue particulière aux étudiants délégués qui assisteront à leur premier congrès de l'ACA. Nous sommes confiants que votre expérience saura enrichir votre talent et potentiel, qui viendra du fait même assurer la croissance solide de l'ACA.

Sous plusieurs aspects, le congrès Edmonton-2003 maintient la tradition des congrès antérieurs de l'Association acoustique canadienne. Ce congrès relativement « petit », qui consiste en seulement deux sessions parallèles, réduit le fréquent problème d'avoir à choisir parmi trop de sessions intéressantes. Elle vous donne donc la chance d'acquérir des connaissances dans des domaines autres que votre domaine de spécialisation. Cette année, nous espérons faciliter l'enrichissement mutuel par l'entremise d'une combinaison casse-croûte/discussion-dirigé en clôture.

Encore une fois un gros merci pour votre présence et participation. Nous sommes confiants que les initiatives du comité organisateur du congrès « Edmonton-2003 » seront maintenues dans le futur et quelles vont promouvoir la vitalité de tous les types de travaux en acoustique au Canada et ailleurs.

Du comité organisateur du congrès Edmonton-2003,

*Corjan Buma*, président

**Proceedings of the Annual Meeting of the Canadian Acoustical Association**

# Acoustics: Bridge to the Future / Acoustique: Un Pont vers L'Avenir

**Les actes de la réunion annuelle de L'Association Canadienne d'Acoustique**

## Table of Contents / Tables des matières

**Vowels II-L2/Perception & Production Modeling**



**ARCHITECTURAL ACOUSTICS**

## NUMERICAL MODELING / EXPERIMENTAL TECHNIQUES



## ENVIRONMENTAL ACOUSTICS



## HEARING CONSERVATION



## NOISE CONTROL

## PSYCHOLOGICAL ACOUSTICS

# CONFERENCE PROGRAM

## TUESDAY, OCTOBER 14, 2003

| | |
|---|---|
| **16:00 – 20:00** | **CAA Board of Directors Meeting** |
| **16:00 – 22:00 (specific time TBA)** | **CSA Standards Meeting** |
| **18:00 – 21:00** | **Opening Registration** |
| **18:00 – 21:00** | **Delegate Reception** |

## DAY 1 – WEDNESDAY, OCTOBER 15, 2003

| | | | | |
|---|---|---|---|---|
| **08:00 – 15:30** | | | **CONFERENCE REGISTRATION** | |
| **08:45 – 09:00** | | | **OFFICIAL CONFERENCE OPENING** | |
| **09:00 – 10:00** | | | **PLENARY SPEAKER: SENATOR TOMMY BANKS** | |
| **10:00 – 10:20** | | | **COFFEE BREAK** | |
| | **10:20 – 12:00<br>SPEECH COMMUNICATION: Vowels I** | | **ARCHITECTURAL ACOUSTICS I** | |
| 1. | A Neural Network Approach To The Question Of The Dimensionality Of The Perceptual Vowel Space | Terrance Nearey, Michael Kiefte | Acoustical Comparison Of Alberta Concert Halls | J.S. Bradley, H. Sato |
| 2. | Vowel Masked Audiograms: Audibility Of Individual Harmonics In Synthetic Monophthongs | Michael Kiefte, Sara Macintosh | Developing A New Measure For Assessing Architectural Speech Security | J.S. Bradley, Bradford Gover |
| 3. | A Review Of Results On Frequency Shifts And Vowel Identification | Peter Assmann, Terrance Nearey | Dealing With Flanking Transmission in Wood-Framed Construction | J.D. Quirt, T.R.T. Nightingale, R.E. Halliwell |
| 4. | Effect Of Phonetic Context On Women's Vowel Area | J. Chesworth, K. Cote, C. Shaw, S. Williams, M. Hodge | On Structure Borne Power Flow From Wood Studs To Direct-Attached Gypsum Board | T.R.T. Nightingale, Katrin Kohler |
| | | | | |
| 5. | Articulatory Compensation For A Bite Block, With And Without Auditory Feedback, In Adults | Melanie McNutt Campbell | On The Simplifications Used In Mobility Models To Predict Structure Borne Power Flow In Wood Stud Walls With Direct-Attached Gypsum Board | T.R.T. Nightingale, Katrin Kohler |
| **12:00 – 13:20** | | | **LUNCH** | |

| | 13:20 – 15:00<br>**SPEECH COMMUNICATION** | | **NUMERICAL MODELING / EXPERIMENTAL TECHNIQUES** | |
|---|---|---|---|---|
| 1. | Silence As A Cue To Rhythm In The Analysis Of Speech And Song | David Gerhard | Analytical Model Of Rail Vibration Induced By Flat Spots On Rolling Wheels | Werner Richarz |
| 2. | Perception Of Timbral Fusion | Dora Chan, Elzbieta Slawinski | Active Noise Control In Non-Diffuse Three-Dimension Enclosures With High Modal Density: Theoretical Studies | Desheng Li, Murray Hodgson |
| 3. | Recognizing Familiar And Foreign Words And Music By Children And Adults: An Examination Of The Critical Period Hypothesis | Elizabeth McFadden, Annabel Cohen | Active Noise Control In Non-Diffuse Three-Dimensional Enclosures With High Modal Density: Experimental Studies | Desheng Li, Murray Hodgson |
| 4. | Effect of Articulation Condition on Children's Acoustic Cues for Bilabial Place | M. Hodge, C. Huang, J. Hegedus, M. Campbell | Multi Channel Data Acquisition And Signal Processing Using MATLAB | S. Bilawchuk, K.R. Fyfe |
| 5. | Comparison Of Intelligibility Measures On Single Word And Spontaneous Speech Tasks For Children With And Without Cleft Palate | C. Gotzke, M. Hodge, J. Daniels | | |
| **15:30 – 17:00** | | | **SITE VISIT TO WINSPEAR CONCERT HALL** | |

## DAY 2 – THURSDAY, OCTOBER 16, 2003

| **08:00 – 15:30** | | | **CONFERENCE REGISTRATION** | |
|---|---|---|---|---|
| **09:00 – 10:00** | | | **PLENARY SPEAKER: MR. DAVE DEGAGNE** | |
| **10:00 – 10:20** | | | **COFFEE BREAK** | |
| | **10:20 – 12:00<br>SPEECH COMMUNICATION: Vowels II-L2/Perception & Production Modeling** | | **ENVIRONMENTAL ACOUSTICS** | |
| 1. | A Longitudinal Examination Of English Vowel Learning By Mandarin Speakers | Murray Munro, Tracey Derwing, Ron Thomson | The Microphone As A Roughness And Wetness Sensor Of The Road In Automotive Applications | A. Trabelsi, B. Heimann, P. Rieth, H. Harting |
| 2. | An Acoustic Study Of The Production Of An Uncommon Vowel Produced By Hungarian-English Bilinguals | Zita McRobbie-Utasi | Investigation Of Dash-8 Run-Up Noise Characteristics For Local Active Noise Control | Ann Nakashima, Murray Hodgson |
| 3. | L2 Vowel Perception: Perceptual Assimilation To What? | Ron Thomson | Application Of DOW-Quash For Limiting Community Noise | Corjan Buma |
| 4. | Fuzzy String Kernel Representations In Speech Processing | Robert Kirchner | Modeled Trends In Environmental Noise Levels As A Function Of The Long-term Meteorology In Alberta | Jason McCrank, Jerome Parkinson |
| 5. | Obtaining The Vocal-Tract Area Function From The Vowel Sound | H. Deng, M. Beddoes, R. Ward, M. Hodgson | Outdoor Noise Propagation | R. Ramakrishnan, C. Novak, H. Ule, S. Penton, R. Gaspar |

| | | | | |
|---|---|---|---|---|
| **12:00 – 13:20** | | | **LUNCH** | |
| | **13:20 – 15:00**<br>**HEARING CONSERVATION** | | **NOISE CONTROL I** | |
| 1. | Detectsound Version 2: A Software Tool For Adjusting The Level And Spectrum Of Acoustic Warning Signals | Y. Zheng, C. Giguere, C. Laroche, R. Lavoie | Case Study: Noise Mitigation For Small Gas Turbine Power Plants | Scott Penton, Tammy Dow |
| 2. | Noise Exposure Of Opera Orchestra Players | J. Lee, A. Behar, H. Kunov, W. Wong | Comparison Of Experimental And Modeled Insertion Loss Of A Complex Multi-Chamber Muffler With Temperature And Flow Effects | C. Novak, H. Ule, R. Ramakrishnan, R. Gaspar |
| 3. | Hearing Protectors Testing And Labeling – What's New? | Alberto Behar | Prediction Of Insertion Loss Of Rectangular Silencers | Ramani Ramakrishnan, Willie Watson |
| 4. | Perception Of Increasing Or Decreasing Signal Intensity And Effects Of Compression By Hearing Aids | M.-J. Palardy, C. Giguere, C. De Segovia | The Influence Of Baffle Orientation On Sound Attenuation Of Dissipative Silencers | Harrison Richarz |
| 5. | Optimal Reverberation Times For Speech Intelligibility For Normal And Hard-Of-Hearing People | M. Ezaki, W. Yang, M. Hodgson | Experimental Studies On The Multi-Channel Active Control Of Complex Sources In The Free-Field | H. Qin, A. Nakashima, D. Li, M. Hodgson |
| **15:00 – 15:20** | | | **COFFEE BREAK** | |
| | **15:20 – 16:40**<br>**PSYCHOLOGICAL ACOUSTICS** | | **NOISE CONTROL II** | |
| 1. | Interaction Of Age And Attention On Driving Performance | M. Motamedi, E. Slawinski, J. McNeil | Effect Of The Ground Surface And Meteorological Conditions On The Active Control Of A Monopole Noise Source. | Ann Nakashima, Murray Hodgson |
| 2. | The Effects Of Attention To Auditory Stimuli On Driver Speed And Lane Weaving Behaviors In A Driving Simulator | Kristen Dugdale, Elzbieta Slawinski | Measurement Of The Effect Of Fittings On Low-Frequency Sound In A Scale-Model Workroom | Galen Wong, Murray Hodgson |
| 3. | How Music Of Differing Rhythmicities And Intensities Affects Driver Performance | Ben Zendel, Elzbieta Slawinksi | The Effect Of Background Noise On Sound Power In Both A Reverberant And Anechoic Environment | R. Ramakrishnan, C. Novak, H. Ule, R. Gaspar |
| 4. | The Auditory And Visual Attentional Blink In Schizophrenia | E. Slawinski, K. Goddard, P. Wass | Directional Characteristics Of An Outdoor Warning Siren | Nigel Maybee |
| **17:00 – 18:30** | | | **CAA/ACA ANNUAL GENERAL MEETING** | |
| **18:45 – 19:15** | | | **BANQUET RECEPTION / COCTAILS** | |
| **19:15 –** | | | **CONFERENCE BANQUET AND AWARDS CEREMONY** | |

# DAY 3 – FRIDAY, OCTOBER 17, 2003

| | | | | |
|---|---|---|---|---|
| **08:00 – 09:00** | | | **CONFERENCE REGISTRATION** | |
| **09:00 – 10:00** | | | **PLENARY SPEAKER: DR. JOS EGGERMONT** | |
| **10:00 – 10:20** | | | **COFFEE BREAK** | |
| | **10:20 – 12:00**<br>**SPEECH COMMUNICATION: More Topics in L2** | | **ARCHTECTURAL ACOUSTICS II** | |
| 1. | Acoustic Indicators Of Spanish-Accented Speech | Amee P. Shah | Acoustical Evaluation Of Berwick Preschool | Wonyoung Yang, Murray Hodgson |
| 2. | Intelligibility Of Foreign-Accented Lombard Speech | Chi-Nin Li | *ClassTalk* System For Predicting, And Visualizing Speech In Noise In Classrooms | Murray Hodgson |
| 3. | The Effect Of Acoustic Information Of Lexical Tones On Non-Native Listeners' Tonal Identifications | Connie So | Experimental Validation Of *Plant Noise* Empirical Prediction Models | Murray Hodgson |
| 4. | Acoustic Cues To Voicing Of Initial Consonants In Mandarin CB Syllables | Chunling Zhang, Terrance Nearey | | |
| 5. | Perceptual Differences In Categorization Of Speech Sounds By Norwegian/English Bilinguals | Audny Dypvik, Elzbieta Slawinski | | |
| **12:15 –** | | | **PANEL DISCUSSION: THE FUTURE OF QUIET?** | |

# PLENARY SPEAKERS

## ABSTRACTS

## THURSDAY

### *CONSIDERATIONS FOR INDUSTRIAL NOISE CONTROL STANDARDS*

**David C. DeGagne, Alberta Energy & Utilities Board**

Anyone who has ever been involved in the process of developing regulatory standards will agree that it is a complex and challenging task. Considerations such as determining where the legislative mandate comes from, what industries are affected, what is the basis for technical requirements, and who needs to be involved are just a few of the many diverse topics that must be resolved.

Using the extensive experience of the Alberta Energy and Utilities Board in conducting such processes this presentation identifies and examines a number of the most significant issues and hurdles that may be encountered and discusses how these might be handled.

Also included in the presentation is a preview of the new Alberta Noise Control Directive and the direction it is taking and what this means to potentially affected parties.

## FRIDAY

### CHILDREN'S POOR UNDERSTANDING OF SPEECH IN NOISE REFLECTS IMMATURITY OF AUDITORY CORTEX.

**Jos J. Eggermont, Ph.D.**
Department of Physiology and Biophysics, Department of Psychology
University of Calgary, Calgary, Alberta

It is well known that even minor high-frequency hearing losses interfere with adequately understanding speech in noisy environments. It is much less appreciated that children under the age of 12 face similar difficulties even when their hearing thresholds are perfectly normal. The ages of 5-12 appear to be the crucial ones in which the processing of masked and degraded speech improve. There are impressive parallels in the structural and functional maturation of the auditory cortex and the improvement of understanding speech in noisy environments. For instance, the maturation of the axons in the supra granular layers continues to improve up to age 12 and concomitantly the N1 component of the auditory evoked potential emerges and matures in mid adolescence. These findings suggest that learning environments should be as free of back ground noise as possible, which has implications for both design of and supervision in these environments.

# A Neural Network Approach to the Dimensionality of the Perceptual Vowel Space

**Terrance M. Nearey[1] and Michael Kiefte[2]**
1) Dept. of Linguistics, University of Alberta, Edmonton, AB, Canada, T6G 2E7, t.nearey@ualberta.ca
2) School of Human Comm. Disorders, Dalhousie University, Halifax, NS, Canada, B3H 1R2, mkiefte@dal.ca

## 1. INTRODUCTION

The question of the dimensionality of the perceptual vowel space for monophthongs has a long history (see [1] for a review). Although three or more formants are typically used to synthesize acceptable vowels, under some circumstances a smaller number of spectral prominences (one or two) can produce acceptable vowel quality. The main question we address is: Can a two-dimensional perceptual space, corresponding roughly to F1 and F2-prime adequately represent the perceptual properties of vowels? F2-prime is believed by some to result from large-scale perceptual integration (e.g., 3.5 Bark 'centers of gravity') of spectral energy in the F2 to F4 range [2]. Others are skeptical of this notion; see [1]. We sketch below a novel method to examine the degree to which a two- or three-dimensional space can accurately represent listeners' perception of a large three-formant vowel continuum. Our results suggest that no two-dimensional representation adequately accounts for listeners' behavior.

In prior work [3], we modeled the categorization of a large (972 stimuli) F1-F2-F3 continuum by 14 speakers of English and 14 speakers of Finnish. Briefly, the stimuli filled a feasible F1-F2-F3 space of an adult speaker in steps of 0.5 Bark on each formant. English speakers responded to each stimulus with one of 11 possible vowel choices, including the rhotic vowel as in the word *her*, which is characterized by a low F3. Finnish speakers responded with one of 8 possible vowel choices, representing the full inventory of short monophthongs in Finnish. Details of the stimuli and procedures are given in [3]. Our prior modeling [3] used several fixed representations of the stimuli. Notably, a two-dimensional F1 by F2-prime representation performed markedly worse than F1-F2-F3 in predicting listeners' categorization of the stimuli via logistic regression. Although the explicit F1 by F2-prime representation of Bladon and Fant [2] is clearly inferior to the three-formant representation, the more general question remains whether some other two-dimensional space is adequate.

## 2. METHOD

The key to our analysis is a neural network architecture that is capable of implementing an optimal, non-parametric, two-dimensional representation of the stimulus space. This model is rooted in an input-layer with a saturated 'dummy-variable' coding of the 972-point stimulus space, i.e. on presentation of stimulus k, the activation of input node $k$ is set to 1.0 and activations of all 971 other nodes are set to 0.0. This input layer fans in to a two-node hidden layer. From this two-dimensional bottleneck, it then 'fans out' again to a group of 11 vowel response output nodes for English and a group of 8 vowel output nodes for Finnish. Each group of output nodes is provided with a softmax transfer function (see [3] for references). This effectively implements a polytomous logistic regression of the response patterns (pooled over listeners) of each language on a common set of two derived stimulus variables. A sketch of the model with a two-dimensional hidden layer is shown in Figure 1.

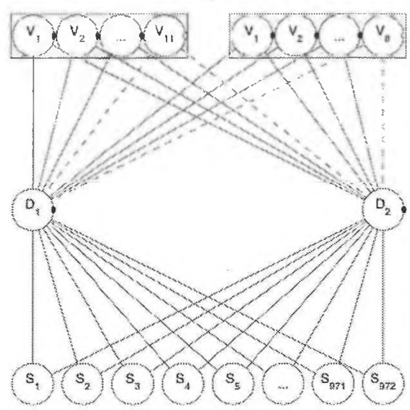



Fig. 1. Sketch of two-dimensional bottleneck neural net. The bottom input layer represents the stimuli. The middle hidden layer represents the two-dimensional reduced space. The top output layer represents listeners' responses.

The weights from the input layer to each hidden unit can be viewed as scores on a single dimension and the trained weights for each stimulus represent the analog of factor scores in multidimensional scaling space of that stimulus. Simulations of the two-dimensional structure were run using random initial weights. To aid convergence, the initial input-to-hidden layer weights were a mixture of 25% random normal deviates and 75% standardized (to zero mean and unit variance) values of F1 to unit $D_1$ or 75% standardized values of Bladon and Fant's F2-prime to unit

$D_2$. (Frequencies were transformed to Bark before standardization). Because of the high dimensionality of the problem and the possibility of 'stalling', 200 different random initializations were run. We report the best results of the 200 starts below. A similar set of analyses was run with a three-dimensional hidden layer. Here, convergence properties were somewhat better, so complete random initialization was used with 200 starts.

## 3. RESULTS

Initial results summarized suggest that two-dimensional bottleneck does not provide a very good account of listeners' categorization, while a three-dimensional hidden layer works quite well.

Table 1. Comparison of goodness of fit of four models. See text.

| *Model* | *rms%* | *Error* | $N_p$ |
|---|---|---|---|
| I. F12p | 9.3% | 434.4 | 51 |
| II. Opt2D | 8.3% | 338.5 | 1995 |
| III. F123 | 5.3% | 204.5 | 68 |
| IV. Opt3D | 4.3% | 167.5 | 2984 |

Table 1 presents a comparison of four models. Model I, *F12p* represents a simple softmax model with two inputs F1 and F2-prime. The architecture of the model is equivalent to the top two layers of Figure 1, with F1 and F2-prime applied to the two input nodes. The column labeled *rms%* shows the rms error of predictions compared to observations when responses are measured in percent. The column labeled *Error* is the error criterion that was optimized in the network fitting process. (This error was calculated on normalized proportions of responses, rather than raw response counts. If counts had been used, as in the modeling in [3], this number would have been approximately 69.5 times the value shown.) The column labeled $N_p$ is the number of (non-redundant) free parameters required to fit the model. Although some of these numbers are quite large, it should be noted that there are 16,524 degrees of freedom in the response data. Models I through IV thus exhaust about 0.3, 12, 0.4 and 18 percent of the available degrees of freedom

Model II represents the optimal two-dimensional solution corresponding to the architecture of Figure 1. Comparing models II and III, we see that there is only a modest gain of about one percentage point in rms error with an additional 1944 degrees of freedom. Model II represents a baseline three-dimensional solution. It is similar to model I except that there are three input nodes, corresponding to the synthesis control parameters F1, F2 and F3. We see that there is about a four-percentage point reduction in rms and roughly a halving of the softmax error compared to model I. This large gain obtains with only 17 more free parameters. Model IV corresponds to an optimal three-dimensional solution. This model is like that of Figure 1, except that a third unit is added to the hidden layer. It is the largest model and fits the best of all.

## 4. DISCUSSION

The fact that the optimal two-dimensional model II is superior to the much smaller two-dimensional model I is not surprising because of the large increase in the flexibility of the model. Similar remarks apply to model IV and model III. Deciding whether the enormous increase in degrees of freedom is justified for the gains observed constitutes a difficult problem in model selection, especially since the problem represents a case of repeated measures categorical data. However, the comparison between models II and III is a much simpler matter. Despite the vastly greater number of degrees of freedom available to model II (which enables an optimal non-parametric, two-dimensional mapping of the stimuli), the two-dimensional bottleneck of the hidden layer apparently makes it impossible to provide a good fit to the data. The fact that model II provides a substantially poorer fit to the responses that the default three-dimensional (F1, F2 and F3) representation of model III leaves little doubt that the dimensionality of the perceptual space underlying vowel perception is greater than two and that no modification of the F1 by F2-prime space can adequately serve as a basis for modeling listeners' categorization.

## REFERENCES


[1] Rossner, B. and J. Pickering. Vowel perception and production. Oxford: Oxford University Press, 1994.

[2] Bladon, A. and G. Fant. A two-formant model and the cardinal vowels. *Speech Transmission Laboratory Quarterly Progress Report*, vol. 1, pp. 1-8. 1978.

[3] Nearey, T. and M. Kiefte. Comparison of several proposed perceptual representations of vowel spectra. *Proc. 15th Int. Cong. Phonetic Sciences,* vol. 1, pp. 1005-1008. 2003.


## ACKNOWLEDGEMENTS


This work was supported by SSHRC. Thanks to U. Fischer, O. Aaltonen and their colleagues at the University of Turku, Finland for their help in the collection of the Finnish data.

# A Review Of Results On Frequency Shifts And Vowel Identification

**Peter F. Assmann[1] and Terrance M. Nearey[2]**
1) School of Behavioral and Brain Sciences, University of Texas at Dallas, Richardson, TX 75083 USA assmann@utdallas.edu
2) Department of Linguistics, University of Alberta, Edmonton, AB T6G 2E7 Canada tnearey@ualberta.ca

## 1. INTRODUCTION

Listeners experience a wide range of variation in fundamental frequency ($F_0$) and formant frequencies in everyday speech communication. Recent studies have shown that speech remains intelligible when the spectrum envelope is shifted up or down along the frequency scale across a fairly wide range. However, intelligibility declines when the shift factor is greater than about 1.5, or less than 0.7 (Fu and Shannon, 1999). In this paper we present evidence from a perceptual experiment and describe a model of vowel identification that accounts for the effects of frequency shifts in terms of listeners' sensitivity to statistical regularities in natural speech. We focus on the effects of upward and downward shifts in spectrum envelope on the identification of vowels in hVd syllables.

Figure 1 illustrates the statistical covariation of $F_0$ and formant frequencies in spoken vowels associated with size differences in the larynx and vocal tract across talkers. The figure shows acoustic measurements of a set of more than 3000 vowels spoken by men, women and children from the north Texas region. There is a moderate correlation between $F_0$ and formant frequencies ($r$=0.45). On average, the change from male to female speech involves an upward shift of 15-20% in formant frequencies, accompanied by an 80-100% increase in $F_0$.

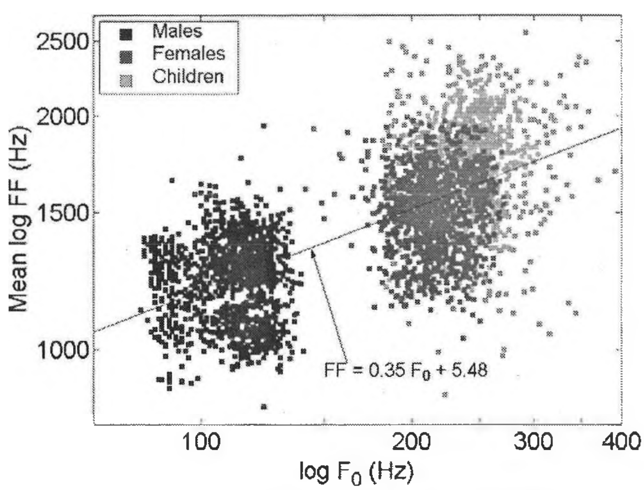



Fig. 1. Geometric mean of formant frequencies (F1-F3) vs. F0 for 11 vowels in hVd words spoken by 10 males, 10 females and 30 children aged 3-7 years (Assmann and Katz, 2000).

Assmann, Nearey and Scott (2002) used a high-quality source-filter vocoder called STRAIGHT (Kawahara, 1997) to investigate the interaction of $F_0$ and formant frequencies in frequency-shifted vowels. The baseline stimuli were /hVd/ words spoken by three adult male talkers. Identification accuracy dropped by more than 30% when the formant frequencies were shifted upward by a factor of 2.0, and by 50% when F0 was scaled by a factor of 4.0. However, when upward shifts in formant frequencies were combined with upward shifts in F0, performance improved, suggesting that listeners are sensitive to the pattern of covariation of F0 and formant frequencies in natural speech.

A pattern recognition model of vowel identification was implemented, using the modeling strategy described by Nearey and Hillenbrand (1999). The model incorporated measurements of the formant frequencies (F1,F2,F3) sampled at the 20% and 80% points in the vowel, combined with the mean F0 and duration. Linear discriminant analysis was carried out to obtain *a posteriori* probabilities of group membership for each frequency-shifted vowel. The training data for the model was a set of 3000+ vowels (examples of each of the 11 vowels from 50 talkers, including 10 males, 10 females, and 30 children from the north Texas region). The model accurately reproduced the overall pattern of results from the experiment, including the decline in identification when F0 and formant frequencies were shifted separately and the improvement when upward shifts in F0 were combined with upward shifts in the formant frequencies.

The experiment reported here extended this approach, applying both upward and downward frequency shifts to vowels from a larger sample of talkers, including adult males, adult females and children.

## 2. METHOD

The stimuli were /hVd/ words (*heed, hid, hayed, head, had, hud, hawed, hoed, hood, who'd, herd*) spoken by two adult males, two adult females, and two 7-year old children. These vowels were selected from a larger sample of vowels recorded from 10 men, 10 women, and 30 children, ages 3, 5, and 7 years from the North Texas region (Assmann and Katz, 2000). The STRAIGHT vocoder was used to resynthesize each vowel in 15 conditions (3 levels of $F_0$ shift factor x 5 levels of spectrum envelope shift factor):

**spectrum envelope scale factor** = 0.6, 0.8, 1.0, 1.5, 2.0

**$F_0$ scale factor** = 0.5, 1.0, 4.0

**Vowel**=/ɪ/, /I/, /ε/, /E/, /Θ/, /℘/, /∈∏/, /A/, /o/, /Y/, /υ/

Vowels were presented diotically over headphones to 14 listeners with normal hearing. Listeners identified the vowels using an 11-category response box on the computer screen. In the main experiment they heard 990 vowels, with all conditions randomly interspersed (11 vowels, 6 talkers, 5 spectrum envelope shifts, 3 $F_0$ shifts).

## 3. RESULTS

Figure 2 displays the interaction between spectrum envelope and $F_0$ shifts as a function of talker group. There are several key aspects of the results. First, upward and downward shifts in spectrum envelope resulted in a decline in identification accuracy, consistent with earlier studies. Second, downward shifts produced smaller effects on children's vowels than those of adults, while upward shifts showed smaller effects on vowels spoken by men compared to vowels of women and children. Third, there was a significant *improvement* in performance when an upward shift in spectrum envelope was combined with an upward shift in F0 for the adult males, consistent our previous study (Assmann et al, 2002). There was also a small but significant improvement in several conditions when a downward shift in spectrum envelope was accompanied by a downward shift in F0 for vowels of women and children.

## 4. DISCUSSION

The experiment replicated earlier studies showing a decline in vowel identification accuracy with upward and downward shifts in spectrum envelope. Adult male vowels were more resistant to upward shifts, while children's voices were more resistant to downward shifts. Children's voices have higher formant frequencies than those of adults, and hence this finding may reflect relatively fixed frequency limits on the shifts in spectrum envelope that preserve intelligibilty.

Shifts in $F_0$ also produced lower identification accuracy, and children's vowels were particularly susceptible to upward shifts in $F_0$. Further support for the perceptual learning account comes from the interaction of $F_0$ and spectrum envelope shifts. For the male vowels, identification improved when an upward shift in $F_0$ accompanied an upward shift in spectrum envelope. For the women and children, improved accuracy was observed in some conditions when downward shifts in $F_0$ were combined with downward shifts in spectrum envelope. Taken together, the results are consistent with the idea that learned relationships between F0 and formant pattern are responsible for the effects of frequency shifts on vowel identification.

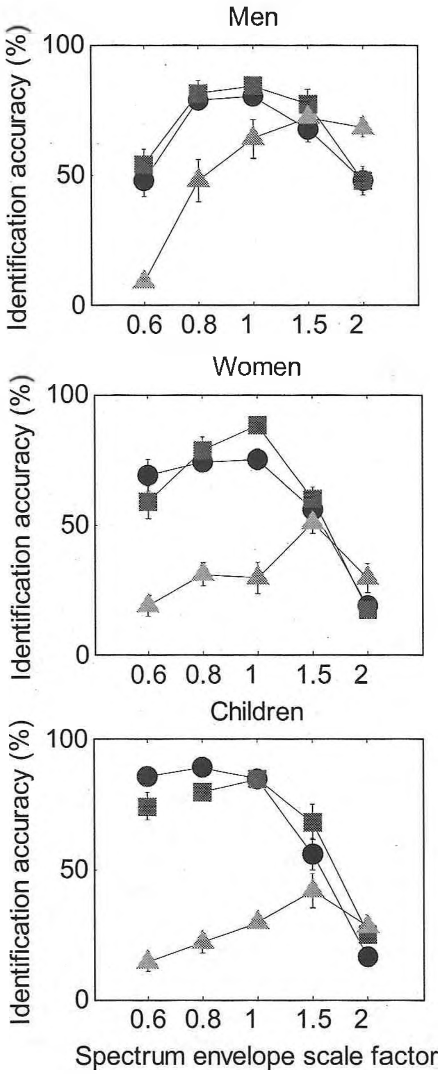



Figure 2: Effects of spectrum envelope and F0 on vowel identification accuracy

## REFERENCES


Assmann, P.F. and Katz, W.F. (2000). Time-varying spectral change in the vowels of children and adults *J. Acoust. Soc. Am.* 108: 1856-1866.

Assmann, P.F., Nearey, T.M. and Scott, J.M. (2002). Modeling the perception of frequency-shifted vowels. *Proceedings of the 7th International Conference on Spoken Language Processing*, pp. 425-428.

Fu, Q-J. and Shannon, R.V. (1999). Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing. *J. Acoust. Soc. Am.* 105: 1889-1900.

Hillenbrand, J.M and Nearey, T.M. (1999). Identification of resynthesized /hVd/ utterances: effects of formant contour. *J. Acoust. Soc. Am.* 105: 3509-3523.

Kawahara, H. (1997). Speech representation and transformation using adaptive interpolation of weighted spectrum: Vocoder revisited. *Proceedings of the ICASSP*, pp. 1303-1306.

# Effect Of Phonetic Context On Women's Vowel Area

**Janine Chesworth, Kim Coté, Colleen Shaw, Sandra Williams & Megan Hodge**
Department of Speech Pathology & Audiology, University of Alberta, Rm. 2-70 Corbett Hall, Edmonton, AB, T6G2G4
(Project of the Canadian Language and Literacy Research Network)

## 1. INTRODUCTION

This study addressed a basic methodological issue for a variable that has been identified as one important acoustic correlate of speech intelligibility. This variable is "vowel quadrilateral area" or VQA. VQA refers to the area enclosed in the quadrilateral formed by a plot of the coordinates for the first and second major vocal tract resonances (formant frequencies) of the corner vowels (/i/, /æ/, /a/ and /u/). Previous studies have used VQA as an acoustic correlate of human perceptual judgments of speech intelligibility. Because VQA has been found to account for between 41 and 53% of the variance observed in speech intelligibility scores[1,2] it has been identified as a variable of interest in acoustic modeling of speech intelligibility. One methodological concern when comparing VQA across studies of similar speakers, and across speakers who differ in stage of speech development, is the phonetic context in which the corner vowels are embedded. It is known that sounds surrounding a vowel influence the vowel's first and second formant frequencies[3,4], abbreviated as F1 and F2. To our knowledge no study has been published that has directly investigated how various surrounding speech sounds, that is, phonetic context, influence the size of VQA in the same sample of speakers. Therefore differences in VQA reported by previous investigators are difficult to compare and interpret because these differences may be due to phonetic context effects as opposed to, or in addition to, speaker differences (e.g., dialect, speaking style, age).

This study addressed the following questions:

1) Is there an effect of phonetic context on the size of F1 by F2 VQA? Based on previous research, it was hypothesized that VQAs would be smaller when the target vowels occurred in less neutral phonetic contexts. This question was addressed by calculating and comparing the F1 by F2 VQAs using a natural log Hz scale for each of four contexts. The log Hz scale was used to reduce the effect of interspeaker vocal tract size differences on vowel formant frequency values[5].

2) Is there an effect of phonetic context on the shape of F1 by F2 VQA? This question was investigated by comparing the size of F1 and F2 extents between the most and least neutral phonetic context. It was hypothesized that F1 and F2 extents for the vowel quadrilateral would be smaller in less neutral phonetic contexts.

## 2.METHOD

### 2.1 Phonetic Context Conditions

Monosyllabic utterances in four phonetic contexts were compared: /hV/, /hVd/, a subset of 24 words from the *Test of Children's Speech or TOCS* (children's speech intelligibility test)[2], and 6 sets of minimally contrastive CVC words, referred to as the MC condition. The stimulus words for the *TOCS* and MC conditions are listed in Appendix A. The /hV/ context was categorized as the most phonetically neutral since no supraglottal articulation is used for /h/. The MC context was categorized as the least phonetically neutral because all target monosyllables had both an initial and final consonant.

### 2.2 Speakers

Ten women between the ages of 20 and 35 who met the following inclusion criteria provided vowel recordings for analysis: 1) no history of speech disorder or treatment; 2) Western Canadian dialect of English as their first language; 3) normal hearing demonstrated by passing a standard hearing screening; 4) non-smokers for the past seven years; and 5) free from colds, sore throats, or any adverse health condition that may have affected their voice at the time of recording.

### 2.3 Vowel Recording

Digital audio recordings (sampling rate = 44 kHz, 16 bit quantization) of each speaker's vowel productions for each phonetic context were made in a quiet environment. Speakers wore a professional, unidirectional head-mounted microphone placed 1.5 inches away from the corner of the mouth. Speakers were familiarized with pronunciation of the /hV/ and /hVd/ printed stimuli prior to recording. Each stimulus word was presented on a computer screen. Eight practice words were presented to familiarize speakers with the task. Speakers then produced the 96 test items (24 items per condition, times 4 conditions) in randomized order. A panel of three judges provided independent verification that the vowel in each recording was a perceptually valid member of the target vowel category.

### 2.4 Formant Frequency Measurement

The F1 and F2 values of the target vowels in each speaker's digital audio files were estimated from wideband spectrograms using CSpeech 4.0[6]. The analyzing bandwidth was at least twice the speaker's highest fundamental frequency (F0) for /i/ and was in the range of 450 to 550 Hz. To isolate the vowel, the cursors were placed on the first and last glottal pulses that excited the first two formants of the vowel segment. The length of the vowel was measured as the distance between the cursors. F1 and F2 values were estimated visually from a 30-millisecond window that was centered at the midpoint of the vowel segment using FFT and LPC power spectra as guides.

## 2.1 Calculation: Vowel Quadrilateral Areas

Mean F1 and F2 frequencies were calculated for the six tokens for each vowel category in each condition, in the log Hz scale. For example, F1 values for a speaker's six TOCS /i/ vowel tokens were averaged to give a mean F1 value. This was repeated for F2 values. These mean F1, F2 coordinates for each corner vowel were used to generate vowel quadrilaterals for each speaker in each phonetic context using a log Hz scale. VQA was calculated using the formula reported by Higgins and Hodge[2].

## 2.2 Calculation: Formant Extents

A formant extent was defined as the difference in log Hz between the smallest and largest formant value, for a given formant, in a given phonetic context condition, for each speaker. For F1, this would be the range covered by the vowel quadrilateral plot on the x axis and for F2, the range of the vowel quadrilateral plot on the y axis.

# 3. RESULTS

## 3.1 Vowel Quadrilateral Areas

Group mean VQAs for each condition are shown in Table 1. As predicted, VQAs are greatest for the neutral context /hV/ and smallest for the two non-neutral contexts (*TOCS* and MC words). A repeated measures 1-way ANOVA revealed a significant effect of phonetic context on VQA ($F(3,27) = 33.0$, $p = .0001$). Results of post-hoc testing indicated that all conditions differed significantly from each other ($p < .008$ – Bonferroni correction) except for the *TOCS* and MC conditions ($p = .015$).

Table 1. Mean VQA for each phonetic context.

| | Mean (log Hz$^2$) | Std. Deviation (log Hz$^2$) |
|---|---|---|
| /hV/ | .544 | .122 |
| /hVd/ | .502 | .109 |
| *TOCS* | .442 | .103 |
| MC | .401 | .862 |

## 3.2 F1 and F2 Extents

Group mean F1 and F2 extents for each condition are shown in Table 2. Statistical testing using 1-way ANOVAS with repeated measures revealed no significant difference across phonetic contexts for F1 ($F(3,27)$ $p = .129$) but a significant difference for F2 ($F(3,27)$ $p < .000$). Post-hoc testing revealed significant differences between the /hV/ and *TOCS* conditions and between the /hV/ and MC conditions ($p < .008$ – Bonferroni correction).

Table 2. Mean F1 and F2 extents for each phonetic context.

| | /hV/ | /hVd/ | *TOCS* | MC |
|---|---|---|---|---|
| *F1 Extent* | | | | |
| Mean (log Hz.) | .992 | .951 | .959 | .931 |
| Std. Dev. | .121 | .102 | .090 | .108 |
| *F2 Extent* | | | | |
| Mean (log Hz.) | .889 | .835 | .787 | .779 |
| Std. Dev. | .133 | .128 | .088 | .067 |

# 4. DISCUSSION AND CONCLUSIONS

As predicted[3,4], VQA was affected significantly by phonetic context. The largest area was obtained in the most phonetically neutral (/hV/) condition and the smallest areas were obtained in the least neutral (*TOCS* and MC) conditions. Change in shape of the vowel quadrilateral from most to least neutral contexts was affected by reduction in both F1 and F2 extents but was significant only for F2. Visual inspection of the VQA plots (not shown here) revealed that increases in F2 frequency values for /u/ contributed most to reduced VQA in non-neutral phonetic context conditions. Future research will test hypotheses about interactions between age and phonetic context on vowel quadrilateral size and shape.

# 5. REFERENCES


1. Weismer, G., Jeng, J., Laures, J., Kent, R., & Kent, J. (2001). Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. Folia Phoniat. et Logopaedica, 53, 1-18.
2. Higgins, C. & Hodge, M. (*2002b*). Vowel area and intelligibility in children with and without dysarthria. J. Medical Speech-Language Pathology, 10(4), 271-274.
3. Hillenbrand, J., Clark, M. & Nearey, T. (2001). Effects of consonant environment on vowel formant patterns. J. Acoust. Soc. Am., 109 (2), 748-763.
4. Stevens, K. & House, A. (1965). Perturbation of vowel articulations by consonantal context: An acoustical study. J. Speech Hear. Res.6, 111-121.
5. Nearey, T. (1992). Applications of generalized linear modeling to vowel data, Proceedings, 1992 International Conference on Spoken Language Processing, 583-586.
6. Milenkovic, P., & Read, C. (1992). CSpeech 4.0, Laboratory version [Computer Program], Department of Electrical and Computer Engineering, University of Wisconsin, Madison, WI.


# 6. APPENDIX A: *TOCS* and MC Stimulus Words

| | /i/ | /ae/ | /a/ | /u/ |
|---|---|---|---|---|
| *TOCS* | | | | |
| | heat | hat | hot | hoot |
| | "D" | badge | hawk | chew |
| | eat | hash | chop | shoe |
| | beat | bad | jaw | sue |
| | seat | hatch | stop | two |
| | sheet | pad | top | zoo |
| MC | | | | |
| | seat | sat | sought | suit |
| | beat | bat | bought | boot |
| | G's | jazz | jaws | Jews |
| | keep | cap | cop | coop |
| | teak | tack | talk | touque |
| | heat | hat | hot | hoot |

# Silence as a Cue to Rhythm in the Analysis of Speech and Song

**David Gerhard**
Computer Science, University of Regina, 3737 Wascana Pkwy, Regina, SK Canada, S4S 0A2. david.gerhard@uregina.ca

## 1. INTRODUCTION

The *rhythm* of a piece of sound is a measure of the existence, duration and repetition of time events in the sound. When dealing with full spectrum instrumental music, these time events are readily identified using spectral analysis techniques such as filter banks. Rhythmic structure is often apparent in percussion instrumentation, for example drums and cymbals, which are normally in different spectral locations than the rest of the instrumentation.

When dealing with human utterances, the spectral filter bank methods are not as successful because the acoustic events that may contribute to a sense of rhythm, are similar in spectral content to the rest of the utterance. Different methods must be employed to identify rhythmic structure. This paper discusses the merit of using silence as a cue to the rhythmic structure of human utterances, specifically speech and song. Identifying the rhythm of a human utterance applies to the problem of music information retrieval, specifically in the instance where a human musical utterance (i.e. song) is used as a query in a musical database. Currently, a string of note directions is commonly used as a search string into the database, without referring to rhythmic information. Rhythmic cues may help in retrieval accuracy. Further, rhythmic information may aid in he differentiation between speech and song, applicable to speech recognition systems as well as speech and music therapy.

This work seeks to characterize the statistical distribution of sound pressure level (*spl*) across an utterance, identifying silence as a value of *spl* below a hysteresis threshold. This work concentrates on English speech and the western musical idiom. Several researchers have identified a 4 Hz power modulation in speech while attempting to differentiate between speech and music [2,3]. Unaccompanied song tends to have a less regular power distribution.

## 2. METHOD

Speech and song development and test data were acquired using human subject testing, where subjects produced vocal utterances based on a series of prompts [REFb]. Each datum consists of an acoustic event of approximately 5 seconds. The development data are used to build a computational model of the feature being tested, and the test data are then applied to the model to verify correct classification of *a priori* unseen data.

The models developed in this work relate to the rhythm of an utterance as evidenced by the distribution of silence in the utterance. Utterances are divided into 0.015s frames, and each frame is characterized as silence if the *spl* is below a hysteresis threshold. The threshold is defined by a 5- to 30-frame window at the beginning of the utterances which is assumed to represent the environmental background noise. To allow fuzziness around the threshold, the hysteresis function shown in Figure 1. is used, where T is the *spl* threshold defined by the beginning of the utterance, and t is the allowable deviation. "0" is a silent frame and "1" is a frame with *spl* above the dynamic threshold.

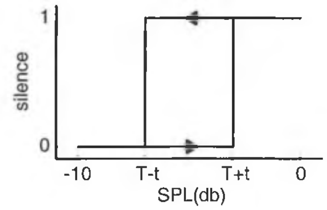



**Figure 1. *spl*-to-silence conversion hysteresis function.**

## 3. RESULTS

Silence cues to rhythm can be as simple as the proportion of silence frames in an utterance, which can be used as a feature to discriminate between talking and singing. Figure 2 shows probability distribution estimations for the proportion of silent frames in an utterance, for both talking and singing. Singing files overall tend to have less silence than talking files, although there is considerable overlap.

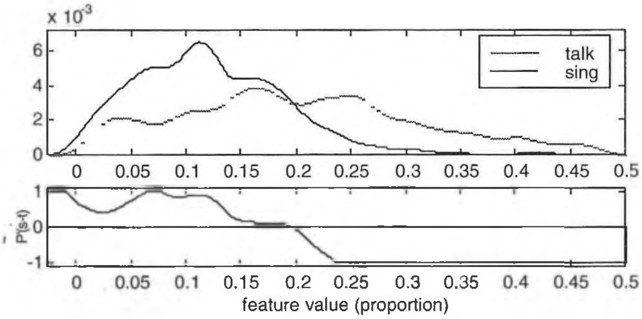



**Figure 2. Probability density estimation comparisons for silence.**

Rhythmic structure is also examined on a case-by case basis to identify characteristics based on silence patterns. Figures 3 to 7 show examples of these characteristics, which are discussed in Section 4. The figures show *spl* over time, with the upper plot of each figure showing the silence metric generated from the hysteresis function.

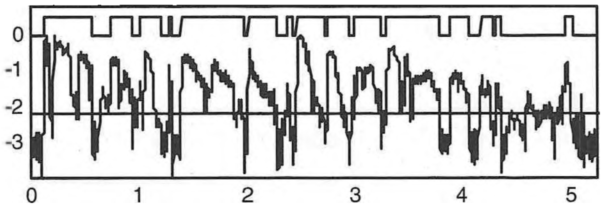



**Figure 3. (*l145*) Typical speech.**

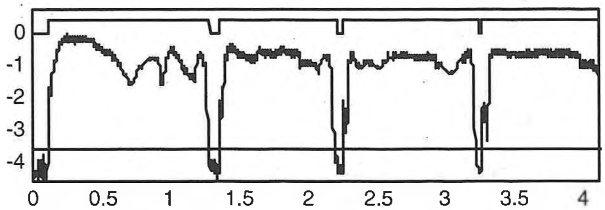



**Figure 4. (*m124*) Typical song**

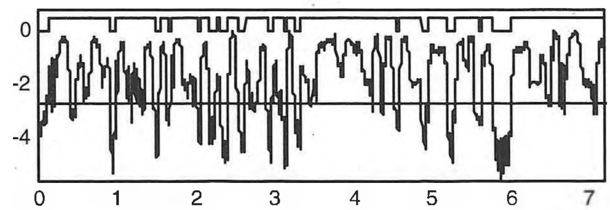



**Figure 5. (*m126*) Song with rapid changes**

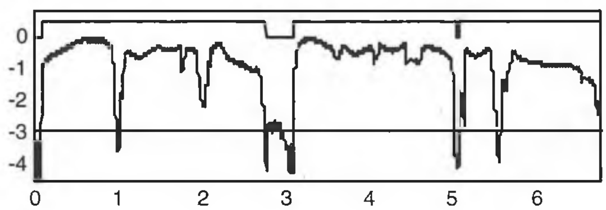



**Figure 6. (*g258*) "O Canada...", sung**

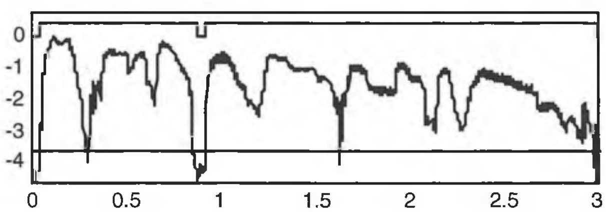



**Figure 7. (*h258*) "O Canada...", spoken**

## 4. DISCUSSION

Silence in human speech utterances typically is the result of two situations. When phonetic stops ('p','t','k') are used in speech or song, a small period of silence precedes the plosive. These silences occur during the course of continuous speech, and are not altered in song. The second situation of silence in an utterance is

Figure 3 shows a typical speech utterance. The rhythmic patterns are not related to any specific underlying beat or tempo. Figure 4 shows a typical sung utterance, where a theme is repeated and a series of silence events can be traced across the utterance. The silence events are well spaced and the *spl* track contains a repeating pattern.

Figure 5 shows a sung utterance where the identification of a rhythmic structure is less easy to define. The sung utterance has rapid changes and while there is rhythmic structure, the distribution of silence events is not sufficient to identify an underlying rhythmic structure. As with any analysis of perceptual events, techniques must be moderated with the domain of the perceptual phenomenon.

Figures 6 and 7 are a comparison of speaking and singing utterances of the same lyric by the same subject. This provides a good example of the use of silence as a rhythmic defining structure. The *spl* track shows a period of silence at 2.7s. This period of silence is the space between the first and second phrases of the utterance "O Canada, our home and native land". The spoken version of the utterance has a much smaller gap between the phrases, showing that the longer gap in Figure 6 is primarily for musical emphasis.

Although this paper did not cover the distribution of the *spl* track itself, it is of note that the tracks have significant differences: In the sung track in Figure 6, there is constancy in the *spl* across a syllable, whereas in Figure 7, the corresponding syllables have more variation in the *spl* track.

The rhythmic structure of a musical utterance, therefore, is characterized by regularly spaced silence events representing phonetic stops or phrase boundaries, and the phrase boundary silence events are typically longer than that in speech utterances. This work may be expanded to include the use of *spl* track and silence event analysis to discover rhythmic characteristics in acoustic items other than human utterances., particularly instrumental music.

## REFERENCES

[1] List, G. (1971). The Boundaries of Speech and Song. in Readings in Ethnomusicology, D.P. McAllester (ed.). New York, Johnson Reprint Co., 253-268.

[2] Karnebäck, S (2002). Expanded Examinations of a Low Frequency Modulation Feature for Speech/Music Discrimination. Proc of ICSLP2002.

[3] Scheirer, E. and Slaney, M. (1997). Construction and Evaluation of a Robust Multifeature Speech/Music Discriminator. IEEE International Conference on Acoustics, Speech and Signal Processing, II:1331-1334.

[4] Gerhard, D. B. (2002). A Human Vocal Utterance Corpus for Perceptual and Acoustic Analysis of Speech, Singing and Intermediate Vocalizations. Journal of the Acoustical Society of America. 112, V: 2264

# Perception Of Timbral Fusion

**Dora Chan and Elzbieta Slawinski**
Department of Psychology, University of Calgary, 2500 University Dr. NW., Calgary Alberta, T2N IN4

## 1. INTRODUCTION

The concept of timbre has been a difficult topic to study in the field of musical acoustics due to its multidimensional nature. This study was aimed at gaining more knowledge on this musical attribute. In particular, this study was aimed at understanding the phenomenon of timbral fusion, wherein one perceives a set of simultaneous frequency components as one complex tone. In order to explore the perception of timbral fusion, pairs of instrumental timbres were perceptually fused or simultaneously sounded together. Listeners were required to make a judgment on whether they heard one or two instruments after listening to each tonal stimulus. Performance differences between musicians and non-musicians were examined. It was hypothesized that musicians would perform better than non-musicians because of their extensive experience with musical tones.

## 2. METHODS

### 2.1 Participants

Thirty-three participants (17-29 yrs) participated in the study for course credit. Musicians (n = 10) had at least 8 years of formal musical training.

### 2.2 Stimuli and Procedure

Stimuli were pairs of simultaneously sounding natural instruments, and simultaneously sounding simple tones. All tones were 300 ms in duration. The steady state portions of three instrumental timbres were used in this study, including clarinet, trombone and harp. Ten selected pitches were considered in this study: Eb4, Ab4, A5, Bb5, B5, C5, C#5, D5, and Eb5. The pairs of stimuli were constructed such that the two timbres began one octave apart (Eb4) then gradually sounded closer together until the two timbres merged as one (Eb5). The combinations of the three timbres yielded three experimental groups: 1) clarinet Eb5 / trombone (10 pitches); 2) harp Eb5/ clarinet (10 pitches), 3) trombone Eb5 / clarinet (10 pitches). The control group consisted of pure tones corresponding to the combination of the 10 pitches in the experimental conditions. The combination of the 10 pitches from the four groups produced 40 different stimuli. Each stimulus was repeated 10 times in random order for a total of 400 trials. These trials were broken into five blocks with 80 trials each.

In the one-hour session, participants were required listen to the 400 signals. Their task was to make a judgment on whether they heard one or two instruments after listening to each tonal stimulus.

## 3. RESULTS

The following results report only a preliminary comparison between musicians and non-musicians. Overall means and standard deviations of the two groups were calculated and compared.

In the control condition, musicians outperformed non-musicians in all signals except for the Eb4 and Eb5 combination. For this signal, non-musicians' mean score was 15.5% ($SD$ = .189) versus musicians' score of 13% ($SD$ = .152).

In the Experimental Condition 1 (Clarinet at Eb5 and Trombone), musicians on average performed better than non-musicians. For two signals, however, non-musicians outperformed musicians. The first signal was Clarinet and Trombone (Eb4) where non-musicians achieved a mean of 84% ($SD$ = .230) and musicians with a mean of 77% ($SD$ = .295). Clarinet and Trombone (Eb5) was the second signal where non-musicians ($m$ = 55.5%; $SD$ = .287) outperformed musicians ($m$ = 46%; $SD$ = .353). Compared to the other signals in this condition, both groups found it harder to perceive the Eb5 combination as two timbres.

In Experimental Condition 2 (Harp at Eb5 and Clarinet), both musicians and non-musicians performed fairly well overall. However, musicians performed consistently better than non-musicians in every signal.

In Experimental Condition 3 (Trombone at Eb5 and Clarinet), musicians again outperformed non-musicians. The differences between the two groups for all signals were quite consistent.

## 4. DISCUSSION

The hypothesis that musicians would perform better than non-musicians was supported. Apparent differences were found in this study. Overall, musicians had higher percentage scores than non-musicians in identifying that there were two timbres in each stimulus. Thus, it was easier for musicians than non-musicians to hear two simultaneous sounding timbres. This may suggest that musical training or experience enhances the perception of timbre. This difference between musicians and non-musicians is in accordance to other studies (e.g., Kendall, 1986).

Pairs of instrumental timbres (Experimental Conditions) were more accurately perceived by both groups than pairs of pure tones (Control Condition). The differences in performance between the experimental conditions and the control condition in this present study can be explained by the characteristics of the waveform. All tones produced by musical instruments are not pure tones but mixtures of pure-tone frequencies or partials (White & White, 1980). The perception of fusion depends on the synchrony of the frequency partials in complex sounds, therefore, the fusion of two instrumental timbres is often harder to perceive because there is a lower probability of synchronicity between all the partials.

Overall, the octave combinations were more difficult for listeners to perceive as two timbres. When tones are separated by an octave they are considered to be musically and perceptually equivalent. Physically, the octave is the only interval in which the harmonics will coincide exactly; therefore, two notes that are separated by octaves cannot create dissonance. In this study, octave combinations were often lower compared to the other pitch combinations. It is interesting to note that the only combinations where non-musicians appeared to perform better than musicians were the Eb4 and Eb5 combinations (i.e., pure tones combinations of Eb4 and Eb5, Clarinet at Eb5 with Trombone Eb4). This result was unexpected; perhaps musicians have more experience with harmonic sounds (i.e., triads and chords), therefore they can fuse the harmonic partials more readily than non-musicians.

## 5. CONCLUSION

The study of timbre as a musical attribute has received much more attention in the recent decades. However, we still do not fully understand its multidimensional nature. Timbral fusion or timbral blending appears to be an important area of study since most music is created by the simultaneously sounding of instruments; however, timbral fusion is still a relatively unexplored area in timbre research. In listening to orchestral music, the perception of fusion also depends on the type conducting. Conductors may manipulate timbre by combining and emphasizing instruments in a certain way. This illustrates that timbral manipulation commonly takes place in the real world, and therefore, in order to understand music in an ecological sense, we must strive to understand timbre.

## 6. SELECTED REFERENCES


Bonfield S. & Slawinski, E. (2002). Targeting timbre recognition abilities of musical participants: when musicians hear what isn't there. *Canadian Acoustics, 30*, 160-161.

Grey, J.M., & Gordon, J.W. (1978). Perceptual effects of spectral modifications on musical timbres. *Journal of the Acoustical Society of America, 63,* 1493-1500.

Kendall, R.A. (1986). The role of acoustic signal partitions in listener categorization of musical phrases. *Music Perception, 4*, 185-214.

Kendall, R.A., & Carterette, E.C. (1993). Identification and blend of timbres as a basis for orchestration. *Contemporary Music Review, 9*, 51-67.

Risset, J-C. & Wessel D.L. (1999). Exploration of timbre by analysis and synthesis. In D. Deutsch (Ed.), *The psychology of music* (2nd ed.) (pp. 113-169). San Diego: Academic Press.

Saldanha, E.L. & Corso, J.F. (1964). Timbre cues and the identification of musical instruments. *Journal of the Acoustical Society of America, 36,* 2021-2026.

Sundberg, J. (1991). *The science of musical sounds.* San Diego: Academic Press.

White, H.E. & White, D.H. (1980). *Physics and music: the science of musical sound.* Philadelphia: Saunders College.

# Recognizing Familiar and Foreign Words and Music by Children and Adults: An examination of the critical period hypothesis

**Elizabeth D. McFadden & Annabel J. Cohen**
University of Prince Edward Island, 550 University Ave., Charlottetown, PE, Canada, C1A 4P3
dmcfadden@upei.ca, acohen@upei.ca

## 1. INTRODUCTION

Acquisition of language is usually an easier task in childhood than in later life (Birdsong & Molis, 2001; Johnson & Newport, 1989; Trehub, 1976; but see Bialystok, 1997). The general finding is consistent with Lenneberg's (1967) proposal of a critical period for language acquisition: typical language development requires extensive exposure to language during childhood.

Comparisons between language and music have revealed many similarities. Both language and music are based on generative rules capable of producing an unlimited number of new sequences (Sloboda, 1985). As they both rely on auditory functions, language and music may share certain neural mechanisms (Zatorre et al., 2002). Finally, both language and music acquisition proceed in systematic stages (Besson & Schön, 2001).

Given the similarities between language and music in terms of generativity, auditory modality and developmental stages, it is of interest to determine whether the linguistic notion of critical period applies to the acquisition of music. Our laboratory has been conducting research on this issue by studying memory and preference in different age groups for music representing popular styles of the last 10 decades. For example, children, pre-adolescents, and young adults were asked to rate popular music from the 10 decades of the 20th century on preference and familiarity (Bailey, 1999; Bailey & Cohen, 2002). Participants later received a surprise recognition task using a test tape containing examples only some of which were previously presented. In support of an early critical period for music acquisition, young children, in contrast to older participants, were uninfluenced by style in immediate recognition of music though performing at beyond chance levels. To account for these results it was proposed that exposure to music of a certain style or grammar early in life establishes a representation of the grammar of that style (Cohen, 2000). Due to declining brain plasticity, this grammar serves the individual throughout life. Consequently music styles that violate the grammar will not be well encoded whereas music styles that reflect the grammar will be well encoded.

Whereas the past work examined music in American (Western) styles, the theory applies also to music of non-Western styles. The non-Western grammar would violate the grammar acquired by Canadian children early in life. The present study investigates the possibility of a critical period for the acquisition of music grammar by comparing recognition of non-Western and Western music.

The argument proposed for examination of the critical period for music also applies to language. It would be expected that if an early critical period for language exists, then recognition of acoustically presented words in a foreign language would be relatively more easily retained by children who are acquiring the rules of language than by adults who had already acquired them. Hence, the present study also investigated the relative ease with which children versus adults carried out a recognition task for words in their own versus an unfamiliar language. According to the critical period hypothesis it was predicted that the difference in the foreign and native recognition scores for music and language would be less than the respective difference within these two acoustical domains for adults.

## 2. METHOD

### 2.1 Stimuli

English words, Hindi words, Western classical music excerpts and non-Western music excerpts were recorded on compact disk. For each of the 4 stimulus categories, there was one presentation track and two test tracks, each with 12 items. The test tracks contained 6 of the original words or excerpts and 6 new but similar examples. The language lists consisted of equal numbers of 2, 3, and 4 syllable words (4 of each). The musical excerpts were instrumental to avoid cues from lyrics and were from either Western culture or a variety of non-Western cultures.

### 2.2 Participants

Adults. There were 19 adults (mean age 28.3 years, SD = 9.5, range 19-46) from the university population approximately half of whom had private music lessons, and 12 children (mean age 10.5 years, SD 1.7, range 8-13) attending a summer camp or friends / acquaintances of the experimenter (E.D.M.). Of the children, 4 had some private music training.

### 2.3 Procedure

All participants were asked to listen to a list of English words after which they were presented with one of two test lists and asked to indicate, for each test-list word, whether or not they had heard it on the original list. The same procedure was carried out with Hindi words, Western musical excerpts, and non-Western musical excerpts.

## 3. RESULTS

As a check on the internal validity of the Western/non-Western music distinction, 12 university students rated the music excepts on a 5-point scale where 1 represented Non-Western music and 5 Western Music. Because several selections received mean ratings in the mid-range, only those excerpts that received extreme ratings were included, leaving 10 rather than 12 items in the music recognition test.

For each participant, the mean percent correct recognition was calculated as shown in Figure 1. The relative ease of recognizing the Hindi words by children as compared to adults is dramatic. The difference does not appear for the music examples.

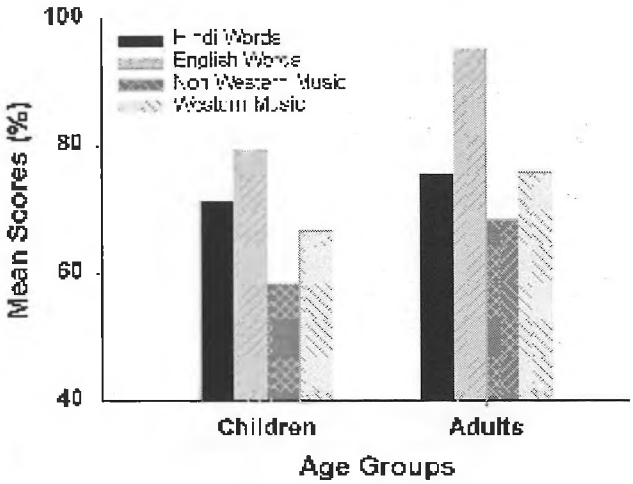



**Figure 1. Mean percent recognition for Western and non-Western words and music for children and adults.**

Mean per cent correct for words and music, for the 2 age groups was entered into an ANOVA with 2 within-subject variables of foreign vs native (Culture), and words vs music (Stimulus Domain), and 1 between groups factor of Age. Separate analyses were also carried out for words, music, and each age group. In the overall analysis, there were main effects of Culture, $F$ (1,29) = 20.60; $p < .0001$, Domain, $F$ (no df shown if same as previous) = 17.43; $p < .001$, and age, $F = 11.36$; $p < .005$. In the analysis of words, the Culture x Age interaction was significant, $F = 4.96$; $p < .05$ as were main effects of Culture, $F = 25.94$; $p < .0001$, and Age, $F = 5.94$; $p < .05$. In the analysis of music, Culture x Age was not significant, the main effect of Culture only approached significance, $p < .064$, but Age was significant, $F = 4.75$; $p < .05$. For children, words were more easily recognized than music, $F$ (1,11) = 5.78; $p < .05$ and Culture approached significance, $p < .09$. For adults words were more easily recognized than music, $F$ (1, 18) = 13.18; $p < .005$. English words and music were more easily recognized than Non-Western counterparts, $F = 27.10$; $p < .0001$ and Culture x Domain was significant, $F = 4.77$; $p < .05$.

## 4. DISCUSSION

The pattern of results for words is consistent with the early critical period hypothesis. Children suffered significantly *less* from linguistic unfamiliarity than did adults. It would be expected that even younger children than those tested here would be less sensitive to the cultural distinctions. In contrast to the age-related performance differences for words, both children and adults suffered relatively equally from unfamiliar music. However, recognition of the familiar music was relatively poor for adults in contrast to that for English words. Future work could examine music that more clearly distinguishes Western and non-Western culture for adults, using popular as opposed to classical genres for the Western music. We have previously proposed the possibility of an additional, later critical period for music due to the social significance of music in adolescence and early adulthood (Bailey & Cohen, 2002). The present data showing less detriment due to cultural variables for music than words in late adolescence are to some extent consistent with the notion of a second critical period for music.

## 5. REFERENCES

Bailey, B. (1999). *Similarities between the acquisition of musical structure and language grammar.* Unpublished honours thesis. Department of Psychology, University of Prince Edward Island.

Bailey, B. & Cohen, A. J. (2002). Acquisition of musical vernacular in children, pre-adolescents, and young adults. *Canadian Acoustics, 20,* 15-155.

Bialystok, E. (1997). The structure of age: in search of barriers to second language acquisition. *Second Language Research, 13,* 116-137.

Birdsong, D. & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language, 44,* 235-249.

Besson, M. & Schön, D. (2001). Comparison between language and music. In R. J. Zatorre & I. Peretz (Eds.), *The Biological Foundations of Music* (pp. 232-258). *Annals of The New York Academy of Sciences, Volume 930.*

Cohen, A. J. (2000). Development of tonality induction: Plasticity, exposure and training. *Music Perception, 17,* 437-459.

Johnson, J. S. & Newport, E. L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21,* 60-99.

Lenneberg, E. H. (1967). *Biological Foundations of Language.* New York: Wiley.

Sloboda, J. A. (1985). *The musical mind.* Oxford: Oxford University Press.

Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development, 47,* 466-472.

Zatorre, R. J., Belin, P. & Penhume, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trents in Cognitive Sciences, 6, 37-46.*

Acknowledgement: The work is supported by a grant from the Natural Sciences and Engineering Research Council (NSERC) to A. J. Cohen.

# Effect Of Articulation Condition On Children's Acoustic Cues For Bilabial Place

**Megan Hodge, Christine Huang, Janka Hegedus and Melanie Campbell**
Department of Speech Pathology and Audiology, University of Alberta, 2-70 Corbett Hall, Edmonton, AB T6G2G4

## 1. INTRODUCTION

Energy patterns in the speech acoustic signal, and how these patterns change over time as speech muscle groups produce the movements of speech, can reveal information about what the speech mechanism is doing during speech production. Acoustic spectral variations are related to the shape of the vocal tract. Cues contained in the speech spectra provide information about the place and manner of articulation of speech sounds[1]. Cues to place of stop articulation are captured in parameters of the burst spectrum (mean or spectral moment 1 and skewness or spectral moment 3) and in second formant (F2) frequency transitions. It is also well established that voice onset time (VOT) varies with place of articulation[2].

The purpose of this study was to determine if these four acoustic cues (stop burst spectral moments 1 and 3, F2 onset and VOT) for bilabial place are stable across articulatory condition, and if these are influenced by the presence of /ɹ/ following the stop. We examined measures of four acoustic cues for two stop consonant places (bilabial and lingua-alveolar) from productions of three children who represent three different articulatory conditions: a girl with normal facial movement and typical speech production (CNM) and two girls with facial paralysis, who differed in the nature of their articulatory compensation. One child had congenital, non-progressive facial paralysis and had undergone facial animation surgery at age 7 years (CP1). Prior to this surgery she used a lingua-dental contact for bilabial stop targets. Following the surgery she had been taught to use a labiodental dental contact for bilabial stop targets. The second child (CP2) had a progressive facial paralysis that was complete by age 14 months. She used a lingua-dental contact for bilabial stop targets.

Previous work[3,4] indicated that all three children (CNM, CP1 and CP2) showed F2 CV onsets that fit the typical pattern, that is, lower F2 onset frequencies for bilabial than for lingua-alveolar stops. Analysis of audio recordings from CNM and CP1 revealed that both children showed the expected VOT pattern, that is, longer duration for lingua-alveolar than bilabial stops in both CV and CɹV contexts. However, results for spectral moments 1 and 3 for CP1 showed the reverse of the expected pattern in. Rather than showing a lower spectral mean and a more positive skewness for bilabial than lingua-alveolar stops, her spectral means for bilabial place were higher and spectral skewness less positive than for lingua-alveolar place in both CV and in CɹV contexts. The child with typical speech production (CNM) showed the expected pattern for burst spectral mean and skewness in CV contexts but showed the reverse pattern, like CP1, in CɹV contexts. In this study, recordings from CP2 were analysed and are compared with previous findings for CNM and CP1.

## 2. METHOD

### 2.1 Speakers

At the time of recording, the three girls ranged in age from 9-10 years and had hearing within or corrected to within normal limits. All had age appropriate receptive and expressive language abilities and school performance. CP1 had a diagnosis of Moebius syndrome associated with complete congenital paralysis of facial muscles, including the lips. CP2 had a diagnosis of early onset facioscapulohumeral muscular dystrophy, a progressive condition that affects muscles of the face and shoulder girdle primarily but can involve other muscle groups (tongue, soft palate, pelvic girdle) as the disease progresses.

### 2.2 Recording of Stimuli

The stimulus set consisted of 16 word pairs minimally contrastive in place of stop consonant articulation. These phonemes were contrasted in initial position in the stimulus words: /p/ versus /t/, and /b/ versus /d/. For 8 of the word pairs, the target sounds occurred in singleton context (CV) with no adjacent consonants. Another 8 word pairs contained the target sounds followed by /ɹ/ (C/ɹ/V). The child produced each of the words in the carrier phrase, "I can say ___ again." Recordings of the word pairs were elicited using an imitative model. The child repeated words presented by the examiner while viewing a semantically relevant picture. The recordings were obtained in a sound booth using a Panasonic 455 S-VHS video camera-recorder unit and an external Sony 150T electret condenser microphone that was positioned approximately 6 cm below the child's chin.

### 2.3 Measurement of Acoustic Cues

CSpeechSP[5] was used to measure the four acoustic cues of interest. To measure VOT and F2 transition onset, digital audio files of the child's recordings of each stimulus phrase were made using a sampling rate of 22 kHz with a quantization size of 16 bits and a low pass cut-off filter of 9 kHz. The target word was extracted from the files with a 50-ms lead. VOT was obtained by measuring the time (ms) from the first transient noise burst for the stop release to the zero-crossing of the first positive-going, large amplitude, well-formed repetitive wave shape indicating vocal fold vibration. F2 transition onset was measured by visually estimating the centre of the second formant on a wideband spectrogram, with reference to a narrowband spectrogram, fast Fourier transform (FFT) spectra, and linear predictive coding (LPC) spectra as needed. Measurements were taken at the first glottal pulse after the stop burst that excited at least the first two formants of the vowel[3]. A 600-Hz analysing bandwidth was used for the wideband spectrogram so that it was slightly greater than twice the child's mean fundamental frequency of 248 Hz[3]. Audio files for spectral moments analysis were prepared as for VOT and F2 measures but using a low pass cut-off filter of 10 kHz. Each word waveform was edited to include the interval from the onset of the first burst to the end of the third pitch period with a 10-ms leader. The spectral moment values for all audio files were calculated using CSpeechSP, which automatically generates FFTs on a series of overlapping 20-ms hamming windows at 10-ms intervals. The initial window was centred on the stop burst and the final window was centred at the end of the third pitch period of the vowel. Burst spectral mean and burst spectral skewness were computed from the resulting probability distributions. Results are reported for the first 20 ms interval, centred on the stop burst.

## 3. RESULTS

Effects of the two independent variables, place and context, were tested statistically using two-way ANOVAs with repeated measures for each acoustic cue. The results are summarized in Table 1. Significance or size of differences in means and the direction of those differences are listed for each cue.

Table 1. Relationship of bilabial and alveolar means for four acoustic cues.

| *Acoustic Cue* | Place | Context | P x C |
|---|---|---|---|
| VOT | B<A ($p$ = .008) | S</ɹ/ ($p$ = .002) | No ($p$ = .161) |
| F2 Onset | B<A ($p$ = .004) | S>/ɹ/ ($p$ = .117) | Yes ($p$ = .035) |
| Spectral Mean $1^{st}$ 20-ms | B<A ($p$ = .04) | S>/ɹ/ ($p$ = .38) | No ($p$ = .99) |
| Spectral Skewness $1^{st}$ 20-ms | B>A ($p$ = .04) | S</ɹ/ ($p$ = .03) | No ($p$ = .49) |

*Note.* B = Bilabial; A = Alveolar; S = Singleton; /ɹ/ = /ɹ/-Context

Significant main effects were found for both place and phonetic context for VOT and F2 onset. Mean VOT and F2 onset values were lower for bilabial than for alveolar place in both contexts. For the first 20-ms segment of the burst, a significant main effect of place was found for moment 1 and moment 3. Moment 1 values were lower and moment 3 values were more positive for bilabial than lingua-alveolar place. Moment 1 did not differ significantly by context but mean values were lower in singleton context for both places of articulation. A significant main effect for context was found for moment 3 values, with lower values in singleton context.

## 4. DISCUSSION AND CONCLUSIONS

The child with normal facial movement and both children with facial paralysis showed the expected pattern of VOT and F2 onset values for bilabial vs. lingua-alveolar stop place in both CV and CɹV contexts. For both children with facial paralysis, these acoustic distinctions in consonant place were greater in /ɹ/-context. However, each of the three children showed a different result for the two stop burst cues. The child with normal facial movement and typical articulation showed the expected pattern of a lower mean and more positive skewness for bilabials in a CV context but the reverse of this pattern in the CɹV context. CP1 showed the unexpected pattern of a higher mean and a less positive skewness for bilabials in both CV and CɹV contexts. CP2 showed the pattern of a lower mean and a more positive skewness for bilabials in both CV and CɹV contexts. These varied results for spectral burst cues suggest that they are sensitive to articulatory differences in stop articulation while VOT and F2 onset appear to be more stable place cues, regardless of articulatory differences. Further research with larger numbers of speakers is necessary to clarify the effect of /ɹ/ on spectral burst place cues in typical speakers and to determine the relative importance of different cues on place identification for speakers with typical and atypical stop consonant articulation.

## 5. REFERENCES


1. Pickett, J. (1999). *The acoustics of speech communication: Fundamentals, speech perception theory, and technology.* Needham Heights, MA: Allyn & Bacon.
2. Cho, T. & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics, 27*, 207-229.
3. Nelson, M.A., & Hodge, M.M. (2000). Effects of facial paralysis and audiovisual information on stop place identification. *Journal of Speech, Language and Hearing Research, 43*, 158-171.
4. Hodge, M., Hegedus, J., Kenny, M., & Souster, C. (2001, June). *Influence of /ɹ/ on burst spectra for stop place.* Poster session presented at the $141^{st}$ meeting of the Acoustical Society of America, Chicago, IL.
5. Milenkovic, P.H., & Read, C. (1997). CSpeechSP, Laboratory Version [Computer Program]. Department of Electrical and Computer Engineering, University of Wisconsin, Madison, WI.

# Comparison Of Intelligibility Measures On Single Word And Spontaneous Speech Tasks For Children With And Without Cleft Palate

**Carrie Gotzke, Megan Hodge & Jason Daniels**
Department of Speech Pathology and Audiology, University of Alberta, 2-70 Corbett Hall, Edmonton, AB, T6G2G4
(Project of the Canadian Language and Literacy Research Network)

## 1. INTRODUCTION

Assessment of intelligibility is often included as part of the evaluation procedure for children with cleft palate. Therefore, development of a reliable and valid means of assessing intelligibility for this population is essential[1]. Furthermore, it is desirable that an intelligibility measure is sensitive to the particular error patterns of the population for which it is intended[2]. The *Speech Intelligibility Probe for Children with Cleft Palate (SIP-CCLP)* was developed to measure single word intelligibility in children with cleft palate[3] based on these children's typical error patterns. Examples of these include sibilant distortions, sonorant for stop substitutions and glottal place substitutions for other for obstruents. The *SIP-CCLP* was designed to evaluate the effect of speech error patterns of children with cleft palate on their ability to make their spoken messages understandable to listeners. The *SIP-CCLP* has three components: 1) administration and recording of target test words from children, 2) presentation of a child's test word recordings at a later time to adult listeners for open- and closed-set word identification tasks, and 3) analysis of listeners' responses to determine intelligibility scores and error profiles. Preliminary evaluation of the reliability and content and construct validity of the *SIP-CCLP* indicated that it has the potential to be a reliable and valid clinical tool[3]. The purpose of this study was to conduct a preliminary evaluation of the criterion validity of a software version of the *SIP-CCLP* using speech samples from young children with and without cleft palate.

The following questions were addressed:

1. What is the relationship between intelligibility scores obtained from *SIP-CCLP* (open-set response task) and a spontaneous speech sample?
2. What is the relationship between speech error patterns obtained from analysis of the children's *SIP-CCLP* productions and a spontaneous speech sample?

## 2. METHOD

### 2.1 Participants

Audio recordings from eight children with typical speech and language development (TS) and four children with repaired complete clefts of the lip and palate (CP) were judged by 35 graduate students in speech-language pathology. Table 1 provides descriptive characteristics for the children. All children had receptive and expressive language skills within normal limits for their chronological age as determined by standardized language testing. The children with typical speech production also had a normal speech mechanism and normal hearing. The children with cleft palate had hearing within normal limits at the time of data recording. Listener judges had normal hearing and English as a first language.

Table 5. Child participant information.

| Child | Age (Mos.) | | Description of Repaired Cleft |
|---|---|---|---|
| TS1 | 37 | female | n/a |
| TS2 | 43 | female | n/a |
| TS3 | 57 | male | n/a |
| TS4 | 57 | male | n/a |
| TS5 | 60 | male | n/a |
| TS6 | 61 | male | n/a |
| TS7 | 65 | male | n/a |
| TS8 | 82 | male | n/a |
| CP1 | 45 | female | Bilateral cleft lip & palate |
| CP2 | 47 | male | Unilateral cleft lip & palate |
| CP3 | 59 | male | Unilateral cleft lip & palate |
| CP4 | 79 | male | Unilateral cleft lip & palate |

### 2.2 Speech Recordings

Recordings were made using a Panasonic AG-196 video camera and a Sony Electret - 150 lapel microphone. Fifteen minute spontaneous speech samples were collected from each child in a sound booth using parallel and interactive play scenarios following procedures described by Shriberg[4]. Item presentation for the *SIP-CCLP* was randomized to create a unique test order for each child. The computer was set up with the monitor and keyboard inside the sound booth and the hard drive outside the sound booth to minimize background noise. Pictures were presented on a 17-inch screen with 800 X 600 resolution. The child was instructed to repeat the name of the picture displayed after the examiner. Four practice words preceded the presentation of the test stimulus words to ensure that the child understood the task. All 123 stimulus words were presented with the examiner modeling the word(s) that went with each stimulus picture.

Digital audio files of the children's spontaneous speech sample utterances and *SIP-CCLP* productions were made using CSpeech[5] with a sampling rate of 22 kHz and 16 bit quantization. These audio files were then converted to .wav files for playback for the listener identification tasks.

## 2.3 Intelligibility Scores

Spontaneous Speech Sample. The first author prepared an orthographic transcription of the speech sample[4] and then randomly selected a section containing 100 consecutive words and few examiner turns[4] for listener identification. During playback, each of the child's utterances was presented in the order of occurrence in the transcript. The orthographic gloss was used as the key against which the listener judge's responses were compared for scoring. Three judges performed the word identification task independently for each child's sample. Their mean number of words identified correctly served as the child's intelligibility score.

SIP-CCLP. For the open-set response task, the *SIP-CCLP* software presented the child's four practice words followed by the 123 stimulus words for listener identification. Listeners were instructed to type in the word that they perceived the child to say. Three judges performed this word identification task independently for each child's recordings. Their mean number of words identified correctly served as the child's intelligibility score. For the closed-set response task, the *SIP-CCLP* software randomly generated the order of presentation for each target contrast judgment. The listener was presented with four choices for each item: the target word and a minimally contrastive foil (in random order), a box to type in a word that was not one of the first two choices, and a box to indicate "can't identify". A contrast item was scored as correct if at least two of three independent listeners selected the target word.

## 2.4 Phonetic Analyses

The first author prepared phonetic transcripts of the 123 SIP-CLLP stimulus words and 100-word speech samples for each subject using the conventions of Shriberg[4]. Tallies of percent consonants correct in each of the manner categories for stops, fricatives, affricates, nasals, glides and liquids were made for the spontaneous speech sample and SIP-CCLP transcriptions for each child. Inter-rater agreement with a second, independent transcriber for a random selection of 20% of the utterances was 85%.

## 3. RESULTS

The group mean intelligibility score for the spontaneous speech sample was 87.3% (SD=6.1%) for the children with typical speech production and 62.7% (SD=20.1%) for the children with cleft palate. Group mean intelligibility scores for the *SIP-CCLP*, based on the open-set response task, were 77.3% (SD=12.8%) for the children with typical speech production and 50.2% (SD= 20.2%) for the children with cleft palate. A significant positive correlation between intelligibility scores from the *SIP-CCLP* and those obtained from a spontaneous speech sample ($r = .90$, $p < .01$) was obtained for the combined groups (N=12). Group results for the phonetic analyses of the spontaneous speech sample and *SIP-CCLP* productions and for the *SIP-CCLP* closed-set response task are shown in Table 1. No significant differences were found between phonetic transcriptions of the *SIP-CCLP* items and the spontaneous speech sample for percent correct stops, fricatives, affricates, nasals or glides. However, a significant difference was found between the two conditions for liquids ($p < .05$). Also there were no significant differences between error analysis of the *SIP-CCLP* closed-set task and phonetic analysis of the spontaneous speech sample for fricatives, affricates, nasals, liquids and glides but there was a significant difference for stops ($p < .05$).

Table 1. Comparison of sound class results (% correct) for phonetic analyses (PA) for the spontaneous speech sample (SSS) and *SIP-CCLP* production and *SIP-CCLP* closed-set (CS) task.

| | Stops | Fric. | Affric. | Nasals | Glides | Liquids |
|---|---|---|---|---|---|---|
| PA: SSS | 85.5 | 81.8 | 77.8 | 93.2 | 96.5 | 82.2 |
| PA: SIPCCLP | 88.6 | 79.5 | 83.5 | 90.4 | 94.0 | 68.7 |
| CS: SIPCCLP | 92.4 | 83.1 | 92.2 | 95.8 | 98.6 | 93.7 |

## 4. DISCUSSION AND CONCLUSIONS

Initial support for the criterion validity of the *SIP-CCLP* is demonstrated by the strong, significant positive correlation between intelligibility scores on the *SIP-CCLP* and the spontaneous speech sample. Higher intelligibility scores on the spontaneous speech sample were expected, given the positive effect of context on listeners' word identification performance. Evidence for criterion validity is also provided by the findings that: 1) with the exception of liquids, there were no significant differences in percent correct scores for the manner sound classes between phonetic transcription of the speech sample and *SIP-CCLP* productions, and 2) with the exception of stops, there were no significant differences in percent correct scores between the speech sample and error patterns on the *SIP-CCLP* closed-set response task. Possible reasons for differences in liquid and stop scores between the two tasks will be explored further and include transcriber and child factors.

## 5. REFERENCES


1. Gordon-Brannan, M. & Hodson, B. (2000). Intelligibility /severity measurements of prekindergarten children's speech. American Journal of Speech-Language Pathology, 9, 141-150.
2. Kent, R.D., Weismer, G., Kent, J.F., & J.C. Rosenbek. (1989). Toward phonetic intelligibility testing in dysarthria. Journal of Speech and Hearing Disorders, 54, 482-499.
3. Hodge, M., Gotzke, C., Connolly, S., McClure, K., Feltz, C. & O'Hare, J. (May, 2003). Measuring speech intelligibility in young children with cleft palate. Presented at the CASLPA conference, St. John's, NL.
4. Shriberg, L. (1986). Programs to Examine Phonetic and Phonological Evaluation Records (P.E.P.P.E.R.) Version 4.0. Madison, WI: University of Wisconsin-Madison.
5. Milenkovic, P., & C. Read. (1992). CSpeech Version 4.0. Laboratory Version. Madison, WI: University of Wisconsin-Madison.

# A Longitudinal Examination of English Vowel Learning by Mandarin Speakers

**Murray J. Munro[1], Tracey M. Derwing[2], and Ron I. Thomson[3]**
1) Dept. of Linguistics, Simon Fraser University, Burnaby, BC, Canada, V5A 1S6, mjmunro@sfu.ca
2) Dept. of Educational Psychology, University of Alberta, Edmonton, AB, Canada, T6G 2G5,
3) Dept. of Linguistics, University of Alberta, Edmonton, AB, Canada, T6G 2E7

## 1. INTRODUCTION

Current models of second language (L2) speech learning were developed to account for various types of difficulties that L2 learners face in the acquisition of specific vowels and consonants. Flege's (1995) speech learning model (SLM), for example, proposes that accurate production of L2 phones ultimately depends on how well learners are able to distinguish them perceptually from similar L1 sounds. According to Major's (1996) Similarity Rate Differential Hypothesis (SDRH), phones dissimilar to L1 sounds are acquired more rapidly than those that are similar. However, a detailed evaluation of the relative difficulties presented by L2 sounds and of their rate of acquisition requires a wide range of data, including data collected at different times duirng L2 learning. Despite the large number of studies of L2 perception and production conducted in recent years, relatively few have examined the acquisition of categories by low-proficiency speakers from a longitudinal perspective. Most studies with a longitudinal component have considered only a limited range of speech sounds or have considered learners over a short period of time. This study addresses the gap in current research by evaluating a broad range of English vowel productions of recently-arrived Mandarin speakers during their first year in Canada.

Mandarin is usually described as having a system of five vowels : /i y u e a/. The English vowels /I U E/, which are missing from this inventory, are known to pose perceptual and production difficulties for Mandarin speakers (see Wang, 2002). We will pay particular attention to these vowels in the analyses of our results.

## 2. METHOD

### 2. 1 Participants

The participants were 20 native Mandarin speakers registered full-time in low-intermediate ESL classes at a local college. They ranged in age from 26-38 years (M = 33.4). None had been in Canada for longer than 4 months at the outset of the study. All passed a pure tone hearing screen. In their ESL program they received little or no explicit pronunciation training, and none received any training specifically on vowels.

### 2.2 Speaking Task

High-quality mini-disc recordings were made during individual sessions in a quiet room. The speakers heard a recorded voice producing the carrier frame, "The next word is ___" and responded by saying, "Now I say _____," completing the sentence with the stimulus items. The target words were /pVt/ and /bVt/ words, and nonsense syllables containing the vowels /i I e E Q u U o A ?/. Each speaker was recorded six times (T1 to T6) at intervals of approximately 8 weeks. Of 2400 possible items, 7 were missing because speakers occasionally failed to repeat a particular item or because of recording or editing errors.

### 2.3 Evaluation

The first two authors evaluated the remaining 2393 tokens in a forced choice randomized listening task in which they heard each word through headphones and identified the closest Canadian English vowel by pressing a labeled button on a computer screen. Items could be replayed up to four times, and a "none" category was available for items that appeared to match no native English category. Multiple blocks were presented over a period of days to reduce fatigue. The listeners had no idea of the time or the speaker of any particular item.

## 3. RESULTS

For each speaker, the total number of correct productions was tallied at each time. Means across all 20 speakers are shown in Figure 1.

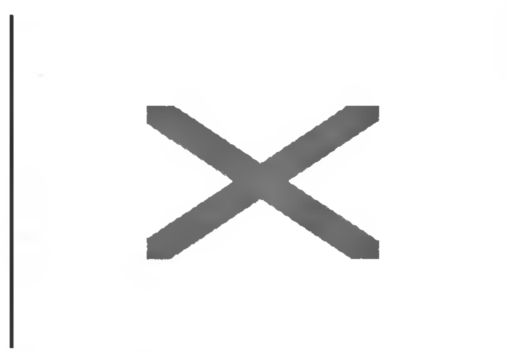


**Fig. 1. Mean correct ID scores (%) pooled over speakers and vowels at each of the eight-week intervals.**

A one-way repeated measures ANOVA on the ID scores with Time (6 levels) as the within-speakers factor yielded a significant effect of Time, $F(5, 95) = 7.66$, $p<.001$. A series of post hoc pairwise comparisons were computed, in which the scores at T1 were compared with the scores at each other time, with $p<.01$ (Bonferroni-adjusted) as the criterion for significance. The mean scores at T4, T5, and T6 were significantly higher than the score at T1, $t(19) = 3.56$, 5.26, and 4.00, respectively. The mean scores at T1 and T2 did not differ significantly; nor did the scores at T1 and T3.

Table 1 shows mean %ID scores on each vowel at T1 and T6, as well as net changes in identifications on each vowel. Despite a general trend toward more accurate productions from T1 to T6, performance on the individual vowels varied considerably.

**Table 1. Scores on individual vowels (%)**

| Vowel | T1 | T6 | Net Change |
|---|---|---|---|
| A | 41 | 71 | 30 |
| I | 19 | 47 | 28 |
| e | 65 | 90 | 25 |
| Q | 56 | 78 | 22 |
| u | 61 | 79 | 18 |
| ? | 62 | 70 | 8 |
| o | 78 | 84 | 6 |
| i | 94 | 94 | 0 |
| U | 63 | 63 | 0 |
| E | 68 | 55 | -13 |

The most dramatic improvements were seen with /A I e/. Of these, /I/, which has no counterpart in Mandarin, was the least accurately-produced vowel at T1 and, despite the improvement, remained so at T6. Throughout the study, incorrect productions of this vowel were overwhelmingly identified as tokens of /E/. Performance on /i U/ remained unchanged, but for /i/, performance was already at near-ceiling levels at T1. Although /U/, the back counterpart of /I/, also does not occur in Mandarin, performance on /U/ appeared to be considerably better overall. The one vowel on which the speakers appeared to show an overall decline in performance was /E/.

## 4.0 DISCUSSION

These results raise a number of interesting issues concerning Mandarin speakers' learning of English vowels over a 1-year period. While it is possible that some of the overall improvement observed here was due to practice or to speaker familiarity with the task, there are at least two reasons for doubting that such factors had a major effect on the outcome: (1) Such factors would most likely be strongest during the early weeks of the study, not in the later weeks, as the statistical analyses showed, and (2) the effects of such factors should be fairly uniform across all vowels, rather than differential as observed here.

The vowel /I/, at least in the contexts used here, seemed to pose particular difficulty for the learners. Although their performance on this vowel improved immensely, it was the least accurately-produced vowel throughout the study. Further work is needed to determine whether the lack of a parallel pattern of difficulty and improvement for /U/ was due to different patterns of perceptual assimilation for the two vowels.

The paradoxical decrease in the speakers' scores on /E/ is also worthy of further examination. Cebrian (2002) concluded that L2 vowel acquisition must be considered in a systemic way, rather than exclusively in terms of individual category acquisition. It should be noted that /E/ is regarded as an allophonic variant of Mandarin /e/. At the outset of the study the speakers may have been able to take advantage of their knowledge of Mandarin vowel categories in order to produce fairly accurate tokens of this vowel. Although we cannot be certain of the reasons for the decline in performance on /E/, it is interesting to note that this decline coincided with improvement on two nearby front vowels (/I/ and /Q/). This suggests that, over the course of the study, the speakers may have been adjusting their perceptual representations of front vowel categories to accommodate the new vowels.

## REFERENCES


Cebrian, J. (2002). *Phonetic similarity, syllabification, and phonotactic constraints in the acquisition of a second language contrast.* Unpublished PhD dissertation, Dept. of Linguistics, University of Toronto.

Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues.* Timonium, MD: York Press.

Major, R. C., & Kim, E. (1996). The similarity differential rate hypothesis. *Language Learning, 46*, 465-496.

Wang, X. (2002). *Training Mandarin and Cantonese speakers to identify English vowel contrasts: Long-term retention and effects on production.* Unpublished PhD dissertation, Dept. of Linguistics, Simon Fraser University.


## ACKNOWLEDGEMENTS


This research was supported by the Social Sciences and Humanities Research Council of Canada. The authors thank H. Li, R. McCallum, S. Morton, and N. Penner for their invaluable assistance. We also appreciate the help of the students and staff of Norquest College in Edmonton.

# An Acoustic Study Of An Uncommon Vowel Produced By Hungarian-English Bilinguals

**Zita McRobbie-Utasi**

Department of Linguistics , Simon Fraser University, British Columbia, Canada, V5A 1S6. mcrobbie@sfu.ca

## 1. INTRODUCTION

One of the uncommon vowels in languages is the low back rounded vowel [Å ]. In Maddieson (1984), only five of the 918 languages surveyed are identified as containing this vowel in their sound system. Because of its phonetic characteristics, the articulation of the [Å ] sound is rather complex. This vowel represents one of the special features of Hungarian speech. It is difficult for foreigners to produce, and it appears to be vulnerable in a languages-in-contact situation. Of the several "markers" that distinguishes immigrant Hungarian speech from that of Hungarian newcomers, the pronunciation of this vowel definitely qualifies as one of those markers. This fact contradicts earlier research asserting that the pronunciation of first generation speakers is completely Hungarian (Kálmán, 1973).

The hypothesis proposed here is based on universal tendencies. Accordingly, it was expected that the direction of pronunciation change is toward the centre of the vowel continuum, i.e., approximating the pronunciation of the vowel [a]. In the discussion that follows it will be argued that this hypothesis may only be partially upheld.

## 2. METHOD

Twenty-four speakers (twelve female and twelve male) participated in the experiment discussed below. Out of the twelve female speakers, two came to Canada between 1946 and 1948; four between 1956 and 1958; all except two (who were under the age of six in 1956) were adults at the time of arrival. All six speakers have worked in an English-speaking environment and had only minimal contact with members of the Hungarian community. The other six female speakers have been closely involved with the Hungarian community. Two of these subjects arrived in Canada in 1957, two in the 70s (1972 and 1975 respectively) and two in the early 80s (in 1981 and 1983 respectively).

The twelve male speakers were also selected on the basis of their association with the Hungarian community. Six of the subjects have worked in an English speaking environment, while the other six have been closely associated with the Hungarian community. The first group of male speakers arrived in Canada in 1957 (four, one of them was under the age of six at the time of arrival) and 1958 (two speakers. Of the six immigrants who were in close contact with the Hungarian community, two arrived in the early 70s (in 1973 and 1974), and four in the late 70s or early 80s (1978, 1982 and 1983 respectively).

Recordings were made in in the Phonetics Laboratory at Simon Fraser University. The subjects read twenty words (containing the vowel under examination) placed in a sentence frame presented to them on randomly sequenced filing cards a total of three times The control tape was prepared by recording of one female and one male speaker (both visitors from Hungary) immediately after their arrival.

First and second formant frequency measurements were obtained (Praat program, Version 3.9.13), in order to identify the quality of the vowel produced by the speakers.

## 3. RESULTS AND DISCUSSION

The acoustic analysis aimed at quantifying the two parameters — tongue height and tongue backing – resulted in acknowledging two separate tendencies in relation to the degree of the speakers' involvement with the Hungarian community. These tendencies are clearly identifiable for female and male speakers alike. The first tendency is observed in connection with those subjects who had minimal or no contact with the Hungarian community, who pronounce the vowel [Å] close to that of the control speakers. On the other hand, the second tendency is related to speakers with strong involvement with the Hungarian community, who appear to produce this vowel with higher tongue position than do the control speakers. Further, those subjects who were children at the time of their arrival in Canada pronounce this vowel with tongue advancement, thus closely approximating the tongue position for the vowel [a]. Figures 1 and 2 present the first and second formant frequency values with respect to those of the control speakers. The arrow points to the formant frequency values obtained from the SB (control) speaker.

First formant frequency (mean) values obtained for female speakers AH'57, AH'70s and AH'80 reveal a tendency to articulate the vowel [å] with a higher tongue position than that of the control speaker. The greatest divergence from the control speaker's pronunciation (F1 = 845 Hz) with regard to tongue height is evident with the A'70 speakers (680 Hz). The three early immigrants appear to be closest to the control pronunciation values (865 Hz, 850 Hz and 840 Hz respectively).

With regard to the measurement (mean) values obtained for the second formant, it can be clearly seen that the early immigrant adult speakers have a comparable tongue position with that of the control speaker : 1200 Hz (SB), 1234 Hz (A'46-'48), 1241 Hz (A'56-'58). The lowest frequency value (1100 Hz) is associated with the two speakers who immigrated in the 1970s and have since been closely associated with the Hungarian community (A'70). The high second formant value (1392 Hz) observed for the two speakers who immigrated to Canada as children reveals a pronunciation approaching that of the vowel [a].

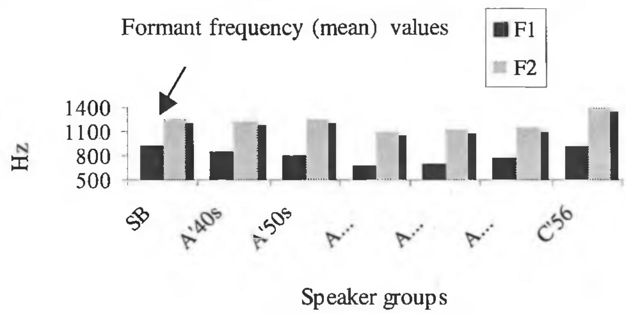



**Fig. 1. Mean formant frequency values (female speakers). SB=speaker from Budapest, A=adult speaker, C=child speaker; the dates indicate the year of immigration.**

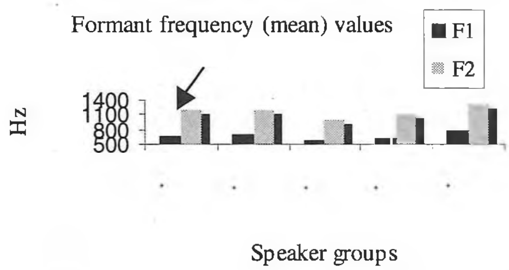



**Fig. 2. Mean formant frequency values (male).**

First formant measurements (mean values) obtained for male speakers indicate a higher tongue position for those in the AH'70s and AH'80s groups relative to the control speaker (SB = 670 Hz; AH'70 = 580 Hz, AH'80s = 610 Hz). Speakers in the A'57-'58 group have a comparable tongue height with that of the control speaker (700 Hz).

Measurement (mean) values obtained for the second formant for those in the A'57-'58 group show a tongue position similar to that of the control speaker (SB = 1185 Hz, A'57-'58 = 1200 Hz). Speakers belonging to the AH'70s and AH'80s group have 997 Hz and 1100 Hz second formant frequency mean values respectively. The measurements (1326 Hz) obtained for the speaker who came to Canada as a child indicate a more advanced tongue position.

The results presented above reveal the following tendencies, valid for both (female and male) speaker groups : (i) the divergence in tongue height is greatest in the pronunciation of the AH'57, AH'70s and AH'80s (female) and the AH'70s and AH'80s (male)speakers, (ii) the divergence in tongue backing is greatest in the pronunciation of the of the C'56 (female) and C'57 (male) speakers, and (iii) the pronunciation of the immigrant speakers having minimal contact with members of the Hungarian community diverges the least from the F1 and F2 averages obtained from the SB speakers.

## CONCLUSIONS

The implications of the experiment reported on here only partially confirm the first hypothesis, i.e., the direction of the pronunciation change is toward the centre of the vowel continuum. This hypothesis may only be upheld for second generation speakers.

With regard to the direction of the sound change and the speaker group leading the change this study has sustained the conclusions reached in the first part of the project (McRobbie-Utasi, 2001).

## REFERENCES


Kálmán, B. (1973). Az amerikai magyarok nyelve *Zborník Filozofickej Fakulty Univerzity Komenského. Philologica* XXV (pp. 199-205).

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

McRobbi—Utasi, Z. (2001). Pronunciation changes in the speech of Hungarian bilinguals. Presented at the *Canadian Association of Hungarian Studies Conference. Québec City.*

# L2 Vowel Perception: Perceptual Assimilation To What?

**Ron I. Thomson**
Dept. of Linguistics, University of Alberta, Edmonton, AB, Canada, T6G 2E7, rit@ualberta.ca

## 1. INTRODUCTION

One of the most influential models of second language speech perception is Flege's (1995) Speech Learning Model (SLM). He argues that during initial stages of second language acquisition, second language (L2) phonological categories that share the same acoustic space as pre-existing first language (L1) categories will map onto those categories. In addition, Flege and Hillenbrand (1987) claim that acquiring a phonetic category that is similar to an L1 category can result in averaging the distance between the L1 and L2 categorical centres.

In addition, a distinction is made in the degree of difficulty learners face in acquiring specific L2 categories (Flege, 1995). Those L2 categories that share perceptual space with an existing L1 category are easiest to learn; those that are distinct from L1 categories or ambiguous are more difficult to learn. While comparisons of L2 productions with native-speaker means have been made (Chen, Robb, Gilbert & Lerman, 2001) few studies have compared L2 productions with productions of similar L1 categories.

This study examines the productions of English /u/ and /U/ categories by Mandarin L1 learners of English. Maddieson (1984) describes the Mandarin vowel inventory as including /u/ but not /U/. SLM would predict that English /u/ will map to an existing Mandarin /u/ category, while /U/, because it does not exist in Mandarin, will require the formation of a new category.

## 2. METHOD

The data used in this study were collected as part of a larger longitudinal study of the development of English language proficiency by recent immigrants to Canada.

### 2.1 Participants

Eight female and two male Mandarin L1 learners of English were selected on the basis of being newcomers to Canada with low English language proficiency. All participants were enrolled in a full-time ESL program at a local college. They ranged in age from 26-39 years.

### 2.2 Data Collection

Recordings of each participant's English productions were made six times over the course of one year, using a minidisc recorder with a sampling rate of 44,100 Hz. Participants were asked to listen to a native speaker's recorded rendition of the target vowels in a /pVt/ frame presented in the carrier phrase "The next word is ____." They had to respond by saying, "Now I say ____." In total, 60 renditions of each vowel stimulus were obtained. To obtain productions of /u/ in a Mandarin context, participants read from a list of ten disyllabic Mandarin words containing the target vowel preceded by an initial voiceless bilabial.

### 2.3 Data Analysis

The author selected the 30 responses to English stimuli that were judged to be closest to the target L2 category for analysis. The Mandarin productions were also checked to insure errors were not made in reading the target words. Using the program Praat, 50 ms sections from the centre of the steady state portion of each English and Mandarin vowel production were selected. Measures of F1 and F2 frequencies were calculated.

All values of F1 and F2 were normalized to the average female values. These were then compared to published female native English speaker productions taken from Hillenbrand, Getty, Clark and Wheeler (1995).

## 3. RESULTS

Mean F1 and F2 values across speakers and their range are provided in Table 1. Production of English /u/ was not significantly different across native speakers (NS) and Mandarin L1 non-native speakers (NNS). There is a significant difference on the F2 values for the English /U/ category.

Table. 1. F1/F2 values of English /u/ and /U/ and Mandarin /u/ and /u/ categories by Mandarin learners of English.

| Target vowel | Speaker L1 | F1 | Range | F2 | Range |
|---|---|---|---|---|---|
| English /u/ | Mandarin | 431 | 371-518 | 1085 | 957-1213 |
| | English | 459 | n/a | 1105 | n/a |
| English /U/ | Mandarin | 531 | 466-607 | 1061 | 948-1279 |
| | English | 519 | n/a | 1225 | n/a |
| Mandarin /u/ | | 354 | 295-416 | 885 | 712-1086 |

As predicted, Mandarin L1 learners of English the English /u/, which is perceptually similar to a Mandarin category. Of greatest interest is that the mean and range of the F1/F2 values for the Mandarin /u/ category is clearly different from the same speakers' productions of English /u/.

Figure 1 below plots the Mandarin speaker productions in a two dimensional space.

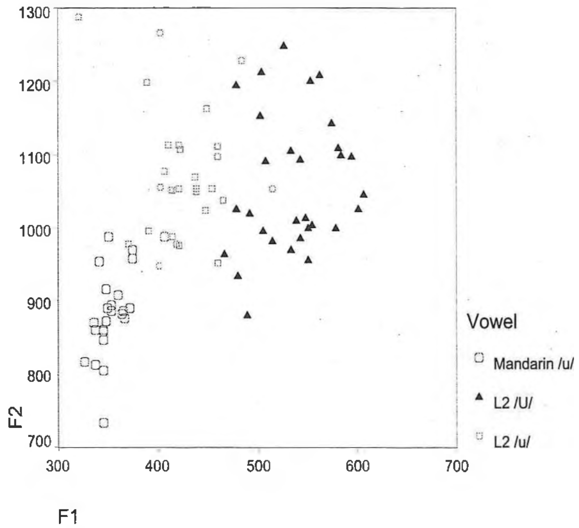



Fig. 1. F1/F2 values of English /u/ and /ʊ/ and Mandarin /u/

Figure 2 illustrates the difference in category centres between NS and NNS productions, comparing these with the centre of the Mandarin /u/ category.

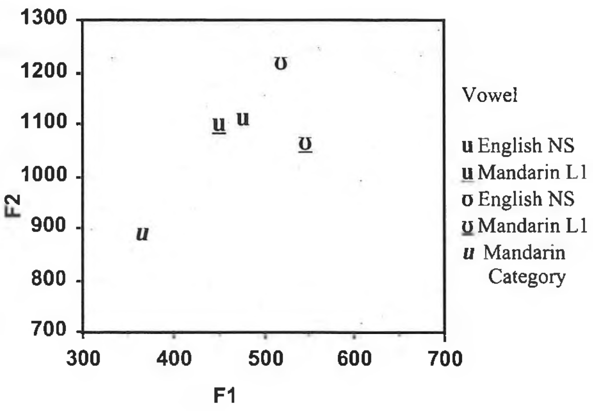



Fig. 2. Mean F1/F2 values of NS and Mandarin L1 production of English /u/ and /ʊ/ and mean of Mandarin /u/.

## 4. DISCUSSION

What is most striking about these results is the difference between the acoustic properties of Mandarin /u/ following a bilabial compared to those produced by the same speakers in response to English /u/ stimuli. Since the means and ranges are different, it is clear that in producing an English /u/, Mandarin speakers do not simply map to an exact replica of the prototypical Mandarin category, as might have been predicted. Rather, it seems they are sensitive to crosslinguistic differences and are developing an English-specific representation.

The small overlap between the upper edges of the Mandarin /u/ productions and the same speakers' English L2 productions suggests that if any categorical mapping is taking place, it is to non-prototypical exemplars of the L1 category; that is, to allophones of the Mandarin category that are closest to the English equivalent. This may explain why acquiring an L2 phonetic category that is similar to an L1 category often results in averaging the distance between the L1 and L2 categorical centres. As the learner is exposed to L2 vowels that are non-prototypical members of the L1 category, he/she still recognizes them as belonging to the L1 category. The frequency of the rare allophones at the L1 category's edge is strengthened by this L2 input. To truly acquire a native-like representation of the English /u/ vowel, however, Mandarin L1 learners need to develop greater sensitivity to differences between similar L1 and L2 categories by noticing L2 exemplars that are less similar to their L1 counterparts. In the case of the English learners in this study, a sensitivity to English /u/ seems evident. The greater difficulty associated with acquiring English /ʊ/ may then be a result of its similarity to English /u/ as much as its closeness to the Mandarin /u/ category.

## REFERENCES


Flege, J.E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.) *Speech perception in linguistic experience: Theoretical and methodological issues*. pp. 229-273. Timonium, MD: York Press.

Flege, J.E. and Hillenbrand, J. (1987) Limits on phonetic accuracy in foreign language speech production. In G. Ioup and S. Weinberger. (Eds.). *Interlanguage Phonology: The acquisition of a second language sound system.* Cambridge: Newbury House. pp. 176-203.

Hillenbrand, J., Getty, L., Clark, M. and Wheeler, L., (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97,* 3099-3111.

Maddieson, I. (1984). *Patterns of Sounds.* Cambridge: Cambridge University Press.

Chen, Y., Robb, M., Gilbert, H. & Lerman, J. (2001). Vowel production by Mandarin speakers of English. *Clinical Linguistics and Phonetics, 15,* 427-440.



## ACKNOWLEDGEMENTS

The author thanks Tracey M. Derwing and M. J. Munro for providing access to speech data used in this study and for feedback on the analysis. R. Kirchner also provided invaluable assistance.

# Fuzzy String Kernel Representations In Speech Processing

**Robert Kirchner**
Linguistics Dept., University of Alberta, Edmonton, AB, T6G2E1, kirchner@ualberta.ca

Many widely used approaches to pattern recognition/machine learning, including neural nets, k-nearest-neighbours classifiers, and support vector machines, have hitherto made little headway in speech processing, largely due to their inability to represent and compute over the variable-length sequential data of speech signals. A new technique, the string (subsequence) kernel, first applied in bioinformatics [1] and text classification [2], and extended to speech recognition by Goddard et al. [3], maps a variable-length input signal to a fixed-length feature array, by taking the inner product of n-gram subsequences. Similarity of signals can then be evaluated, by any of the above approaches, in the kernel space. In this presentation, two variations on Goddard's approach are considered and evaluated: a string kernel using fuzzy rather than absolute k-means clustering; and a kernel in which the feature counts are preserved as waveforms rather than scalars, to address the reverse mapping problem.

## 1. BASIC TECHNIQUE

Goddard vector quantizes a cepstral representation of the speech signal to a string of prototypes by k-means clustering. Assuming for illustrative purposes a mere three prototypes, all possible 2-grams are represented in a 3×3 matrix. The string {21313}, for example, contains the 2-grams {21} and {13}, and {31}. The corresponding cells of the matrix receive a value of 1 for each occurrence in the string ({13} occurs twice). Non-contiguous subsequences (e.g. {2...3}) are counted as well – allowing for detection of similarity notwithstanding interruption in the string – albeit with a value that exponentially decays with the distance between the elements. Multiple values within a cell are summed; and 0 is assigned to all other cells. The string thus maps to the kernel representation in Figure 1.

| 1st / 2nd | 1 | 2 | 3 |
|---|---|---|---|
| 1 | 0.3679 *=exp(-1)* | 0 | 2.1353 *=1+1+exp(-2)* |
| 2 | 1 | 0 | 0.4177 *=exp(-1)+exp(-3)* |
| 3 | 1 | 0 | 0.3679 *=exp(-1)* |

**Figure 1.** 2-gram string kernel representation of the prototype string [21331].

In sum, this technique proceeds from the intuition that *two strings are similar to the extent that they contain the same subsequences*. Thus, the string kernel for {213} will have similar values to {21313}, even though the original strings are of different lengths, as it contains many of the same subsequences; while {22222} will have radically different values. This approach can be extended to n-grams simply by assigning values to an n-dimensional array rather than a matrix. Goddard reports that on a multi-voice (Spanish) digit recognition task, with 32 prototypes, an SVM classifier using this kernel outperforms a discrete HMM.

## 2. FUZZY K-MEANS CLUSTERING

A possible alternative to Goddard's vector quantization step would be to construct multiple string kernels for a given input, one for each frequency (or quefrency) channel. This approach proved unworkable, presumably due to its failure to detect patterns relating subsequences in one channel to simultaneous events in other channels. Vector quantization, by comparison, provides a useful 'gestalt' of each frame. On the other hand, vector quantization loses intra-prototype differences in the signal. Furthermore, frame A may be somewhat similar to frame B, but highly dissimilar to frame C; but under vector quantization, unless A and B happen to fall within the same prototype, the greater distance AC vs. AB is lost. It was hypothesized that recognition in the kernel space would improve, using fuzzy k-means clustering. In this fuzzy kernel approach, each frame receives fuzzy membership threshold scores for all the prototypes, and these strings of prototype scores are then mapped to 2-gram kernel space. That is, the kernel consists of multiple channels, one for each prototype, encoding counts of 2-grams such as 'membership > 0.8 precedes membership > 0.1 in prototype 6.' The resulting representation provides a gestalt of the frame; but because it encodes fuzzy membership of frames to multiple prototypes, this kernel is less lossy.

To test this hypothesis, recordings were made of 6 English speakers, 4 male and 2 female, saying 100-200 tokens each, in random order, of the English digits, for a total of 1000 tokens. In Matlab, the sound files were segmented into isolated digits, and converted into spectral and cepstral form with the RastaMat toolbox [4], an implementation of RASTA-PLP filtering [5]. Cepstral representations were initially used, as in [3]; but as this method achieved recognition rates below 30% on the training data, cepstra were abandoned in favour of spectral frames. Absolute and fuzzy k-means clustering (with 32 prototypes) and mapping to string kernel form were each implemented in Matlab [6]. 66% of the data were used for training, and 33% as test data. Similarity between kernels was measured as the negative exponential of their Euclidean distance, and a k-nearest-neighbour classifier was applied to these the similarity scores from these kernel representations.

Contrary to the hypothesis, Goddard's string kernel with absolute k-means clustering outperforms the fuzzy kernel (Table 1).

| | 2-gram | 3-gram |
|---|---|---|
| Goddard's kernel | 83 | 91 |
| Fuzzy kernel | 74 | 79 |

**Table 1.** Percent correct on knn classification (k=8) for test set

A possible explanation for this result lies in the ubiquity of the features on which the kernel is based, the fuzzy prototype membership threshold scores. In attempting to extract similarities or differences beyond prototype labels, the fuzzy kernel method, it seems, gets swamped with slight similarities, to the point that critical differences are obscured. This result is somewhat disturbing, as it suggests that the relatively poorly understood vector quantization process is not merely a convenient way of reducing the dimensionality of the data, but a crucial part of getting patterns to emerge from the data under this approach. Further research is required to determine whether some variant use of fuzzy clustering in string kernels, resulting in less ubiquitous features, might yield higher recognition scores.

## 3. REVERSE MAPPING WITH WAVEFORM KERNELS

While Goddard's string kernel appears to work well for perceptual classification/recognition, to develop this approach into a general model of speech processing, i.e. including synthesis, it must be possible to reconstruct the original (prototype string) representation from the kernel. While this issue is perhaps not of immediate concern to engineers in the speech recognition industry, it is of concern to phonological theory, as it has been suggested in a growing body of research, e.g. [7-11], that a quasi-exemplar-based model of speech processing affords an elegant account of a range of lexical frequency effects in phonological and grammatical patterning, as well as an explicit learning algorithm whereby patterns of phonetic variation become entrenched as phonological constraints. Unfortunately, the Goddard kernel is not reversible (John Goddard, p.c.). A single occurrence of a contiguous subsequence cannot be distinguished from multiple occurrences of a non-contiguous subsequence, once the counts have been summed.

A technique for preserving the component counts notwithstanding summation is to treat them not as scalars, but as amplitudes of a complex waveform. Specifically, for each subsequence {AB} where A occurs in frame $f$ in the vector quantized representation, a fixed-length sine wave of frequency $f$ is created. The amplitude of this wave is the count of the (contiguous and non-contiguous) occurrences of this subsequence in the vector quantized representation. Now, if there are multiple occurrences of {AB}, with A occurring in several different frames, the corresponding waves can be summed to a single complex waveform for cell {AB}; the affiliation of A to its original timeframe(s) is retained in the component frequencies of the waveform. To map the string kernel back to the vector quantized representation, one merely has to identify these frequencies, by Fourier analysis, and to assign each prototype with a particular frequency to the corresponding frame. An approximation of the original spectrogram can then be obtained by replacing each prototype label with the spectral values of the corresponding prototype centre, as computed during the vector quantization process.

To test whether the representation of subsequence counts in waveforms rather than scalars reduces accuracy in recognition, waveform kernels for the training and test data were constructed, as described above. For reasons of computer memory limitations, only 2-gram kernels were evaluated. Results of knn classification, compared to those of Goddard's kernel, are not yet available as of the date of submission of this summary paper. However, preliminary results on the training data suggest that the waveform kernel representation achieves recognition rates at least as high as those of Goddard's kernel.

## References


[1] C. Watkins, Kernels from matching operations, CSD-TR-98-07, Royal Holloway, U London, 1999

[2] H. Lodhi, C. Saunders, J. Shawe-Taylor, N. Cristianini, Text classification using string kernels, JMLR 2:419-444, 2002

[3] J. Goddard, A. Martinez, F. Martinez, H. Rufiner, A comparison of string kernels and discrete hidden Markov models on a Spanish digit recognition task, JASA 112:5, pt. 2, 2002

[4] Dan Ellis, RastaMat Toolbox, http://labrosa.ee.columbia.edu/matlab/rastamat.

[5] H. Hermansky, N. Morgan, RASTA processing of speech, IEEE Transactions Speech & Audio Proc, 2:4, 578-589, 1994

[6] Robert Kirchner, Exemplar Toolbox, http://www.ualberta.ca/~kirchner/.

[7] K. Johnson, Speech perception without speaker normalization, in Talker Variability, Academic Press, 1997.

[8] R. Kirchner, Preliminary thoughts on phonologization within an exemplar-based speech processing system, UCLAWPL 6, 1999.

[9] J. Bybee, Phonology and language use, Cambridge U. Press, 2001.

[10] J. Pierrehumbert, Word-specific phonetics, Laboratory Phonology VII, 101-140, Mouton de Gruyter, 2002.

[11] R. Kirchner, C. Fraser, Modelling phonologization within an exemplar-based lexicon, poster, 3rd Intl. Conf. Mental Lexicon, Banff, Oct 7, 2002.

# Obtaining The Vocal-Tract Area Function From The Vowel Sound

**Huiqun Deng, Michael P. Beddoes, Rabab K. Ward, Murray Hodgson***
Electrical and Computer Engineering Department, *Mechanical Engineering Department
University of British Columbia, 2356 Main Mall, Vancouver, BC, Canada V6T 1Z4
huid@ece.ubc.ca, mikeb@ece.ubc.ca, rababw@icics.ubc.ca, hodgson@mech.ubc.ca

## 1. INTRODUCTION

Vocal-tract area functions (VTAFs) are needed in speech synthesis, speech recognition, and the detection of the vocal-tract shape. The vocal-tract area function can be measured using X-ray or MRI methods. But, both methods are time-consuming and not convenient. It has long been desired to obtain the vocal-tract area function from the speech signal.

It is shown that the VTAF can be derived from the vocal-tract filter (VTF) assuming the wall of the vocal tract is rigid, the length of the vocal tract is known, and the lip or the glottal reflection coefficient is 1 [1,2]. In Atal's method for deriving the VTAF from the VTF, the vocal tract is assumed to completely close at the glottis, and to be terminated with characteristic impedance at the lip opening [1]. In Wakita's method, the vocal tract is assumed to be terminated with characteristic acoustic impedance at the glottal end, and with zero acoustic impedance at the lip opening [2]. However, these assumptions about the boundary conditions cannot be satisfied all the times. The glottal reflection coefficient is time varying, because the glottis opens and closes periodically during voicing. The lip radiation impedance can only be approximated as zero at low frequencies, and can be characteristic impedance only when the lip opening is connected with a reflectionless tube.

Accurate estimation of the VTAF requires the VTF estimated from a vowel signal should not contain the influence of the glottal wave, and the influence of the non-ideal glottal and lip boundary conditions. The method for eliminating the influence of the glottal wave on the VTF estimation from a vowel sound signal is developed in [3]. In this paper, we investigate the effect of non-ideal glottal and lip boundary conditions on the estimation of the VTAF.

## 2. THE VOCAL-TRACT FILTER

The acoustic effect of the vocal tract can be modeled using a multi-sectional cylindrical tube, with each section having the same length and different cross-sectional area (Fig.1). The signal flow diagram from the glottal wave $U_g$ to the lip volume velocity $U_{lip}$ can be represented in terms reflection coefficient $r_i$ and the delay D in each section, $r_g$ (the glottal reflection coefficient), and $r_{lip}$ (the lip

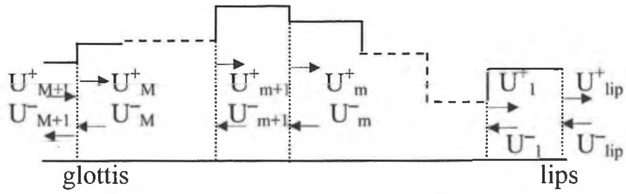



Fig. 1. The acoustic tube model of the vocal tract.

reflection coefficient), as shown in Fig. 2 [3], where $r_{lip}$ is:

$$r_{lip} = (\rho c / A_1 - Z_{lip}) / (\rho c / A_1 + Z_{lip}) \quad (1)$$

where $Z_{lip}$ is the lip radiation impedance, $A_1$ is the cross-sectional area of section 1, and $\rho c$ is the characteristic impedance of the air. $r_g$ is defined as:

$$r_g = (Z_g - \rho c / A_M) / (Z_g + \rho c / A_M) \quad (2)$$

where $Z_g$ is the acoustic glottal impedance, $A_M$ is the cross-sectional area of section M,

The transfer function from the glottal wave to the lip volume velocity is an all-pole filter, which is denoted as $TF_{GL}$. It is time varying due to the time varying $r_g$. The all-pole filter estimated from speech signals using LPC [1] is usually called as vocal-tract filter. But, it is actually an averaged version of the $TF_{GL}$, and is some different from what is required in the estimation of the VTAF. In Atal's method, the VTF used for estimating the VTAF is defined to be the complex ratio of the total volume velocity at the lips to the total volume velocity at the backend of the vocal tract, i.e., $U_{lip}/(U^+_M+U^-_M)$. The $TF_{GL}$ is identical to the VTF in Atal's method, only if rg=1. In Wakita's method, the VTF is defined to be the complex ratio of the total volume velocity at the lips to the volume velocity entering the glottis from the trachea, i.e., $U_{lip}/U^+_{M+1}$, assuming $Z_{lip}=0$, i.e., $r_{lip}=1$. The $TF_{GL}$ is identical to the VTF in Wakita's method, only if $r_{lip}=1$.

## 3. VOCAL-TRACT BOUNDARY CONDITIONS AND VOCAL-TRACT AREA FUNCTION ESTIMATION

In order to see the effect of non-ideal $r_g$ and $r_{lip}$ contained in the $TF_{GL}$ on the estimation of the VTAF, we synthesize the $TF_{GL}$ for a given VTAF, and different $r_{lip}$'s and $r_g$'s, and use the synthetic $TF_{GL}$ to estimate the VTAF. The difference between the estimate and the given VTAF is

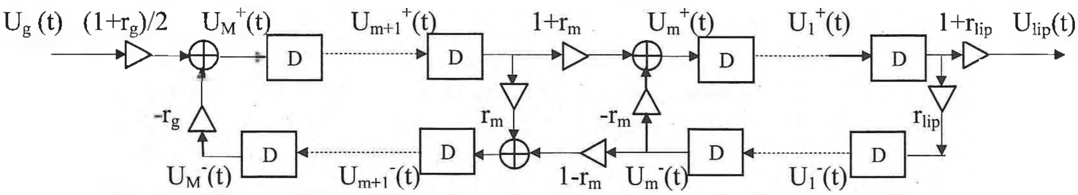



Fig. 2. The signal flow diagram from the glottal wave to the lip volume velocity.

the effect of the non-ideal $r_g$ and $r_{lip}$. The VTAF of /ɑ/ measured using the magnetic resonance imaging method [4] is used in synthesizing the impulse response of the $TF_{GL}$. From the synthetic impulse response, we estimate the VTAF under different assumptions about the vocal-tract boundary conditions. If the vocal-tract boundary conditions are assumed to be $r_g=1$ and $r_{lip}$ is arbitrary, the vocal-tract reflection coefficients are estimated using method 1:

$$r_m = - k_{m-1} \qquad m=1,2,\ldots\ldots,M-1 \qquad \text{(method 1)}$$

where $r_m=(S_m-S_{m+1})/(S_m+S_{m+1})$, $S_m$ is the cross-sectional area of section m, m increases from the glottis to the lips, $k_m$ is defined in Eq. (11) in [2]. This method is derived through matching Atal's acoustic filtering process [1] to Wakita's mathematical filtering process of the VTF [2]. If the vocal-tract boundary conditions are assumed to be $r_{lip}=1$ and $r_g$ is arbitrary, the vocal-tract reflection coefficients are estimated using method 2:

$$r_m = k_{m-1} \qquad m=1,2,\ldots\ldots,M-1 \qquad \text{(method 2 [2])}$$

where m increases from the lips to the glottis [2]. Let $S_1=1$, the normalized cross-sectional areas, which form the VTAF, are then obtained:

$$S_{m+1}=S_m(1- r_m)/(1+ r_m) \qquad m=1,2,\ldots\ldots, M-1 \qquad (3)$$

The frequency responses of the synthetic $TF_{GL}$'s and the VTAF estimates derived from the synthetic $TF_{GL}$'s using methods 1 and 2 for /ɑ/ are shown in Fig. 3. The bandwidths of the resonance of the VTF are damped, which represent the energy loss in the $TG_{GL}$, if $r_g$ or $r_{lip}$ is small. The influence of the glottal loss and lip loss on the estimation of the VTAF can be seen comparing the estimates with the given VTAF used for synthesizing the VTF. Our results show that the VTAF can be recovered from the $TF_{GL}$ using method 1, if the $TF_{GL}$ corresponds to $r_g=1$; or, using method 2, if the $TF_{GL}$ corresponds to $r_{lip}=1$. Method 1 works well only if the glottal reflection coefficient is one, not being affected by different lip reflections. Method 2 works well only if the lip reflection coefficient is one, not being affected by different glottal reflections.

## 4. DISCUSSION

Although method 2 is not sensitive to $r_g$, it is sensitive to $r_{lip}$. $r_{lip}=1$ (i.e. $Z_{lip}=0$) is true for ka<<1, where $k=2\pi f/c$, a is the lip opening radius. Therefore, only low sampling rate ($F_s$< 7 kHz) should be used in method 2. The sampling rate $F_s$ determines the number of sections of the tube model: $M=2LF_s/c$, where L is the length of the vocal tract, and c is the sound speed. Thus, method 2 cannot obtain detailed structures of the VTAF.

Method 1 is not subject to $r_{lip}$. Therefore, it can work well over wide frequency range, and can allow high sampling rate. Thus, method 1 can obtain more detailed structures of the VTAF than method 2. For more accurate estimation of the VTAF, the $TF_{GL}$ used in method 1 should be estimated from the speech signal recorded during closed phases of the glottis, and the speech signal should be recorded in a reflectionless tube connected to the lip opening.

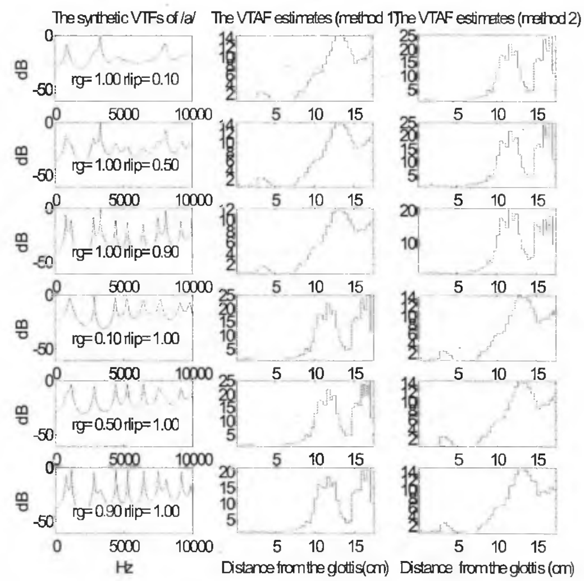



Fig. 3. The synthetic $TF_{GL}$'s and the estimates of the VTAF.

## REFERENCES

[1] Atal, B. S. and L. Hanauer, "Speech Analysis and Synthesis by Linear Prediction of the Speech Wave", *J. Acoust. Soc. Amer., Vol. 50, Number 2 (part 2), 1971, p 637-655.*

[2] Wakita, H., "Direct Estimation of the Vocal Tract Shape by Inverse Filtering of Acoustic Speech Waveforms", *IEEE Trans. Audio Electroacoust. Vol.AU-21,*1973, p *417-427.*

[3] Deng, H., M. P. Beddoes, R. K. Ward and M. Hodgson, "Estimating the Derivative of the Glottal Wave and the Vocal-tract Filter from a Vowel Sound Signal", IEEE PACRIM'03 Conference on Communications Computers, and Signal Processing, Aug. 28-30, 2003, Victoria, Canada.

[4] Story, B. H. and R. Titze, "Vocal tract area functions from magnetic resonance imaging", *J. Acoust. Soc. Amer., Vol. 100,1995, p 537-554.*

# Acoustic Indicators Of Spanish-Accented English

**Amee P. Shah**
School of Communication Sciences and Disorders, McGill University
1266 Pine Avenue West, Montreal, Quebec H3G 1A8.
Email: amee.shah@mail.mcgill.ca

## 1. BACKGROUND

The speech of non-native speakers of a language or "foreign-accented" speech is usually characterized by presence of acoustic deviations from the norm that may pose a challenge to the listeners' perception. Depending on the degree of speech deviation (among other factors), the listeners' perception may be affected to varying degrees, ranging from a mild awareness that the speaker is "foreign-sounding" or "foreign-accented", to a complete breakdown in overall communication due to the reduced intelligibility of the speech.

The present study attempted to address the distinct speech characteristics of Spanish-accented speakers and their influence on the native listeners' perception of accentedness. Previous research (MacDonald, 1989; Ortega-Llebaria, 1997) involved contrastive phonological analyses of speech characteristics in an attempt to identify the speech deviations that were related to perceived accentedness. Only a few studies have employed acoustic measures in studying Spanish-accented English (e.g. Magen, 1998; Backman, 1978; Flege and Eefting, 1987; Flege , Munro and Skelton, 1992). The above phonological and acoustic studies led to the hypotheses motivating the present study, in that Spanish-accented speech is characterized by deviations in temporal patterns, and that these deviations account, at least in part, for the perception of "foreign-accentedness".

Two studies were designed to test the hypotheses. Study 1, a perception study, was used to obtain accentedness ratings from native listeners as they listened to multisyllabic words (and sentences) produced by native as well as the non-native (Spanish-accented) speakers. Study 2, a production study, tested the hypothesis that Spanish-accented speakers' temporal patterns, i.e. segmental durations, differ from those of native English speakers in multisyllabic words. A third part, a set of analyses testing the correlation of the range in segmental durations of the Spanish-accented speakers and the range in accentedness ratings assigned by the native listeners, was attempted to test the hypothesis that the temporal deviations in the nonnative (Spanish-accented) speech are related to the perception of foreign-accentedness.

## 2. STUDY 1

In this perception study, the objectives were: a) to establish the relative global accentedness of the nonnative speakers in this study, and b) to determine the correlation between ratings of accentedness in multisyllabic word and overall sentence condition in order to test whether the accentedness on the multisyllabic words were representative of the global accentedness across the nonnative speakers.

Ten native speakers of American-English judged accentedness on a nine-point rating scale where 1=least-foreign accented, and 9=most-foreign accented. The stimuli consisted of randomized samples of eight multisyllabic words that had been spliced from sentences that were read by 22 nonnative speakers (Spanish speakers of English) and five native speakers of American-English.

The ratings of listeners on each of the eight words were pooled and the typical (median) native listener ratings of accentedness on each word were used as the speakers' scores in the correlational analyses with the production measures as described later. Additionally, results showed that all of the eight multisyllabic words were good predictors of global accentedness, as they correlated strongly ($\rho$=0.82) with previously-tested accentedness ratings on overall sentences by the same listeners. Hence, these multisyllabic words were the focus of the following analyses.

## 3. STUDY 2

In this production/acoustic study, the above-mentioned eight multisyllabic (3, 4, and 5 syllable) target words, as recorded by 22 nonnative and the five native speakers of American English were analyzed for temporal acoustic differences. The segments measured included overall word duration, unstressed vowel duration, ratios of stressed to unstressed (S/U) vowel duration as a measure of lexical stress, Voice Onset Time (VOT) of word-initial, voiceless stop consonants and closure duration in intervocalic flaps/stops. Based on phonological descriptions of Spanish-accented English and Spanish phonological rules (MacDonald, 1989; Ortega-Llebaria, 1997), predictions were made for each of the acoustic segments measured. The Spanish-speakers of English (nonnative group) were expected to produce longer overall word durations due to an expected slower rate of speech. Voiceless stop VOT values were expected to be smaller in the nonnative group since voiceless stops in Spanish are unaspirated. Flap/closure duration of intervocalic /t/ was expected to be greater in the nonnative group, approximating durations for stops. Vocalic segmental durations were expected to be approximately equal for stressed and unstressed syllables; thus S/U ratios were expected to be equal to 1 in the nonnative group, as predicted by the syllable-timed stress pattern of Spanish phonology.

Results showed that the nonnative speakers' productions were, on average, significantly longer than the native productions [$t(56)=-2.15$, $p<0.04$]. The unstressed vowel durations were, on average, proportionately longer for the nonnative group compared to the native group. Unlike the native group, the nonnative group failed to make a sufficient temporal distinction between stressed and unstressed vowels, with S/U ratios, on average, varying from 1.00 to 2.05. VOT, on average was shorter in the nonnative group, on average, compared to the native group. Flap/closure durations were, on average, longer in the nonnative group, compared to the native group.

As expected, the nonnative speakers, as a group, showed greater variability in each of the measured segmental durations as compared to the native group. This variability in word duration lent itself well to test its correlation with the range in accentedness ratings by the native listeners (as obtained from Study 1).

## 4. CORRELATION ANALYSES

Each of the above segmental measures were rank-ordered and correlated with ranked median ratings of accentedness on words, using a Spearman rank-order correlation. The objective of this analysis was to determine the relationship of the temporal deviations in the nonnative group with the perceived accentedness, as judged by the native speakers.

Findings indicated that overall word duration was correlated with perceived accentedness with low to moderate strength ($\rho=0.04$ to 0.56), although only one word, "committee", yielded a statistically significant relation. Stressed to unstressed vowel ratio, a measure of lexical stress, correlated positively with perceived accentedness for all but one word, although, only the correlation for the word "economic" was significant ($\rho=0.80$). Deviations in VOT also correlated positively with accentedness ($\rho=0.26$ to 0.36) and so did deviations in flap/closure durations ($\rho=0.29$ to 0.59). While none of the VOT-accentedness correlations was statistically significant, three of the five correlations of flapped /t/ and accentedness were highly significant.

## 5. GENERAL DISCUSSION

The overarching goal of this research was to examine the temporal speech characteristics of native Spanish speakers of English-as-a-second-language that relate to the native English-speaking listeners' perception of accentedness. Previous studies were restricted to either describing the phonological characteristics or a analyzing a select few acoustic measures of this nonnative group. Multiple acoustic parameters as they relate to perceived accentedness have not yet been studied in a correlational design. The present study tapped several temporal characteristics of multisyllabic, words, selected on the basis of previous phonological descriptions of Spanish-accented speech, as they relate to a measure of perceived accentedness. Two studies were designed to test the hypotheses that temporal deviations characterize Spanish-accented speech, and that these deviations are related to the perception of accentedness as judged by native speakers of English. The multisyllabic words used in the two studies had been found to be good predictors of global accentedness, i.e. accentedness on overall sentences. Results showed systematic temporal deviations in the nonnative group compared to the native group for each of overall word duration, unstressed vocalic duration, stressed to unstressed vowel ratios, VOT for voiceless, word-initial stops as well as for closure duration in intervocalic stops/flaps. However, overall group differences were small. Moreover, correlations tested with each of these segmental deviations with accentedness ratings showed only low to moderate strength of relationship. It is hypothesized that, while each of the measures alone does not strongly predict accentedness, some combination of these temporal deviations may account, at least in part, for native listeners' judgments of perceived accentedness.

## REFERENCES


Backman, N. (1978). Intonation errors in second-language pronunciation of eight Spanish-speaking adults learning English. Paper presented at the Fifth Annual Congress of Applied Linguistics, Montreal.

Flege, J.E. & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English, Speech Communication, 6, 185-202.

Flege, J.E., Munro, M.J., & Skelton, L. (1992). Productions of the word-final English /t/-/d/ contrast by native speakers of English, Mandarin, and Spanish. Journal of Acoustical Society of America, 92 (1), 128-143.

MacDonald, M.G. (1989). The influence of Spanish phonology on the English spoken by United States Hispanics. In Hammond, R., & Bjarkman, P. (Eds.), American-Spanish Pronunciation: Theoretical and Applied Perspectives (pp. 215-237). Washington D.C.: Georgetown University Press.

Magen, H.S. (1998). The perception of foreign-accented speech. Journal of Phonetics, 26, 381-400.

Ortega-Llebaria, M. (1997). An explanatory intelligibility test for Spanish accented English. Doctoral dissertation, Department of Linguistics, Indiana University.


## ACKNOWLEDGEMENTS


This work was partially supported by the "Doctoral Dissertation Research Grant" (BCS 0001862) from the National Science Foundation to Amee Shah and NIDCD grant (DC00323) to Dr. Winifred Strange.

# Intelligibility of Foreign-Accented Lombard Speech

**Chi-nin Li**
Dept. of Linguistics, Simon Fraser University, 8888, University Drive., B.C., Canada, V5A 1S6. clia@sfu.ca

## 1. INTRODUCTION

The Lombard reflex occurs when normal-hearing speakers modify their vocal effort while speaking under noisy conditions (Lane & Tranel, 1971; Van Summers, Pisoni, Bernacki, Pedlow, & Stokes, 1988). Speech produced in noisy conditions has been reported to be different from speech produced in quiet conditions in a number of ways: increase in vowel duration, a shift in the first two formants, decrease in speaking rate, and increase in fundamental frequency (Junqua, 1996). Also, Lombard speech is more intelligible than normal speech when presented at an equal level of masking noise (Van Summers et al., 1988).

Because research associated with the Lombard reflex has been conducted mainly with native speakers, little is known about its impact on foreign-accented speech. The purpose of this experiment is to examine the effects of masking noise on the intelligibility of Lombard speech produced by non-native English speakers using a sentence-transcription task and a sentence-verification task.

## 2. METHOD

### 2.1 Speech material

A list of 24 true and 24 false statements were recorded by 12 Hong Kong Cantonese speakers (6 female, 6 male) and a comparison group of native speakers of Canadian English (6 female, 6 male) in a sound-treated room. All speakers passed a pure-tone hearing screen. They produced in a conversational manner the list of sentences_once under each of two speaking conditions: quiet and_noise. Under the noise condition, cafeteria–like masking noise of 70 dB was fed through headphones to the speakers. No noise was presented to the speakers in the quiet condition. Stimuli were digitized at a 22.05 kHz sampling rate with 16-bit amplitude quantisation.

Four different sentences (2 true, 2 false) produced in each of the speaking conditions were selected from each of the 24 speakers for a total of 192 utterances. From each speaker, the four specific sentences produced in the quiet condition (quiet) were identical to those produced in the noise condition (Lombard). Selected Lombard and quiet speech samples were mixed with the same type of masking noise used in the production task at a constant signal-to-noise (S/N) ratio (noisy), and were presented to listeners with unmasked sentences (clean). This resulted in four different speech conditions: (1) Clean quiet (Q-N), (2) Noisy quiet (Q+N), (3) Clean Lombard (L-N), and (4) Noisy Lombard (L+N). Four separate stimulus sets were prepared. Each set consisted of the full complement of 48 test items (24 true, 24 false). For each set, individual speakers were represented twice, once producing a true sentence and once producing a false sentence. In addition, each of the stimulus sets was balanced for speakers' native language and gender, and the speech conditions. The root-mean-square (RMS) amplitude for each of the 192 stimulus sentences was equated.

### 2.2 Listeners

The listeners were eight female native speakers of Canadian English. They were born and raised in Vancouver, British Columbia. Their mean age was 21.5 years, and all passed a pure-tone hearing screen. The listeners in groups of two were randomly assigned to listen to one of the four sets of stimuli.

### 2.3 Procedure

Stimuli were presented via headphones at a comfortable listening level to the listeners using a custom response-collection program. The listeners transcribed the stimulus sentences in standard orthography. They then verified the truth value of the statement by clicking one of three buttons shown on a computer screen ("True", "False", or "Unknown").

## 3. RESULTS

### 3.1 Transcription Scores

Scores were assigned to each sentence by computing the percentage of words that were correctly transcribed. Figure 1 illustrates the mean values for the sentences presented under the four speech conditions. In each of the four speech conditions, the Cantonese speakers' sentences were less correctly transcribed than those of the native productions. A two-way repeated measures ANOVA revealed significant effects of Native Language of Speakers (NL), $F\ (1,7) = 54.99$, $p < 0.01$, and Speech Condition (SC), $F\ (3,21) = 15.06$, $p < 0.0001$. Post hoc analyses revealed that noisy utterances (Q+N and L+N) received significantly lower scores than did clean utterances (Q-N or L-N), $ps < 0.05$, and that listeners' scores in the condition of L+N were significantly higher than those in the condition of Q+N, $ps < 0.05$. No significant difference was found between the conditions of L-N and Q-N, $ps > 0.05$.

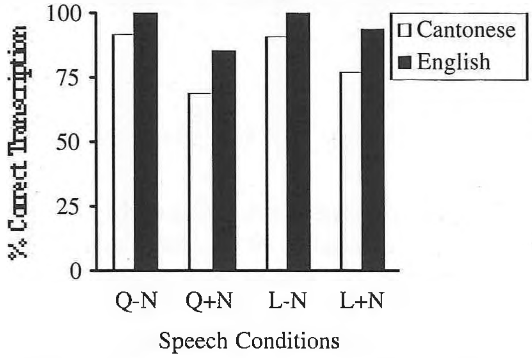



**Fig. 1. Percent correct transcription scores for the stimulus sentences presented in the four speech conditions.**

3.2 Verification Scores

The verification scores were determined by summing the number of correct true and false responses. The choice of "Unknown" was considered as incorrect. The mean values are given in Figure 2. Cantonese speakers' sentences were correctly verified less frequently than were the native productions in all speech conditions. A repeated measures ANOVA yielded significant effects of NL, $F\ (1, 7) = 68.03$, $p < 0.0001$, and SC, $F\ (3, 21) = 27.14$, $p < 0.0001$, as well as a significant NL x SC interaction, $F\ (3, 21) = 3.24$, $p < 0.05$. For the non-native productions, statistical analyses revealed that clean sentences (Q-N and L-N) were correctly verified more often than were noisy sentences (Q+N and L+N), $ps < 0.05$. Other pairwise comparisons (L-N vs. Q-N and L+N vs. Q+N) were not significant, $ps > 0.05$. For the English speakers, utterances presented in Q+N received significantly lower scores than did those presented in Q-N, L-N or L+N, $ps < 0.05$. Nevertheless, there were no significant differences among the three speech conditions, Q-N, L-N, or L+N, $ps > 0.05$.

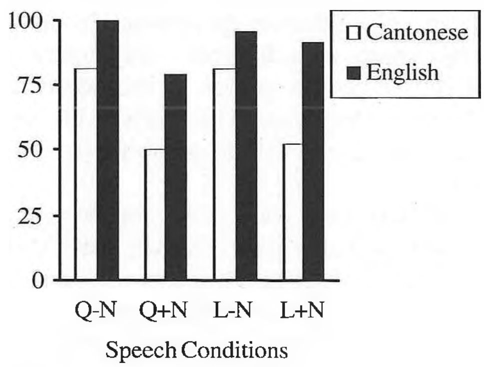



**Fig. 2. Percent correct verification scores for the stimulus sentences presented in the four speech conditions.**

# 4. DISCUSSION

Cantonese speakers' utterances were found to be less intelligible than were the native English productions. The sentences spoken by the Cantonese speakers were less correctly transcribed and verified than were the native-produced sentences in each of the four speech conditions.
Unmasked sentences were better perceived than were noisy sentences in most cases. Utterances presented in the 'clean' conditions (Q-N and L-N) were more correctly transcribed than were sentences presented in the 'noisy' conditions (Q+N and L+N), indicating that the masking noise degrade the intelligibility of both the Lombard and quiet speech. This pattern was again observed in the verification scores for Cantonese speakers' sentences, but not for English speakers' productions. For the English speakers, there was no significant difference in the verification scores between their noisy (L+N) and clean (L-N) Lombard speech.

For both groups of speakers, their noisy Lombard speech (L+N) was correctly transcribed more frequently than were the noisy quiet sentences (Q+N). The improvement in intelligibility is likely due in part to the characteristics of the Lombard speech. For the verification task, such a pattern was found for the native English speakers, but not for the Cantonese speakers. No significant difference in verification scores was found between the non-native Lombard speech and quiet speech when presented in noisy conditions. The discrepancy in patterns for the intelligibility of the non-native noisy sentences may be due to the different nature of the two tasks involved, as it has been suggested that the verification measure appears to be a coarse evaluation of intelligibility (Pisoni, & Dedina, 1986).

# REFERENCES

Junqua, J-C. (1996). The influence of acoustics on speech production: A noise-induced stress phenomenon known as the Lombard reflex. *Speech Communication, 20,* 13-22.

Lane, H., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research, 14(4),* 677-709.

Pisoni, D., & Dedina, M. (1986). Comprehension of digitally encoded natural speech using a sentence verification task (SVT): A first report. *Research on speech perception (No. 12).* Bloomington, IN: Indiana University.

Van Summers, W., Pisoni, D., Bernacki, R., Pedlow, R., & Stokes, M. (1988). Effects of noise on speech production: acoustic and perceptual analyses. *Journal of the Acoustical Society of America, 84(3),* 917-928.

# ACKNOWLEDGEMENTS

The author would like to thank Dr. Murray Munro for his advice on this research.

# The Effect of Acoustic Information of Lexical tones on Non-native Listeners' Tonal Identifications

**Connie Kwok Lai So**
Dept. of Linguistics, Simon Fraser University, Burnaby, BC, Canada, V5A 1S6 klso@sfu.ca

## 1. INTRODUCTION

This paper reports a perception study in lexical tones by non-native speakers of Mandarin. The role of acoustic information of lexical tones in a perception training is examined.

Although studies (Leather, 1990; Wang et al., 1999) have shown that auditory-training improves non-native listeners' tonal identification, persistently perceptual confusions of several tone pairs are still observed after a considerable amount of training. For example, a high falling tone is misidentified as a high level tone. This implies that learners have not yet learned the lexical tones.

Learners' tonal confusions might be reduced further if the acoustic information of tones, such as intrinsic duration-pattern and pitch contour, could be implemented and emphasized during training. In general, learners in those studies involving auditory-training do not receive explicit explanations for a given contrastive tone pair when the response is incorrect. However, studies in tonal perception (Gandour, 1984; Massaro et al., 1985), have clearly demonstrated that acoustic cues (e.g., pitch height and contour) are essential for the identification of tones. Accordingly, employing acoustic cues of lexical tones during training may have a fruitful result. To verify the above assumption, the present study examines the effect of using acoustic information of lexical tones as feedback on non-native listeners' performance during a computer-based perception training of Mandarin tones.

Mandarin has four lexical tones, each of which has a unique pitch pattern: Tone 1 – high level, tone 2 – mid-rising, tone 3 – low rising, and tone 4 – high falling. In the present study, two pairs of tones (tone 1 vs. tone 4 and tone 2 vs. tone 3) were investigated, since learners of Mandarin have great difficulties producing and perceiving these tone pairs (Shen, 1989; Sun, 1998).

## 2. METHOD

2.1 Participants

The listeners were 24 non-native speakers with no prior knowledge of Mandarin, ranging in age from 18 to 37 years ($M$ = 24.5 years). They were randomly assigned to one of two groups: control vs. experimental.

2.2 Material

The present study used a total of 13 CV syllables, each of which has a lexical meaning when associating with one of the four tones. All stimuli were produced by two native Mandarin speakers (1 male & 1 female) who were born and raised in Beijing. Their speech samples were correctly identified (100%) by another female native Mandarin speaker born and raised in Beijing. These digitized stimuli were used for a hypermedia authoring computer program, Mandarin Tonal Tutor (MTT), to carry out the experiment. The MTT program has two versions: Simple feedback (SF) and Detailed feedback (DF). The listeners in the control group used the SF version, and those in the experimental group were trained with the DF version. The MTT program consisted of four phases: an introduction, a pre-test, seven training exercises, and a post-test.

2.3 Training

The training session was about 30-45 minutes. In the SF version, the listeners were merely shown that the answer was right or wrong (i.e., no acoustic information was provided). In contrast, the listeners using the DF version received acoustic information by means of both visual and auditory feedback whenever they made an error. The feedback came in three forms: (a) text describing the essential acoustical characteristic differences, (b) pitch graphs, and (c) audio files for the contrastive tonal pairs.

## 3. RESULTS & DISCUSSION

Listeners' scores for the post-test were higher than those of the pre-test. However, the performance for the listeners in the two groups was different. On average, listeners who used the DF version showed an increase of 21.61% in the post-test, whereas those who received the SF version only had an increase of 4.95% in the post-test.

Listeners' perception scores for the pre- and post-tests were submitted to a mixed ANOVA, with Group (control vs. experimental group) as a between-subjects factor, and Test (pre-test vs. post-test) as a within-subjects factor. The statistical analysis revealed a significant effect of Test on the scores, $F(1, 22) = 48.77$, $p < .01$, indicating that the listeners had a higher score in the post-test than the pre-test. However, the effect of Group did not approach significance, ($p > .05$). Furthermore, there was a significant Group x Test interaction, $F(1, 22) = 17.78$, $p < .01$, indicating that the listeners who were trained with the DF version significantly outperformed those who used the SF version in the post-test (see Fig. 1).

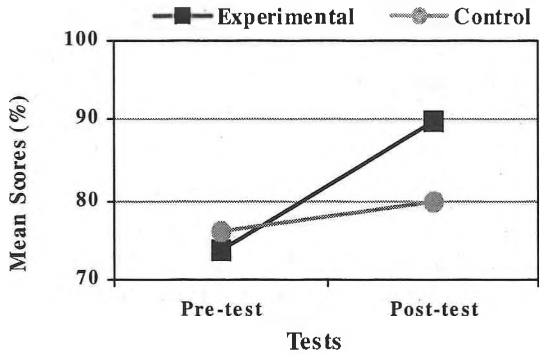



**Fig. 1. Mean scores (%) of the listener groups for the pre- and post-test.**

Identification errors made by the listeners in the two groups in the pre- and post tests were also examined. Their errors are shown in Figure 2 (for the control group) and in Figure 3 (for the experimental group).

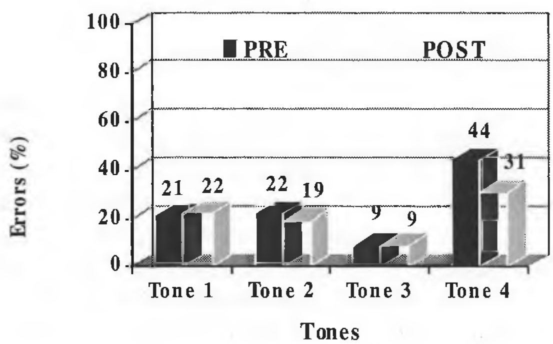



**Fig. 2. Errors (%) for the control group in the pre- and post-tests**

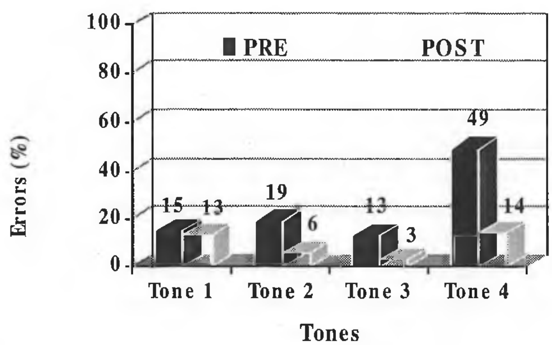



**Fig. 3. Errors (%) for the experimental group in the pre- and post-tests**

As shown in the figures, the listeners in the experimental group show dramatic and consistent improvement in the post-test, from 2% to 35% (Fig. 2). In contrast, the same pattern of improvement is not observed in the performance for the listeners in the control group; their errors reduce from 0% to 13% only (Fig. 3).

## 4. CONCLUSION

The present study examines the role of acoustic information in the perception training of Mandarin tones. The results indicate that using acoustic information as feedback during a short-span training significantly improves listeners' performance in tonal identifications, and that the pattern of improvement is consistently observed for each lexical tone. This suggests that feedback with the acoustic information during training assists non-native learners to learn the lexical tones, since the contrastive characteristics of the tone pairs are shown to them. A further examination in listeners' speech productions is desirable to find out if these learners are able to produce the tones according to the acoustic information that they have received during training.

## REFERENCES


Gandour, J. (1984). Tone dissimilarity judgements by Chinese listeners. *Journal of Chinese Linguistics, 12*(2), 235-261.

Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers, In J. Leather & A. James (Eds.), *New Sounds 90* (pp. 72-89). Amsterdam: University of Amsterdam.

Massaro, D. W., Cohen, M. M. & Tseng, C. Y. (1985). The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese. *Journal of Chinese Linguistics*, *13*(2), 267-289.

Shen, X. (1989). Toward a register approach in teaching Mandarin tones. *Journal of the Chinese Language Tearchers Association, 24*(3), 27-47.

Sun, S. (1998). *The development of a lexical tone phonology in American adult learners of standard Mandarin Chinese*. Honolulu, HI: University of Hawai'i Press.

Wang, Y., Spence, M., Jongman, A., & Sereno, J. (1999). Training American listeners to perceive Mandarin tones. *The Journal of Acoustical Society of America*, 106(6), 3649-3658.


## ACKNOWLEDGEMENTS


This research was supported by the Social Sciences and Humanities Research Council of Canada.

# Perceptual Differences In Categorization Of Speech Sounds By Norwegian/English Bilinguals

**Audny T. Dypvik and Elzbieta B. Slawinski**
Dept. of Psychology, University of Calgary, 2500 University Drive, N.W., Alberta, Canada, T2N 1N4

## I. INTRODUCTION

The ability to process auditory information is critical for speech perception. Deficits directly affect an individual's well-being and ability to participate in society. When a second language is introduced, bilinguals, and especially late bilinguals, perceptual abilities change with the degree and intensity of exposure to the second language.

Various acoustic cues serve to precipitate the perceptual distinction between phonemic contrasts. Spectral and temporal cues are used to discriminate the phonemic contrast between, for example, prevocalic /r/ and /l/ for English listeners (Polka and Strange, 1985). When consideration is taken for differing perceptual abilities of bilinguals, it has been found that bilingual Japanese/ English listeners with early and extensive exposure to English had improved perceptual ability in a phonemic distinction (Slawinski and MacNeil, 1994).

Perceptual categories are revealed in stop consonants. These are separated by well-defined phonemic boundaries and can be defined by listeners presented with a series of stops varying in equal physical steps relative to voicing. One such study (Elman, 1977) focused on perceptual switching in English and Spanish bilinguals. When the listeners heard the test stimuli in either an English or a Spanish context, the placement of category boundaries changed. Not only did the induction of a particular "language set" serve to alter the identification of stimuli by the listeners, but the degree of bilingualism interacted with the extent of perceptual switching effect (eg. Elman, 1977; Bohn and Flege, 1993).

Synthetic stimuli as well as natural speech stimuli have been implemented in speech studies. When Elman (1977) attemped to replicate his findings using synthetic stimuli, a reliable tendency toward a shifted perceptual boundary was not observed. Nonetheless, the noted absence of secondary voicing cues in their synthetic stimuli (such as spectral characteristics of the release burst and changes in the fundamental frequency following articulatory release) may have increased the difficulty for the bilinguals who conceivably place greater reliance on such cues. Since that time, synthetic stimuli have been improved.

Norwegian is one of the few European languages that is not an intonation language but a tone language (Moen, 1993). Norwegian has a distinct phonemic categorization for certain syllables, prominent among these being /v/ and /w/. This and a shortage of available data for the Norwegian language instigated the conceptualization of the current study. A revolutionary departure for speech synthesis was utilized - an articulatory synthesizer.

## 2. OVERVIEW OF APPROACH

In the past, the primary means of synthesizing speech was spectral reconstruction as exemplified by the Klatt synthesizer (Klatt, 1980). Recently, a real-time articulatory-based speech synthesizer has been developed and is currently in the process of being refined (Hill, Manzara and Taube-Schock, 1995). The "tube synthesizer", as it is called, together with supporting tools and components, allows one to create precisely controlled stimuli.

Energy in the form of pressure variations is injected into one end of the tube model. This corresponds to air being forced through the vocal folds causing them to vibrate. The energy is then spectrally modified by the resonant characteristics of the oropharyngeal and nasal cavities using an acoustic waveguide (in Hill, Manzara and Taube-Schock, 1995). The Distinctive Region Model (DRM) (Carre, 1992) was used as a basis to model the oropharyngeal cavity.

The tube model has 8 regions (vocal tract sections), which closely relate to human articulators, primarily the lips, tongue, and teeth. In contrast to older synthesizers, the control of speech is achieved by directly controlling the articulators as opposed to the formants.

A tool called MONET was developed to edit speech postures (roughly related to phonemes), interpolation rules, and timing data. MONET was used to create rules that combine basic postures into articulatory sequences. Once developed, these rules were used to create parameter tracks that drive the synthesizer. MONET can handle diphone combinations of postures as well as triphone combinations. Triphone combinations enable one to model co-articulatory effects.

In the current study, the tube synthesizer was used to create a series of preliminary stimuli for the phonemic distinction between /v/ and /w/. These phonemes have particular articulatory postures. We created 9 postures which included the 2 endpoints of /v/ and /w/ plus 7 intermediate postures equally spaced and linearly interpolated.

Using these postures, we generated 4 series of stimuli. The first series used diphone transitions describing silence to the interpolated sequence of /w/ to /v/ postures to the following vowel /i/. This series also included frication energy injected at the tube section representing the teeth. Frication amplitude in this series depended upon the size of the orifice at the teeth. The second series was similar to the first except that no frication energy was injected. The third series used a triphone transition describing silence to the interpolated sequence of /w/ to /v/ postures to the following vowel /i/. Frication was included in the series. The fourth series was similar to the third except that frication was not injected. These stimuli will be presented to participants via headphones in an anechoic chamber (Industrial Acoustics Company, Inc.). Analysis of speech production will follow analysis of perception.

## 3. DISCUSSION

It is anticipated that the combination of examining categorization of English phonemic perception in the little-studied language of Norwegian with the use of real-time articulatory speech synthesis will lead to some exciting results.

Bilingual immigrants do indeed face unique challenges in every-day living. They perceive acoustically identical stimuli differently depending on the language being processed at the time, or "language set", and the extent of this perceptual difference is related to the timing and extent of exposure to the second language. We propose that because /v/ and /w/ are allophones in the Norwegian language, English phonemic boundaries of bilingual listeners between /v/ and /w/ will vary from those of English speakers.

To date, use of the "tube synthesizer" has revealed a very real clarity of articulation. As well, spectrograms already created from these synthesized-by-rules speech stimuli appear similar to those created from natural speech stimuli. A more natural approach to speech production is introduced when articulatory postures are used to dictate sound synthesis rather than formant manipulation. Validation of this approach will require substantiation of formerly-achieved results illustrating perceptual discrimination differences of certain phonemic boundaries by bilinguals.

## 4. REFERENCES


Bohn, O-S. and Flege, J.E. (1993). Perceptual switching in Spanish/English bilinguals. Journal of Phonetics, 21, 267-290.

Carre, R. (1992). Distinctive regions in acoustic tubes. Speech production modeling. Journal d'Acoustique, 5, 141-159.

Elman, J.L., Diehl, R.L. and Buchwald, S.E. (1977). Perceptual switching in bilinguals. Journal of Acoustical Society of America, 62(4), 971-974.

Hill, D.R., Manzara, L. and Taube-Schock, C-R. (1995). Real-time articulatory speech-synthesis-by-rules. Presented at AVIOS, San Jose, CA, USA.

Klatt, D.H. (1980). "Software for a cascade/parallel formant synthesizer". Journal of Acoustical Society of America, 67, 971-995.

Moen, I. (1993). Functional Lateralization of the Perception of Norwegian Word Tones – Evidence from a Dichotic Listening Experiment. Brain and Language, 44, 400-413.

Polka, L. and Strange, W. (1985). Perceptual equivalence of acoustic cues that differentiate /r/ and /l/. Journal of Acoustical Society of America, 78(4), 1187-1197.

Slawinski, E.B., and MacNeil, J.F. (1994). The relationship between perceptual and productive discrimination of English /r/ and /l/ sounds by Japanese speakers. International Journal of Psycholinguistics, 10(3), 1-23.


## ACKNOWLEDGEMENTS


Thank to Dr. David R. Hill, Dr. Leonard Manzara and Craig-Richard Taube-Schock, MSc., for their expertise, generosity of time, and use of their speech synthesizer.

Dr. Greg Shaw added valuable support in programming.

# Developing a New Measure for Assessing Architectural Speech Security

**J. S. Bradley and B. N. Grover**
Institute of Research in Construction, National Research Council, Montreal Road, Ottawa, K1A 0R6, Canada

## Introduction

This paper describes the process and problems of developing a new measure of architectural speech security. Such a measure is required to more accurately rate the probability of a listener outside a room being able to overhear conversations from within the room. Previous work has considered various levels of speech privacy, where some speech is intelligible [1]. However, speech security usually implies that none of the overheard speech is intelligible, or in some cases it is not even audible.

One can describe 3 levels of speech security. The first level would be when only a very small percentage or none of the overheard words are intelligible. Even when no words are intelligible, it is often still possible to recognize the rhythm or cadence of the speech. Finally, the highest level of speech security would be when all speech sounds from the adjacent space are completely inaudible.

Speech privacy and speech security have been related to signal/noise (S/N) type measures (where the signal is the speech from the adjacent space). The simplest would be a difference of A-weighted speech and noise levels. More sophisticated measures such as the Articulation Index (AI), and its more recent replacement the Speech Intelligibility Index (SII) (ANSI S3.5), are known to be better related to speech intelligibility within rooms. They more correctly weight the importance of S/N ratios at different frequencies and more accurately combine the various frequencies. These frequency weightings may not be optimum for speech security situations, and these measures are not ideal at very low levels of speech intelligibility.

## Experimental Procedure

In this work, subjects rated simulated speech security situations. The subjects sat in a sound-isolated room and heard speech sounds, modified to simulate transmission through various walls, and combined with typical ambient noises. The speech and noise sounds were spatially separated and were precisely measured at the listener's head position in an acoustically dead environment.

Many variables will influence the intelligibility of overheard speech, including: talker gender and voice characteristics, speech material, voice level, wall transmission loss characteristics, ambient noise spectrum shape and level, listener hearing sensitivity, and other listener characteristics. Many of these were determined to be of less importance in pilot tests. Only subjects with negligible hearing loss were included.

In the main intelligibility experiment, 30 subjects each listened to 340 test sentences. The phonetically balanced and low predictability Harvard sentences were used [2], and 5 different sentences were used for each physical condition. Each condition was one of 68 combinations of varied: ambient noise, wall transmission loss characteristics, and S/N ratio. The conditions were chosen so that intelligibility ranged from 0 to 100%. A second experiment was intended to determine the thresholds of: (a) audibility of any speech sounds, (b) audibility of the cadence of the speech, and (c) the intelligibility of the speech. In this experiment the 20 best listeners from the first experiment each listened to 160 sentences. Again there were 5 sentences for each condition, and a range of ambient noise and wall transmission loss values. However, in this experiment conditions had, on average, much lower S/N values so that they included situations where no speech sounds were audible to the listeners.

## Evaluation of Measures of Intelligibility

Figure 1(a) plots intelligibility scores versus measured SII values in the test sound fields. To simplify the plot, the results were averaged over all subjects. Although intelligibility scores increase with SII as expected, at SII=0 intelligibility is not zero. Therefore, SII (and AI) cannot be used to describe conditions for high levels of speech security which would correspond to acoustical conditions below SII=0, where SII is not defined. Figure 1(b) shows that differences of A-weighted levels are not limited in this way but are much less accurately related to intelligibility scores.

An example of a more successful measure is shown in Figure 1(c), which plots the same intelligibility scores versus a S/N ratio measure that included the same frequency weightings as the SII measure.

## Speech Security Threshold Measurements

The 6 graphs of Figure 2 show the results of evaluations of the 3 types of thresholds and the effects of different weightings of the importance of each frequency band. Each graph shows the percentage of subjects with responses indicating: at least one word is intelligible (a) & (d), the cadence of the speech is audible (b) & (e), and some speech sounds are audible (c) & (f). In graphs (a)-(c) results are plotted against SII-weighted S/N ratios and in graphs (d)-(f) against LF-weighted S/N ratios, having greater emphasis on the lower frequencies.

If one considers the threshold to be when 10% of the subjects respond, the threshold for intelligibility is reached at -18.5 dB, for detection of cadence at -24 dB, and for audibility at -27 dB on the SII-weighted S/N ratio measure. Thus, complete speech security (inaudibility) requires speech levels to be about 8.5 dB lower than for the threshold of intelligibility. This would correspond to a

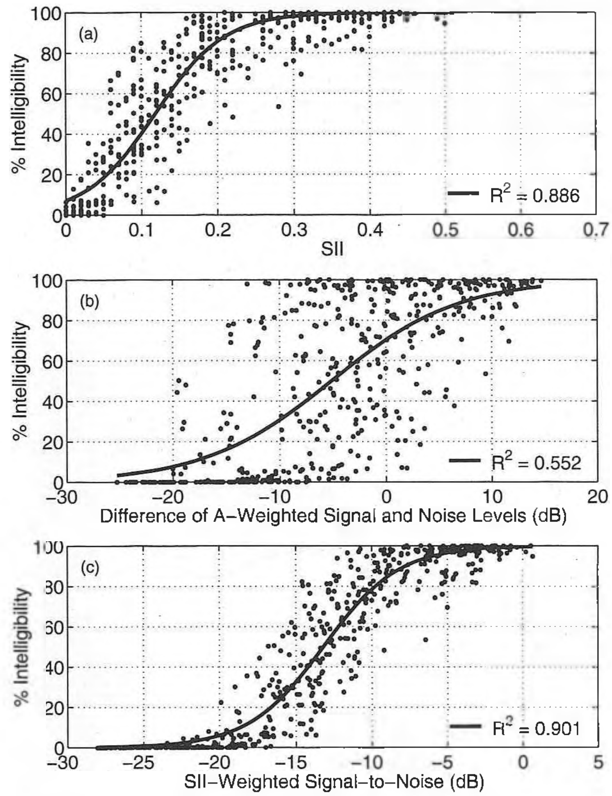



*Fig 1. Speech intelligibility scores vs. (a) SII, (b) difference of A-weighted levels, and (c) SII-weighted S/N ratio.*

substantially better wall transmission loss. Although the SII-weighted S/N ratios predict the intelligibility threshold reasonably well, the LF-weighting provides more accurate estimates of the thresholds of cadence and audibility.

## Conclusions

Existing measures of speech intelligibility and speech privacy are not adequate for evaluating the speech security of closed offices and meeting rooms.

The optimum frequency weighting for predicting the onset or threshold of intelligibility is different from that for predicting the threshold of the audibility or the cadence of speech sounds.

Speech security must be statistically described in terms of the percentage of listeners able to hear or understand speech from adjacent spaces.

Complete speech security, where speech sounds are totally inaudible, would require substantially better sound isolation of meeting rooms than is required for eliminating word intelligibility in adjacent spaces.

## References


1. Cavanaugh et al., J. Acoust. Soc. Am. 34 (4), 475-492 (1962).
2. IEEE Trans. On Audio and Electroacoust., 227-246 (1969).

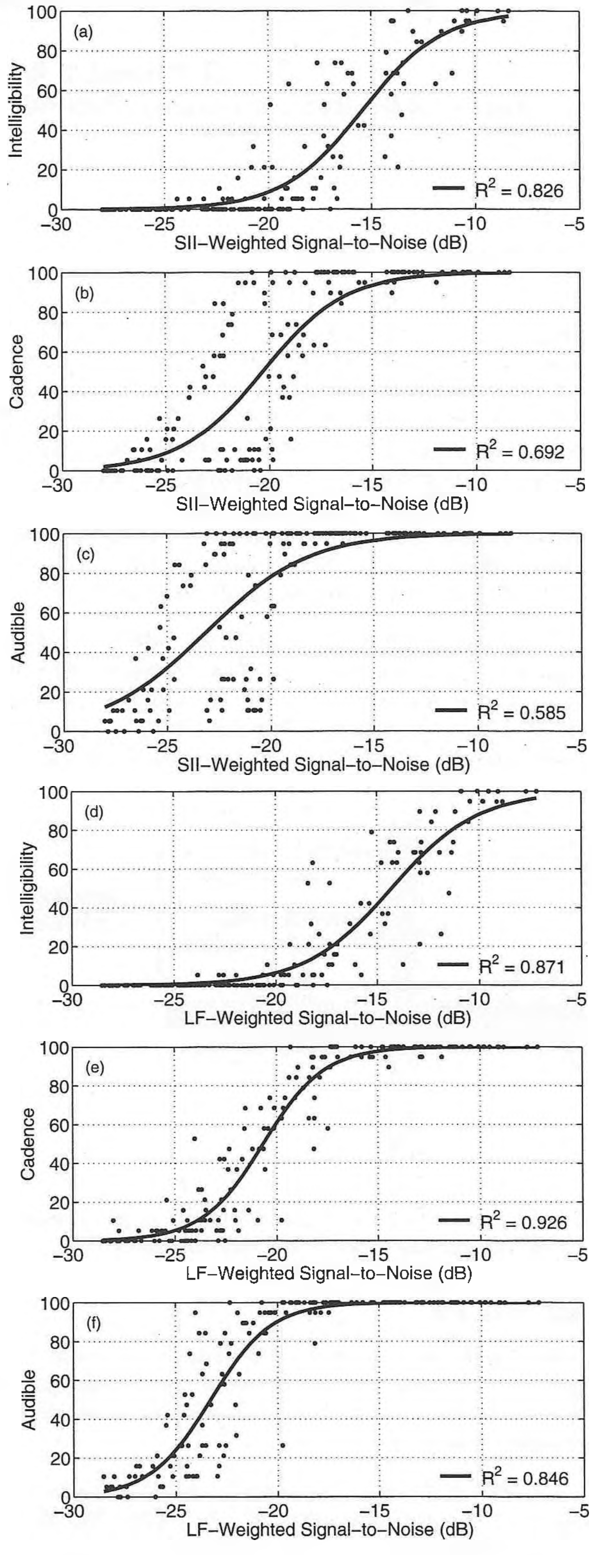




*Fig 2. Percentage of subjects detecting some intelligibility (a) & (d), cadence (b) & (e), and audibility (c) & (f) vs. SII- weighted S/N (a)-(c), and LF-weighted S/N (d)-(f).*

# Dealing with Flanking Transmission in Wood Framed Construction

**J. D. Quirt, T. R. Nightingale and R. E. Halliwell**
Institute of Research in Construction, National Research Council, Montreal Road, Ottawa, K1A 0R6, Canada

## 1. INTRODUCTION

The sound isolation between adjacent dwellings in multi-family buildings is often much less than would be expected from the STC rating of the nominally-separating wall or floor, due to structure-borne transmission of vibration via the junctions of wall and floor assemblies. Recent studies[1-3] have both identified key paths for this structure-borne transmission, or "flanking", and established the performance for a variety of construction modifications to control such transmission. This paper will present the trends evident from the available data, to provide guidance for practical solutions and indicate some remaining challenges for designers and consultants.

For adjacent rooms in a building, direct transmission through the separating partition, plus "flanking" (structure-borne vibration involving other surfaces of the rooms and transmitted via the junctions between these surfaces) both contribute to the transmitted sound power. Their combined effect is the Apparent Sound Transmission Loss.

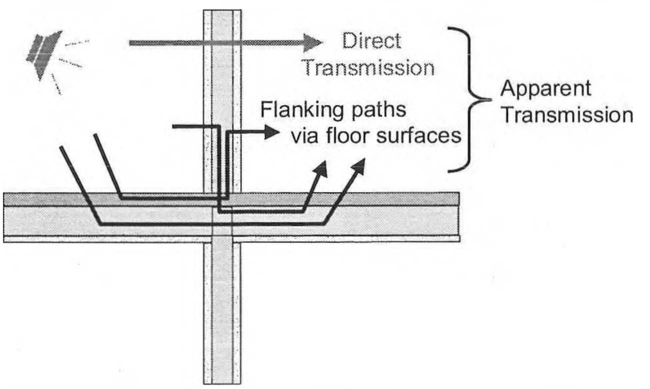



Figure 1: Sound transmission paths between adjacent rooms are shown on a conceptual sketch of the wall/floor junction.

## 2. RESULTS

The trends are similar for transmission of airborne and impact noise; this paper focuses on the former. Although many surfaces contribute to flanking transmission, some transmit more energy than others.

For vertical transmission (where direct transmission is through the floor/ceiling assembly) the main flanking path involves transmission via the sub-floor in the room above and wall surfaces with directly attached gypsum board in the room below. For horizontal transmission (where direct transmission is through the separating wall assembly, as shown in Figure 1) the main flanking path involves transmission via the sub-floor in each room. The range of typical flanking effects with a bare sub-floor of plywood or OSB is illustrated in Figures 2 and 3.

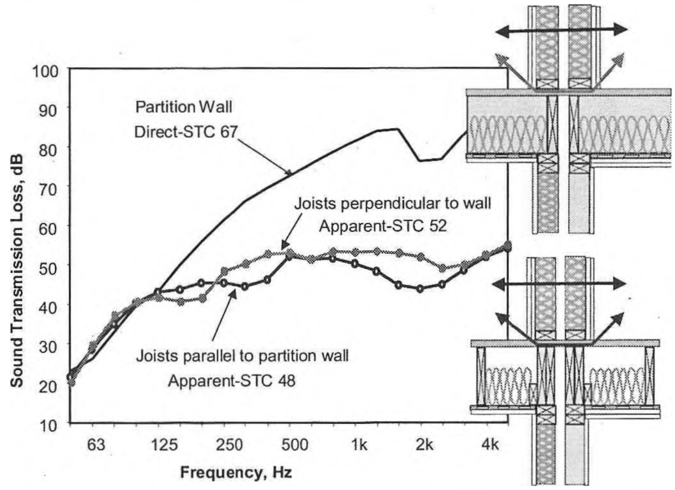



Figure 2: Apparent Sound Transmission Loss with a double wood stud wall (STC 67) and bare OSB sub-floor in the rooms.

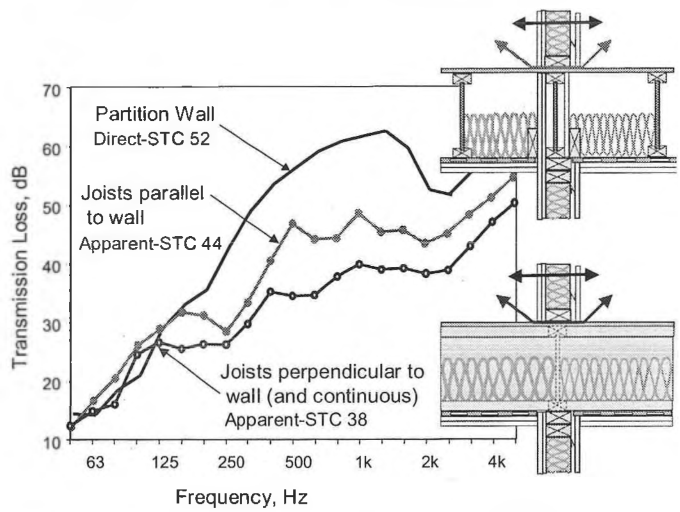



Figure 3: Apparent Sound Transmission Loss with a single wood stud wall (STC 52) and bare OSB sub-floor in both rooms, for two configurations of wood-I floor joists.

Flanking reduces the isolation between side-by-side rooms over most of the frequency range of interest, and reduces Apparent STC (which indicates what occupants will perceive) by nearly 20 dB in some cases. The flanking is primarily due to continuity of the floor joists and/or the sub-floor across the wall/floor junction[1,3].

In practice, such connections are required for 3- or 4-storey multifamily buildings, and even for row housing in areas of high seismic risk. Hence the practical problem shifts to design changes to reduce the flanking.

For horizontal transmission, most of the flanking energy is transmitted via the sub-floors; hence modifying the floor surfaces is the obvious remedy. An example with 25 mm gypsum concrete over the OSB is shown in Figure 4.

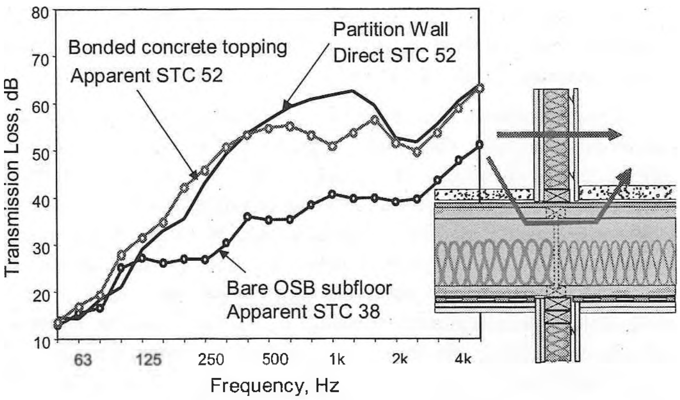



Figure 4: Apparent Sound Transmission Loss with bare sub-floor of 16 mm OSB, or with addition of a gypsum concrete topping.

Addition of the gypsum concrete topping increased the Apparent Transmission Loss at all frequencies where the flanking was significant. Except in a few bands near 1 kHz (the coincidence dip for this topping) the transmission via the floor became negligible, so the Apparent TL was limited only by direct transmission through the partition wall.

For cost-effective construction, transmission must be tested for a variety of floor toppings. To measure sound transmission via the floor-floor path, the direct transmission through the wall was blocked using a covering assembly. The results (listed in the table as "Flanking STC") were normalized to the area of the separating wall, for direct comparison with the overall Apparent STC.

| Type of topping | Flanking STC* | Change |
|---|---|---|
| Bare 16 mm OSB sub-floor | 39 | |
| 19 mm oriented strand board (OSB), stapled to OSB sub-floor | 48 | +9 |
| 25 mm gypsum concrete bonded to OSB sub-floor | 57 | +18 |
| 38 mm gypsum concrete floating on resilient foam pad over sub-floor | 60 | +21 |

* *Flanking transmission (floor-to-floor path). Joists and sub-floor continuous across wall/floor junction, and wall blocked.*

Note that somewhat different results would be expected for the same toppings when installed on different floor assemblies, especially in the case of bonded toppings, which depend significantly on joist orientation[3].

For vertical transmission (where the main flanking path is via the OSB or plywood sub-floor in the room above and wall surfaces with directly-attached gypsum board in the room below), a reduction of 1-3 in the Apparent STC could be ascribed to flanking, as shown in Figure 5. The variation could be due to specific construction details, or experimental uncertainty. In the test facility, flanking was suppressed except for one test wall; in a normal building with flanking via several walls, the Apparent STC could be reduced more.

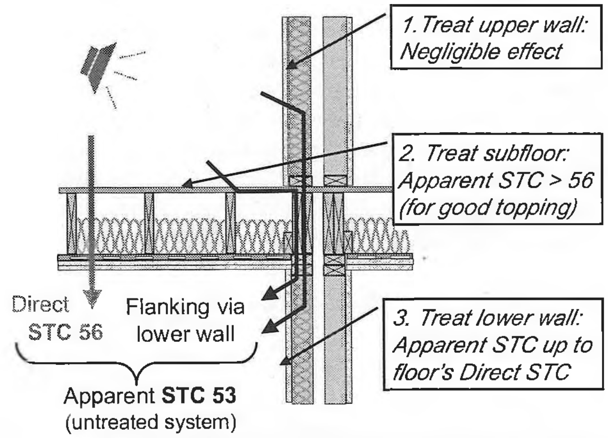



Figure 5: Flanking paths and typical effect on the Apparent STC, with one layer of directly-applied gypsum board on the flanking wall, and a sub-floor of one layer of OSB or plywood.

Because the wall-wall path transmits little flanking energy, the two effective treatments are to modify the sub-floor or the wall below. Mounting the wall's gypsum board on resilient channels increased the attenuation for this flanking path by ~10 dB, making the flanking insignificant. Adding a topping reduced transmission via the flanking path, but also improved attenuation of direct transmission through the floor/ceiling system. As toppings also suppress horizontal flanking, they may be more cost-effective.

## 3. CONCLUSION AND REFERENCES

A study of flanking transmission in wood frame construction has shown that for airborne excitation the floor/wall junction in multifamily buildings provides serious structural flanking. This can be controlled by systematic changes to the floor and wall assemblies.

1. R.E. Halliwell, J.D. Quirt, T.R.T. Nightingale, "Construction details affecting flanking transmission in wood framed multifamily dwellings," Proceedings INTERNOISE 2002
2. T.R.T. Nightingale, R.E. Halliwell, J.D. Quirt, "Vibration response of floors and the effectiveness of toppings to control flanking transmission", Proceedings INTERNOISE 2002
3. T.R.T Nightingale, R.E. Halliwell, J.D. Quirt, (2002) "Flanking Transmission in Multi-Family Dwellings Phase II: Effects of Continuous Structural Elements at Wall/Floor Junctions", *Research Report 103, IRC, NRC Canada*

This work was supported by a consortium that included: Canada Mortgage and Housing, Forintek Canada, Marriott International, Owens Corning, TrusJoist, and USG.

# On Structureborne Power Flow from Wood Studs to Direct-Attached Gypsum Board


T. R. Nightingale[1] and Katrin Kohler[2]

1) Institute of Research in Construction, National Research Council, Montreal Road, Ottawa, K1A 0R6, Canada

2) Department of Biophysics, University of Applied Science Stuttgart, Schellingstrasse 24, D-70174 Stuttgart, Germany


## 1. INTRODUCTION

This paper, the second of two [1], examines the effect of screw location and number on the power flow from a wood stud to directly attached gypsum board. It is necessary to evaluate how an offset distance between the point of an applied force and the fastening point affects the power transmitted from the stud to the gypsum board because if the wall has resilient channels drive points and fastening points will not be aligned. Earlier studies had always considered that the points would be aligned [2].

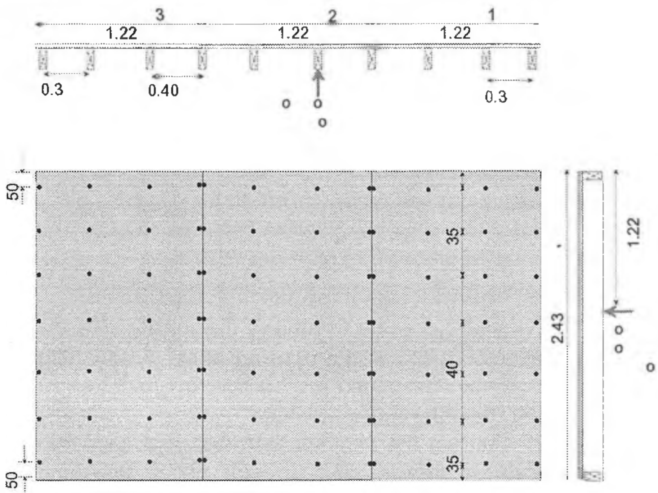



*Figure 1: Sketch of the construction under investigation.*

Figure 1 shows that the wall evaluated in this paper consists of a single layer of 16 mm Type X gypsum board attached to 35 x 85 mm clear western red cedar studs, spaced 406 mm on centre. A single point force was applied to one of the studs. The number and the location of the screws securing the gypsum board to the excited stud were systematically changed and the transmitted power evaluated.

## 2. E ALUATION METHOD

It is not practical to measure the power transmitted by a junction directly – indirect evaluation is necessary. Statistical energy analysis (SEA) may be used if both of the connected elements satisfy the conditions of a subsystem – modes are spaced equally in the frequency band and create a uniform energy density proportional to the damping. This allows one to write,

$$VLD_{12,SEA} = 10\,log\left(\frac{m_2}{m_1}\frac{\eta_2}{W_{12}}\frac{E_1}{\omega}\right) \qquad (1)$$

where subscript 1 indicates source, 2 receiver, and m is mass, W is transmitted power, E is energy, and ω is angular frequency. The equation shows the measured VLD is inversely proportional to the transmitted power, $W_{12}$, and proportional to the energy contained in the source subsystem, $E_1$. This approach will be used in this paper. Mobility models are used to obtain an expression for the transmitted power, $W_{12}$.

To accurately evaluate the power flow it is necessary to determine what elements of the wall form the source and receive subsystems. The source subsystem is the excited stud. The receiving subsystem is more complicated because it may be one or all three sheets of gypsum board. To define the receiving subsystem the vibration level of the gypsum board was measured using a scanning laser vibrometer

Figure 2 shows that at 315 Hz the energy level of plate 2 is significantly greater than that of plates 1 and 3, which are not directly attached to the excited stud. The same trend is exhibited for frequencies above 250 Hz. Below this frequency the vibration levels are significantly more uniform and at about 80 Hz modal patterns are clearly evident. Thus, for 250 Hz and above the receiving subsystem is effectively that of gypsum board plate 2, while below this frequency it would be best to treat the receiving subsystem as being all three gypsum board sheets plus all studs except for the excited one.

The VLD between the excited stud and plate 2 of the gypsum board is used to gage the power flow through the screws.

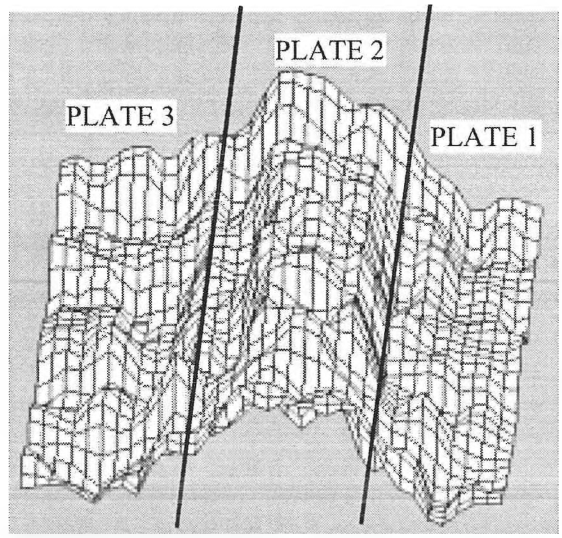



*Figure 2: Surface velocity map of the gypsum board at 315 Hz when excited by the single stud shown in Figure 1. Plate 2, which is excited by the stud, is clearly identifiable by the higher levels.*

## 3. EFFECT OF SCREW LOCATION AND NUMBER

It is has been assumed that the transmitted power is proportional to the number of screws [2]. This implies two things. First, the VLD for a single screw will be independent of screw location. Second, when there are two or more screws, the motion of each is incoherent. Figure 3 clearly shows that with the screw at location "D" (immediately opposite from the source) there is significantly lower VLD (hence significantly more transmitted power) than any other position. VLD ranking indicates less power is transmitted as the fastening point is moved away from the source.

To explain this it is necessary to examine equation 1 and vibration level of the stud in more detail. If the stud were an ideal subsystem, the energy, $E_1$, would be spatially uniform, but highly damped systems, or ones with low mode counts, can give rise to localized energy and the space average result may not be meaningful when there are one or a few point connections.

Figure 4 shows the measured VLD for screw locations B, C, D and E when the energy of the stud near the screw is normalized to a common level. There is considerably better agreement between the sets of data indicating that at points B, C, D and E the ratio of energy, $E_1$, to transmitted power, $W_{12}$, is resasonably invariant.

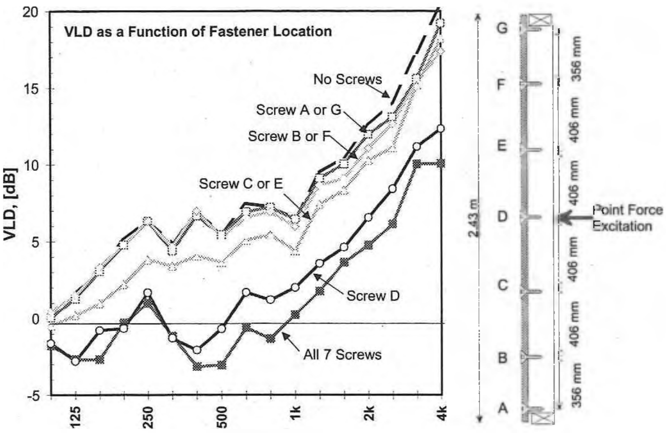



*Figure 3: VLD as a function of the location of a single screw attaching the gypsum board to the excited stud Also shown are the cases with all 7 screws installed and the case with no screws.*

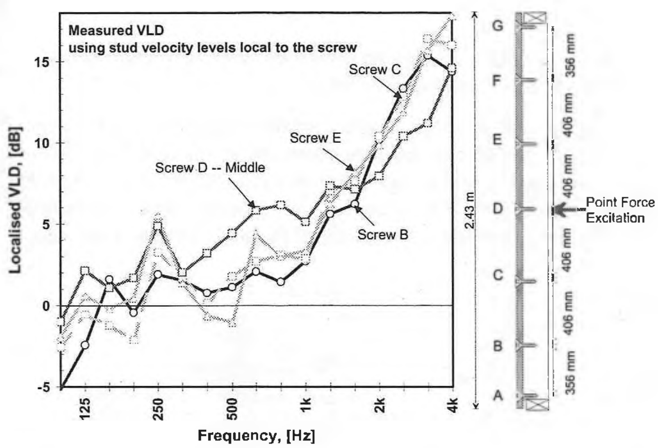



*Figure 4: VLD as a function of the location of a single screw. VLD's computed using stud velocity levels local to points of fastening rather than the space-averaged results used in Figure 3.*

It is important to note that the VLD (of Figure 3) does not become very large when there are no screws even though the head and sole plates of the wall are cut on either side of the excited stud. This suggests that although there are no screws there may still be points where the stud and gypsum board touch and transmit power although less effectively than if there were a screw. This prevents evaluation of points A, F, and G. However, the mobilities at all seven of the screw locations were similar which would suggest that for these points the ratio of energy to transmitted power should be invariant, too.

Having demonstrated that if the energy of the stud is spatially uniform then the transmitted power is reasonably independent of location the next step is to evaluate the assumption that the total power transmitted is proportional to the number of screws. This has been previously observed [2] for all frequencies for which the bending wavelength in the gypsum board (the least stiff element) is less than twice the screw spacing.

Figure 5 shows the measured change in VLD due to removing four (F, E, D and B) of the seven screws. Assuming the stud energy is uniform, the VLD should increase by 10log(7/3) or 3.7 dB. Inspection of Figure 3 and Figure 4 as well as inspection of the stud velocity levels indicates that of the three points A, D and G, only D contributes significantly. Similarly, only E, D, and C contribute significantly when all seven screws are installed. Thus, the predicted change in VLD will be approximately 10log(3/1) or 5.1 dB. Figure 5 shows there is good agreement for frequencies above about 160 Hz. Below 160 Hz one should not expect good agreement because the stud and gypsum board will be effectively line-connected because the bending wavelength in the gypsum board is more than twice the screw spacing (406 mm).

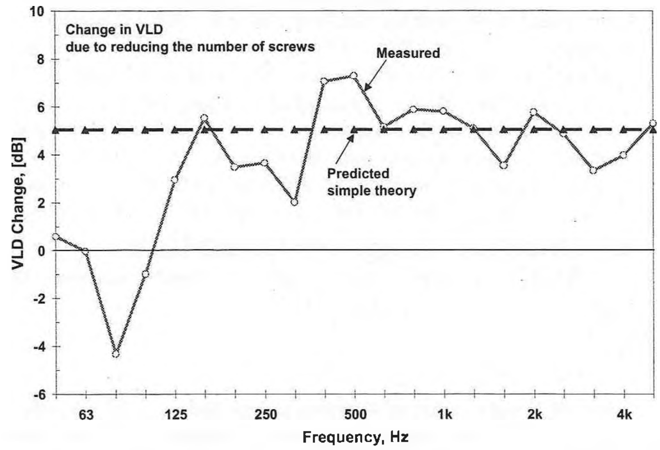



*Figure 5: Change in VLD due to reducing the effective number of screws from 3 to 1.*

## 4. DISCUSSION AND CONCLUSIONS

This paper has shown that a wall consisting of a single layer of gypsum board attached to a single row of wood studs acts as one or more subsystems depending on frequency. Above about 250 Hz, there are three subsystems. The ratio of transmitted power to source (stud) energy is reasonably independent of screw location, which is to be expected as the measured mobilities do not change appreciably if the distance from the edge was greater than 50 mm.

Because the stud is highly damped and has low modal density (i.e., poor subsystem approximation) the energy will not be uniform and the power flow will not be the same at all points. Having only one drive point (point source) represents an extreme case but illustrates the need to recognise that there will be less power transmitted by screws far from the source. Consequently, it may be necessary to include the effect of near field vibration levels in the models.

## 5. REFERENCES

1 T.R.T. Nightingale, Katrin Kohler, (2003), Canadian Acoustics, Vol. 31(3).

2 R.J.M. Craik, R.S. Smith, (2000) "Sound transmission through lightweight parallel plates," Applied Acoustics, Vol. 61, pp. 247-269.

# On the Simplifications Used in Mobility Models to Predict Structureborne Power Flow in Wood Stud Walls with Direct-Attached Gypsum Board

**T. R. Nightingale[1] and Katrin Kohler[2]**

1) Institute of Research in Construction, National Research Council, Montreal Road, Ottawa, K1A 0R6, Canada

2) Department of Biophysics, University of Applied Science Stuttgart, Schellingstrasse 24, D-70174 Stuttgart, Germany

## 1. INTRODUCTION

The mobility approach is a comparatively simple method to compute structure borne power flow between plates and beams coupled at one, or more, well defined points. Implicit to these expressions are a number of assumptions. Simplifications are also made when modelling wood stud wall systems. Using ordinary point force mobilities Craik and Smith [1] predicted with good accuracy the structural power flow across a wood stud wall with direct-attached gypsum board on both sides when the screws on either side of the stud were aligned. This type of alignment cannot be assumed if the wall has resilient channels. Thus it is necessary to evaluate the mobility assumptions to determine the most appropriate method to model walls with resilient channels. In this paper, the first of two, assumptions regarding the vibration response of the stud and gypsum board in isolation are evaluated. The second paper [2] examines the power flow from the stud to the gypsum board as a function of number and location of the screws. The paper begins with a brief review of the mobility expressions.

## 2. MOBILITY MODEL AND EXPRESSIONS

Power flow from a beam (stud) to a point-connected plate (gypsum board) can be written as,

$$W_{12} = N\frac{v_o^2\Re(Y_2)}{|Y_1+Y_2|^2} \tag{1}$$

where N is the number of fastening points, and $v_o$ is the velocity of the source. $Y_1$ is the mobility (inverse of impedance) for the source (stud), $Y_2$ is the mobility for the receiver (gypsum board), given by,

$$Y_1 = \left(2\rho bhc_B(1+i)\right)^{-1} \tag{2}$$

and,

$$Y_2 = \left(8\sqrt{(B\rho h)}\right)^{-1} \tag{3}$$

where ρ is the bulk modulus, h is the thickness, b is the width, B is the bending stiffness, and $c_B$ is the bending wave speed. Equations 2 and 3 are for point forces located far from an edge of a semi-infinite system. In the limit that the excitation point is at an edge these equations must be multiplied by 4 and 8/3.5, respectively. The equations are strictly valid only for systems that behave as if they are infinitely thin – ones for which there is no local deformation at the drive point and there is no deformation of the volume due to the applied force

## 3. ASSUMPTIONS MADE DURING APPLICATION

<u>All fasteners are center-located</u>: It has been suggested [1] that for practical purposes the mobility of the gypsum board at all screw locations can be approximated by the mobility of a point near the center of a large plate. To test this assumption the mobility of a sheet of 16 mm type X gypsum board was measured at 9.5, 19, and 50 mm from an edge as well as at the sheet center.

Figure 1 indicates that each mobility curve approaches the theoretical value (equation 3) for a center location asymptotically but at a different frequency. A point farther from the edge satisfies the assumption at a lower frequency than a point that is closer. Screws into the stud at the top and bottom of the sheet are typically 50 mm from the edge and can be considered to be center located for frequencies above 800 Hz. Between 315 and 800 Hz the edge-location assumption works best. There are few modes in the gypsum board below 315 Hz and the mobility estimates become unreliable and are overestimated. Screws at a butt joint, which are typically 9.5 mm from the edge, should be considered edge-located throughout the building acoustics frequency range.

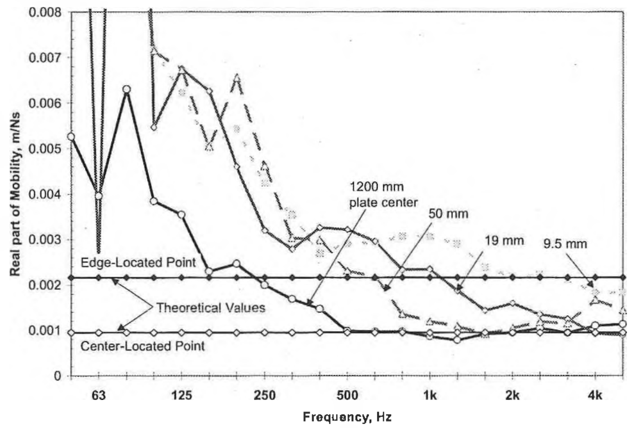



*Figure 1: Measured gypsum board mobility as a function of the distance from the edge of the sheet.*

In the mid and high frequencies the gypsum board mobility is not overly sensitive to location when the point is at least 50 mm from an edge. A similar trend was observed for the stud. Thus, the power flow from the stud to the gypsum board should be reasonably independent of location if the stud velocity is uniform.

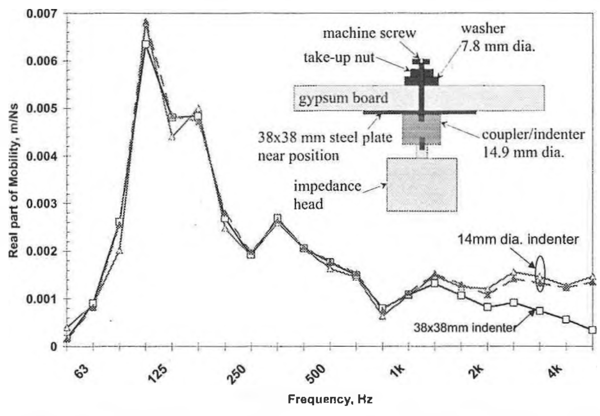



*Figure 2: Measured gypsum board mobility for two indenter sizes.*

<u>Mobility is independent of contact area</u>: Thin plate/beam theory assumes that the area of the drive point is infinitely small. Consequently, mobility expressions based on thin plate/beam theory may not be applicable if measured mobilities show a strong dependence on the size of the drive point, i.e. size of the indenter. To examine the sensitivity the drive point mobility of 16 mm gypsum board was measured using two indenter sizes – a 38 mm×38 mm ×2 mm plate (38 mm corresponds to the stud width) and a 14.9 mm dia. brass cylinder.

Figure 2 shows that the drive point mobility of gypsum board is a function of the contact area above about 1250 Hz where the real part of the mobility with a larger contact area (steel plate indenter) is about 3 dB lower than with the smaller 14.9 mm dia. indenter. Equation 1 indicates that less power will be transmitted from the stud to gypsum board for the larger drive point. This is consistent with that predicted from the advanced mobility theory of Petersson [3]. For a real wall this effect will reverse as the size of the indenter becomes significant compared to the spacing between excitation points.

Plates and beams do not deform volumetrically: This assumption states that the velocity - in magnitude and phase - is the same on both sides of the element. It is implicit in all expressions derived from thin plate/beam theory, e.g., equations 2 and 3.

Figure 3 shows a significant VLD across the depth of a stud for frequencies above about 2000 Hz that increases with frequency. The VLD peak at 630 Hz and 800 Hz between positions A1 and B1near the drive point might be caused by the local deformation and/or near field of the source. Figure 3 shows the assumptions of thin beam theory are not satisfied indicating that the theoretical value given by equation 1 will be a poor estimate above 2000 Hz. This is shown in Figure 4. For gypsum board the VLD was effectively zero and the mobility predicted by equation 2, which is shown in Figure 1, provides an accurate estimate in the mid and high frequencies.

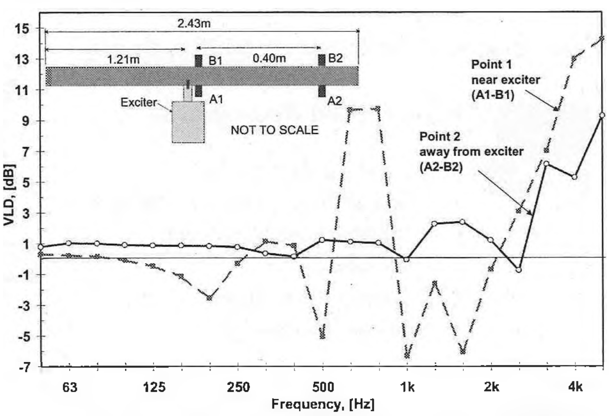



*Figure 3: Measured VLD across the depth of a 38x85 mm red cedar stud*

Mobility is not appreciably affected by the boundary conditions of the plate and beam: Because ordinary mobilities assume that the plate or beam is infinite, the effect of boundary condition is assumed negligible. To assess the effect, the average center mobility of the studs is compared for two boundary conditions. First, the stud is resiliently supported, which approximates the free-free conditions upon which the theory (equation 2) is based. Second, the stud is installed in the test frame but without attached gypsum board. This boundary condition is not free-free and is probably between clamped and simply supported. We restrict the mobility comparison to frequencies above 160 Hz where the point connection assumption is valid [2].

Above 630 Hz, the real part of the mobility for the two boundary conditions is quite similar suggesting that for mid and high frequencies differences in boundary condition are not important. The very close agreement above 2000 Hz may be due to the volumetric deformation where the mobility is largely determined by the behaviour near the drive point. Below 630 there is a noticeable difference – the stud mobility tends to be greater when installed in the wall and agrees better with theory (equation 2). This might seem counter intuitive but it should be recognised that installing the stud in the frame significantly increases the damping which diminishes the importance of individual modes which is important in the low frequencies where the there are few, and perhaps no, modes in some bands.

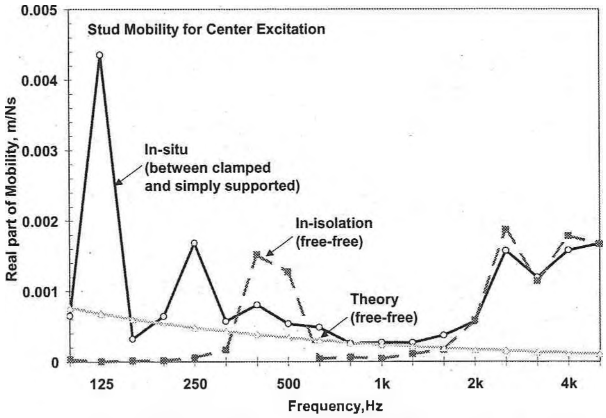



*Figure 4: Effect of boundary conditions on the mobility of 35x85mm red cedar studs.*

## 4. DISCUSSION AND CONCLUSIONS

The review of the primary assumptions used in ordinary mobility models of structure borne power flow indicate there will be a limited frequency range where the models based on thin plate theory can be applied. For the western red cedar studs considered here the upper frequency limit is determined by volumetric deformation of the stud. Volumetric deformation may not be important for other wood species; especially ones that have a comparatively high shear modulus or have knots. Other species and quality grades will be examined in subsequent phases of this project. Advanced mobility theories [3,4] account for the effect of volumetric and local deformation.

For fastening points located at least 50 mm from an edge the mid and high frequency mobility can be reasonably approximated by a center-mobility and that boundary conditions should have little effect. A low mode count prevented examination in the low frequencies.

The high frequency mobility of gypsum board is a function of the area of the drive point. Consequently, the gypsum board mobility measured with a small indenter may agree well with theory (Figure 1) but might be significantly different from the in-situ mobility seen by the stud where there the area over which the force is applied may be considerably larger (Figure 2). The effect of contact area on the mobility wood studs and gypsum board should be investigated further.

## 5. REFERENCES

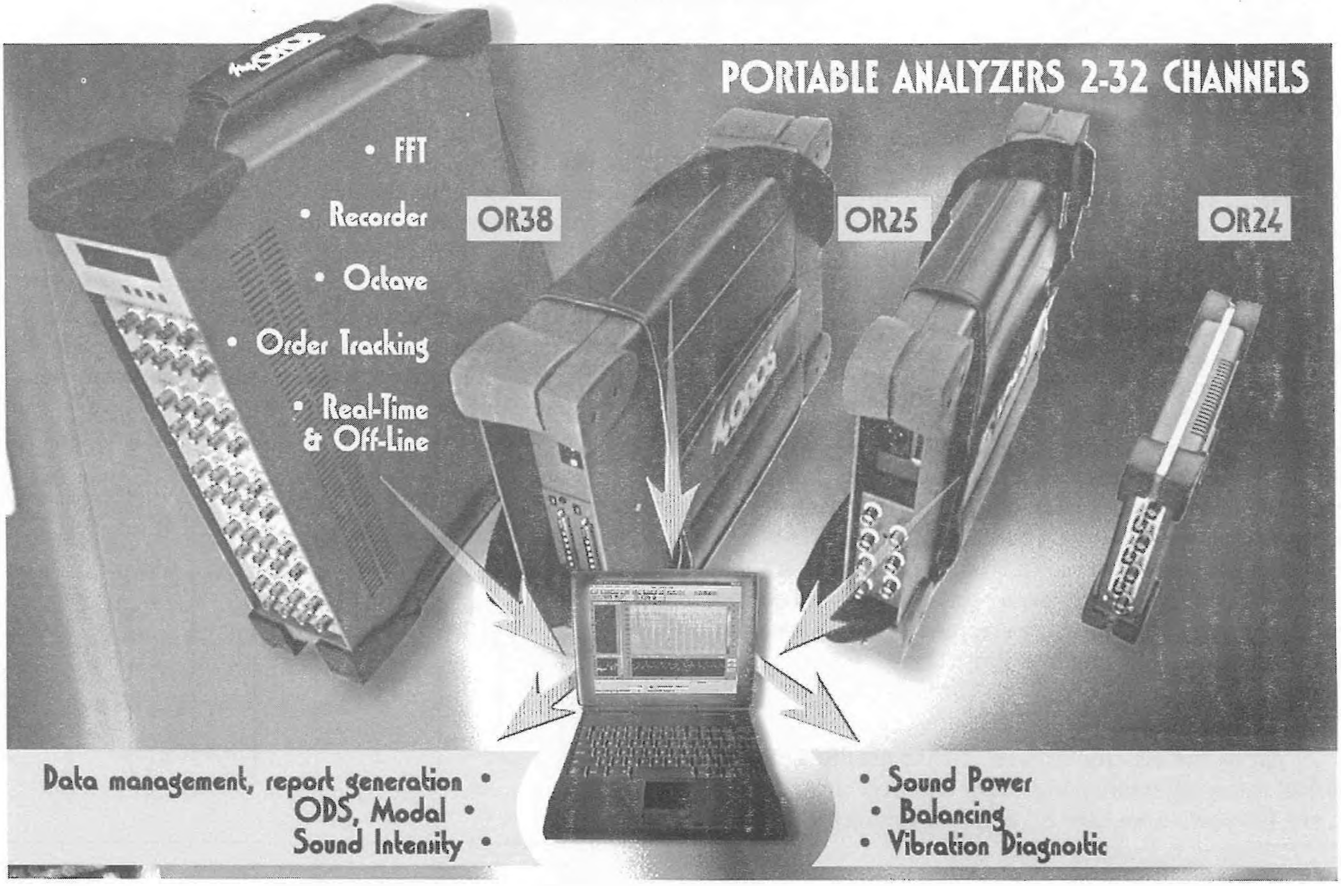

1 R.J.M. Craik, R.S. Smith, (2000) Applied Acoust., Vol. 61, pp. 247-269.

2 T.R.T. Nightingale, Katrin Kohler, (2003) "On structure borne power flow from wood studs to direct-attached gypsum board," Canadian Acoustics Vol.31 No. 3.

3. B.A.T Petersson, M. Heckl, (1996) J.S.V. Vol. 196(3), pp. 295-321.

4 B.A.T Petersson, M. Heckl, (1999) J.S.V. Vol. 224(2), pp. 243-266.

# WANT MORE CONFIDENCE IN YOUR MEASUREMENTS?




# THEN ASK FOR THE COMPLETE RANGE OF OROS ANALYZERS & SOLUTIONS

• **Need a complete solution, guaranteed by a recognized name in the NV field?** Just choose an OR24, OR25 or OR38 analyzer and select the application package you need.

• **Think you may need more in the future?** With the OROS range, you can expand your hardware as well as add software modules at any time.

• **Can't take the time to repeat your tests?** The new OR38 records the original time domain data on the internal disk from 1 to 32 channels, and can also run several analyzers simultaneously. This unprecedented flexibility also ensures never losing your data.

OROS Analyzers and Solutions cover a wide range of measurement needs in the noise and vibration field.
For structural, rotating as well as acoustics measurements, OROS offers complete solutions.

| | Channel count | Frequency band | Interface | Calculation | Recording | Power | Weight |
|---|---|---|---|---|---|---|---|
| OR24 | 2, 4 | 20kHz | PCMCIA | DSP | PC disk | AC, external DC-Batt | 2kg |
| OR25 | 2, 4, 8, 16 | 20kHz | PCMCIA | DSP | PC disk | AC, DC, Batt | 5kg |
| OR38 | 8, 16, 24, 32 | 40kHz | LAN | DSP or PC | Internal or PC disk | AC, DC, Batt | 7kg |

OROS

**Dalimar Instruments Inc.**

At your service since 1986

Montreal : Tel. : (450) 424-0033 Fax : (450) 424-0030 • Toronto : Tel. : (905) 707-9000 Fax : (905) 707-5333

**E-mail: info@dalimar.ca • WEBSITE: WWW.DALIMAR.CA**

design by genco oil - france: +33 476 992 999

# ACOUSTICAL EVALUATION OF BERWICK PRESCHOOL

**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 the major concern in classroom acoustics, since speaking and hearing are the most important modes of communication in teaching and learning.

Findings [1,2] indicate that academic achievement - in particular reading skills - are vulnerable to the effects of chronic noise exposure. The general pattern of adverse psychological stress reactions is associated with chronic exposure to noise among children [3].

The teachers in many preschool and child care - as well as, in the Berwick Preschool studied here - frequently commented on how noisy the classrooms were. For many who work in preschool classrooms, there is a tendency to simply accept the fact that young children can be noisy. Regardless of the many reasons, it is clear that there is a fair amount of noise in preschool classrooms. Individuals working in early childhood classrooms tend to tolerate the noise as the "price of doing business".

## 2. EVALUATION

All of the classrooms were architecturally identical and had same classroom equipment. The floor area was 73 $m^2$, and the volume was 269 $m^3$ in the classroom.

### 2.1 Objective measurement

As shown in Figure 3.3, all of the classrooms had $RT_{mid}$'s that exceed the 0.6 seconds limit of ANSI S12.60-2002 for core learning spaces with enclosed volumes less than 283 $m^3$. The variation of the RTs in the different classrooms may be caused by the different furniture layouts and the toys, as well as usual experimental variations.

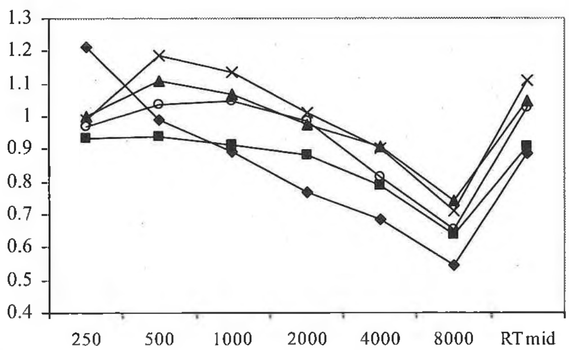



Fig 1. Measured reverberation times (♦:Rm1, ■ :Rm2, ▲: Rm3, ×:Rm4, ○:Rm5)

The background noise levels exceeded the 35 dB(A) limit of ANSI S12.60-2002 in all of the classrooms.

Figure 3.9 shows the range and average of the sound levels for each teacher and classroom. Average sound levels during various activities differed for the different classrooms and teachers. This difference may be caused by the different noise levels made by children, and teacher's different teaching styles. Considering the microphone's position in the measurement, children in the classroom were exposed the sounds over 70 dB(A), and teachers were exposed to sounds over 83 dB(A) for 5 hours a day.

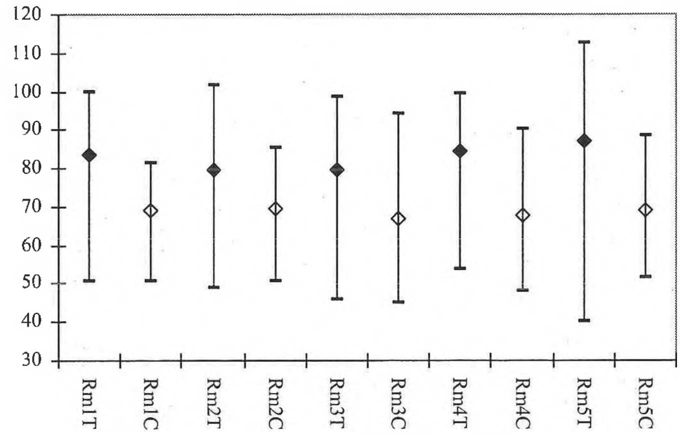



Fig 2. Average sound levels during the school activities (♦: teacher, ◊:classroom)

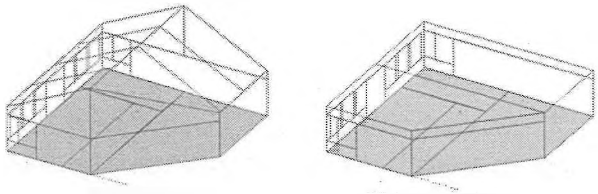
### 2.2 Subjective Survey

Eleven of a total of thirteen teachers in the Preschool responded to the questionnaire. The listening environments in the classrooms were considered less acceptable than the non-listening environments.

Teachers were asked about their perception of the consequences of a poor listening environment in the classrooms. 'Increased fatigue' was frequently experienced by most teachers in the classrooms. Teachers were also asked to assess the interference with their ability to hear because of different sources of noise inside and outside the classrooms. In particular, children's talking was considered as a major source of noise in the classrooms.

## 3. SIMULATION

Basically, acoustical treatment of the classrooms involved controlling the material and the volume of the classroom. The basic materials for the floor, wall, and ceiling were selected based on the current classroom materials, and each component was changed to a different material, one by one. CATT room acoustical simulation software was used for prediction.



(a) Current (b) Lowceiling

Fig 3. Classroom models

Table 1. Classroom configurations

| Control | Name | Area ( $m^2$ ) by material | | | | | |
|---|---|---|---|---|---|---|---|
| | | Rubber sheet | Thin carpet | Plywood | ¾" porous | Ceiling tile | Glass |
| No treatment | Current | 52.2 | 21.0 | 177.2 | - | - | 11.2 |
| Floor | Floor | 73.2 | - | 177.2 | - | - | 11.2 |
| | Carpet | - | 73.2 | 177.2 | - | - | 11.2 |
| Wall | Wall | 52.2 | 21.0 | 144.8 | 32.4 | - | 11.2 |
| Ceiling | Ceiling | 52.2 | 21.0 | 95.4 | 81.8 | - | 11.2 |
| Volume $178m^3$ | Lowceiling1 | 52.2 | 21.0 | 147.5 | - | - | 11.2 |
| | Lowceiling2 | 52.2 | 21.0 | 95.4 | 52.2 | - | 11.2 |
| | Lowceiling3 | 52.2 | 21.0 | 95.4 | - | 52.2 | 11.2 |

Figure 4.6 shows the predicted RTs in the 125 Hz to 4000 Hz octave bands, as well as $RT_{mid}$. Both changing to higher absorptive material and reducing volume were effective at decreasing the RT in the classroom, to acceptable values.

The speech level of the speaker is a major component of speech intelligibility. However, it is inversely related to the total absorption of the room. Adding absorption decreases both speech levels and RT.

Figure 4.8 shows the RASTI with background noise in unoccupied classroom. None of them exceeded 0.75 which corresponds to excellent speech intelligibility. Therefore, the classroom does not have good conditions for speech even when it is unoccupied. The RASTI was simulated in more realistic condition with classroom activity noise as shown in Figure 4.9. With the higher background level, the RASTI was higher when the RT was longer in general.

## 4. DISCUSSION

None of the classrooms in the preschool was acceptable according to the criteria relevant to this study: ANSI S12.60-2002. Based on the in-class sound levels, teachers apparently always talked with loud voices during the class. The sound levels to which the teachers were exposed were close to typical occupational noise limits. Teachers agreed that the non-acoustical environments in the classrooms were fair, but the acoustical environments had problems.

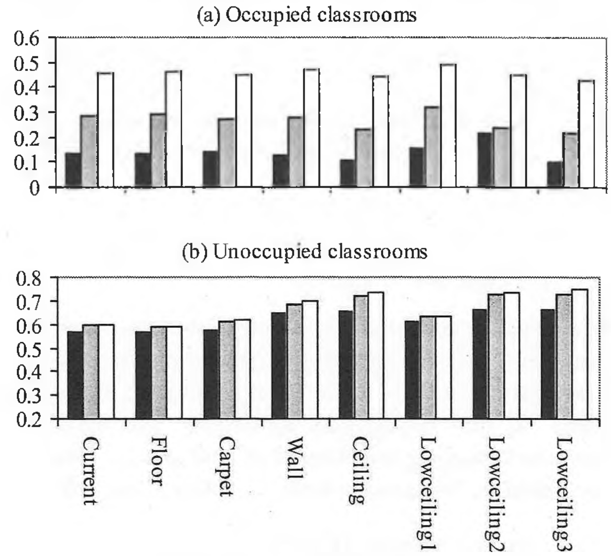



Fig 4. RASTI (■: normal voice, ▒: raised voice, □: loud voice,)

In noisy environment, the speech level and volume of the space were dominant factors for speech intelligibility; however, in quiet environment, the absorption was more effective than speech level or volume. Ceiling heights are critical as well. While the existing heights created interesting looking spaces, they were problematic in terms of noise levels and reverberation.

Decreasing the volume of the classrooms would be the most effective solution. Installing a suspended acoustical ceiling would be an option. They decrease the RT and increase RASTI.

## REFERENCES

1. E.E.Manlove, T. Frank, and L. Vernon-Feagans, 'Why should we care about noise in classrooms and child care settings?', *Child & Youth Care Forum*, **30**(1), 55-64, 2001

2. L. E. Mazwell and G. W. Evans, 'The effects of noise on pre-school children's pre-reading skills', *J. Environ. Psych.*, **20**, 91-97, 2000

3. G.W. Evans, S. Hygge, and M. Bullinger, 'Chronic noise and psychological stress', *Psych. Science*, **6**(6), 333-338, 1995

4. Accredited Standards Committee S12, Noise, America National Standard, *Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools*, ANSI S 12.60-2002

## ACKNOWLEDGEMENTS

The authors wish to express their thanks to the Berwick Preschool for the opportunity to test; thank es pecially due to Ms. Diane Burger, director of the preschool for her close cooperation. Likewise we are grateful to the teachers in the preschool.

# Classtalk System For Predicting And Visualizing Speech In Noise In Classrooms

**Murray Hodgson**
Dept. of Mechanical Engineering and School of Occupational and Environmental Health,
University of British Columbia, 3rd Floor 2203 East Mall, Vancouver, BC, Canada, V6T 1Z3
hodgson@mech.ubc.ca

## 1. INTRODUCTION

The objective of the work reported here was to develop a classroom speech-intelligibility prediction tool, applicable to typical classrooms, which is accessible to non-specialist users. It was required to be simple, fast, accurate and interactive, running on standard PC computers. The system developed - a Windows system - is called *ClassTalk*.

## 2. FUNCTIONALITY

*ClassTalk* allows a classroom to be modeled, sound-field descriptors to be predicted, and for the classroom and predicted quantities to be visualized on the computer monitor. In real time, the user can 'walk-through' the classroom on the monitor, exploring the sound field and associated speech quality. Using *ClassTalk*, the user can define the physical and acoustical characteristics of a classroom and its noise sources. *ClassTalk* takes into account occupant sound absorption and student-activity noise, both crucial to the accurate prediction of classroom acoustics. The classroom floor-plan is visualized on the monitor, along with the speech- and noise-source positions, and a receiver icon indicating the receiver position. The user can 'walk-through' the classroom at will by moving the receiver icon. At every walk-through position, in real time, five outputs - speech level, background-noise level, signal-to-noise level difference, Speech Transmission Index (*STI*) and speech intelligibility - are calculated and displayed. Contour maps of the five predicted quantities can be plotted. Also calculated and displayed are reverberation times for both the unoccupied (i.e., as designed by the architect) and occupied (i.e., as experienced by the occupants) classroom. The classroom floor-plan display can be printed, along with the input data, in the form of a one-page prediction report. The values of the five predicted quantities at the receiver-grid positions can be exported to a file for further processing.

The input characteristics of the classroom that must be entered are its description, dimensions, sound-absorbing features, and the coordinates and sound-power levels of the sources. Student-activity noise can be considered or not, as desired. Sound-absorbing features include carpeting, acoustical treatment of the walls and ceilings, upholstered seating and students, who absorb propagating sound. Sources can be male or female speakers, three types of projector or ventilation outlets. At any time, the classroom data can be edited - for example, to introduce control measures. Control can be achieved by modifying the room geometry, adding surface absorption, and reducing ventilation-outlet and equipment output levels. The modified classroom and the new values of the five acoustical quantities are displayed. The classroom can be 'walked-through' with real-time update of the visual display. New noise contours can be drawn.

## 3. PREDICTION MODELS

*ClassTalk* uses novel, simplified empirical models to predict total A-weighted speech levels (*SLA*) and noise levels - *BGNA*, due to ventilation (*VNA*) and student-activity (*SANA*) noises, as well as 1000-Hz early-decay times (*EDT1*). Noise levels at a receiver position are calculated by summing the contributions of the individual noise sources. Individual speech- and noise-level contributions are determined from the predicted sound-propagation curves describing the rate of decrease of levels with increasing distance from a source, as well as from their absolute levels. The empirical models were developed from a database of 'typical' university classrooms measured when unoccupied and occupied. These account for speaker voice level (and how it adapts to the prevailing acoustical conditions), noise-source output levels, classroom shape and absorption (including occupants), as well as student-activity noise, as they occur in real classrooms. Full details of the prediction models can be found elsewhere [1, 2, 3].

## 4. INPUT DATA

### 4.1 Sound-Absorbing Features

Entering the extent of the sound-absorbing features in the classroom involves inputting the percentage of the side walls and of the ceiling covered with acoustical treatment, as well as the percentage floor area that is carpeted. Since students absorb sound, the number of students inside the classroom is specified; the resulting absorption is determined from the 1000-Hz absorption per student [4]. Seating can be defined as hard or acoustically-absorbent upholstered seating. The total 1000-Hz surface absorption coefficient is found by summing the contributions of the untreated and sound-absorbing surfaces [1].

## 4.2 Speech and Noise Sources

The typical ranges of total A-weighted output sound-power levels of male and female speakers were determined from published data [5]. Those of the four types of noise source - overhead, slide and LCD projectors, and of ventilation outlets - were determined from extensive sound-intensity measurements [6]. From the ranges, values corresponding to four output levels (referred to as quiet, normal, raised and loud) were assigned.

# 5. EXAMPLE PREDICTIONS

Figure 1 shows *ClassTalk* prediction results for a large classroom, before and after treatment. Student-activity noise was not considered. The figures show contour maps of speech intelligibility (*SI*o), and the full *ClassTalk* visual displays, for the two cases. They include the unoccupied and occupied reverberation times, and the values of the predicted quantities at the walk-through position indicated by the square receiver icon.

The classroom has dimensions of 24 m by 22 m by 6 m high, and contains 400 occupants. Before treatment, it has non-absorptive seats, walls and ceiling; the floor is 80% carpeted. A female instructor, speaking in a normal voice, is located 2.5 m from the front wall. An overhead projector, with normal output level, is located near the instructor. An LCD projector, with normal output level, hangs from the ceiling in the middle of the classroom. Three ventilation outlets, with loud output levels, are located on one side-wall. Speech levels vary with distance from the instructor from 55 to 39 dB. Noise levels vary from 49 dB near the noisy ventilation outlets, to 39 dB far from them. Signal-to-noise level differences vary from 12 dB near the instructor to -9 dB near the ventilation outlets. Speech intelligibility varies from 94% ('good' quality) near the instructor to 39% ('bad' quality) near the vents. Reverberation times vary from 1.1 to 2.8 s in the unoccupied classroom, and from 0.6 to 1.7 s in the occupied classroom.

Acoustical treatment of this classroom aimed to reduce noise, and to control reverberation to near optimal values [7]. It consisted of HVAC noise control (reducing the ventilation outlets to 'normal' output levels), applying sound-absorptive materials to 50% of the ceiling and side walls, and installing sound-absorbing upholstered seating. After treatment, speech levels vary with distance from the instructor from 50 to 36 dB (reductions of 3 to 5 dB). Noise levels vary from about 35 dB near the overhead and video projectors - now the dominant noise sources - to 28 dB far from noise sources (reductions of 10 to 13 dB). Signal-to-noise level differences vary from 18 dB (an optimal value) near the instructor, to 6 dB under the video projector, and to 8 dB near the ventilation outlet farthest from the instructor (increases of 6 to 17 dB). Speech intelligibility varies from 94 to 96% ('very-good' quality) throughout the room (increase of 3 to 55 %). Reverberation times vary from 0.6 to 0.8 s in the unoccupied classroom, and from 0.5 to 0.7 s in the occupied classroom. The sound absorption - in particular, the upholstered seating - has reduced the sensitivity of the reverberation times to occupancy.

# 6. CONCLUSIONS

*ClassTalk* achieves the objective of developing a classroom prediction tool that is accessible to the non-specialist. A demo version is available from www.flintbox.ca.


[1] M. R. Hodgson, *J. Build. Acoust.* **8**(1) 1-14 (2001).

[2] M. R. Hodgson, R. Rempel and S. M. Kennedy, *J. Acoust. Soc. Am.* **105**(1) 226-233 (1999).

[3] M. R. Hodgson, *J. Acoust. Soc. Am.* **112**(2) 568-575 (2002).

[4] M. R. Hodgson, *J. Acoust. Soc. Am.* **106**(4) 1810-1819 (1999).

[5] *American National Standard Methods for Calculation of the Speech Intelligibility Index,* ANSI S3.5-1997.

[6] R. Wang, Project report, U. British Columbia (2001).

[7] M. R. Hodgson and E.-M. Nosal, *J. Acoust. Soc. Am.* **111**(2) 931-939 (2002).

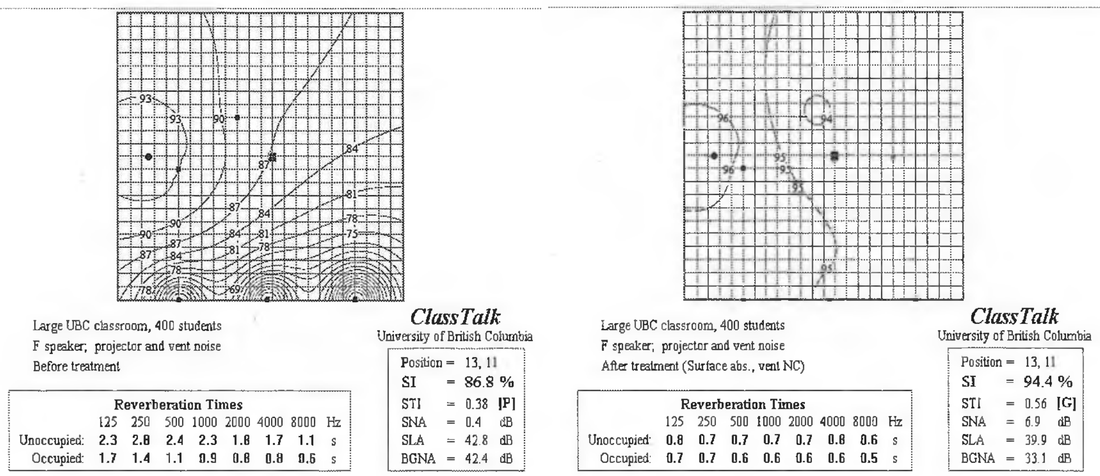




Figure 1. *ClassTalk* predicted sound-field and speech-quality descriptors for a large classroom before (left) and after (right) acoustical treatment. Classroom and treatment details are in the text.

# Active Noise Control In Non-Diffuse Three-Dimensional Enclosures With High Modal Density: Theoretical Studies

**Desheng Li and Murray Hodgson**
School of Occupational & Environmental Hygiene and Department of Mechanical Engineering, University of British Columbia, 3rd Floor, 2206 East Mall, Vancouver, BC, Canada, V6T 1Z3 - Email: desheng@mech.ubc.ca

## 1. INTRODUCTION

In this paper, the active control of low-frequency noise in rooms is investigated. A novel design method called the Image-GA method is proposed. In this method, a noise field is modeled using the image-source method (Guo and Hodgson, 1999), and Genetic Algorithms (GAs) and the quadratic optimization method are combined to optimize locations of control loudspeakers and error microphones. The introduction of the image source model allows the consideration of different acoustical characteristics of walls. Generally speaking, global control is very difficult, if not impossible, to realize in a large enclosure. The local control strategy, which can only assure sound cancellation at error microphones, is not very useful in practice. Considering that workers usually work only in a certain area of a workroom, a new control strategy called "locally-global" control, using the reduction of the acoustical potential energy in the target area as the cost function, is proposed. Numerical investigations were reported in this paper. Experimental results will be reported in next paper.

## 2. METHOD

2.1 Image-source model

In a free field, the sound pressure emitted by a point source can be expressed as

$$p(\mathbf{X},\mathbf{X}') = q\frac{i\omega\rho_0 e^{-ikR}}{4\pi R} = qZ_R \tag{1}$$

where $q$ is the complex source strength, $Z_R$ is the complex acoustical transfer impedance from the source position to the receiver position, $\omega$ is the angle frequency, $k$ is the wavenumber, $\rho_0$ is the air density, $\mathbf{X}$ is the source position $(x,y,z)$, $\mathbf{X}'$ is the receiver position $(x',y',z')$, and R is the distance between the source and the receiver.

When a rigid wall is present, the sound at a receiver position will be the sum of the direct sound and reflections from the wall. To model a reflection, one may assume there is an image source located symmetrically on the other side of the wall. Thus, the sound pressure becomes the sum of the noise from the original and image sources.

For a room with low sound absorptive walls, the sound pressure can also be calculated using Eq. (1). In this case, the acoustical transfer impedance can be written as

$$Z_R(\mathbf{X},\mathbf{X}') = \frac{i\omega\rho_0}{4\pi}\sum_{p=0}^{1}\sum_{r=-\infty}^{\infty} \beta_{x1}^{|n-q|}\beta_{x2}^{|n|}\beta_{y1}^{|l-j|}\beta_{y2}^{|l|}\beta_{z1}^{|m-s|}\beta_{z2}^{|m|}\frac{e^{-ik|\mathbf{R_p}+\mathbf{R_r}|}}{|\mathbf{R_p}+\mathbf{R_r}|} \tag{2}$$

2.2 Placement optimization of control loudspeakers

The strong ability of Genetic Algorithms in dealing with complicated problems has been shown in many research areas. Therefore, Genetic Algorithms are employed here as a tool to find the optimal placement of the control loudspeakers and error microphones.

For each searched configuration of control loudspeakers, the control output is optimized using the quadratic optimization method. As is known, the minimization of the sound field at discrete error sensor locations does not guarantee the best results in terms of sound reduction throughout the enclosure. Therefore, some global error criterion is preferred. In a small enclosure such as vehicle cabins, global control can be relatively easy to realize because of the low modal density in the low-frequency range. However, for a large-sized room, the modal density will be high even over the low-frequency range. This makes it very difficult to realize sound attenuation throughout the enclosure with a limit number of control sources. Actually, in a large workroom, workers usually work in a certain area. Hence, one may use the acoustical potential energy in the area as an error criterion. This control strategy can be called "locally–global" control.

Since a linear system is considered, the acoustic potential energy can be expressed as the sum of the primary and secondary sources

$$E_p = \mathbf{q_c^H A_E q_c} + \mathbf{q_c^H b_E} + \mathbf{b_E^H q_c} + \widetilde{E} \tag{3}$$

By minimizing the acoustical potential energy, one can obtain the vector of optimum control source outputs. The difference between the acoustic potential energy levels before and after control is used as the cost function in the Genetic Algorithms to evaluate the fitness of the control source configuration. The higher the difference, the higher is the fitness value assigned to the configuration. At the end of the search, one optimum configuration, or at least a sub-optimum one, will be obtained.

2.3 Placement optimization of error microphones

After the optimum configuration of control sources is obtained, the optimum position of the error microphones are searched in the target area using Genetic Algorithms. For each searched configuration, the optimum control output is calculated by minimizing the sum of the squared sound pressures at the error microphones

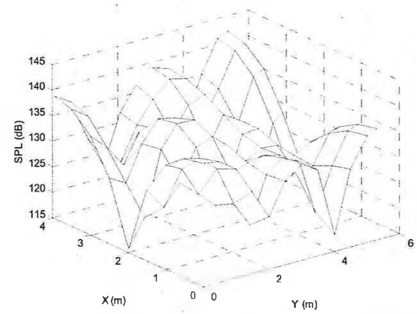
Then the reduction of acoustic potential energy in the target area under this optimum control output is used as the cost function in Genetic Algorithms. At the end of the search, the optimal, or at least a sub-optimal, configuration of error microphones can be found.

## 3. NUMERICAL RESULTS AND ANALYSIS

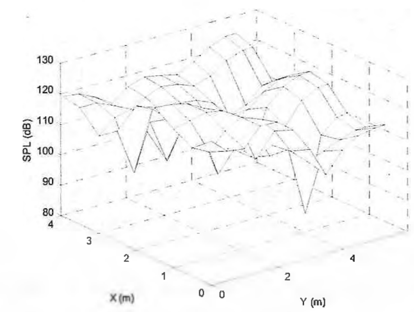
In the numerical investigation, a 2 by 2 control system was designed for a room of 5.3 m long, 3.95 m wide and 2.72 m high. The reflection coefficients were assumed to be, respectively, 0.938 for the four walls, 0.99 for the floor, and 0.812 for the ceiling. One point source placed in one of the corners at (0,0,0) generates noise at 100Hz. The source strength is $0.5m^3/s$. The control area is set to be the whole plane at z=1.60m, the typical height of human ears. The element dimension is 0.5 m, both in the x direction and in the y direction. Theoretically, an infinite number of image sources should be used in the calculation. However, in practice, a finite number has to be used. The image number was determined using the criterion in the literature (Guo and Hodgson, 1999), which is 60 in the research.

As is known, when a control source is placed within a half-wavelength distance from a noise source, global control can be realized. However, this is impossible in many practical situations. To make this research more meaningful and general, the clearance constraint of half-wavelength in at least one direction was imposed. Without this clearance constraint, after optimization, the control sources tend to be located at the same position as the noise source.

The optimum positions for the control sources are (2.42, 0.85, 0.91) and (1.15, 2.22, 0.54). The two error microphones are optimally placed at (2.29, 1.20, 1.60) and (1.91, 3.42, 1.60). Figures 1 (a) and (b) show the sound field before and after control, respectively. From this figure, one can see that a significant sound reduction is achieved at most positions using the optimally designed 2 by 2 control system. The reduction of acoustical potential energy in the target area is 13.3 dB. However, after control, the sound pressure level increases at a few positions which had relatively low noise levels before control. This often happens when the global control strategy is used in control system design.




(a) Before control




(b) After control

Figure 1: Control performance of the optimal control system

## 4. Conclusions

In this paper, active control of low-frequency noise in three-dimensional rooms was investigated. A novel design method called the Image-GA method and a new control strategy called "locally-global" control were proposed. The numerical results show that, through optimal design, significant sound reduction can be achieved in the target area. The experimental validation will be reported in next paper.

## REFERENCES:


Guo, J. and Hodgson, M. (1999), "Investigation of active noise control in non-diffuse sound fields", Proceedings of ACTIVE 99, 621-632.


## ACKNOWLEDGEMENTS


This work is supported by the UBC Bridge Program and NIH of the United States.

# Active Noise Control In Non-Diffuse Three-Dimensional Enclosures With High Modal Density: Experimental Studies

**Desheng Li and Murray Hodgson**
School of Occupational & Environmental Hygiene and Department of Mechanical Engineering, University of British Columbia, 3rd Floor, 2206 East Mall, Vancouver, BC, Canada, V6T 1Z3 - Email: desheng@mech.ubc.ca

## 1. INTRODUCTION

In the previous paper, active control of low-frequency noise in three-dimensional rooms was theoretically investigated. The simulation results showed that, through optimal design, significant sound reduction can be achieved. In this paper, the experimental studies are reported. Experiments were done with optimal and non-optimal systems in a rectangular room.

## 2. EXPERIMENTAL RESULTS AND ANALYSIS

The room for the experiments is 3.95 m wide (x direction), 5.3 m long (y direction), and 2.72 m high (z direction). The ceiling is treated with mineral-fiber acoustic tile. The floor is made of vinyl tiles on concrete. The walls are double-plasterboard stud construction. The Filter-x LMS controller is used. A loudspeaker is placed in the corner (0,0,0) as the noise source. It produces a 100 Hz pure tone noise. The target control area is the whole plane at 1.60m above the floor, the typical height of human ears. The measurement grid had a 0.5 m interval in the x and y directions. The reflection coefficients of the room surfaces are estimated by measuring the reverberation time, which is 0.938 for the walls, 0.99 for the floor, and 0.812 for the ceiling. The background noise from the ventilation system is around 65 dB.

An optimal 2 by 2 control system designed using the Image-GA method (Li and Hodgson, 2003) was implemented in the room. Figure 1 shows the configuration of the optimal control system. The positions for the two control loudspeakers are (2.42, 0.85, 0.91) and (1.15, 2.22, 0.54), respectively. The two error microphones are optimally placed at (2.29, 1.20, 1.60) and (1.91, 3.42, 1.60), respectively.

First, the control signals and error signals were recorded as shown in Figure 2. From this figure, one can observe that, once the control was on, the system converged very quickly and stably. The residue signals in the error microphones are background noise. However, the objective is not just to attenuate the noise at error sensor positions. To investigate the control performance in the whole target area, the sound field was measured before and after control, respectively. The results are shown in Figure 3. The horizontal and vertical axes correspond to the length and the width of the room, respectively. One can see that after control, sound reduction is achieved in most parts of the target area. The maximum sound reduction of 33.2 dB is achieved at (2.0, 3.5, 1.60), a position close to the second error microphone. The average sound reduction is 7.8 dB. It can also be observed that, at some positions where original noise is low, the noise increases. This is because that in the design, the "locally-global" control strategy (Li and Hodgson, 2003) is employed. With this control strategy, the control system will tend to attenuate noise at positions where is loud and increase noise a little bit at positions where is quiet.

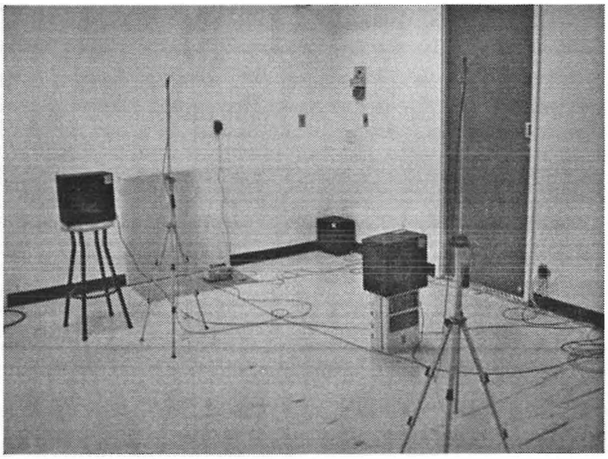


Figure 1: Configuration of the optimal control system

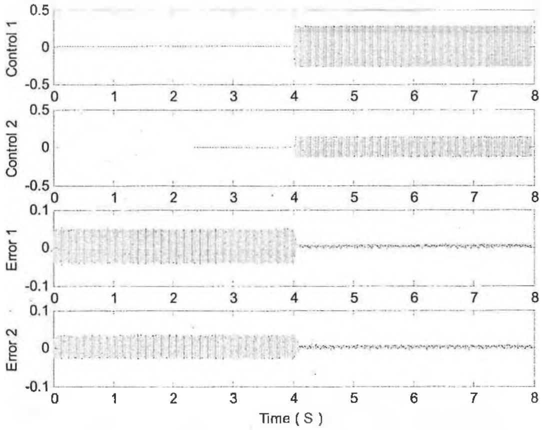
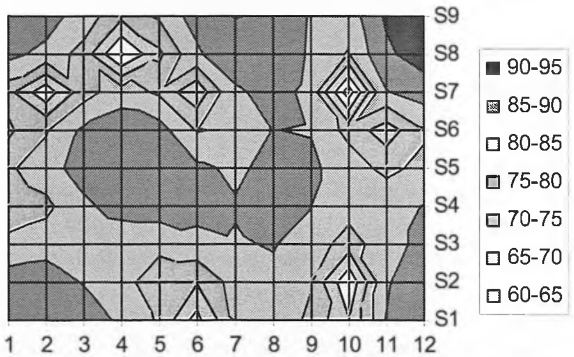
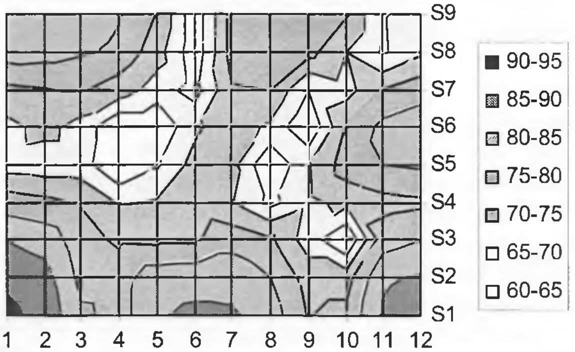



Figure 2: Control signals and error signals




(a) Before control




(b) After control

Figure 3: Sound fields before and after control

To show the effectiveness of optimal design, the experiment was also done with a randomly configured control system, which is shown in Figure 4. The positions for the two control loudspeakers are (2.15, 1.50, 0.91) and (0.75, 3.26, 0.54). The two error microphones were placed at (2.46, 2.95, 1.60) and (2.20, 4.50, 1.60).

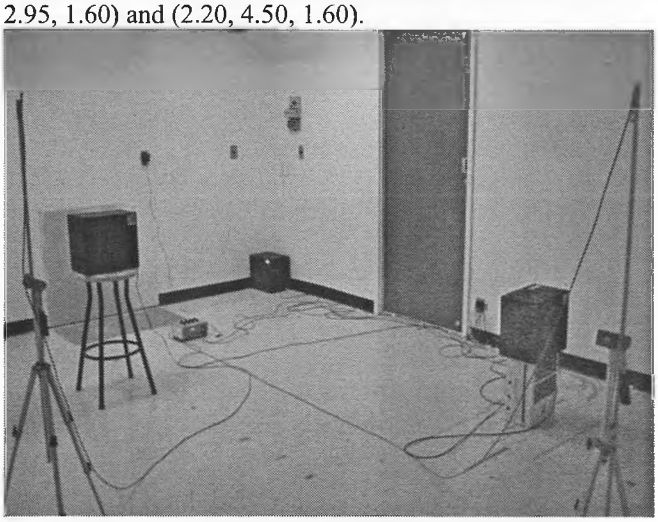


Figure 4: Randomly configured control system

Again, the noise was measured in the target area. The sound field after control is shown in Figure 5. One can observe that, with the randomly configured system, noise is increased at most positions instead of being attenuated. The average increase is 9.7 dB.

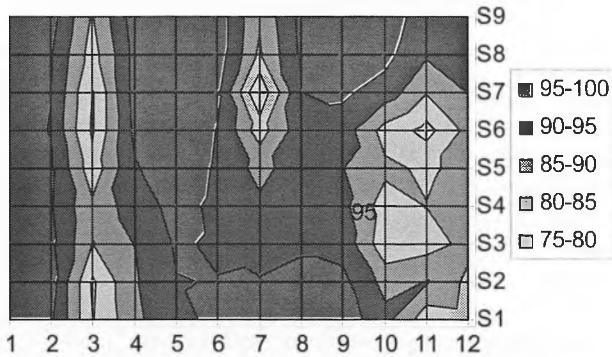



Figure 5: Sound field after control with a randomly configured control system

# 3. CONCLUSIONS

In this paper, the experiments were reported on active noise control in a three-dimensional room. A 2 by 2 control system was designed using the Image-GA method, and implemented. Experimental results showed that, with the optimally designed control system, significant sound reduction can be achieved even in a three-dimensional room with high modal density. To further show the effectiveness of the optimal design, a randomly configured control system was also implemented. With this system, the noise was increased instead of being attenuated. The research showed that active control technology can be used to attenuate the low-frequency noise in a three-dimensional room. But the optimal design is critical to ensure good sound reduction.

# REFERENCE


Li, Desheng and Hodgson, Murray (2003) "Active noise control in non-diffuse three-dimensional enclosures with high modal density: theoretical studies", Acoustic week in Canada, Edmonton.


# ACKNOWLEDGEMENT


The work is supported by UBC Bridge Program and NIH of the United States. The authors would like to thank Mr. Haoming Qin and Mr. Tim Luu for their help in the experiment.

# Multi Channel Data Acquisition and Signal Processing Using MATLAB

**S. Bilawchuk and K.R. Fyfe.**
**aci** Acoustical Consultants Inc.,
Suite 107, 9920-63 Ave, Edmonton, Alberta, Canada, T6E 0G9. www.aciacoustical.com

## 1. INTRODUCTION

With recent advances in digital data acquisition, multi-channel data acquisition systems for acoustics and vibrations are more commonplace. There has been an industry-wide shift from closed box systems with limited configuration options to PC based systems with numerous software driven data acquisition choices. These systems, however, are still relatively expensive and invariably contain a certain degree of constraints for both data acquisition and signal processing. Utilizing engineering computational program software such as MATLAB [1,2] enables individuals to customize their multi-channel data acquisition systems and use a variety of external hardware, all for less money than off-the-shelf systems. This paper is intended to discuss some of the highlights and pitfalls with using MATLAB for digital data acquisition and signal processing.

## 2. DATA ACQUISITION

The Data Acquisition Toolbox in MATLAB is designed to work with several hardware devices and offer complete user control over virtually all data-acquisition parameters. Variables such as sample rate, sample duration, and triggering can be easily controlled. Probably the simplest data acquisition device is the soundcard installed on most PC's. This can be used to acquire two channel data, with a sample rate up to 44.1 kHz with 16-bit resolution. The two most important things to be aware of with the sound card are the fact that it is A/C coupled (i.e. high pass filtered at about 2 – 3 Hz), and that it has preferred sample rates (i.e. 8 kHz, 10.025 kHz, 22.05 kHz, and 44.1 kHz). Acquiring data at sample rates other than these produced unpredictable/un-expected results. It is also important to consider the ability to sample dual channels simultaneously. Tests with a common 16-bit SoundBlaster sound card showed that the time lag between the acquisition on each channel translates to a much higher frequency than the 44.1 kHz maximum sample rate, and thus will not cause a problem for signal processing between the two channels.

Various other data acquisition cards are supported by MATLAB. All work for this paper was done using a Measurement Computing PC-CARD_DAS16/16-AO which has either 16 single-ended (common ground) or 8 differential analog inputs as well as several outputs and digital connections. The connections to the card (single or double) are VERY important and dependent on the type of transducers being used and the type of measurements made (consult the user manual). Given the varied types of transducers used for typical acoustic/vibration sensory work, it was found that differential connections were appropriate. Although the card has a maximum sample rate of 200 kHz for a single channel, as the number of channels increases, the maximum sample rate goes down by (essentially) the ratio of channels. This is because the card uses a multiplexer to route each channel to a single A/D converter. As the number of channels increases, this process slows down. More expensive cards have multiple A/D converters, which allow for maximum sample rates in each channel regardless of the number of channels.

Using MATLAB, the data is acquired as a large matrix for each record block. The block length is dependent on the sample time and duration. Although these can be set to virtually any value, it is common to use radix 2 ($2^n$) numbers to take advantage of faster FFT algorithms [3]. Once the data is acquired, it must be stored or processed. Refer to the Appendix for sample data acquisition code. The variable *t* stores the time vector for the entire sample length, and the variable *d* has the data stored in a matrix with 1 column per channel.

As with any data acquisition system, an appropriate analog anti-aliasing low-pass filter is required. There are numerous filter choices available with specific data acquisition needs dictating which filter is appropriate. For the purposes of this paper, an external, switched capacitor filter was used. One key advantage to this filter is that the cutoff frequencies are controlled digitally by a variable clock frequency. This can be used in conjunction with the variable frequency output of some data acquisition cards to provide an infinitely variable set of frequencies. For example, when choosing to acquire the data at a particular sample rate, the computer program could automatically select the appropriate filter cutoff frequency and adjust it accordingly.

## 3. SIGNAL PROCESSING

Once the data has been acquired, it needs to be processed to obtain the desired form. MATLAB has various functions/operations built in for signal processing. Windowing, FFT analysis, cross and auto power spectrum calculations, coherence, averaging, $L_{eq}$ measurements, integration and differentiation are just some of the key operations which can be done. Each of these operations are relatively simple, as illustrated in the sample code for calculating the cross spectrum and coherence between two channels, in the Appendix.

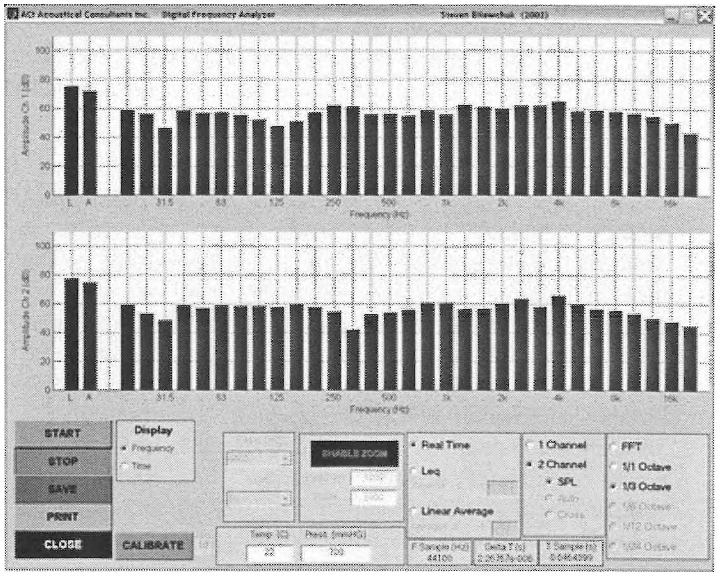
Similar operations such as FFT, windowing, and averaging require equivalent amounts of code. The key advantage is that the entire process is completely custom and additional features (such as additional channels) are minimal to incorporate.

One important aspect to consider when using MATLAB is the process involved in acquiring the data, processing it, and displaying it. For a "real-time" application each of these processes needs to be done in succession before the next sample block can be obtained. Current versions of MATLAB do not allow for simultaneous operations such as acquiring the next signal while the previous one is being processed. Thus, the "real-time" capabilities of the entire process are limited by the computer processor speed and, more importantly, the efficiency of the code. Most applications have been found to be essentially "real-time" but the process does slow down as more channels are added and more complicated computations are performed. For most measurements, where long duration averaging is required, this has not been a concern. If a continuous data block is required, the entire process can be expedited by not analyzing and graphing the results for each data block sampled and waiting until the end.

## 4. GRAPHICAL USER INTERFACE

One of the most important features that makes MATLAB viable for data-acquisition and signal processing is the ability to write programs with a graphical user interface (GUI). Having data acquisition parameters, and plotting displays in the common "Windows" format enables much more power, flexibility, and "user-friendliness". This paper will not digress toward the various methods and pitfalls of creating GUI programs, rather the purpose is to state that with a few extra steps, a more powerful program for data acquisition can be obtained.

Figure 1 shows a simple program written for multi-channel data acquisition. This program is available for download at www.aciacoustical.com




## 5. REFERENCES


1) MATLAB Data Acquisition Toolbox Version 2.0 Users Guide, The Mathworks, www.mathworks.com
2) sample MATLAB code can be downloaded at www.aciacoustical.com
3) Frequency Analysis, Brüel and Kjær. 1987, K. Larsen & Son A/S


## APPENDIX

The following is a sample of MATLAB code used to acquire a single data block:

| | |
|---|---|
| `ai = analoginput('mcc',1);` | *sets up connection to data acquisition card* |
| `ch = addchannel(ai,[0 1]);` | *adds channels 0 and 1 resulting in a total of 2 channels* |
| `set(ch, 'inputRange',[-5 5]);` | *sets input range to ± 5V* |
| `set(ai,'samplerate',1000);` | *sets sample rate to 1000 Hz* |
| `ai.samplespertrigger = 1000*10;` | *sets sample length to 10 seconds* |
| `start(ai);  [d,t] = getdata(ai);` | *starts data acquisition and stores data into variables d (data) and t (time)* |
| `stop(ai); delete(ai);` | *stops data acquisition and deletes data in "ai" before acquiring next sample block* |

The following is a sample of MATLAB code used to perform the FFT calculation on 1 of the 2 channels

| | |
|---|---|
| `Ch_1 = fft(d(:,1))/blocksize*2;` | *FFT for Ch1 with scale correction* |
| `Ch_1(1) = Ch_1(1)/2;` | *Correction for DC Offset* |
| `Ch_1 = Ch_1(1:(length(Ch_1)/2 + 1 ));` | *Storing only the positive frequency components* |

The following is a sample of MATLAB code used to calculate auto and cross spectra as well as coherence:

| | |
|---|---|
| `G_11 = G_11 + conj(Ch_1) .* Ch_1;` | *auto power spectrum for Ch1 (added to previous value)* |
| `G_22 = G_22 + conj(Ch_2) .* Ch_2;` | *auto power spectrum for Ch2 (added to previous value)* |
| `G_12 = G_12 + conj(Ch_1) .* Ch_2;` | *cross power spectrum between Ch1 & Ch2 (added to previous value)* |
| `G_21 = G_21 + conj(Ch_2) .* Ch_1;` | *cross power spectrum between Ch2 & Ch1 (added to previous value)* |
| `H1 = G_12 ./ G_11;` | *Frequency response function (H1)* |
| `H2 = G_22 ./ G_21;` | *Frequency response function (H2)* |
| `coh = real(H1 ./ H2);` | *Coherence between Ch1 & Ch2* |

# INVESTIGATION OF DASH-8 RUN-UP NOISE CHARACTERISTICS FOR LOCAL ACTIVE NOISE CONTROL

**Ann Nakashima and Murray Hodgson**
Dept. of Mechanical Engineering and School of Occupational and Environmental Hygiene
The University of British Columbia, 3rd Floor, 2206 East Mall, Vancouver, BC, Canada, V6T 1Z3
e-mail : naka@mech.ubc.ca

## 1. INTRODUCTION

Recordings of noise from a deHavilland Dash-8 aircraft, provided by Air Canada Jazz, were made at Vancouver International Airport in August 2002. The measured narrow-band and 1/3 octave band spectra were calculated and analyzed in order to determine the blade-passing frequencies (BPF) and noise directivities for three different engine settings: idling, 50% power and full power. Decibel subtraction of the idling engine spectra from the 50% engine power spectra was performed to estimate the power spectra of a single engine. The results were compared to determine if and how an ANC system must be changed to obtain the best noise attenuation for the different run-up conditions.

## 2. RUN-UP NOISE MEASUREMENTS

The run-up measurements were performed on a clear summer night with low wind. Receiver microphones were positioned on a circle of radius 40m (measured from the center of the aircraft) at 20° increments. The area directly behind the aircraft was excluded due to excessive wind from the propellers, and because of safety concerns. Four different 1/2" free-field microphones in combination with a conditioning amplifier or sound level meter were used to record the data at the various positions. Portable Sony PCM-M1 DAT recorders were used to capture the data at a sampling rate of 48 kHz. In total, noise was recorded at 15 positions around the aircraft for three conditions: both engines idling, both engines at full power, and the right engine at 50% power while the left engine was idling.

## 3. SPECTRAL ANALYSIS

The BPF of the Dash-8 at idling engine power was found to be 18 Hz. A spike in the spectra at this frequency was clearly visible at all positions except those directly to the front of the aircraft, which were the farthest away from the propellers. The noise contribution at the BPF was not significant enough to produce a spike in the 16 or 20 Hz bands of the 1/3 octave band spectra. Thus, at idling engine power, the noise contribution at the BPF was not significant enough to dominate the overall noise spectra. Given also that the harmonics were weak or insignificant, it can be concluded that idling engine noise does not contain enough tonal noise for ANC to be effective.

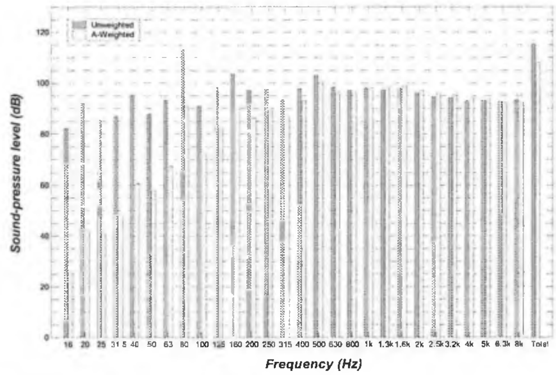



Fig. 1. 1/3 octave band spectrum of Dash-8 noise recorded at the position closest to the right propeller.

At full engine power, the BPF was found to be 80 Hz. The noise at this frequency reached as high as 120 dB at the positions closest to the right propeller. At those positions, the noise at the BPF was clearly dominant over the harmonics. The noise contribution at the BPF is also clearly visible in the 1/3 octave band spectra, as can be seen in Fig. 1. At other positions, the amplitudes of the first and second harmonics were close to or exceeded the amplitude of the BPF. In particular, the noise spectra measured at the front of the aircraft were dominated by high frequency noise (from higher harmonics).

Given the strong tonal noise components in the spectra, it may be possible to control full-power engine noise with ANC. The presence of strong harmonics suggests that it would be of interest to consider using ANC to control not only the noise at the BPF, but the noise at the first and second harmonics as well.

## 4. NOISE DIRECTIVITY PATTERNS

The noise directivity for both of the engines at idle settings is shown in Fig. 2. The lowest levels were behind the left propeller, at about 80 dBA; the loudest levels were to the front of the aircraft, at about 100 dBA. Both the unweighted and A-weighted directivity patterns resemble that of a dipole source, with the strongest radiation to the front of the plane, and low levels to the sides, as expected for a fan-like source [1]. However, unlike a perfect dipole source, the directivity is asymmetrical.

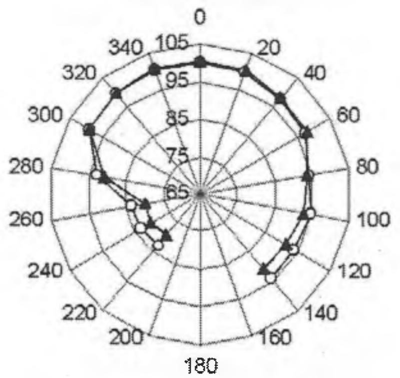



Fig. 2. Noise directivity of the Dash-8 with both engines idling. O= unweighted data, ▲ = A-weighted data.

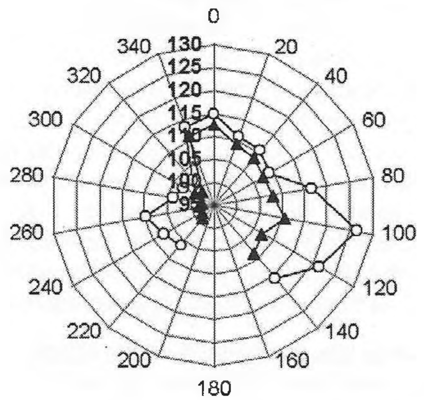



Fig. 3. Noise directivity of the Dash-8 with both engines at full power. O = unweighted data, ▲ = A-weighted data.

With both of the engines at full power, the directivity is as shown in Fig. 3. Clearly, it is very different from the idling results; however, the noise levels to the right of the aircraft are still louder than those to the left. The unweighted directivity pattern resembles the clover-leaf pattern characteristic of a quadrupole source, though asymmetrical, with the strongest lobe to the right of the aircraft peaking at 127 dB.

It has been determined experimentally [2] that an ANC system is most effective when it is placed facing the lobe of strongest directivity from the primary source. Thus, with the engines running at full power, an ANC system would likely be most effective if it was placed behind the right propeller (at 100°). The noise directivity of the Dash-8 was also found to be different from that of the Beechcraft 1900D [3]; this suggests that the position of the ANC system must also be optimized differently for different aircraft.

## 5. ESTIMATING THE NOISE RADIATION OF A SINGLE PROPELLER

It is also of interest to understand how a single propeller acts as a noise source; however, it was not possible to run one engine of the Dash-8 at above idling power with the other engine turned off. To obtain an approximation of the spectra of a single engine operating at 50% power, the idling engine spectra were decibel-subtracted from the right engine at 50% power/left engine idling spectra.

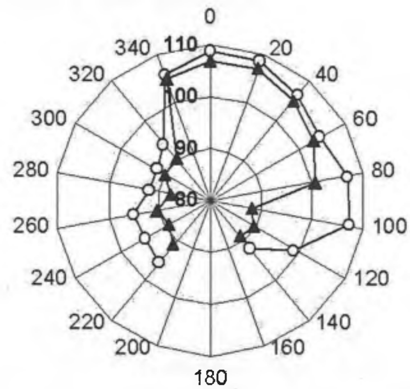



Fig. 4. Estimated noise directivity of one propeller engine at 50% power. O = unweighted data, ▲= A-weighted data.

The decibel-subtracted narrow-band spectra were virtually identical to the original spectra. The 1/3 octave band spectra show differences only at very low frequencies (<40Hz). The idling engine noise thus does make a small contribution to the noise of an engine operating at 50% power at very low frequencies, but not significant enough to change the total unweighted or A-weighted levels.

The directivity of the decibel-subtracted data is shown in Fig. 4. This gives an approximation of the directivity of a single engine operating at 50% power. The radiation is approximately dipole with the strongest headings towards 0° and 20°, and the weakest towards 300° and 140°.

Since the noise of both engines running at 50% power was not recorded, and it was not possible to record the noise of one engine operating at full power, the noise directivities for one propeller and two propellers cannot be compared directly. It can only be concluded that the directivity of two propeller engines operating at full power is approximately quadrupole, and that the directivity of a single propeller engine operating at 50% power is approximately dipole. The configuration of an ANC system thus must be changed accordingly for these two run-up conditions.

## REFERENCES

1. G. Lauchle, J. MacGillivray and D. Swanson. "Active control of axial-flow fan noise", J. Acoust. Soc. Am. **101**(1), 341-349, (1996).
2. A. Nakashima, M. Hodgson and J. Guo. "Preliminary investigation of active control for dipole noise sources", ACTIVE 2002 proceedings, (2002).
3. P. Germain, J. Guo and M. Hodgson. "Experimental study of propeller-aircraft run-up noise", Canadian Acoustics **29** (2), 21-27, (2001).

**APPLICATION OF DOW-QUASH TO LIMITING COMMUNITY NOISE**

**Corjan Buma**
aci Acoustical Consultants Inc., #107, 9920 - 63 Avenue Edmonton, AB, T6E 0G9

# 1 INTRODUCTION

Annoyance with community noise occurs in all degrees of severity and in all contexts (rural, urban, etc). The degree of annoyance can usually be anticipated based on familiarity with "average hearing" of the "typical" human listener. Yet there are occasions where an individual claims, for example, sleep disturbance when experience suggests that there is no obvious reason for this to be occurring. With regard to potential noise mitigation the task is then first to assess the disturbance claim, and, if deemed valid, develop feasible solutions. This paper outlines the application of a relatively new noise control product to provide a "path"-based solution (as opposed to treating the noise "at-source" or "at-receiver").

In the current study the human noise disturbance issue was complicated by the presence of a peregrine falcon: the falcon's nest was in the immediate proximity of where one of the most-feasible noise mitigation options would normally be installed. Noise was not deemed a concern with regard to the falcon: the falcon had been returning annually to this location immediately adjacent to the offending noise source for many years and had regularly produced offspring without any apparent detrimental effect due to noise. Of great concern was whether relocating the falcon's nest would upset its behavioral patterns, including producing offspring. (The peregrine falcon is considered an endangered species and as such is protected by law.)

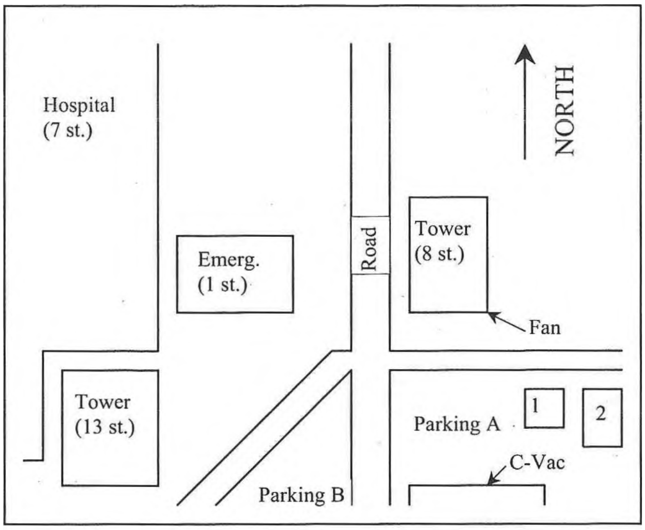



Figure 1 – Area Plot Plan

Figure 1 shows a macroscopic view of the study area. Of note are the large teaching Hospital, the 13-storey research tower with mechanical penthouse, a central-vacuum outlet recessed in a loading dock in close proximity to the noise-affected Residence and the restaurant exhaust fan. In the quietest night-time hours traffic flows on the Roadway typically reduced to one vehicle pass-by every minute or so. The Hospital had a major roof-mounted noise source (exhaust unit) on its south wing which was significantly shielded from the noise-affected Residence (Fig. 1, "2") by the Research Tower. The Hospital also had numerous additional ventilation intakes and outlets on the upper storey of its east face as well as its Emergency ward at grade on the east side. The central vacuum system of the south high-rise did not stop at the same time each night; it had a generally low-frequency signal that was clearly audible across the Parking Lot. The Restaurant exhaust fan on the north high-rise was mounted at essentially the height of the second floor and initially ran continuously.

Figure 2 indicates the position of the falcon's nest relative to the Research Tower's exhaust-fan outlet louvres.

# 2 MEASUREMENT PROGRAM, RESULTS

Three separate sets of night-time measurements were conducted. The first involved logging the 1/3-octave and broadband sound levels in 30-second intervals for an entire night-time. In addition, localized spot measurements were conducted at other nearby noise sources. This effectively verified that the primary noise offender was a series of east-facing exhaust fans atop the Research Tower, all other candidate noise sources being masked by other area noise sources. Surprisingly, while the central-vacuum was intuitively suspected as a key noise-offender, the Complainant ruled this out: it had been suspected that the *cessation* of this noise source was resulting in a change of noise climate to the degree of waking the Complainant.

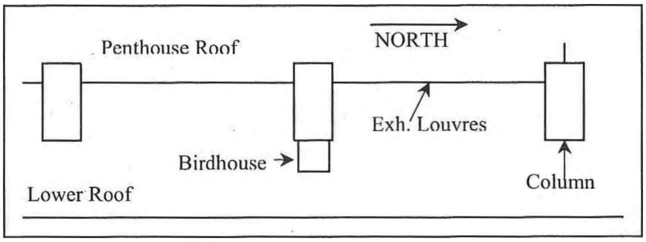



Figure 2 – Penthouse Area (Plan)

The second and third measurement sets involved short-term 1/3-octave sound level logging while ALL penthouse noise sources atop the Research Tower were progressively shut off and returned to normal operation. One such round was measured on the upper-most level of Parking "B" (Figure 1), the final set being measured near the eastern edge of Parking "A". The measurements atop Parking "B" confirmed the Penthouse noise sources as the primary noise offender, while those on Parking "A" were done to assess the noise impact relative to the City Noise Bylaw. These latter measurements indicated a resultant sound level of 51 dBA while all penthouse noise sources were running and a nominal 47 dBA with all Penthouse noise source shut off. Given the relative distances involved, it was determined that the Penthouse noise sources just met the City Bylaw allowable sound level of 50 dBA at the residential property line. Thus the options available to the Owner of the Research Tower were to (a) indicate to the noise-affected Resident that further action was not required or (b) implement some degree of noise mitigation despite not being required by law. The latter was chosen.

The upper (solid) line on Figure 3 indicates the 1/3-octave spectrum measured immediately near the Penthouse exhaust systems on the Research Tower. There is a local maximum centered on the 800Hz and 1000Hz bands. The second trace (dashed) on Figure 3 is the combined effect of all Penthouse noise sources in operation, as measured on Parking "A". It is evident that the local maximum was still evident. The lowest (dotted) trace is the spectrum with all Penthouse noise sources off.

## 3 NOISE-CONTROL TREATMENT

The optional treatments were (1) mitigation at source by means of re-configuring the ductwork, introducing silencers and possibly obtaining a lower-noise type of exhaust fan and (2) constructing a noise shield externally in front of the set of exhaust louvres.

Several considerations covering relative costs, relative effectiveness for noise reduction, and the impact on the falcon played into the decision-making. It was considered that re-configuring the ductwork-and-exhaust-fans within the Penthouse could run as much as two-to-three times the expense of constructing a noise-shield. The advantage of re-configuring the exhaust systems was that it potentially could avoid any interference with the falcon and thus could be implemented without time restrictions. For the noise shield, while likely being less labor-intensive than the ductwork option, it would likely require relocating the bird-house or somehow integrating the birdhouse into the noise shield. There was concern that the "change-of-scenery" in the immediate vicinity of the bird-house could adversely affect the falcon in its nesting, feeding and parenting habits. Also, it was deemed necessary to have a noise shield completed by 15-March, the usual time the falcon could be expected back. A relatively minor concern was that of visually integrating any new construction into older architecture directly exposed to public view. From the standpoint of noise control, it was deemed that either solution could provide the necessary degree of attenuation.

Mainly for reasons of cost, the Owner favored the noise-shield option. Therefore, a meeting was called between Building-staff, the acoustical consultant and a Provincial Wildlife biologist. It was determined that introduction of the noise shield and relocation of the bird-house would very likely *not* adversely affect the falcon. A detailed design for the wall was developed, submitted for approvals and subsequently built.

In this instance it would be quite feasible to construct a noise shield that provides barrier-type noise attenuation and forego any sound-absorbing lining. However, given the sensitivity that precipitated the study and the relatively small cost of adding the lining, it was decided to include a liner directly facing the exhaust-fan louvres. Traditionally, one would automatically opt for fibrous-based core material, usually wrapped in thin plastic to withstand effects of wind, water and winter and protected by expanded-metal mesh or, at minimum, wire mesh. However, since DOW-QUASH, a relatively new poly-ethylene based cellular product that provided maximum sound attenuation in the preferred frequency range, can be left directly exposed, it was determined to be the liner-of-choice.

Upon completion a follow-up visual inspection of the wall was done and a few spot measurements taken at grade. The noise reduction realized was a decrease for the "all systems on" condition by 4 dBA (from 51 to 47 dBA); thus the net sound level had been reduced to the "penthouse off" condition measured during earlier measurements. Subjectively, at Parking "A" it was necessary to listen intently to distinguish the exhaust-fans sound. Indications are that the falcon has continued its usual life-cycle patterns as though nothing has changed.

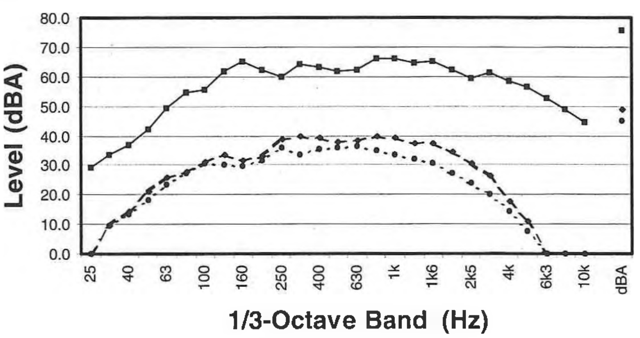



Figure 3 - 1/3-Octave Spectra

# Modelled Trends in Environmental Noise Levels as a Function of Long-Term Meteorology in Alberta

**Jerome Parkinson and Jason McCrank**
HFP Acoustical Consultants Corp., Suite 1140, 10201 Southport Road SW, Calgary, AB , Canada, T2W 4X9
info@hfpacoustical.com

## 1. INTRODUCTION

Environmental noise surveys in rural settings are commonly conducted in Alberta. A frequently encountered scenario is that of a listener located near an industrial facility with no other sources of continuous noise nearby. When surveys are conducted to measure the noise level occurring at the listener's location, it is found that the noise level varies as a result of changes in meteorological conditions.

The current investigation looked at the predicted noise level for listeners located 1 km from a noise source with a sound power spectrum of a typical natural gas compressor station. Hourly meteorological data for a 5 year period was used to predict 43,824 hourly Leq noise levels at listener locations. The predictions were made for listeners 1 km away from the noise source in 4 cardinal wind directions. Meteorological data from 6 distinct air sheds was investigated. The purpose of the exercise was to investigate long-term trends in the hourly noise levels as a function of meteorology.

## 2. METHOD

An analytical algorithm was used to model the propagation of noise from the source to the listeners and to predict the noise levels at the listener locations. A predictive model had several advantages over real long-term measurements. Most real world noise surveys require extensive analysis to isolate the noise level of a single source from other sources of noise. With modelled noise levels, it was possible to analyze the idealized scenario of a listener near an industrial noise source with no other source of noise to contribute to or contaminate the noise level. With a model it was feasible to predict five years of hourly sound levels. It would not have been realistic to collect five years of real hourly sound levels. Finally, it was feasible to predict hourly noise levels in four different directions for locations in 6 different air sheds.

All predictive algorithms face the question of how well they model reality. It would have been ideal to compare the predicted noise levels of this investigation to real world data. However, it would have been impractical to acquire the data that would have been required for such a comparison.

"The Propagation of Noise from Petroleum and Petrochemical Complexes to Neighbouring Communities" algorithm as developed by the CONCOWE Special Task Force on Noise Propagation was used to predict the noise levels. This algorithm is specific to petroleum sector facility noise which is the most common source of environmental noise concern in Alberta. It was also selected as the algorithm for this study because it accounts for the influence of meteorological stability category.

Hourly meteorological data was obtained from Environment Canada's Meteorological Services Division. Hourly wind speeds, wind directions, relative humidities, and temperatures were obtained for five years from six air sheds encompassing the Edmonton Namao, Calgary, Red Deer, Medicine Hat, Edson, and Peace River airports. The corresponding Pasquill stability classes for each hour in each air shed were obtained from Alberta Environment.

## 3. RESULTS

A set of hourly Leq dBA noise levels was generated for a listener located 1 km from the noise source in four cardinal wind directions (north, east, south, and west) for the six air sheds. The result was 24 sets of data that each encompassed 43,824 hourly Leq noise levels. The data was assessed using various statistical and graphical tools to identify trends and draw conclusions.

## 4. DISCUSSION

### 4.1 Total data set

In all sets of data, it was found that the Leq's of nighttime noise levels (10pm to 7am) were higher than daytime noise levels. The nighttime Leq's were on average 1.0 dBA higher than daytime noise levels.

Figure 1 illustrates a typical distribution of hourly Leq's at a listener location. In this example, the lowest hourly noise level was 21.2 dBA and the highest hourly noise level was 40.8 dBA in this example. The standard deviation of all hourly noise levels in the data was 4.5 dBA. The variation in hourly noise levels is relatively large.

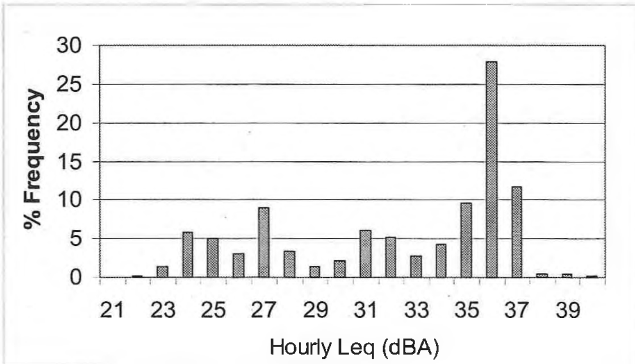



Figure 1. Frequency of representative hourly Leq noise levels as a percentage of the total hourly Leq noise levels for a listener 1 km east of a typical compressor station in the Peace River air shed.

In the data shown in Figure 1, the standard deviation in the daytime (7am to 10pm) hourly noise levels was 4.6 dBA. The standard deviation in the nighttime (10pm to 7am) hourly noise levels was 4.2 dBA. These standard deviations were typical of all locations and air sheds. In all locations for all air sheds the standard deviation in the nighttime noise levels was less than in the daytime noise levels. This indicates that the variation in nighttime noise levels is less than the variation in daytime noise levels.

In all six air sheds, it was found that the Leq noise levels for a listener to the east of the noise source was higher than for a listener to the north, south, or west of the source. This trend was observed in the total noise levels, the daytime noise levels, and the nighttime noise levels. As Alberta has a predominant westerly wind, this observation was not surprising.

4.2 Data from 'representative nights'

Real surveys of industrial environmental noise generally focus on nighttime noise levels because this is when the other masking sound is lowest, listeners are more commonly trying to rest, and regulations for noise levels are most strict. Conscientious noise surveys are done under meteorological conditions which are likely to capture a sample of the noise levels which listeners are concerned about. The Alberta Energy and Utilities Board, which regulates energy sector environmental noise levels, refers to noise levels which represent the noise of concern as 'representative noise levels'. An investigation was conducted into the variations and trends in Leq nighttime noise levels that occur during meteorological conditions which would be considered 'representative'.

The conditions that define 'reprentativeness' are not precisely defined. However, the general assumption is that a wind direction which puts the listener directly downwind or crosswind of the noise source is likely to be representative. Periods of high wind speed are also not considered representative because it is practically impossible to take quality noise measurements in such conditions. For the current analysis, a nighttime Leq noise level was considered to be representative if at least 7 hours in the 9 hour nighttime period met the condition of being downwind, partially downwind, or precisely crosswind from the source to the listener and the wind speed was less than 20 km/hr.

Figure 2 shows a typical distribution of representative nighttime noise levels (9 hour Leq). In this example, the lowest representative nighttime noise level was 29.4 dBA while the highest was 40.6 dBA. The standard deviation for these representative nighttime Leq's was 1.2 dBA. The distribution indicates that it is more common for a representative nighttime noise level to fall below the average representative nighttime noise level. The variation in representative nighttime noise levels is relatively small. All 24 data sets showed very similar trends.

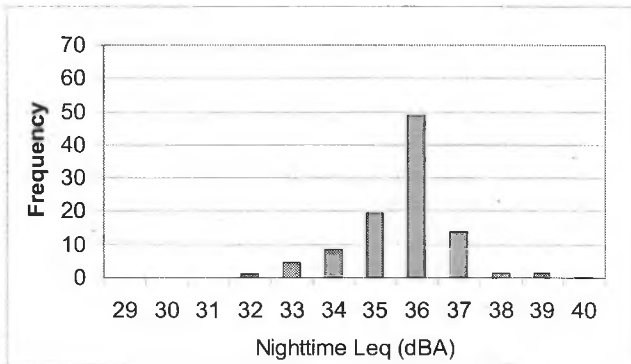



Figure 2. Frequency of representative nighttime Leq noise levels as a percentage of the total number of representative nights for a listener 1 km east of a typical compressor station in the Peace River air shed.

It was found in all air sheds that a listener to the east of the noise source was up to 22 times more likely to experience "representative conditions" than a listener to the north, south, or west of the noise source. People living to the east of an industrial facility will experience high industrial noise levels far more commonly than people living in other directions from the facility in Alberta.

## REFERENCES

Manning, C.J. (1981) "The Propagation of Noise from Petroleum and Petrochemical Complexes to Neighbouring Communities", CONCOWE's Special Task Force on Noise Propagation, (CONCOWE is the oil companies European organization for environmental, health, and safety).

# Detectsound Version 2: A Software Tool For Adjusting The Level And Spectrum Of Acoustic Warning Signals

**Y. Zheng, C. Giguère, C. Laroche, C. Sabourin**
Programme d'audiologie et d'orthophonie, Université d'Ottawa, Ottawa, Ontario K1H 8M5

## 1. INTRODUCTION

Improper uses of warning signals are quite common in the workplace. In a study carried out in a steel plant [1], 93 different conditions where warning signals are used were surveyed. Results revealed that in 40% of the cases, the signals were not properly adjusted to effectively warn of safety hazards when hearing protection was worn. Another quite common problem was to deliberately set the warning signals at excessively high levels. Other researchers have reported that the signal level could be 20 dB above necessary in some cases [2].

Organizations focusing on health and safety in the workplace have proposed some theoretical models for designing warning signals [3], but none provides a comprehensive solution to this problem. Detectsound Version 1 released in 1991 [4] seems to be the first of practical software tool for predicting the detectability of acoustic warning signals in real-life conditions, taking into account the hearing status of the target population, the background noise in the workplace and the wearing of hearing protection.

Since it was developed, Detectsound has been used to design safer warning signals and to propose modifications for existing signals in term of spectral content and overall sound pressure level. Despite its success, the needs for new features and improvements have been emerging in the past decade. This article presents a comprehensive revision of Detectsound to update the main algorithm and to expand the available options to account for the hearing status of workers.

## 2. METHOD

In Version 2, an algorithm reflecting the direct estimation of masked thresholds is adopted, instead of the loudness estimation procedure used in Version 1. As proposed by Moore et al. [5], the estimation of masked thresholds is based on the background noise spectrum and the frequency selectivity data (K & ERB) for the target worker. The equation is:

$$P_s = K\int_0^\infty N_{(f)} W_{(f)} df \qquad (1)$$

where $P_s$ is the power of the signal at the detection threshold in noise, $N(f)$ is the masker spectrum, $K$ is the detection efficiency constant, and $W(f)$ is the auditory filter shape of the worker. The latter can be described by:

$$W(g) = (1+pg).exp(-pg) \qquad (2)$$

where $g = | f - f_0|/f_0$ is the normalized frequency, $p = (4\times f_0)/ERB$ is the slope of the auditory filter, $f_0$ is the centre frequency of the filter, $f$ is the frequency (Hz), and ERB (Hz) is the Equivalent Rectangular Bandwidth of the auditory filter.

This improvement avoids the complexity and estimation errors arising from loudness estimation procedures and facilitates individualized estimation. $W(f)$ and $K$ can describe the frequency selectivity data of either a specific individual or worker, or a statistical population of workers.

The latest normative data on hearing sensitivity shift [6] (ISO 1999-1989) and frequency selectivity change with hearing loss have been integrated into Detectsound. ISO 1999 allows predicting the effect of long-term noise exposure and aging on hearing loss for otologically normal population. Frequency selectivity can also be estimated based on published normative data [7,8].

A software tool, Detectsound Version 2, reflecting the algorithmic improvements over the previous version has been developed, complete with a new and improved graphic user interface [9].

## 3. RESULT

Individual hearing status can now be accounted for in Detectsound Version 2. Figure 1 gives an example of Detectsound's application to analyzing the functional requirements of a specific worker Y. The measured hearing status for this worker is shown in Table 1. The background noise is machinery room noise at a level of 87.2 dBA, and no hearing protection is used.

Table 1. Measured hearing status of worker Y (Left ear)

| Frequency (Hz) | 250 | 500 | 1000 | 2000 | 3000 | 4000 |
|---|---|---|---|---|---|---|
| THR(dB HL) | 6 | 3 | 2 | 6 | 6 | 18 |
| ERB (Hz) | 57 | 93 | 173 | 321 | 487 | 1013 |
| K (dB) | 3.9 | 3.4 | -0.7 | 1.0 | -1.2 | - |

The predicted optimal range (Design window) of warning signal levels at various frequencies is shown in Figure 1 for the above worker Y. To facilitate recognition of warning signals, a level of 10-15 dB above masked thresholds has been proposed [3]. Thus, in Detectsound, the lower and upper boundaries of the design window are defined to be 12 dB and 25 dB above the estimated masked thresholds respectively, with a maximum of 105 dB SPL to prevent overly loud sounds. The frequency components falling within the design window are optimally adjusted. If the number such components is three or more, the signal is considered to be effective for worker Y.

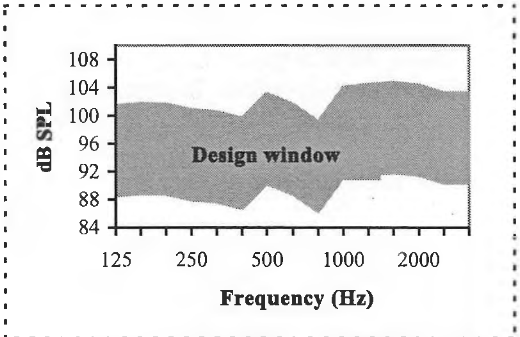



Figure 1: Warning signal design window applicable to worker Y

Detectsound Version 2 can also be used to find optimal signalization solution for several workers sharing a common work area, or for populations of different hearing status. A comparison conducted between the two versions of Detectsound also suggests that Version 2 gives out more accurate predictions [9]. Figure 2 illustrates the comparison result. The target population consists of 6 normal hearing subjects, and the background noise at white noise of 80 dB SPL. The observed data are measured masked thresholds from an independent validation study.

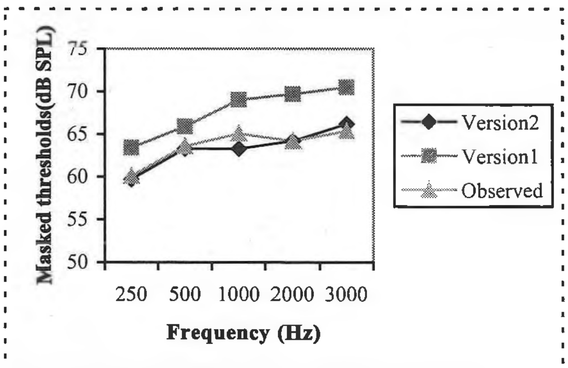



Figure 2. Prediction accuracy for two versions of Detectsound

An average overestimation of 4 dB is found for Detectsound Version 1 across frequencies from 250 to 3000Hz, while Version 2 carries an average underestimation of only 0.3dB.

# 4. DISCUSSION

Detectsound Version 2 provides a more complete solution for warning signalization in the workplace. If the hearing status of a specific worker is fully known, the analysis result is tailored to meet the functional requirements of that worker. The analysis result can also be made to suit the needs of a population of workers if the hearing status of this population is considered. Detectsound Version 2 can also make predictions based on estimated hearing thresholds and frequency selectivity data, and this expands its application to situations where the hearing status of the workers are unavailable (such as planning a new plant). The predictions by Version 2 are also more accurate.

One current limitation of Detectsound Version 2 is that it cannot be applied when the background noise spectrum is unspecified. The integration of Detectsound with a sound propagation model inside industrial plants is proposed. The latter could predict the noise distribution within a workplace given the noise sources, and provide Detectsound with the required background noise data needed for the design of acoustic warning signals.

[Work funded by an NSERC Collaborative Health Research Project Grant, Canada].

# 5. REFERENCES


1. Tran Quoc, H., Hétu, R., and Laroche, C. (1992). "Computerized assessment and prediction of the audibility of sound warning signals for normals and hearing-impaired individuals", in M. Matilla and W. Karwowski (eds.), *Computer applications in ergonomics occupational safety and health* (Elsevier, Amsterdam), pp. 105-112.

2. Patterson R.D., & Milroy, R. (1980). "The appropriate sound level for auditory warnings on civil aircraft", J. Acoust. Soc. Am. 67, S58.

3. ISO 7731 (1986). "Danger signals for work places -- Auditory danger signals", International Organization for Standardization, Geneva.

4. Laroche, C., Tran Quoc, H., Hétu, R., and McDuff, S. (1991). ""Detectsound": A computerized model for predicting the detectability of warning signals in noisy workplaces," Applied Acoustics, 32: 193-214.

5. Moore, B.C.J., & Glasberg, B.R., (1987). "Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns", Hearing Res., 28:209-225.

6. ISO 1999 (1989). "Acoustics – Determination of occupational noise exposure and estimation of noise-induced hearing impairment", International Organization for Standardization, Geneva.

7. Laroche, C., Josserand, B., Tran Quoc, H., Hétu, R., and Glasberg, B.R. (1992). "Frequency selectivity in workers with noise-induced hearing loss," Hearing Research 64: 61-72.

8. Moore, B.C.J., & Glasberg, B.R., (1996). "A revision of Zwicker's loudness model", Acta Acustica, 82, 335 –345.

9. Zheng, Y. (2003). "Development of new "Detectsound" – A computerized model for adjusting the level of acoustic warning signalization in workplaces", Master's Thesis, University of Ottawa, Canada.

# Noise Exposure of Opera Orchestra Players

**Jason Lee[1,2], Alberto Behar[1], Hans Kunov 1,2 and Willy Wong1,2**
1) Sensory Communication Group, Institute of Biomaterials and Biomedical Engineering, University of Toronto, 4 Taddle Creek Road, Toronto, ON, Canada, M5S 2G9
2) The Edward S. Rogers Sr. Dept of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, ON, Canada, M5S 3G4

## 1. INTRODUCTION

There is an abundant literature concerning the noise exposure of classical orchestra players. These studies assess risk of hearing loss of musicians playing in concert hall, using results from measurements of noise exposure levels and the hearing threshold of the musicians. Most of these studies deal with orchestra players playing in a concert hall. Most conclude that the musicians are not overexposed and do not experience significant hearing loss (E.g., see Eaton and Gillis, 2002; Kahari et. al., 2001).

However, for musicians playing in an orchestra pit, as in the case of opera and ballet, their noise exposure is expected to be higher due to the confined space of the pit. A study of the Finnish National Opera concluded that most musicians were over exposed, contrary to the studies performed in concert hall (Laitinen et. al., 2003).

The objective of this study is to examine the noise exposures of orchestra players from the Canadian Opera Company (COC) during rehearsals and performance of two operas, Madama Butterfly by Puccini and The Italian Girl in Algiers by Rossini.

## 2. METHOD

A survey was conducted on 73 musicians Measurements of $L_{eq}$ were taken during either a rehearsal, .a dress rehearsal or a performance, which takes about three hours. $L_{EX}$ was then calculated by normalizing the $L_{eq}$ for a one-year exposure, assuming that the duration of their activity with the company is 300 hr,

We performed the measurements using five Quest Type Q-300 dosimeters. The "Slow" setting and A-weighting were used. The range was set to measure between 40 and 110 dBA. The microphone was attached on the shoulder of the musician as far as possible from the head to minimize sound reflections, according to the procedures in the CSA Standard Z107.56 – 94.

The noise exposure criterion adopted in this study was a daily 8 hours exposure limit of 85 dBA and a 3 dB exchange rate. This criterion is used by the Federal Government and several provincial agencies in Canada, as well as by the International Standard Organization (ISO).

## 3. RESULTS

3.1 Average $L_{eq}$ of different musical instruments

Table 1 shows the averages of $L_{eq}$ obtained from different instruments in the two operas. In general, highest noise exposure was found among the brass, followed by the woodwinds, and then the strings. The conductor has one of the lowest exposures among the musicians.

Comparing the two operas, higher noise exposure was measured in Madama Butterfly. A factor that appears to increase noise exposure is the proximity of the brass and woodwind section. For example, the second violins were the only instruments that have higher noise exposure in The Italian Girl in Algiers, as they are positioned in front of the woodwinds in this opera, and far from the brass and woodwinds in Madama Butterfly.

| | Madama Butterfly | The Italian Girl in Algiers |
|---|---|---|
| Violin 1 | 84.8 | 82.8 |
| Violin 2 | 85.7 | 86.5 |
| Viola | 88.3 | 85.8 |
| Cello | 88.7 | 81.4 |
| Double Bass | 88.2 | 83.7 |
| Trumpet | 93.7 | 91.4 |
| Trombone | 90.3 | N/A |
| Horn | 91.7 | 89.9 |
| Piccolo / Flute | 91.7 | 87.4 |
| Clarinet / Base Clarinet | 88.6 | 86.8 |
| Oboe / Bassoon | 88.3 | 84.6 |
| Percussion | 87.6 | 79.8 |
| Cymbal | 87.4 | N/A |
| Conductor | 83.3 | 81.3 |
| **Total** | 89.3 | 86.4 |

Table 1. Average $L_{eq}$ of instruments

3.2 Average $L_{EX}$

According to the Personnel Manager of COC, the musicians play for 300 work hours per year for that company. The normalized yearly noise exposure level ($L_{EX}$) was calculated using the formula $L_{EX} = L_{eq} + 10 \log t/T$ (1), where t = 300 and T = 2000 (the yearly equivalent of a daily work period of 8 hours, used in the ISO document).

Table 2 shows the $L_{EX}$ of different instrument after averaging the $L_{eq}$ between the two operas. The $L_{EX}$ of all instrument groups is below the safety limit 85 dBA. Therefore, we can conclude that the musicians are not at risk of noise-induced hearing loss from playing in the COC.

| | Average $L_{eq}$ | $L_{EX}$ |
|---|---|---|
| Violin 1 | 83.9 | 75.7 |
| Violin 2 | 90.0 | 81.8 |
| Viola | 87.3 | 79.1 |
| Cello | 86.4 | 78.2 |
| Double Bass | 86.3 | 78.1 |
| Trumpet | 92.7 | 84.5 |
| Trombone | 90.3 | 82.1 |
| Horn | 90.9 | 82.7 |
| Piccolo / Flute | 90.0 | 81.8 |
| Clarinet / Base Clarinet | 87.8 | 79.6 |
| Oboe / Bassoon | 86.9 | 78.7 |
| Percussion | 85.2 | 77.0 |
| Cymbal | 87.4 | 79.2 |
| Conductor | 82.4 | 74.2 |
| **Total** | 88.1 | 79.9 |

Table 2. Average $L_{EX}$ of different instruments

## 4. DISCUSSION

The noise exposure obtained in this study considers only the activity of the musicians with the COC, i.e., rehearsals, dress rehearsals and performances. Individual rehearsals, and activities with other companies were not taken into account. Although the noise exposure was found to be below the safety limit, the combination of other activities could result in exposure higher than the limit, posing risk of hearing loss.

We recommended that a hearing protection program be instituted where musicians are made aware of the potentially hazardous noise levels, provided with hearing protection (musician earplugs), and educated about the care and use of earplugs. In addition, musicians should undergo bi-annual audiometric tests to ensure that the hearing protection measures are effective and that their hearing has not deteriorated.

## REFERENCES


Eaton, S. and Gillis, H (2002). Review of orchestra musicians' hearing loss risks. Canadian Acoustics, 30(2): 5 – 12. .

Kahari, K. R., Axelsson, A., Hellstrom, P. A. and Zachau, G. (2001). Hearing assessment of classical orchestral musicians. Scandinavian Audiology. 30(1): 13 – 23.

Laitinen, H. M., Toppila, E. M., Olkinuora, P. S. and Kuisma, K. (2003). Sound exposure among the Finnish National Opera personnel. Applied Occupational and Environmental Hygience, 18(3): 177 – 182.


## ACKNOWLEDGEMENTS


We would like to acknowledge Graham Greenland, Elaine Lui and Gerry Fung from the Institute of Biomaterials and Biomedical Engineering for help in collecting the data. Ian Cowie, from COC, was instrumental with the organization of the tests. Finally, we would like to thank the musicians that wore the dosimeters during the tests. Without their cooperation and enthusiasm, this study could not have been performed.

# Hearing Protectors Testing And Labeling – What's New?

**Alberto Behar**

Institute of Biomaterials and Biomedical Engineering, University of Toronto, 4 Taddle Creek Rd., Toronto, ON, M5S 3G9, Canada - behar@sympatico.ca

## INTRODUCTION

Two issues are associated when dealing with hearing protectors: how to measure their attenuation and what to do with the result of the measurement. Procedures for how to perform both operations have been in place for many years, normalized by national and international standards. In this paper we will review some of the existing standards and the new tendencies that are now under development.

## Normalization Documents

## International Organization for Standardization (ISO)

All ISO standards that deal with attenuation of hearing protectors are under the number ISO 4869. Right now, the following standards have been published:

ISO 4869 – 1: Acoustics – Hearing Protectors – Part 1: Subjective method for the measurement of sound attenuation. It deals with testing the attenuation on human subjects, measuring the threshold of hearing with and without protectors.

ISO 4869 – 2: Acoustics – Hearing Protectors – Part 2: Estimation of the effective A-weighted sound pressure levels when hearing protectors are worn. It describes three different procedures for the estimation of the sound level resulting from the use of the protector, using three different methods.

ISO/WD 4869 – 3.2 (Working Draft): Acoustics – Hearing Protectors – Part 3: Simplified method for the measurement of insertion loss of ear-muff type protectors for quality inspection purposes. This is the working draft for testing of insertion loss using an Acoustic Test Fixture.

ISO/TR 4869 – 4 (Technical Report): Acoustics – Hearing Protectors – Part 4: Method for the measurement of effective sound pressure levels for level dependent sound restoration ear-muffs. Sound restoration ear-muffs amplify the external noise when bellow a certain level, so the person can hear as if he didn't wear the protector. As the sound level increases, the amplification provided by the muff decreases and the system switches off when the level reaches the pre-set limit. This Technical Report describes the measurement procedure and the instrumentation to be used.

ISO/WD 4869 – 7.4 (Working Draft): Acoustics – Hearing Protectors – Part 7: Subjective method for measurement of sound attenuation – Subject-fit method.

Similar to ISO 4869 – 1, this test is performed using subjects. However, subjects are naive (not familiar with the use of protectors). Fitting of the protectors during the test is done by the subjects, without external help. This is a procedure similar to the Method B in the ANSI S12.6-1997 Standard.

## 2.2 American National Standard Institution (ANSI)

There are two measurement procedures in the ANSI S12.6-1997 Methods for Measuring the Real-Ear Attenuation of Hearing Protectors. Method A is similar to the experimenter-supervised fit from the previous version of the Standard (ANSI 12.6-1984). It yields results that are the upper limits of performance that can be obtained using highly experienced subjects under the direct supervision of the experimenter. Method B, on the contrary, relays on subjects that are naive in the use of hearing protectors. Their only information regarding the fitting of the protectors is that on the manufacturers' package. It approximates the attenuation that has could be obtained in the real-life by users involved in hearing protection programs. The ANSI Standard applies to only passive hearing protectors.

## 2.3 The Environmental Protection Agency (EPA)

The EPA produced in 1979 the document EPA (1979) Noise Labeling Requirements for Hearing Protectors. Federal Register, Vol. 42, No 190, 40 CFR Part 211, 56139-56147. by which every protector sold in the USA should be labeled using the Noise Reduction Rating (NRR). The NRR is a single number rating, calculated from the measured attenuation values. When subtracted from the measured ambient noise in dBC, it results in the noise level in dBA of the protected ear. The NRR is probably the most known characteristic of a protector. However, it is also well known fact that its value is overly optimistic. Up to the point that OSHA had recommended that the NRR be derated from 25% up to 70% for certain types of protectors.

On March 2003, EPA hold a two-days public meeting to collect information regarding revising the Federal Regulation 40 CFR Part 211 regarding the effective rating and labeling of protectors. Participants included manufactur-

ers, researchers and government. All protectors were considered including the active and passive non-linear devices, passive, active noise reduction, sound restoration systems, and communication systems/radios. It is expected that EPA will take into account
the material that was presented during the sessions and will come with a new and improved method for rating and labeling the protectors.

## 2.4 Canadian Standard Institution (CSA)

The CSA Standard CSA Z94.2-02 "Hearing Protection Devices – Performance, Selection, Care and Use" was issued in January 2002. Contrary to the ANSI and the ISO standards, it does not contains provisions regarding testing of the protectors. As the title indicates, it is mostly a document for the users, rather than for the measuring laboratories.
As per the attenuation testing, the standard specifies that tests should be performed as per ANSI Standard S12.6, Method B, that is a Real-ear attenuation at threshold (REAT) procedure, where protectors are fit by the subjects that are persons not familiar with the use of protectors (naïve subjects). Results from the test are used to compute the Single Number Rating (Subject Fit 84$^{th}$ Percentile), abbreviated SNR($SF_{84}$ ) that provides a nominal 84% protection confidence interval (i.e., 84% of the users in a well-run hearing conservation program are expected to receive at least that much protection).

Provisions for writing Hearing Protection Program are also provided, as are details of the different types of hearing protectors and their advantages and applications. It touches subjects such as sound attenuation, attenuation at frequency extremes, double protection, overprotection, etc.

For the selection of the protectors, the concept of the Grade, calculated from the SNR($SF_{84}$ ) is included.

# Perception Of Increasing Or Decreasing Signal Intensity And Effects Of Compression By Hearing Aids

**M.-J. Palardy, C. Giguère and C. De Segovia**

Programme d'audiologie et d'orthophonie, Université d'Ottawa, 451 Smyth Road, Ottawa, Ontario K1H 8M5

## 1. INTRODUCTION

Signal-intensity changes provide important cues to evaluate the distance and motion of sound sources in the environment. In one study [1], signal intensity cues facilitated the perception of distance for both close and far sound sources. In another study [2], the listener's movement towards a sound source was also shown to help the evaluation of distance. For some people, like the blind, it is of utmost importance to perceive spatial sound information in order to gain independence in their mobility [3]. However, if an individual with a functional visual impairment also suffers from a hearing loss, his/her security might be threatened by the loss of some or all the essential spatial sound information. For example, when a car is moving closer, sound intensity increases, whereas when the car is moving away, sound intensity decreases.

Hearing aids are often equipped with non-linear noise reduction algorithms to increase the signal/noise ratio. Compression functions are also used to compensate for loudness recruitment. A typical compression hearing aid will amplify soft (or far) sounds and reduce loud (or closer) sounds. These strategies, useful for speech communication, may constitute a hazard for deaf-blind people [4]. For example, the increase in sound intensity of an approaching car may be lessened by hearing aid compression. This issue remains rather unexplored [5].

The purpose of this study was to determine whether compression algorithms in hearing aids make intensity changes more difficult to perceive than if no compression (linear aid) is used. The study also aims to determine whether the effect of a hearing aid on the perception of intensity changes can be predicted by the compression ratio. The signal was a car horn, which could increase or decrease in level, presented in silence or constant traffic noise.

## 2. METHOD

2.1 Subjects

Twenty subjects with normal hearing, aged 22 to 29 years, were recruited among students at the University of Ottawa. Selection was made according to the following criteria: a) air conduction hearing threshold $\leq$ 15 dBHL (*cf.* ANSI S3.6-1996) between 0.25 and 6 kHz bilaterally; b) normal tympanograms, and c) negative otologic history.

2.2 Materials

The car horn signal was extracted from a CD library of environmental sounds. It was a complex periodical signal of 1-sec duration and constant in intensity. The principal spectral components were at 700, 840 and 1045 Hz. Using this basic signal, a bank of increasing and descending car horn signals with different intensity changes were generated using MATLAB software. Figure 1 illustrates an increasing signal with a delta of 6 dB (difference in level between end and beginning of signal).

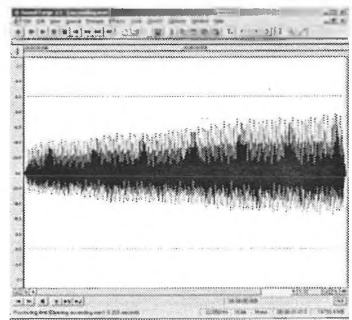


**Figure 1**: Car horn signal with increasing level (delta = 6 dB)

For listening conditions of the car horn in background noise, a 2-sec traffic noise recording was taken the CD library of environmental sounds. Using MATLAB, the bank of increasing and descending car horn signals was mixed with the constant traffic noise at a signal/noise ratio of + 6 dB.

The ascending and descending car horn signals, in silence and traffic noise, were then processed through a simulated hearing aid using MATLAB. The simulation was carried out for a "compression" hearing aid and a "linear" hearing aid. The simulated compression corresponded to an AGC-I hearing aid with a compression threshold of 45 dB SPL, a compression ratio of 3:1, and attack and release times of 3 ms and 100 ms respectively. A 1:1 ratio was used to reproduce the condition with linear amplification.

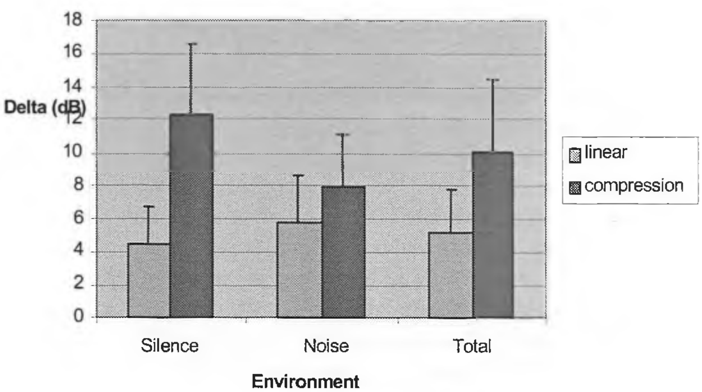
2.3 Procedure

All stimuli were presented at 65 dB SPL. For all subjects, the minimum detectable car horn signal-intensity change was measured using a forced-choice adaptive procedure for the two directions of signal change (increasing or decreasing), the two hearing aid simulations (compression or linear), and the two backgrounds (silence or traffic noise). In each experimental condition, the listeners received 20 stimuli. The first stimulus had a delta signal level change of 8 dB for conditions without compression and 16 dB with compression. Depending on whether or not the subject perceived the intensity change in the car horn, the next stimulus had a smaller or greater delta level change. The signal-intensity threshold was determined as the average delta level change (in dB) for the last 10 trials.

An adaptive A-B comparison procedure was also used to find the equivalence in the increasing (or decreasing) sensation between the two hearing aid conditions. The subjects received 20 pairs of stimuli. One of the stimulus was the car horn signal with hearing aid "compression" and its delta level change was always 15 dB. The second stimulus was the car horn signal with "linear" processing. The delta level change of the latter was varied adaptively from trial to trial to reach equivalence with the reference signal processed with compression. The starting delta was 5 dB. The ordering of the two stimuli was random from trial to trial. Subjects were tested in 4 experimental conditions (ascending in noise, ascending in silence, descending in noise and descending in silence). In each condition, the comparaison threshold was determined as the average delta level change (in dB) for the last 10 trials.

## 3. RESULTS AND DISCUSSION

Statistical analyses of the results revealed that there is a significant difference ($p < 0.05$) for the effect of hearing aid processing. Over all subjects and conditions, the average threshold in level change for the "linear" condition was of 5.1 dB, whereas for the condition "with compression", the threshold was of 10.1 dB. Thus, simulation of a hearing aid




with a compression ratio of 3:1 made intensity changes 1.98 times more difficult to perceive than with a simulated linear hearing aid. This is illustrated in Figure 2.

**Figure 2**: Significant results for various environmental background conditions (noise/silence), with or without compression

Crossed effects processing × background also displayed a significant difference (Figure 2). In silence, perception of ascending or descending car horn signals was more difficult in the compression (12.3 dB) condition than linear (4.5 dB) amplification by a factor of about 2.75. Thus, in silence, the effect of compression can be predicted by the hearing aid compression ratio (3:1 in the present study). This is because hearing aid compression is completely controlled by the signal in silence, and thus the car horn level increases or decreases are diminished according to the compression ratio. In noise (Figure 2), the level change threshold is also higher with compression (7.8 dB) than with linear (5.7 dB) amplification, but by a factor of only 1.37. This is because compression is now controlled by the whole sound (signal + noise), and the constant noise has the effect of linearizing the gain of the hearing aid (less gain changes as the signal increases or decreases).

Perceiving ascending or descending signals was easier in noise (6.8 dB) than in silence (8.4 dB). There was also a significant difference in the crossed effect background × direction of level change. In silence, perception is as easy in the ascending (8.6 dB) as the descending condition (8.2 dB), while in noise perception in ascending condition is much easier (5.2 dB) than in the descending condition (8.3 dB).

Statistical analyses carried out with the adaptive A-B comparison procedure revealed essentially the same results. In silence, the adverse effect of compression on the perception of signal level changes can be predicted by the compression ratio of the hearing aid.

## 4. CONCLUSIONS

The results confirm the hypothesis that signal compression in hearing aids make signal-intensity changes more difficult to perceive than linear amplification. However, the effect depends strongly on background noise. In silence, the effect of compression is most pronounced, and it can be predicted by the compression ratio of the hearing aid. In noise, the effect is less pronounced. It appears that the most difficult situation for perceiving signal-level changes with compression hearing aids is when the compression ratio is high and the signal/noise ratio is also high. In contrast, the adverse effects of compression are minimized when the compression ratio is low and the signal/noise ratio is also low, assuming constant noise. These results need to be confirmed with subjects with sensorineural hearing losses.

## 5. REFERENCES


[1] Ashmead, D.H., LeRoy, D. & Odom, R.D. (1990). Perception of the relative distance of nearby sound sources. *Perception in Psychophysics,* 47(4), 326-331.

[2] Ashmead, D.H., Davis, D.L. & Northington, A. (1995). The contribution of listener's approaching motion to auditory distance perception. *Journal of Experimental Psychology: Human Perception and Performance,* 21(2), 239-256.

[3] Wiener, W.R., *et al.* (1997). The use of traffic sounds to make street crossings by persons who are visually impaired. *Journal of Visual Impairment and Blindness*, 91(5), 435-445.

[4] Leroux, T., *et al.* (2001). *The problem of spatial perception in individuals with auditory and/or visual impairments.* Oral presentation at the XXVth World Congress of the International Association of Logopedics and Phoniatrics, Montréal, 5-9 August.

[5] Giguère, C. *et al.*, (2002) *Évaluation auditive de la distance.* Revue de littérature par le GILA – Groupe d'intérêt en localisation auditive.

# The Influence of Baffle Orientation on Sound Attenuation of Dissipative Silencers

**Harrison Richarz**
175 Rue Sherbrooke Ouest, Montreal, QC, H2X 1X5 harrison.richarz@mail.mcgill.ca

## 1. INTRODUCTION

Noise transmission in HVAC systems is often controlled with the installation of "dissipative silencers" in the air-ducts. For rectangular ducts, the preferred arrangement consists of several baffles. The greater the length of the baffle, the greater amount of acoustic energy absorbed. Two other parameters control the sound absorption: the thickness of the baffle and the size of the air space between the baffles (1,2,3). Another geometric parameter, namely the ratio of the cross-sectional area to the perimeter also influences sound absorption. Although this has been known for a considerable period of time, no one seems to have examined this effect in a systematic manner to date. This experimental study reports some interesting results for ducts with square cross-sections and un-conventional baffle arrangements.

## 2. METHOD

The noise reductions afforded by 10 different fiberglass baffle configurations were measured. Each configuration consisted of rigid, acoustic grade fibreglass board cut to appropriate size. These were installed at one end of a 122cm long 30cm square steel duct. An ILG fan was used as a sound source. This arrangement is not as sensitive to radiation loading are loudspeakers that are often used in model experiments. The test section was attached to a 91cm long duct that terminated in a reverberant space. The insertion loss was taken to be the difference between the sound pressure level with and without silencer elements.

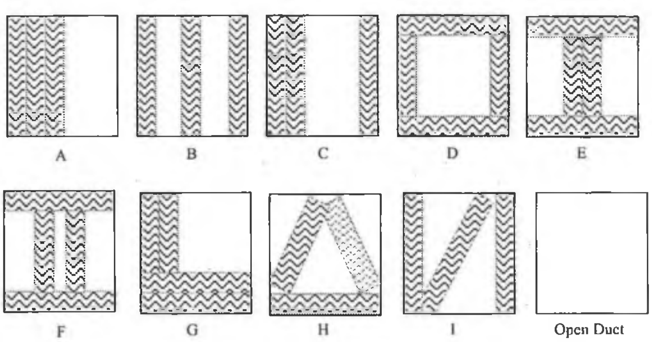



Figure 1. Schematic of test configurations

## 3. RESULTS

All measured data exhibited poor attenuation at low frequencies, rising to a maximum, and then decreasing again at higher frequencies. The measured data was very repeatable; deviations from the mean were within +/- 0.5 dB.

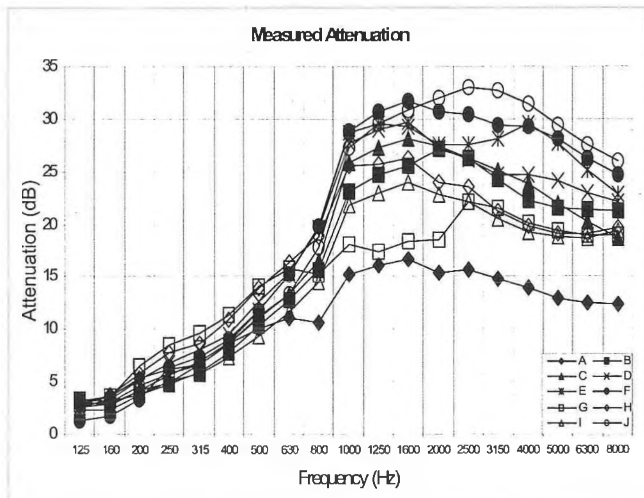



Figure 2. Measured Insertion Losses

For design purposes it is important to select the frequency (***f****peak*) where maximum attenuation (***IL****max*) is required. This is because the attenuation band of most silencers is not easily widened. Suppose the maximum attenuation $L_0$ occurs at $f_0$., and the levels and frequencies in the neighboring bands are ($L_1$,$0.8f_0$) and ($L_2$, $1.25f_0$). Introducing $z=(f/f_0 -1)$ one can generate three equations for the coefficients of the quadratic: $L(z)=a+bz+cz^2$. The quadratic has a maximum at $z=-0.5b/c$ from which $IL_{peak}=L_0-0.25b^2/c$, and $f_{peak}=f_0(1-0.5b/c)$ follow.

Table 1

| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| fmax (Hz) | 1623 | 2117 | 1681 | 1653 | 3838 | 1708 | 2620 | 1697 | 2603 |
| ILmax (dB) | 17.0 | 27.4 | 28.2 | 29.7 | 29.7 | 31.9 | 22.3 | 23.9 | 33.0 |
| Dh | 0.20 | 0.12 | 0.20 | 0.20 | 0.14 | 0.10 | 0.20 | 0.11 | 0.11 |
| λ/Dh | 1.03 | 1.32 | 1.00 | 1.01 | 0.65 | 1.96 | 0.64 | 1.87 | 1.17 |
| ℒ/Dh | 2.95 | 4.92 | 2.95 | 2.95 | 4.43 | 5.91 | 2.95 | 5.60 | 5.35 |
| Ao/A | 0.50 | 0.50 | 0.50 | 0.44 | 0.44 | 0.44 | 0.44 | 0.54 | 0.49 |
| Ad/Ao | 1.00 | 0.50 | 1.00 | 1.00 | 0.50 | 0.33 | 1.00 | 0.33 | 0.46 |
| Pl/Pd | 0.33 | 0.80 | 0.67 | 1.00 | 0.67 | 0.75 | 0.50 | 0.62 | 0.76 |

# 4. DISCUSSION

Conservation of energy suggest that

$$p^2(x)A_d/\rho c = p^2(x+dx)A_d/\rho c + \alpha p^2 P_l dx/\rho c$$

Here $A_d$ is the duct area, $P_l$ the lined perimeter of the duct and $\alpha$ a measure of the amount of energy absorbed. The differential form is has a solution $p^2(x)=p^2(0)e(-\alpha P_l x/A_d)$. Noting that $10\log[p^2(0)/p^2(x))]$ is the attenuation, it follows that $IL=bP_l\mathcal{L}/A_d$ which can be expressed in terms of geometric parameters of the air passage: $IL=B\ P_l/P\mathcal{L}/D_h$, where $D_h$ is the hydraulic diameter the open air-passage. Now B is not likely to be a constant, as the attenuation depends on the amount of energy dissipated in the sound absorbing material. The absorption coefficient of fibrous material is a function of the thickness of the material. The factor B takes the form: $60\ t_{eff}^{\ 0.65}$; $t_{eff}=(A/Ao-1)D_h\ P/P_l$ when IL is a maximum ($IL_{max}$).

The theory of dissipative silencers suggests that the peak frequency and the inter-baffle spacing (***h***) are related $B=\dfrac{2hf_{peak}}{c}$ where $c$ is the speed of sound and $f_{peak}$ is the frequency for which peak attenuation is realized. The present configurations differ from the typical baffle arrangement, and an effective inter-baffle spacing $h_{eff}$ is required. A review of the measured data shows that the attenuation curves are not all self- similar. This suggests that it may not be possible to extend the concept of $h_{eff}$ to non-parallel baffle configurations. It is likely that $h_{eff}$ will correlate with the hydraulic diameter $D_h$ and square root of the individual air-duct cross-sections. In fact the data for both fall into three distinct groups. and $c/(D_h f_{peak})$ falls into three distinct groups

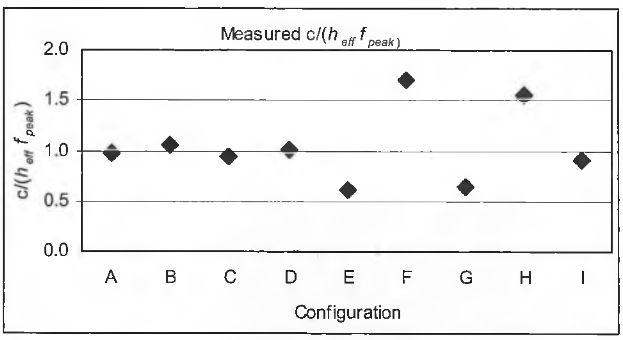



Figure 3. Pattern of $c/(h_{eff}f_{peak})$ ratios

Configurations that are of the parallel baffle type group tend to take on values near unity, whereas configurations with more complex geometry take on higher or lower values. The results are shown in figure 3 use the parameter

$$heff = 0.5(D_h + A_d^{\ 1/2})$$

In view of the limited number of configurations, it is not possible to establish a trend. It does appear that configurations that contain multiple passages with different wall treatment (Configurations F and H) have $c/(h_{eff}\,f_{peak})$ ratios greater than unity. Those with L-shaped lay-out (E,G) have $c/(h_{eff}f_{peak})$ less than unity.

# 5. SUMMARY

Several silencer configurations suitable for installation in rectangular ducts have been tested. It was found that the peak insertion loss could be described by a general expression that takes into account several geometrical parameters that describe the configuration:

$$IL_{max} = 60\ (A/A_o-1)P_l/P^{\ 0.35}\ L/D_h\ (dB)$$

This value occurs at

$f_{peak}=c/h_{eff}$; $heff=0.5(D_h+A_d^{\ 1/2})$ : Similar ducts
$f_{peak}=1.5c/h_{eff}$; $heff=0.5(D_h+A_d^{\ 1/2})$ : L shaped baffles
$f_{peak}=0.6c/h_{eff}$; $heff=0.5(D_h+A_d^{\ 1/2})$ : different ducts

All silencer configurations tested had the same length, making it impossible to separate out end-effects, which are due to diffraction. As the magnitude of the end effects is normally small (2-4 dB), one would expect that the above formalism will not change significantly, when this effect is accounted for.

# 6. ACKNOWLEDGEMENTS

The author would like to thank Mr. A. Della Penna President of Coolbreeze Air Conditioning and Heating Inc. for making available the test facility and instrumentation, and Dr. Sharbaf Ebrahimi for his guidance and encouragement. This paper is based on Experimental Investigation of the Effect of the Configuration of Fibreglass Baffles on Sound Attenuation in a Rectangular Duct, an extended essay in partial fulfillment of the requirement of the author's International Baccalaureate (2001).

# 7. REFERENCES

1) Beranek, L. L.. *Noise and Vibration Control* McGraw-Hill, New York, NY: 1971.

2)Morse, P.M. *Vibration and Sound* Acoustical Society of America, Woodbury, NY: 1976.

3)Munjal, M.J. Acoustics of Ducts and Mufflers Wiley Interscience, 1987

EXPERIMENTAL STUDIES ON THE MULTI-CHANNEL ACTIVE CONTROL OF COMPLEX SOURCES IN THE FREE-FIELD

**Haoming Qin, Ann Nakashima, Desheng Li and Murray Hodgson**
School of Occupational and Environmental Hygiene and Dept. of Mechanical Engineering
The University of British Columbia, 3rd Floor, 2206 East Mall, Vancouver, BC, Canada, V6T 1Z3
e-mail : qinhming@yahoo.com

# 1. INTRODUCTION

In this experimental study, single and multiple channel control of monopole and dipole primary sources in the free-field were investigated. In previous experiments of active control, most studies have concentrated on controlling a single-tone low-frequency signal; however, many real low-frequency noise sources also emit harmonic frequencies. In order to relate the experimental studies to more useful practical applications, such as the active control of propeller aircraft noise, the effectiveness of a control system on two-tone signals emitted from both a single primary source and two primary sources was investigated.

# 2. EXPERIMENTAL SETUP

The experiments were performed in a $4.7x4.2x2.2m^3$ anechoic chamber at the University of British Columbia. A total of 56 points were measured at the height of the control system (1.3m) with a resolution of 0.5m, as shown in Fig. 1. 1-, 3- and 4-channel control was used on single dipole and single monopole primary sources, as well as two monopoles. Using computer simulations, the distance between the primary source and control speakers (d), the distance between the control speakers and the error sensors (t) and the distance between adjacent control channels (s) were optimized for maximal noise reduction. The test signals were pure tones of 100Hz and 200Hz. In total, six different experiments were performed, as summarized in Table 1.

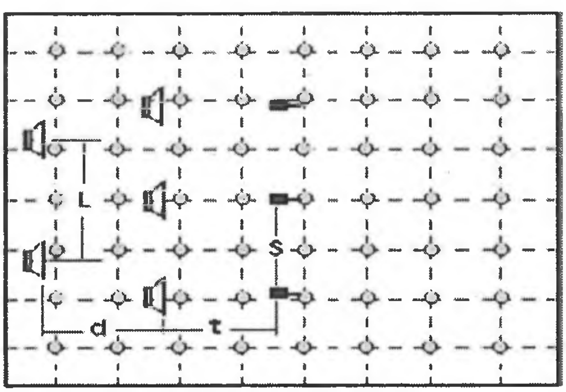



Fig. 1. Measurement points at the height of the control system.

Table 1. Summary of active control experiments.

| No. | Noise | Control | Frequency |
|---|---|---|---|
| 1 | 1 dipole | 1 channel | 100 Hz |
| 2 | 1 dipole | 3 channels | 100 Hz |
| 3 | 1 monopole | 1 channel | 100 + 200 Hz |
| 4 | 1 monopole | 3 channels | 100 + 200 Hz |
| 5 | 2 monopoles | 3 channels | 100 + 200 Hz |
| 6 | 2 monopoles | 4 channels | 100 + 200 Hz |

# 3. RESULTS

For active control of a dipole noise source, it has been found experimentally that an active control system is most effective when it is placed in the direction of the strongest primary noise directivity [2]. The dipole experiments were thus performed with a maximum of the dipole radiation facing the control system. Fig. 2 and Fig. 3 show the control results for a dipole source using a pure tone of 100Hz with one-channel control and three-channel control, respectively. For both cases, d=0.8m and t=0.6m were used, and in the three-channel case s=0.9m was used. Using one-channel control, the attenuation achieved in most areas of quiet zone was more than 12dB, with the maximum attenuation being 30dB. Using the three-channel system, the attenuation in most areas of the quiet zone increased to 15dB, with the maximum being 36dB.

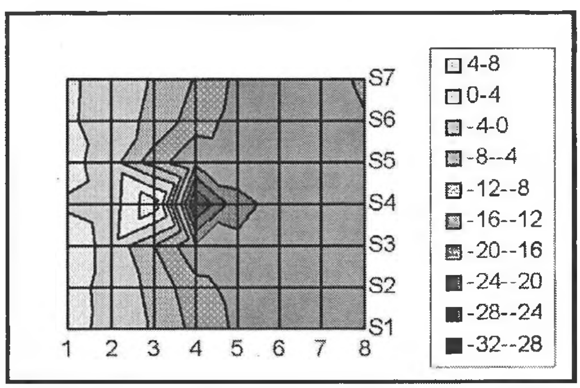



Fig.2. One-channel control of a single dipole at 100Hz.

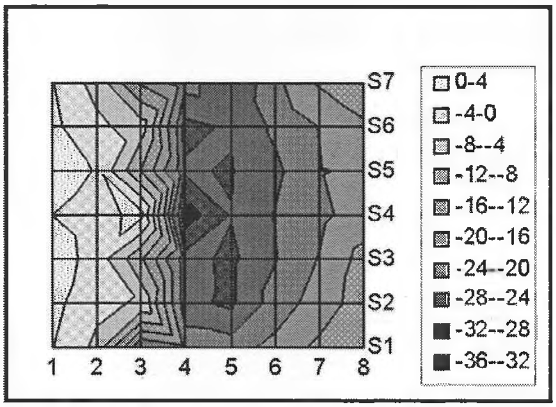



Fig. 3. Three-channel control of a single dipole at 100Hz.

The control results for a single monopole emitting two tones of 100 and 200Hz are shown in Fig. 4 (one channel) and Fig. 5 (three channels). In the one-channel case, d=0.8m and t=0.6m was used. Because of the two-tone signal, the size of quiet zone created by one channel control is very small. In most areas, only about 3dB of attenuation was achieved. For the three-channel system, using d=0.6m, t=1.0m and s=0.9m, the quiet zone is expanded considerably, with attenuation of up to 27dB in the center.

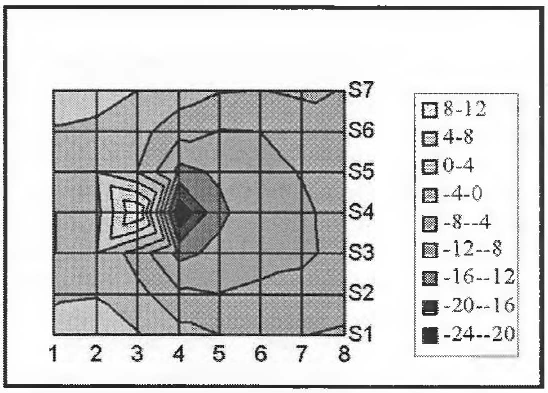



Fig. 4. One-channel control of a single monopole emitting a two-tone signal of 100 and 200Hz.

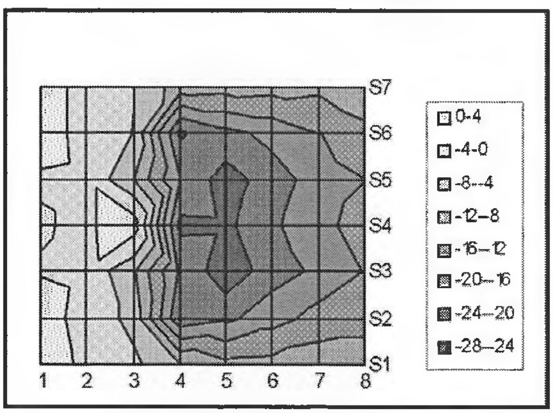



Fig. 5. Three-channel control of a single monopole emitting a two-tone signal of 100 and 200Hz.

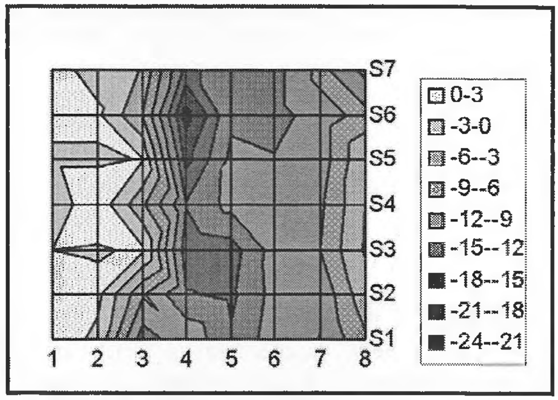



Fig. 6. Three-channel control of two monopole sources spaced 1.2m apart, emitting a two-tone signal of 100 and 200Hz.

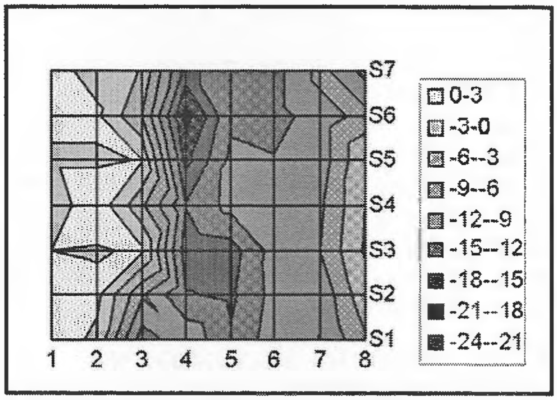



Fig. 7. Four-channel control of two monopole sources spaced 1.2m apart, emitting a two-tone signal of 100 and 200Hz.

The performance of the same three-channel control system on two monopoles, spaced 1.2m apart, emitting a two-tone signal of 100 and 200Hz is shown in Fig. 6. Comparing the results to those shown in Fig. 5, the effectiveness of the control system is reduced, as only about 3dB of attenuation was achieved in most areas. Increasing the number of control channels to four improved increased the attenuation to about 6dB as shown in Fig. 7.

## 4. CONCLUSION

The experimental results clearly show that the performance of an active control system can be improved by using multiple control channels. A multi-channel system not only increased the attenuation of the control area, but also increased the size of quiet zone. It has also been shown that more complex noise sources can be controlled by using a multi-channel system, suggesting that it may be possible to control real noise sources such as propeller aircraft.

## REFERENCES

1. J. Guo, J. Pan and C. Bao, "Actively created quiet zones by multiple control sources in free space " J. Acoust. Soc. Am. **101**(3), (1997).

2. A.Nakashima, J. Guo and M. Hodgson, "Preliminary investigation of active control for dipole noise sources", ACTIVE 2002 proceedings (2002).

# EFFECT OF THE GROUND SURFACE AND METEOROLOGICAL CONDITIONS ON THE ACTIVE CONTROL OF A MONOPOLE NOISE SOURCE

**Ann Nakashima and Murray Hodgson**
Dept. of Mechanical Engineering and School of Occupational and Environmental Hygiene
The University of British Columbia, 3rd Floor, 2206 East Mall, Vancouver, BC, Canada, V6T 1Z3
e-mail : naka@mech.ubc.ca

## 1. INTRODUCTION

Until recently, most studies of active noise control (ANC) have focussed on noise reduction in free- or half-space or in an enclosure; moreover, the medium is assumed to be non-refracting. In order to study the effectiveness of an ANC system in an outdoor environment, it is essential to include meteorological and realistic ground effects. In this research, the Green's function parabolic equation method (GFPE) developed by Gilbert and Di [1] was used to predict the sound field of a monopole primary source and a single monopole control source at 80Hz in one dimension for several atmospheric conditions. The results indicate the feasibility of using an active system to control the noise radiated from a Dash-8 propeller aircraft during engine run-up tests; 80Hz was chosen because it is the blade-passing frequency of the Dash-8.

## 2. ATMOSPHERIC AND GROUND CONDITIONS

In the prediction of long-range outdoor sound propagation, the effects of a refractive atmosphere must be taken into account. Refraction of sound waves is caused by temperature and wind-speed gradients. Temperature gradients can be described as a lapse or an inversion. During a temperature lapse, there is often a decrease of 5°C in the first meter above the ground and a further decrease of 3 to 5°C in the next 100m. A weak temperature inversion can be represented as an increase of 0.9°C/100m, a medium inversion as 1.8°C/100m and a strong inversion as 3.6°C/100m [2].

The variation in the wind speed is greatest near the ground surface, and decreases with increasing height. A realistic profile for the effective sound speed at height z is $c_{\text{eff}}(z) = c_0 + b\ln(z/z_0 + 1)$, where $c_0$ is the sound speed at the ground, $z_0$ is the roughness length of the ground surface in m, and b is a parameter describing the refraction in m/s [3]. For the results shown here, $c_{\text{eff}}(z)$ was calculated using b = 1 m/s for downwind, b = -1m/s for upwind, $z_0 = 10^{-4}$m for hard ground, and $z_0 = 10^{-2}$m for soft ground.

Single-channel ANC predictions were performed for hard ground and soft ground (Z = 14.65+13.63i), for weak, medium and strong temperature inversion conditions in the presence and absence of a downwind, as well as for temperature lapse conditions in the presence and absence of an upwind.

## 3. RESULTS

The GFPE code was modified to include a single active control channel using equations described by Guo and Pan [4]. The model was validated for single-channel ANC predictions by comparing the results obtained using the GFPE model with Guo's and Pan's model for a non-refractive atmosphere above a reflective plane. Predictions were done up to a horiztonal distance of 10km, using source and receiver heights of 4.5m. For all of the predictions, the control source and error sensor were positioned in a line at the height of the primary source. The control source was placed at a distance of 30m from the primary source, and the error sensor was placed 30m from the control source.

The results for temperature inversion and lapse conditions above a reflective ground surface are shown in Fig. 1. In the case of no refraction, about 13dB of attenuation was achieved in the far field. The temperature inversion caused fluctuations in the attenuation of ± 2 to 3dB. About 10dB of attenuation was achieved on average under weak and strong inversion conditions. Under medium version conditions, about 13dB of attenuation was obtained on average. In these cases there were significant spatial variations. The temperature lapse did not cause fluctuations in the attenuation, but the ANC system was less effective for this case overall, achieving only 5dB of attenuation.

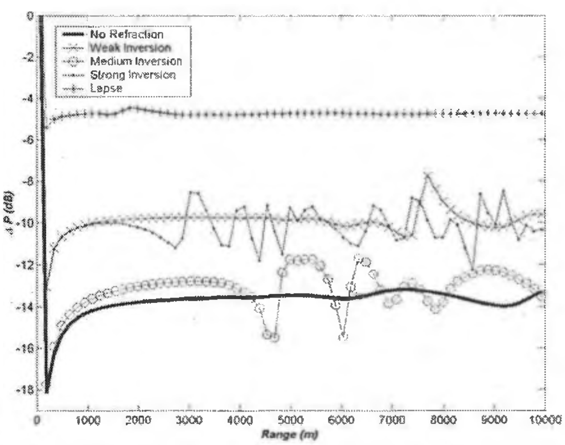



Fig. 1. ANC results for a monopole at 80Hz in the presence of temperature gradients above reflective ground.

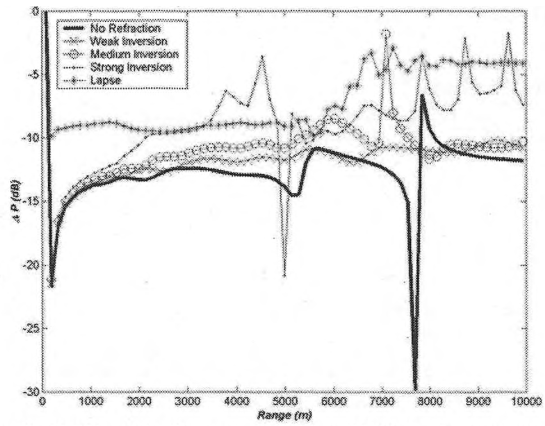



Fig. 2. ANC results for a monopole at 80Hz in the presence of temperature gradients above soft ground.

For temperature inversion and lapse conditions above soft ground, the results are shown in Fig. 2. The attenuation for a non-refractive atmosphere is also shown, to show the effect that the soft ground has on the control results (compare with Fig. 1). About 11 to 12dB of attenuation was achieved over the entire range, except at 7.7km where there was a sharp dip. Similar attenuation was obtained under weak inversion conditions. With the medium inversion, 11 to 12dB of attenuation was obtained on average, but a sharp peak (up to -2dB) occured at 7km. Under strong inversion conditions, there were large fluctuations in the attenuation in the far field, with the average attenuation being about 5dB. Under temperature lapse conditions, about 10dB of attenuation was achieved up to 5km; beyond the 5km point, only about 5dB was obtained.

Fig. 3 shows the results for temperature inversion with downwind conditions, as well as for temperature lapse with upwind conditions, in the presence of a reflective ground. The results for the different degrees of temperature inversion are almost identical, showing 4dB of attenuation along the entire range. The control system is ineffective under temperature lapse conditions in the presence of an upwind; the sound field increased by 2 to 7dB.

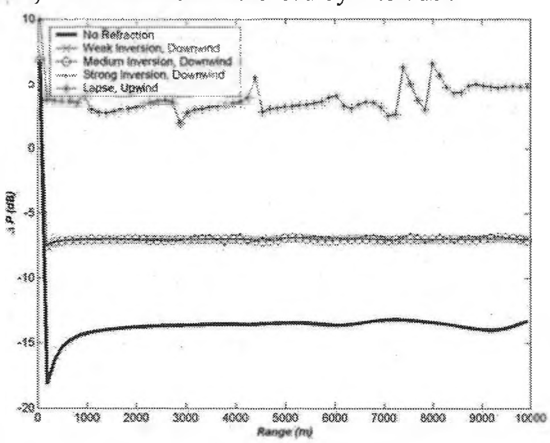



Fig. 3. ANC results for a monopole at 80Hz in the presence of temperature and wind-speed gradients above reflective ground.

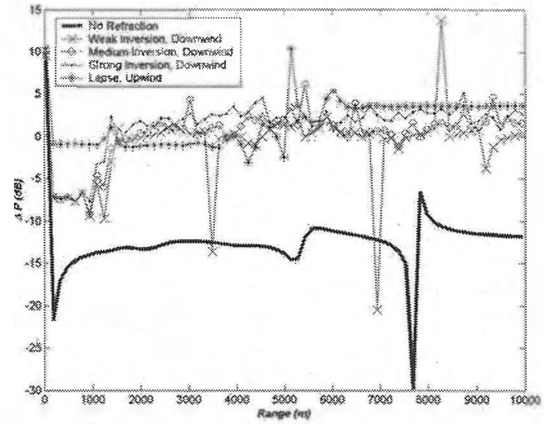



Fig. 4. ANC results for a monopole at 80Hz in the presence of temperature and wind-speed gradients above soft ground.

Fig. 4 shows the results for the most complex conditions, combining temperature and wind-speed gradients above soft ground. The control system did not achieve any attenuation at large distances. The weak inversion with downwind, and lapse with upwind, caused a few peaks and dips in the attenuation, while the field was relatively constant over the range for the medium and strong inversions with downwind conditions.

## 4. DISCUSSION

Clearly, atmospheric refraction and soft ground have a significant effect on the performance of a single-channel ANC system. It is most difficult to achieve attenuation under temperature lapse conditions, especially in the presence of an upwind. However, under such conditions the noise levels near the ground have already been signicantly attenuated due to the upward refraction of sound waves. It is thus more important to concentrate on attenuating the noise under temperature inversion and downwind conditions. Using single-channel control over soft ground, a quiet zone (10dB of attenuation or more) was achieved under weak temperature inversion conditions, but the control system was less effective for stronger temperature inversions. Future work should investigate multi-channels ANC system.

1. K. E. Gilbert and X. Di. A fast Green's function method for one-way propagation in the atmosphere. J. Acoust. Soc. Am. **94**(4), 1993.
2. T. F. W. Embleton. Tutorial on sound propagation outdoors. J. Acoust. Soc. Am. **100**(1), 1996.
3. E. M. Salomons. Computational atmospheric acoustics. Kluwer Academic Publishers, 2001.
4. J. Guo and J. Pan. Effects of reflective ground on the actively created quiet zones. J. Acoust. Soc. Am. **103**(2), 1998.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge Dr. Xiao Di of Penn State University for providing the GFPE source code, and as well as the valuable input received from Dr. Mike Stinson, Dr. Keith Attenborough and Dr. Sharam Taherzadeh.

# MEASUREMENT OF THE EFFECT OF FITTINGS ON LOW-FREQUENCY SOUND IN A SCALE MODEL WORKROOM

**Galen Wong and Murray Hodgson**
Dept. of Mechanical Engineering and School of Occupational and Environmental Health,
University of British Columbia, 3rd Floor 2203 East Mall, Vancouver, BC, Canada, V6T 1Z3
Email: galenw@interchange.ubc.ca

## 1. INTRODUCTION

The spatial variation of low-frequency sound was measured in a scale model workroom when empty, fitted with large fittings, and fitted with small fittings. Although the effects of fittings on high-frequency sound have been studied extensively [1,2], its effects on low-frequency sound have not generally been investigated. The results of this study will be used to determine the accuracy of prediction models to account for fittings in active noise control (ANC) prediction. Previous prediction work with ANC has not taken into account the effect of fittings [3,4].

## 2. EXPERIMENTAL SETUP

The scale model room was set up at 1/8th scale to model a rectangular workroom with dimensions of 30mFS x 15mFS x 7.5mFS (FS = Full Scale equivalent value). The floor was unpainted concrete. The walls were made of 10mm plywood, varnished on the inner surface while the roof consisted of 3mm plywood, varnished on the outer surface. The walls and roof were supported by a metal frame.

The large fittings were 2mFS cubes constructed of 6mm plywood, varnished on the outside. Ten cubes were evenly distributed around the room, leaving an area of about 8mFS around the source unobstructed. The small fittings were solid wood blocks, with dimensions of 1mFS x 0.65mFS x 0.65mFS. Forty blocks were evenly distributed around the room, also leaving an area of about 8mFS around the source unobstructed.

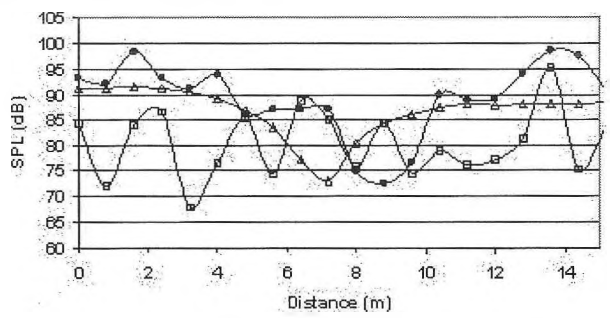
A single 100mm loudspeaker was used as a source in the scale model tests. It was placed in a 120mm x 120mm x 200mm enclosure, with the speaker set in 80mm from the top face of the enclosure. For measurements of the steady-state level, the speaker was positioned facing one corner of the scale model, in an effort to obtain the maximum excitation of the modal response in the scale model.

Sound pressure level measurements were made along a line down the length of the scale model room. Also, two lines along the width of the room were measured, one close to the source, and one far from the source. To ensure that the line measured was not along a nodal line, the lines were carefully chosen and measurements were made along lines with many antinodes in the empty room. This resulted in choosing slightly different lines for each of the frequencies measured. The receiver microphone was placed at a height of 1.6mFS for all measurements.

## 3. RESULTS

The scale model workroom was tested using pure tones at 31.5, 63, and 125HzFS under three different conditions – empty, fitted with large fittings and fitted with small fittings. The sound pressure level versus distance away from the source was compared at each of the three frequencies measured, for the three different room conditions. Fig. 1 shows an example of the sound pressure level variation with distance away from the source measured along the width of the scale model room.

As shown in Fig. 1, the addition of fittings caused the variation in low-frequency sound pressure level in the room to change significantly from the empty workroom. In order to more easily and directly compare the differences between each of the room conditions, a linear regression line was fitted to the data and the residual standard deviation ($S_{y|x}$) was calculated. These values are listed in Tables 1, 2 and 3 corresponding to measurements made along the length of the workroom, the width close to the source and the width far from the source, respectively. The reverberation times (RT) were also measured for the three different room cases; their values are listed in Table 4.




● - Empty △ - Small Fittings □ - Large Fittings

Fig. 1. SPL versus distance measured along the width of the scale model room close to the source, at 31.5HzFS.

Table 1. Residual standard deviation (in decibels) of the sound pressure level at 31.5HzFS, 63HzFS and 125HzFS, measured along the length of the workroom.

| | 31.5HzFS | 63HzFS | 125HzFS |
|---|---|---|---|
| Empty | 7.0 | 7.3 | 7.8 |
| S Fitted | 3.5 | 7.1 | 5.7 |
| L Fitted | 7.7 | 7.1 | 5.1 |

Table 2. Residual standard deviation (in decibels) of the sound pressure level at 31.5HzFS, 63HzFS and 125HzFS, measured along the width of the workroom close to the source.

| | 31.5HzFS | 63HzFS | 125HzFS |
|---|---|---|---|
| Empty | 7.4 | 6.1 | 6.8 |
| S Fitted | 4.9 | 3.3 | 5.8 |
| L Fitted | 6.7 | 8.5 | 5.7 |

Table 3. Residual standard deviation (in decibels) of the sound pressure level at 31.5HzFS, 63HzFS and 125HzFS, measured along the length of the workroom far from the source.

| | 31.5HzFS | 63HzFS | 125HzFS |
|---|---|---|---|
| Empty | 8.3 | 5.1 | 3.4 |
| S Fitted | 7.6 | 6.9 | 5.3 |
| L Fitted | 6.4 | 4.8 | 4.5 |

Table 4. Reverberation times (in seconds, full scale) measured at 31.5HzFS, 63HzFS and 125HzFS, for the three different room cases.

| | 31.5HzFS | 63HzFS | 125HzFS |
|---|---|---|---|
| Empty | 6.3 | 7.6 | 7.2 |
| S Fitted | 6.1 | 6.6 | 6.3 |
| L Fitted | 3.5 | 5.0 | 5.7 |

## 4. DISCUSSION

Along the length of the scale model workroom, the most spatial variation occurs when the room is empty at 125HzFS, with a $S_{y|x}$ of 7.8dB. The addition of small fittings generally decreased the spatial variation compared to the empty room, but it was highest at 63HzFS. At 31.5HzFS, the small fittings decreased the $S_{y|x}$ to 3.5dB. With the large fittings, at 31.5HzFS, the $S_{y|x}$ increased slightly by 0.7dB, while at 125HzFS it was decreased by 2.7dB.

Close to the source and along the width of the workroom, spatial variation was highest at 31.5HzFS when empty. Small fittings decreased the $S_{y|x}$ by 2.5dB at 31.5Hz and 2.8dB at 63Hz. Large fittings decreased the $S_{y|x}$ slightly at 31.5 and 125HzFS, but increased it by 2.4dB at 63HzFS.

In the empty room, along its width and far from the source, the $S_{y|x}$ was highest at 31.5HzFS with a value of 8.3dB. At 125HzFS in the empty room, the $S_{y|x}$ was quite low, at 3.4dB. The addition of fittings decreased the $S_{y|x}$ at 31.5HzFS by 0.7dB and 1.9dB for the small and large fittings, respectively. At 63 and 125HzFS the fittings tended to increase the $S_{y|x}$.

Overall, the small fittings tend to smooth out the spatial variation in sound pressure level, but at distances far from the source, large fittings lowered the variation more than the small fittings.

The fittings decreased the overall RT in the room compared to the empty room; the large fittings more than the small fittings. At 31.5HzFS, the small fittings did not decrease the RT significantly compared to the empty room, but at 63 and 125HzFS it decreased the RT by about 1s. The large fittings decreased the RT considerably – by over 2.5s at 31.5 and 63HzFS and by 1.5s at 125HzFS. The monotonic decrease in RT due to the introduction of fittings (from empty to small fittings to larger) compared to the empty room does not tend to follow changes seen in the $S_{y|x}$ for the different room conditions.

## 5. CONCLUSIONS

The spatial variation in the sound pressure level in a scale model workroom is changed significantly with the addition of fittings in the room. Overall, small fittings tended to decrease the amount of variation in sound pressure level in the room. Small fittings tended to decrease the variation in sound pressure level more than large fittings, close to the source, while far from the source large fittings decreased the variation more than small fittings. The RT does not correlate with the amount of variation of the sound pressure level measured in the room.

## REFERENCES


Hodgson, M.R. (1983). Measurement of the influence of fittings and roof pitch on the sound field in panel-roof factories. Appl. Acoust., 16(5), 369-391.

Hodgson, M.R. (1999). Effective fitting densities and absorption coefficients of industrial workrooms. Acustica, 85, 108-112.

Li, D., Cheng, L., Gosselin, C.M. (2002). Active control of sound radiation into enclosures using structural error sensors. Proceedings of the 5th FSI, AE & FIV+N Symposium, ASME International Mechanical Engineering Congress and Exposition.

Li, D., Cheng, L., Gosselin, C.M. (2002). Active sound radiation control of cylindrical structures using piezoelectric actuators. Proceedings of the Annual Meeting of the Canadian Acoustical Association.

# Directional Characteristics Of An Outdoor Warning Siren

**Nigel Maybee**
HFP Acoustical Consultants Corp., Suite 1140, 10201 Southport Road SW, Calgary, AB T2W 4X9
Nigel@hfpacoustical.com

## 1. INTRODUCTION

A utility company installed an outdoor warning siren at one of its facilities to alert the surrounding community in the event of an emergency situation. The selected unit was a rotating, electromechanical siren supplied by a manufacturer of signaling devices. A single test of the system was conducted after it was installed and it was observed that siren noise was not clearly audible in some parts of the warning area. The siren is capable of various modes of operation and the utility company wanted to determine which operation mode would maximize the audibility of the siren throughout the warning area. Because of the potential public disturbance caused by further testing of the siren, a theoretical study of the siren noise was conducted to analyze the characteristics of the various operating modes.

## 2. SIREN OPERATION MODES

The siren is comprised of a single horn mounted on a motor that rotates the horn at a uniform speed. The speed of rotation is relatively slow (*i.e.* a few rpm) and is adjustable within a small range. The siren's electromechanical mechanism produces a steady primary tone when it is energized. The sound power of this tone is sufficient for the alarm to be normally audible up to a distance of about 3 kilometres if the siren is located in a semi-urban or rural area and is mounted well above the ground. The siren noise may be presented as a steady tone or as two types of modulating tones (*i.e.* fast and slow). The provision of modulating tones appears to be a feature designed to improve the urgency and 'attention-getting' aspects of the alarm. The various operating modes of the siren include selection of the type of tone emitted (*i.e.* steady or modulating) and the rotational speed of the horn.

The modulating tones are produced by periodically de-energizing and energizing the siren's electromechanical mechanism. For the fast mode, the unit cycles off and on at a rate of about 17 cycles per minute. For the slow mode, the unit cycles off and on at a rate of about 6 cycles per minute. During the 'off' interval of each cycle, the siren tone steadily decreases in both sound level and frequency. During the 'on' interval, the tone increases in sound level and frequency until the primary tone is again achieved. Figure 1 illustrates the sound of the 3 types of tones produced by the siren.

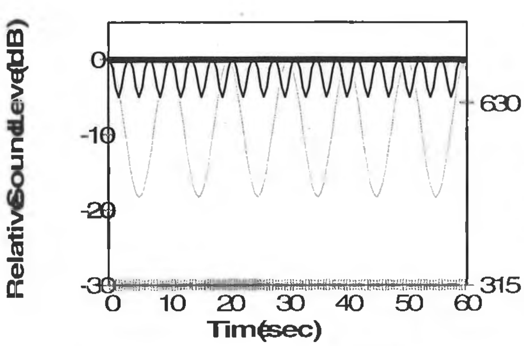



**Figure 1. Siren modes: Steady tone ▬▬ ; Fast modulating tone ▬ ▬ ; Slow modulating tone – – – – –.**

## 3. DIRECTIONAL CHARACTERISTICS

The manufacturer's data identifies the directional characteristics of sound radiation from the siren horn. Directivity plots of the radiation patterns were calculated [1] from the manufacturer's data. Radiation patterns for the 3 types of tones produced are shown in Figure 2. These patterns correspond to the sound radiation from the siren without rotation of the horn. The steady tone plot shows a single-lobe pattern directed along the axis of the horn. The fast and slow modulating tone plots show single-lobe patterns which correspond to the upper and lower limits of the envelope of each modulating tone. These are also directed along the horn axis.

Rotation of the horn provides 360 degree coverage of the siren tone. Since audibility of the alarm is its most important aspect, the directivity patterns of the maximum sound levels produced by the rotating horn were calculated. Maximum sound level results for each type of tone at two of the selectable rotational speeds are shown in Figure 3. For the steady tone mode, the sound level of the lobe remains constant as the horn axis sweeps through the circle. This results in a uniform distribution of the maximum sound level around the siren. At any location in the coverage area, the sound level of the tone is observed to increase to the maximum and then decrease with the passage of the lobe, repeating at the same rate as the rotational speed. This mode provides the best coverage (in terms of audibility) for a siren centrally located in a warning area.

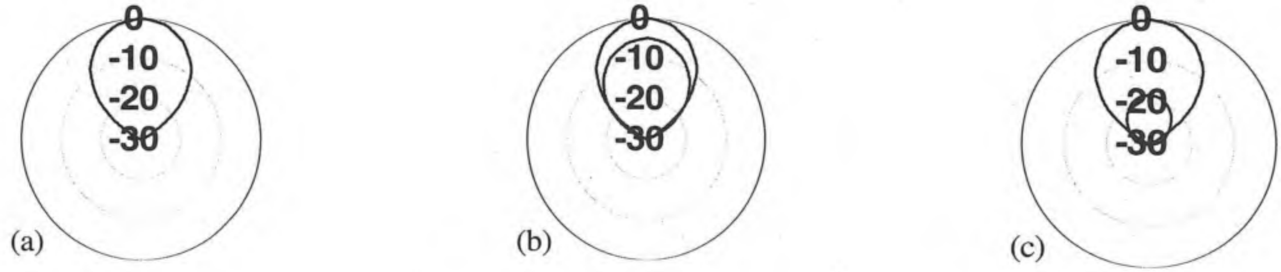



**Figure 2. Siren directivity without rotation: (a) Steady tone; (b) Fast modulating tone; (c) Slow modulating tone.**

For the modulating tone modes, the sound level of the lobe changes as the horn axis sweeps around the circle. This results in maximum sound levels that vary with angular position around the siren and causes the siren to display directional characteristics. These directional characteristics become most pronounced when the rotational frequency of the horn approaches the modulation frequency of the tone, which occurs for this siren when it operates in the slow modulation mode.

When the rotational speed is fast, the rotational frequency is equal to the modulation frequency. It is apparent from the directivity patterns in Figure 3(c) that the poorest coverage occurs when the frequencies of the modulation tone and the rotational speed are equal. In the areas of poorest coverage, the maximum sound level of the siren diminishes to the minimum sound level produced by the modulating tone.

## 4. CONCLUSION

These directional characteristics are not identified in the product literature for the siren and may not be apparent to the manufacturer. Moreover purchasers, possibly having minimal opportunity to test the siren because of potential community disturbance, may intuitively think that the modulating tone mode is the best way to operate the siren, but would be unaware that this mode could result in uneven coverage of the warning area.

## REFERENCES

Kinsler, L.E. et.al. "Fundamentals of Acoustics", (page 180) John Wiley & Sons, 1982.

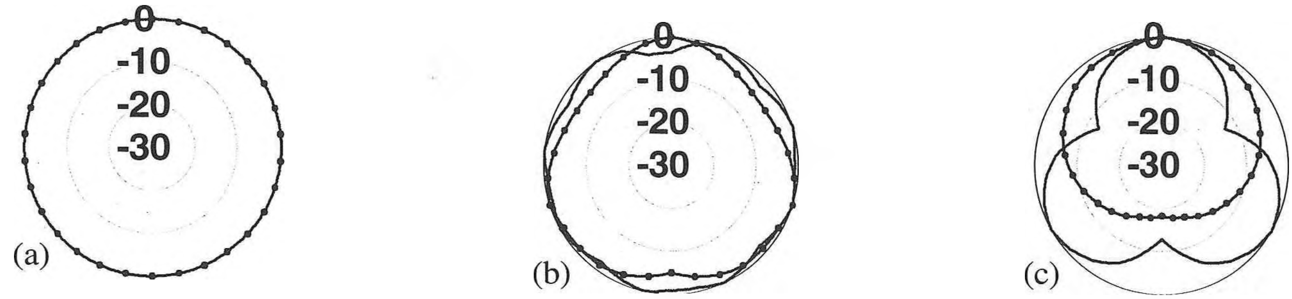



**Figure 3. Maximum siren sound levels with slow rotational speed ▬, and fast rotational speed —— :(a) Steady tone; (b) Fast modulating tone; (c) Slow modulating tone.**

## Interaction of Age and Attentiion on Driving Performance

**Mona Motamedi and Elzbieta B. Slawinski**
Dept. of Psychology, University of Calgary, 2500 University Dr. N.W. Calgary, AB, Canada, T2N 1N4

# 1. INTRODUCTION

The Canadian elderly population will increase in the next forty-five years and with it there will be a marked increase in the number of elderly drivers (65+ ) (Slawinski and MacNeil, 2002). It has been established that driving abilities gradually deteriorate with age (Szlyk, 1995). The auditory environment influences the performance of young and elderly drivers. The effect is dependent on the level of attention of the driver to the environment (Slawinski and MacNeil, 2002).

Safe and effective driving necessitates detection of important information embedded in a background of continuously changing masking sounds, called distracters. (Slawinski and MacNeil, 2002). This acoustic driving environment can affect driving performance as it forces the driver to shift or divide his/her attention between the road environment and the distracter.

This study investigates the effects of age and attention levels to auditory distracters on ability to detect external vehicular warning signals. Results show that the level of attenuation to a warning signal (target signal) decreased (intensity increased) when participants were attending to the distracter and simultaneously to the warning signals. These observations were true for young and elderly drivers. The effect of age was manifested in a decrease in the level of attenuation for both the attended and unattended conditions obtained for elderly participants.

# 2. METHOD

2.1 Participants

While data collection is still in progress, the results up to date are shown here with the final results to be reported at a later date. Sixteen younger drivers, represented as Group A (18-31 years old, M = 22.5; SD = 4.39) and nineteen elderly adult drivers, represented as Group B (65-79 years old, M =69.07; SD = 4.41) served as participants.
Analysis of data was done on those who met criteria as described in section 2.2 (fourteen younger drivers and fifteen elderly drivers). All participants had normal hearing as per age. No neurological problems were reported.

2.2 Stimuli

The detection tasks were carried out using a Hewlett Packard (Vectra) platform. Stimuli were presented using custom software and a TDT (Tucker-Davis Technologies) hardware. Two stimuli were used in the detection task: a warning signal (car horn) with a 300 ms duration simultaneously with a continuous distracter (a narrated story). The intensity level of the story was held constant at 67 dB SPL in both conditions.

The initial attenuation of the target stimuli (car horn) was 40 dB; subsequent attenuation levels were chosen by an adaptive psychophysical procedure (3 up and 1 down) as described by Slawinski and MacNeil ( 2002).

A ten item multiple choice test (questionnaire) was administered following the attended condition. Eligibility criterion to be included in analysis was such that a minimum of three questions had to be answered correctly.

2.3 Procedure

Participants were tested individually in an IAC anechoic chamber. Sound was presented binaurally via Sony stereo headphones. Participants were asked to detect a stimuli (car horn) while a narrated story was played. The detection task included two intervals and the car horn would randomly occur in one of the intervals. Participants were asked to indicate which interval contained the warning sound by pushing one of two corresponding buttons on a response console. An adaptive tracking procedure (3 correct responses followed by an increase in attenuation and 1 incorrect response followed by a decrease in attenuation) was used. Each presentation was composed of 9 reversals [changes in direction of curve (see Figure 1)] after which the subsequent presentation started with an intensity of 10 dB higher than the final intensity level of the previous presentation. For each presentation the first four reversals had a 5 dB step of increase and decrease in intensity while the last five reversals had a 2 dB step. Average of means of three presentations was used as a threshold value for a particular condition.

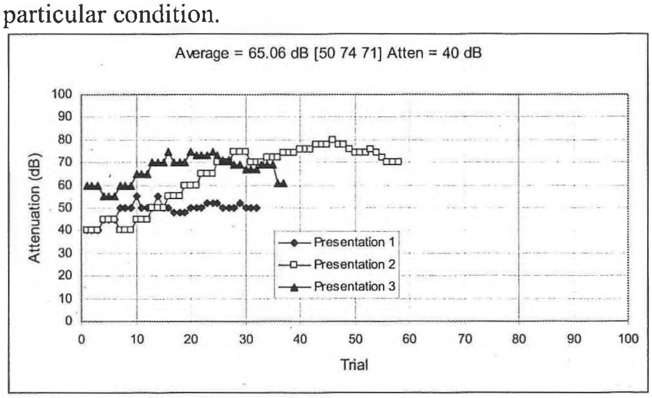



*Fig. 1.* An example of the output generated for the adaptive tracking procedure

For each participant the detection task was repeated twice: In one detection task, the "attended condition", the participant was asked to attend to the story. In the other detection task, the "unattended condition", the participant was asked not to attend to the story and focus only on the warning signal. The order of the two detection tasks was randomized among the participants so that half the participants were given the "attended" detection task first while the other half were given the "unattended" detection task first.

A mean threshold (the last 4 reversal) was calculated for each of the three presentations of the detection task in the two conditions. The mean threshold of detection (attenuation level) for each participant and condition was computed as the average of the three mean thresholds.

## 3. RESULTS

The experiment showed a dependence of the obtained attenuation level on the level of attention. For group A, the mean attenuation level obtained for the "attended condition" was 56.58 ± 12.92 dB while the one for the "unattended condition" was 64.27 ± 2.43 dB. For group B, the mean attenuation level obtained for the "attended condition" was 39.02 ± 23.09 dB while the one for the "unattended condition" was 49.19 ± 22.17 dB. Figure 2 shows the levels of attenuation for both groups for the two conditions (attended and unattended).

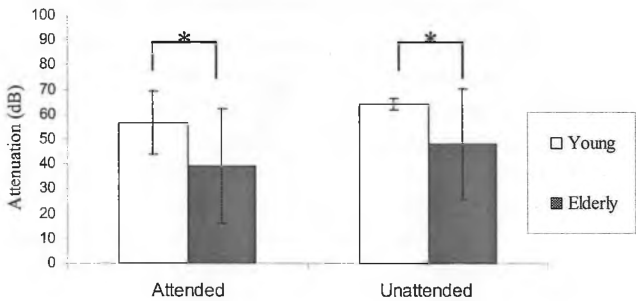



*Fig. 2.* Young vs. elderly participants attenuation levels in the attended and unattended conditions. (* Significant difference)

A one way ANOVA conducted for the attended condition showed a significant difference between the threshold means obtained for young participants and those obtained for elderly participants $F(1,27) = 5.95, p < .05$. A significant difference was also found between young and elderly participants when comparing the threshold means in the unattended condition $F(1,27) = 6.72, p < .05$. Levene's test for homogeneity of variances showed a significant difference between the variances of elderly and younger participants in both attended $F(1,27) = 9.64, p < .05$ and unattended conditions $F(1,27) = 30.5, p < .001$.

## 4. DISCUSSION

This study confirmed the hypothesis that the detection of vehicular warning sounds depends on the level of attention to the surrounding acoustic environment. Attenuation levels for young and elderly participants decreased when they were asked to attend to the narrated story compared to the condition when they were asked to not attend to the story.

This study also showed that there is a significant difference between the attenuation level obtained for young and elderly drivers. This result supports our hypothesis that detection of vehicular warning sounds in an acoustic environment decreases with age. The results are consistent with those obtained for Slawinski and MacNeil (2002) in which the effect of age on detection of distracters was studied.

Our study also shows that the ability to divide attention diminishes with age. A significant difference was found in the variability of the attenuation levels obtained between young and elderly participants. As is evident from Figure 2, elderly participants have significantly larger variability in both attended and unattended conditions compared to the young participants.

Another very interesting result can be seen in the large variability of the results in the case of the "Attended" condition. For young participants the standard deviation obtained was ± 2.43 dB for the "unattended" condition, while the one obtained for the "attended" condition was ± 12.92 dB. This shows that there is a great variability in how people divide their attention in the "attended" condition. This result, however, was not shown in elderly participants as the standard deviation for the "unattended condition" was ± 22.17 dB while the one for the "attended condition" was ± 23.09 dB. This can be attributed to the results obtained by Slawinski and MacNeil (2002) showing the decrease in the ability to detect vehicular warning sounds embedded in background noise.

## 5. REFERENCES


Slawinski, E.B.; McNeil, J. F. (2002). Age-Related changes on the Influence of Music on Driving Performance in the Detection of Vehicular Warning. *The Journal of Psychomusicology, 18*(2), *In Press.*

Szlyk, J., (1995). Relative effects of age and compromised vision on driving performance. *Human Factors 37*(2), 430-436.


## ACKNOWLEDGEMENTS


The authors wish to thank Dr Jane MacNeil and Samer M. Adeeb for their support with the experiment.

# THE EFFECTS OF ATTENDING TO AUDITORY STIMULI ON DRIVER SPEED AND LANE WEAVING BEHAVIORS IN A DRIVING SIMULATOR

**Kirsten Dugdale[1] and Elzbieta Slawinski[2]**
1) Dept. of Psychology, University of Calgary, Alberta, Canada, kldugdal@ucalgary.ca
2) Dept. of Psychology, University of Calgary, Alberta, Canada, eslawins@ucalgary.ca

## 1. INTRODUCTION

Research concerning the effects of various stimuli on driving behaviors has become more important as the number of drivers on the road has increased 64% since 1970 ("Aggressive Driving," 2002). Studies have ranged from the effects of alcohol on signal detection, divided attention and reaction speed (Gengo, Gabos, Straley & Manning, 1990) to the effects of fatigue (e.g. Lal & Craig, 2002) and recently, driving behavior while using a cellular phone (McKnight & McKnight, 1993). Music has also been studied in regards to its effects on driving, and has been found to reduce driver stress in highly frustrating traffic congestion (Wiesenthal, Hennessy & Totten, 2000).

Intuitively, all of these factors involve an attentional component, as does driving. Attention can be divided between many tasks, however with a limited capacity (Wiesenthal et al.). The amount of cognitive resources available for attention to each of these tasks decreases as the number of tasks increases (Wiesenthal et al.). Therefore, performing such tasks as conversing on a cellular phone or listening to an engaging auditory stimulus while driving increases the number of cognitive tasks the driver must perform, and decreases the amount of attention that can be paid to cognitively demanding tasks such as visual tracking and motor reactions.

The present study attempts to determine whether or not attending to an auditory stimulus while driving in a simulator will have an effect on drivers' speed and lane-weaving behaviors. A comedy sketch is used in the effect condition, as it is assumed that because of the pleasant nature of the stimulus, participants will be more likely to attend to the stimulus than they would be to a neutral, non-engaging stimulus (control condition). It is hypothesized that drivers attending to an engaging auditory stimulus will drive a significantly different speed than drivers listening to a nonsense recording of scrambled words. It is also hypothesized that drivers listening to the engaging stimulus will engage in significantly different lane-weaving behaviors than drivers listening to the scrambled words.

## 2. METHOD

Participants were 47 undergraduate students enrolled in at least one psychology class at the University of Calgary. Participation was voluntary through a bonus credit system. Participants were randomly assigned to the control condition (listening to scrambled words while driving) or the effect condition (listening to a comedy sketch while driving). There were 29 female participants enrolled in the study and 18 male participants. Participants' driving experience ranged from one to 20 years (M=5.6) and participants' age ranged from 18 to 39 (M=21.6). All participants were administered a hearing test before commencing the study, and all participants displayed average or above average hearing. Eight participants reported being prone to motion sickness while reading in a moving vehicle. Due to technical problems, data from two participants was unable to be collected, and four participants terminated participation in the study due to motion sickness after commencing the driving portion of the study. Final analysis was based on data from 41 participants: 20 in the effect condition and 21 in the control condition.

This experiment was carried out using the University of Calgary Driving Simulator (UCDS). The UCDS is comprised of a Saturn SL1 positioned in front of a wraparound screen. An elaborate connection of computers collected data from the driving scenario and sent data to the screens in front of the car [3]. Participants listened to the auditory stimuli through speakers in the car while they were driving a predetermined course in the simulator, and were asked to pay attention to the stimulus as they drove.

Participants in the effect condition listened to a Jerry Seinfeld comedy sketch, which was recorded using a male voice, and did not contain a laugh track. Upon completion of the drive, these participants were asked to complete a questionnaire regarding the content of the comedy sketch, to ensure they were attending while driving. Participants in the control condition listened to the same male speaker, with the exact same content in the stimuli, however the words were scrambled, though still spoken with prosody.

## 3. RESULTS

The simulator data consisted of measures taken every 0.33 seconds. Speed data, measured in meters per second, was averaged to compute the overall speed for the entire drive for each participant. An alpha level of 0.05 was used for all statistical tests. A one-way analysis of variance found that participants in the control (scrambled word) condition ($M = 20.134$, $SD = 2.377$) drove significantly faster than participants in the effect (comedy) condition ($M = 18.337$, $SD = 1.780$), $F(1,39) = 7.46$, $p = 0.009$.

**Table 1. Speed behavior results (meters per second).**

| | M | SD |
|---|---|---|
| Control | 20.134 | 2.377 |
| Effect | 18.337 | 1.780 |

Data on lane position was also collected every 0.33 seconds. In order to obtain an average lane-weaving statistic, the absolute difference between each measure was calculated, and theses difference scores were then averaged for each participant. A one-way analysis of variance determined that there was no significant difference in lane-weaving behaviors between the control group ($M = 0.07670$, $SD = 0.02291$) and the effect group ($M = 0.06769$, $SD = 0.01589$), $F(1, 39) = 2.12$, $p = 0.154$.

**Table 2. Lane-weaving behavior results (average lane position).**

| | M | SD |
|---|---|---|
| Control | 0.07670 | 0.02291 |
| Effect | 0.06769 | 0.01589 |

## 4. DISCUSSION

It is assumed that participants in the control condition did not attend to the stimulus, as it consisted of nonsense, scrambled prose. Participants most likely would have started the drive by attending to the prose (because they were asked to), however when they realized that it consisted of nonsense, would have stopped attending. Participants in the effect condition had to complete a questionnaire about the content of the comedy sketch following the drive, and if participants answered fewer than 80% of the questions correctly, their data would not have been included (however this did not occur).

As hypothesized, the average speed of drivers listening to the comedy sketch was significantly different than the speed of the drivers listening to the scrambled words, as drivers listening to the comedy drove slower. This could be attributed to the fact that since the drivers attended to the comedy stimulus and not the scrambled words, the amount of attentional resources available for the driving task becomes decreased. Since the drivers have an increased attention load, their ability to concentrate on the speed of the vehicle becomes decreased, perhaps leading to the decreased speed.

The hypothesis that lane-weaving behavior would also be affected by attending to an auditory stimulus was not supported. This could be due to the fact that while driving simulators replicate speed control accurately, they are deficient at replicating "lane keeping" behavior in comparison with actual on-road driving (Reed & Green, 1999).

## REFERENCES


Aggressive Driving. (2002). Retrieved April 2, 2003, from http://www.hopkinsmedicine.org/hse/memos/TSAFE/Aggressive%20Driving.pdf.

Gengo, F. M., Gabos, C., Straley, C. & Manning, C. (1990). The pharmacodynamics of ethanol: effects on performance and judgment. Journal of Clinical Pharmacology, 30, 748-754.

Lal, S .K. L. & Craig, A. (2002). Driver Fatigue: Electroencephalography and psychological assessment. Psychophysiology, 39, 313-321.

McKnight , A. J. & McKnight, A. S. (1993). The effect of cellular phone use upon driver attention. Accident Analysis and Prevention, 25(3), 259-265.

Reed, M. P. & Green, P. A. (1999). Comparison of driving performance on-road and in a low-cost simulator using a concurrent telephone-dialing task. Ergonomics, 42, 1015-1037.

Wiesenthal, D. L., Hennessy, D. A. & Totten, B. (2000). The influence of music on driver stress. Journal of Applied Social Psychology, 30(8), 1709-1719.


## ACKNOWLEDGEMENTS


Thanks to the Canadian Foundation for Innovation (CFI) as well as the staff of the Cognitive Ergonomics Laboratory at the University of Calgary and all participants who took part in this study.


## AUTHOR NOTES


This work was conducted while Kirsten Dugdale was a student at the University of Calgary. For more information about the driving scenario and UCDS setup, please send requests to kldugdal@ucalgary.ca.

# How Music of Differing Rhythmicities and Intensities Affects Driver Performance

**Benjamin Rich Zendel, Elzbieta Slawinski, and Linda Pearson**
Dept. of Psychology, University of Calgary, 2500 University Dr. N.W., Alberta, Canada, T2N 1N4

## 1. INTRODUCTION

Music has an effect on driving. Style of music has been shown to effect speed & control activity of drivers (Konz & McDougal, 1968). Volume/intensity of music has been shown to affect response times in driving-type behaviors (Beh & Hirst, 1999). Music with differing "arousal potentials" affects speed of driving (North & Hargreaves, 1999). Music has been found to affect stress levels of drivers (Wiesenthal, Hennessy, & Totten, 2000). Finally, speed, swerve, and red-light violations, all have been related to music tempo (Brodsky, 2001).

Research has shown that any sort of variation in music can affect arousal levels. Tempo (Kellaris & Kent, 1995), style (Gowensmith & Bloom, 1997), and 'musical mood' (Thompson, Schellenberg, & Husain, 2001) all have an effect on the arousal levels of listeners of music. Driving behaviors often change in different states of physiological arousal. Fatigued, and therefore less aroused drivers, tend to have more collisions, and make more driving errors then non-fatigued drivers (Dureman & Boden, 1972 & McBain 1970). People who are sleepy tend to do much worse at driving than someone who is fully alert (Dureman & Boden, 1972). It is has also been shown that too much arousal can lead to decreased performance in any task (Anderson, 1995). The concept that there is an ideal amount of arousal for optimal performance on a task is called the 'Yerkes-Dodson law.' Optimal performance occurs with intermediate levels of arousal. Music affects arousal levels and arousal levels have an effect on driving performance.

The current study examines musical intensity (volume), and rhythmicity and its impact on drivers' speed. Past research has shown that these variables do affect driver performance. No study has looked at the interaction of these variables. We expect that the variables will interact; rhythmicity should have a differing affect on driving at different intensity levels.

## 2. METHOD

All participants were given two questionnaires: susceptibility to simulator sickness and demographics. Participants were also given a chance to test drive the simulator.

The experiment occurred in three separate segments. Each of these segments was further divided into three zones. Road types varied from rural to highway to urban throughout the experiment.

A different music type was randomly played in each zone. This created a total of nine musical zones. After each segment of three zones, the participant was given the option of getting out of the car to stretch and get some water.
The two independent musical variables were intensity (volume) and rhythmicity. Volume was manipulated at three levels: quiet (65 dB), medium (75 dB) and loud (85 dB). Rhythmicity was also manipulated at three levels, low (50bpm; afro-Cuban percussion), medium (100bpm; rock and roll percussion) & high (200bpm; electronica percussion). The percussion track of the music was the only variation in the rhythmicity variable. The music itself stayed constant. The two manipulations created a total of nine musical variables (3 intensity by 3 rhythmicity).

The order of the music was counter balanced, so that an each of the nine musical variables occurred in each driving zone an equal number of times during the entire study. This allows us to average out the effect of differences due to road type. The differences in road type are only meant to be used as a means to increase generalizability.

Each zone took about one minute to complete. Participants were told to drive as they would normally. They were also told that if they felt uncomfortable at all during the study to exit the simulator.

## 3. RESULTS

The Speed data was analyzed using a 3 (volume: high, medium, low) by 3 (rhythmicity: high, medium, low) within subject's analysis of variance. Speed scores were measured by computing the average speed for each music type. Speed was measured in meters per second. The measure was again sampled three times per second. All scores of 0 m/s (i.e. the person was stopped), were removed from the data.

No significant effects or interactions were found in the original data. When the speed data was adjusted to lower variance a significant interaction was found. Variance was lowered by eliminating the differences between zones. The overall average speed was calculated (20.78 M/S). The average speed for each zone was also calculated (Z1=24.76 M/S, Z2=20.99 M/S, Z3=13.93 M/S, Z4=27.71 M/S, Z5=17.40 M/S, Z6=14.40 M/S, Z7=23.52 M/S, Z8=14.14 M/S, Z9=30.13 M/S). These zones varied tremendously due to differences in the roads themselves, not due to music differences. To eliminate this variance, each average was rounded to the nearest whole number, and the difference was found between that average and the overall average. That difference was then applied to every score from that zone. This lowered the variance between the scores due to road type, without changing the effect of music on speed. Each zone had three measures of each music type within the average. The analysis of this data is below.

The main effect of volume was not significant, $F(52,2)=1.984$, $p=0.148$. The main effect of rhythmicity was also not significant, $F(52,2)=0.704$, $p=0.499$.

The interaction between volume and rhythmicity was significant, $F(104,4)=2.93$, $p=0.024$ (Figure 1). Using simple main effects it was found that there was no difference between rhythmicity at high, $F(52,2)=2.65$, $p=0.080$, and medium, $F(52,2)=0.33$, $p=0.968$, volume levels. There was a difference in speed scores between the rhythmicity conditions only in the low intensity condition, $F(52,2)=3.707$, $p=0.031$. This difference was because there was a difference between high rhythmicity and low rhythmicity in the low volume condition, $t(26)=2.675$, $p=0.013$ (20.30 vs. 21.84). T-tests between high rhythmicity and medium rhythmicity, and medium rhythmicity and low rhythmicity were not significant, $t(26)=1.188$, $p=0.245$ and $t(26)=1.576$, $p=0.127$. In the low intensity condition rhythmicity was positively related with driver speed.

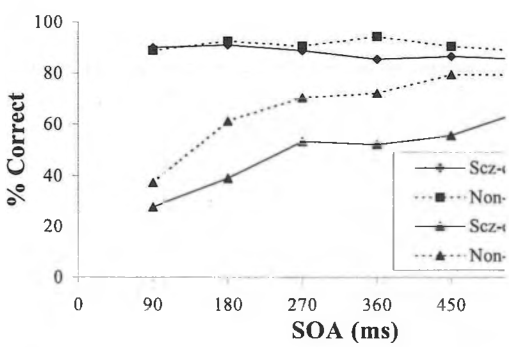



**Fig. 1. Driver speed as a function of the musical rhythmicity and intensity**

## 4. DISCUSSION

Music does affect driving. Past research, the current study, and theory point to this fact. The current study found that rhythmicity has an effect on driver speed in low intensity musical situations. This finding can be explained via driver arousal levels. Music that is louder increases arousal levels more than quieter music. In the louder musical conditions arousal levels are already maximized by musical intensity. Rhythmicity is therefore unable to increase arousal levels further. This results in no significant speed differences in the medium and high intensity conditions. In the low intensity condition (quiet), the rhythmicity is able to affect arousal levels, because the intensity does not increase arousal to its highest level. This allows rhythmicity to have an effect on driver speed.

## REFERENCES


Anderson, K.J. (1994). Impulsivity, Caffeine, and Task Difficulty: A Within Subjects Test of the Yerkes-Dodson Law. *Personality and Individual Differences, 16*(6), 813-829.

Beh, H.C. & Hirst, R. (1999). Performance on Driving-Related Tasks During Music. *Ergonomics, 42*(8), 1087-1098.

Brodsky, W. (2002). The Effects of Music Tempo on Simulated Driving Performance and Vehicular Control. *Transportation Research Part F, 4*(4), 219-241.

Dureman, E.I., Boden, C.H. (1972). Fatigue in Simulated Car Driving. *Ergonomics, 15*(3), 299-308.

Gowensmith, W.N. & Bloom, L.J. (1997). The Effects of Heavy Metal Music on Arousal and Anger. *The Journal of Music Therapy, 34*(2), 33-45

Kellaris, J.J., Kent, R.J. (1994). An Exploratory Investigation of Responses Elicited by Music Varying in Tempo, Tonality, and Texture. *Journal of Consumer Psychology, 24*(4), 381-401.

Konz, S. & McDougal, D. (1968). The Effect of Background Music on the Control Activity of an Automobile Driver. *Human Factors, 10*(3), 233-244.

McBain, W.N. (1970). Arousal, Monotony & Accidents in Line Driving. *Journal of Applied Psychology, 54*(6), 509-519.

North, A.C. & Hargreaves, D.J. (1999). Music and Driving Game Performance. *Scandinavian Journal of Psychology, 40*, 285-292.

Thomson, W.F., Schellenberg, E.G. & Husain, G. (2001). Arousal, Mood, and the Mozart Effect. *Psychological Science, 12*(3), 248-251.

Wisenthal, D.L., Hennessy, D.W. & Totten, B. (2000). The Influence of Music on Driver Stress. *Journal of Applied Social Psychology, 30*(8), 1709-1719.


## ACKNOWLEDGMENTS

We would like to thank the Canadian Foundation for Innovation (CFI), and all the staff and volunteers in the Cognitive Ergonomics Lab at the University of Calgary.

# The Auditory And Visual Attentional Blink In Schizophrenia

**Elzbieta B. Slawinski, Kim M. Goddard, and Peter J. Wass**
Dept. of Psychology, University of Calgary, 2500 University Dr. N.W., Calgary AB, T2N 1N4

## 1. INTRODUCTION

Schizophrenia is a severe and persistent psychiatric illness, consisting of positive and/or negative symptoms resulting in marked social and occupational dysfunction. More importantly, schizophrenia is also associated with cognitive dysfunction. These cognitive impairments are present in childhood, are almost fully developed by the time of the first psychotic episode, and tend to remain relatively stable over time in the absence of intervention.[1] The range of cognitive impairments is extensive, and includes such domains as executive functioning, learning and memory, and attentional dysfunction.[1, 2] Understanding the nature and extent of cognitive deficits in schizophrenia is important because they have been shown to predict functional outcome for these individuals; that is, whether they are able to secure paid employment, maintain relationships, and live independently.[2]

Attentional dysfunction is considered a core cognitive deficit in schizophrenia,[3] and to date, numerous studies have examined attentional processes.[3-5] However, apart from studies utilizing sensory gating paradigms, the majority of these investigations have concentrated on visual-spatial types of attentional processes. Whether potential impairments exist in other types of attentional processes (e.g., temporal attention) and encompass other modalities (e.g., audition) is not fully known.

Experimentally, the temporal nature of attention can be studied with rapid serial visual or auditory presentation techniques (RSVP, RAP). In these procedures, a series of stimuli (letters, pictures, tones, etc) are presented in rapid succession at the same location, and participants must identify 1 or 2 prespecified targets. Studies using these procedures have consistently found that having to attend to the first target impairs the ability to identify a second target (probe) for ~ 500 ms - a phenomenon known as the *attentional blink* (AB).[6-8]

The AB has been shown to be influenced by extrinsic factors including stimulus complexity, task difficulty, and learning strategies.[7, 10] In addition, intrinsic factors such as aging, attention deficits, and focal brain lesions have been found to increase the magnitude of the AB.[11-13] Therefore, it was anticipated that brain disease, i.e., schizophrenia, would also affect the AB. Accordingly, the purpose of this study was to examine both auditory and visual temporal attention in individuals with schizophrenia, using attentional blink techniques.

## 2. METHOD

Participants: Participants were 31 DSM IV-diagnosed individuals (29 males; 2 females) with chronic schizophrenia recruited from an in-patient psychiatric hospital who were age- and gender-matched with 31 healthy control participants recruited from the University of Calgary and surrounding community. Participants ranged in age from 20 to 50 years.

Procedure and Stimuli: After training, participants were presented with 168 RAP streams (11 tones/sec) consisting of 25 equally loud tones ranging from 1000 to 2490 Hz. All tones were 85 ms in duration, separated by a silent 5 ms interstimulus interval. Targets to be named were 1500 (low) and 2500 (high) Hz tones increased in intensity by approximately 10 dB SPL above stream items. Both targets were present on half of trials, balanced across stimulus onset asynchronies (SOAs) of 90, 180, 270, 360, 450, 540, and 630 ms. In the visual task, participants were presented with 168 RSVP streams (11 lines/sec) consisting of 25 sequentially-presented lines at orientations ranging from 30° to 150°. Lines were 15 ms in duration, separated by a blank interval of 75 ms. Targets to be named were thicker lines of 45° (right), 90° (vertical), and 135° (left). Both targets were present on 1/2 of the trials, balanced across SOAs of 90, 180, 270, 360, 450, 540, and 630 ms.

## 3. RESULTS

Schizophrenia patients had a more pronounced AB for both visual and auditory stimuli compared to healthy controls and performance differed statistically across every SOA ($p$'s < .01). In addition, a calculation of the overall magnitude of these ABs (difference in *% incorrect* between control condition and experimental conditions averaged across SOAs within tasks for each group) was significantly larger for this group ($p$'s < .05). Specifically, AB magnitudes for schizophrenia versus non-schizophrenia are: Auditory: 38.5% vs. 22.3%; Visual: 36.8% vs. 27.4% (data not shown). Auditory and visual ABs are shown in Figures 1 and 2, respectively.

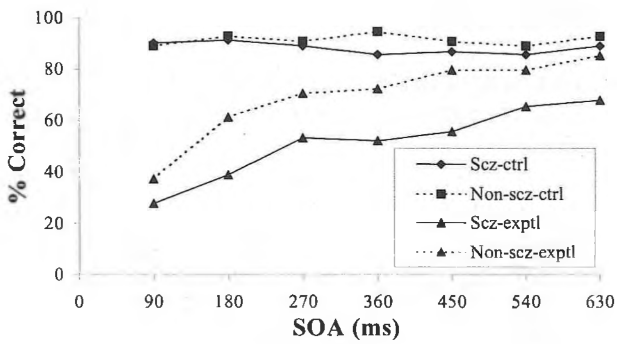



**Figure 1. Auditory AB in schizophrenia and non-schizophrenia**

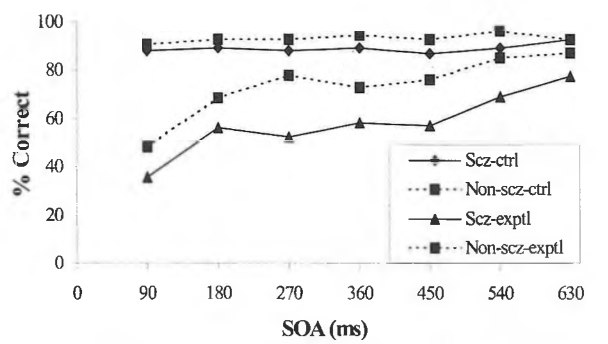



**Figure 2. Visual AB in schizophrenia and non-schizophrenia**

## DISCUSSION

As anticipated, the presence of schizophrenia affected the AB; schizophrenia patients produced larger (deeper) and more protracted (longer) ABs when compared to those of healthy controls. These results suggest that schizophrenia patients are impaired on measures of temporal attention. Furthermore, because AB differences were found for both auditory and visual stimuli, the temporal attention impairment in schizophrenia appears to be modality independent; that is, it extends across modalities.

Group differences across cognitive tasks are common in studies of schizophrenia and it has been proposed that this is indicative of a generalized cognitive deficit in this population. However, differential impairments in some tasks (e.g., learning and memory, attention) are also quite robust.[1, 2] It may be that schizophrenia reflects specific cognitive deficits superimposed upon a generalized cognitive deficit. In support of this latter claim, post-mortem studies have revealed cytoarchitectural abnormalities in the schizophrenia brain, where there are too many neurons present in deeper cortical layers (1 to 3), and too few in more superficial layers (4 to 6). The consequences of such abnormalities is noisy and inefficient communication between and among cortical areas, which at the most basic level, manifests itself as a reduced ability to discriminate signal from noise. By extension, it seems probable that complex tasks, which would rely on more extensive communication between cortical sites would be most prone to impairments, while simple tasks might escape performance deficits.

Although our investigation was not specifically directed at the question of differential versus generalized deficits, our findings are nevertheless broadly consistent with the above interpretation of schizophrenia. For example, while there appears to be a trend for the schizophrenia patients' performance to be slightly poorer in both control conditions – a relatively simple task – their performance is statistically equivalent to that of the control group. Conversely, in the much more difficult dual-task (experimental) conditions, performance was substantially poorer. Thus, temporal attention is *generally* impaired when considering that dual-task deficits extend across modalities, but is *differentially* impaired when dual-task performance is compared to single-task performance.

In conclusion, individuals with schizophrenia demonstrated a reduced ability to discriminate two – but not one – target(s) in a stream of distractors, independent of modality. These findings extend the previous research of visual-spatial attention deficits by demonstrating similar impairments in the temporal domain, providing the task at hand is cognitively demanding. To the extent that functional consequences can be extrapolated from experimental studies, then the finding of deficits in the temporal allocation of attention likely has implications for those everyday activities requiring the processing of multiple, nearly-simultaneous stimuli under distracting conditions, such as driving.

## REFERENCES


Green, M.F. (1998). *Schizophrenia from a neurocognitive perspective.* Allyn & Bacon: Toronto.

Green, M.F. (1996). *American Journal of Psychiatry, 153*, 321-330.

Braff, D.L. (1993). *Schizophrenia Bulletin, 19*, 223-259

Winterer, G., Ziller, M., Dorn, H., Frick, K, et al. (2000). *Clinical Neurophysiology, 111*, 837-849.

Goddard, K.M., Isaak, M.I., & Slawinski, E.B. (1997). *Journal of the Acoustical Society of America, 101*, 3125.

Maki, W.S. & Padmanabhan, G. (1994). *Psychonomic Bulletin & Review, 1*, 499-504.

Raymond, J.E., Shapiro, K.L., & Arnell, K.M. (1992). *Journal of Experimental Psychology: Human Perception & Performance, 18*, 849-860.

McAuliffe, S.P. & Knowlton, B.J. (2000). *Perception & Psychophysics 62*, 187-195.

Slawinski, E.B. & Goddard, K.M. (2000). *Canadian Acoustics, 29*, 3-13.

Hollingsworth, D.E., McAuliffe, S.P., Knowlton, B.J. (2001). *Journal of Cognitive Neuroscience 13*, 298-305.

Rizzo, M., Akutsu, H., & Dawson, J. (2001). *Neurology*, 57(5), 795-800.

Schatz, J. (1998). *Schizophrenia Research, 30*, 41-49.

Weinberger, D. & Lipska, B. (1995). *Schizophrenia Research, 16*, 87-110.

## Abstracts of Papers

### Day 1 – Speech Communication: Vowels I

**Articulatory Compensation For A Bite Block, With And Without Auditory Feedback, In Adults**

Melanie McNutt Campbell; Speech Pathology and Audiology; University of Alberta; 2-70 Corbett Hall; Edmonton, Alberta T6G 2G4

The purpose of this study was to investigate directly the role of auditory feedback in articulatory compensation for a bite block. Six adults with normal hearing and speech produced 16 tokens of each of the vowels /i/, /I/, and /œ/ in a "hVt" context within a "say hVt again" carrier phrase in every one of four conditions: unperturbed; with a customized bite block creating a 20-mm interincisal distance; while listening to 80 dB pink masking noise presented through earphones; and in a bite-block-plus-masking condition. $F_2$ and $F_1$ measurements were obtained from each token. We assessed the degree of compensation for the bite block by comparing formant values in the three test conditions with those in the normal condition and with predictions from computer simulation of the effect of a bite block if there were no compensation. Intensity and fundamental frequency measurements were also made to monitor any potentially confounding voice source changes. The results suggest that adults compensate substantially, but not completely, for a bite block and that auditory feedback plays a role in that compensation. The implication is that auditory feedback may play a role in on-line monitoring of adult speech production.

### Day 1 – Architectural Acoustics I

**Acoustical Comparison of Alberta Concert Halls**

J.S. Bradley and H. Sato (Institute for Research in Construction, National Research Council, Montreal Rd., Ottawa, K1A 0R6)

Modern room acoustics measurements, as described in ISO 3382, have been made in four Alberta concert halls: the two Jubilee Auditoria in Calgary and Edmonton, the Jack Singer Hall in Calgary, and the Winspear Centre in Edmonton. The halls are compared in terms of the average values of some measures as well as in terms of within hall variations of these parameters. The quite different acoustical characteristics of the wide fan-shaped Jubilee auditoria are compared with the other two more rectangular halls. Within each hall the effects of variable acoustics features are illustrated. The value of objective measurements to quantify the differences among the halls will be discussed.

### Day 1 – Numerical Modeling / Experimental Techniques

**Analytical Model Of Rail Vibration Induced By Flat Spots On Rolling Wheels**

Werner Richarz, Ph.D., FASA, Aercoustics Engineering Ltd. 50 Ronson Dr. Toronto, ON M9W 1B3

The vibrations of rail are usually unwanted by-products of train motion. Train motion is usually ignored because of the mismatch between the train speed and the phase velocity in the rail. This is an over-simplification at low frequencies for most modern train rails. Recent interest in detecting and identifying specific wheel flats on subway cars via measurement of rail vibrations has identified a need for a more comprehensive analytical model that deals with source motion. The conventional model for a resiliently supported rail has been extended to deal with moving sources. The closed form solution of the analytical model permits the construction of model signatures for a variety of wheel flats and other sources of rail vibration (e.g. rail and wheel roughness). The formalism can deal with realistic train configurations consisting of several cars, each with several axles. This allows one to test algorithms that are designed to detect wheel-flats. The predictions of the model correlate well with measured data.

### Day 2– Environmental Acoustics

**The Microphone As A Roughness And Wetness Sensor Of The Road In Automotive Applications**

Ahmed Trabelsi, Bodo Heimann (Mechatronic center of Hanover, Appelstr. 11, Hannover, 30167, Germany), Peter Rieth & Hans-Joachim Harting (Continental AG, Jaedekamp 30, Hannover, 30419, Germany)

The maximal friction coefficient is one of the most important factors for the braking procedure of a vehicle. The prediction of its value requires the knowledge of the different members of the interaction: the wheel, the vehicle, the roughness and the wetness of the road. The approach mentioned in this abstract is based on close proximity tire/road noise measurements.

Spectral analysis has shown, that characteristic frequency intervals can be defined to describe the roughness of the road. Other frequency intervals are more important to detect the wetness of the road. Further measurements have shown the relationship between the water film height over the road and the noise level. These measurements were combined with friction studies and have indicated the potential of the noise information to predict the friction.

**Outdoor Noise Propagation**


Ramani Ramakrishnan[1], Colin Novak[2], Helen Ule[2], and Robert Gaspar[2]

1. Ryerson University, Department of Architectural Science, 350 Victoria St.; Toronto, ON
2. University of Windsor, Mechanical, Automotive and Materials Engineering, 401 Sunset Ave.; Windsor, ON


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. There are a number of unanswered concerns that remain to be addressed as far the accuracy of ISO-9613 procedures. Two typical industrial sites will be chosen to study the accuracy of the ISO-9613 standard. Two softwares, Cadna and SPM, will be used to predict the noise levels at a few typical receptor locations around the industrial plant. The sound power levels, known a priori, of the noise sources at the industrial sites will be used in the modeling softwares. The sound pressure levels, both near-field and far-field, will be measured over a two-week duration. The sound levels will be expressed both as octave band sound pressure levels and averaged A-weighted overall levels. The measurements will be compared to the predicted sound pressure levels. The accuracy of the ISO-9613 procedures can thus be verified.

## Day 2– Hearing Conservation

**Optimal Reverberation Times For Speech Intelligibility For Normal And Hard-Of-Hearing People**


Maki Ezaki, Wonyoung Yang, and Murray Hodgson (School of Audiology and Speech Sciences, 5804 Fairview Ave, Vancouver, BC Canada V6T 1Z3).


It is well recognized that the acoustical environment is a critical factor for people, especially those with hearing impairment. In rooms for speech, such as classrooms, speech intelligibility is the main concern. Speech intelligibility is related to signal-to-noise ratio and is inversely related to the reverberation time. The optimal reverberation time for speech intelligibility for normal-hearing people has been studied extensively in the past decades, though there is some controversy about the interpretation of the results. In the present study, the optimal reverberation times for both normal and hearing-impaired people were studied using auralization and an idealized virtual test room. A clinical speech hearing test was processed through the simulated room with various reverberation times and noise levels. Subjects had moderate to severe high-frequency sensorineural hearing loss; as a control group, normal-hearing subjects were also tested. The results for normal and hearing-impaired people were compared.

## Day 2– Noise Control I

**Case Study: Noise Mitigation For Small Gas Turbine Power Plants**


Scott Penton, P.Eng., and Tammy Dow, B.Sc.Eng. (Rowan Williams Davies & Irwin Inc.(RWDI), 650 Woodlawn Road West, Guelph, Ontario, N1K 1B8)


A case study of environmental noise impacts from a small 50 MW gas turbine-fired power plant, located in the U.S. Midwest, is provided. Formerly located in a remote rural area, encroaching residential development resulted in adverse noise impacts. Initial noise mitigation measures implemented by the operator proved to be ineffective. RWDI was retained to conduct a detailed noise impact assessment and develop mitigation options. After installation, a follow-up acoustical audit was performed. This presentation will examine the key sources of noise, guideline limit development, mitigation development, and the results of the post-construction audit.

**Prediction Of Insertion Loss Of Rectangular Silencers**


Ramani Ramakrishnan[1] and Willie Watson[2]

1) Ryerson University, Department of Architectural Science, 350 Victoria St.; Toronto, ON, Canada

2) NASA Langley Research Center, Hampton, VA, USA


Considerable research has been undertaken to evaluate the insertion loss of simple duct silencers that use passive and porous materials such as fiberglass or mineral wool. However, the acoustic performance of these silencers has been evaluated assuming simplistic models. These models were extended to apply a realistic description of porous materials such as fiberglass or mineral wool, commonly applied in HVAC system ducts. The current research focuses on extending the simple one-dimensional models to two-dimensional models. These extensions would assist in evaluating the performance of baffles arranged in a staggered fashion within the ducts. Theoretical results of the finite element model will be presented. Comparison with one-dimensional will also undertaken and salient results of the comparison will also be presented. The impact of staggering baffles inside the duct will be presented.

## Day 3 – Speech Communication: More Topics in L2

### Acoustic Cues To Voicing Of Initial Consonants In Mandarin CV Syllables

Chunling Zhang and Terrance M. Nearey (U. Alberta, 4-32 Assiniboia Hall, Edmonton AB T6G 2E7)

Previous small-scale experiments suggest that Mandarin, a tone language, $F_0$, exhibits complex interactions with phonetic distinctions. Furthermore, there may be important differences from the interaction patterns observed in non-tone languages like English. Thus, J. Kingston and R. Diehl [Language, 70 419-454] review evidence that aspirated stops in English CV syllables lead to an increase in $F_0$ of the following vowel, while Mandarin aspirated stops show a slight decrease. We report here progress on a larger scale study designed to examine certain cue patterns more systematically. Specifically, we have conducted an experiment to investigate (1) the pattern of acoustic cues to voicing of initial consonants in Mandarin, including VOT, aspiration, tone, as well as $F_0$ and $F_1$; (2) whether and how lexical tone interacts with these acoustic cues. The experiment involves 20 native speakers of Beijing Mandarin (ten male and ten female) and three replications of 220 different CV syllables per speaker including 16 consonants, 5 vowels and 4 tones. Pilot work by Zhang and Nearey on a few subjects from this experiment suggests that Mandarin does indeed show different patterns of acoustic cues to voicing of initial consonants than does English, but the differences are not quite as previously reported. We will report on a larger number of subjects and we will discuss certain methodological issues, such as choice of measurement methods, sample size and appropriate statistical techniques, that arise in a study of this nature.

## Day 3 – Architectural Acoustics II

### Experimental Validation Of *Plant Noise* Empirical Prediction Models

Murray Hodgson, School of Occupational and Environmental Hygiene *and* Department of Mechanical Engineering,
University of British Columbia, 3rd Floor, 2206 East Mall, Vancouver, BC V6T 1Z3, Canada

The *Plant Noise* system predicts noise levels and reverberation times in typical industrial workrooms from values of the input prediction parameters. The objective of this paper was to evaluate the empirical models experimentally. Noise levels and reverberation times were measured in octave bands from 125-400 Hz in a number of different 'typical' workrooms without and with sound-absorbing ceiling and wall treatments. These were compared with predicted results to evaluate the models. For treated workrooms, treatment absorption coefficients were estimated. In the case of the fitting densities, which are not known a-priori, predictions were made for a range of values. Values which gave a best fit with prediction were thus determined and were generalized into an empirical fitting-density prediction method. Information was also obtained on how to make accurate predictions in the case of workrooms with non-flat, inhomogeneous roofs.

**Hearing Conservation Manual. Council for Accreditation in Occupational Hearing Conservation** Alice H. Suter (www.caohc.org). 2002. ISBN 0-9723143-0-X. 312 pgs. 12 Chapters and17 Appendices. U$A 55.00.

The Council for Accreditation in Occupational Hearing Conservation (CAOHS) has been, since 1973, a leader in providing standards for occupational hearing conservation programs. Another important activity of the CAOHS has been training and certification of occupational hearing conservationists (HC). In that respect, the Hearing Conservation Manual has been an important tool for teachers and students. This is the forth edition of the Manual, the third being published in 1993. It contains important revisions and updates, especially in the field of legislation and instrumentation.

The author, Alice Suter, is well known in the field of the occupational hearing conservation, having been on the staff of EPA, OSHA and NIOSH. Her impressive curriculum includes publications, presentations and fellowships in several professional associations such as the Acoustical Society of America and the American Industrial Hygiene Association.

The Manual encompasses every topic that the HC needs to know. The first chapter deals with the mission and role of the HC, and the following three chapters deal with the effects of noise on hearing, the anatomy and physiology of the ear and the hearing disorders in general. Each of the chapters is presented at a generic level, without getting into too many details that may not interest the reader.

The fourth chapter – Introduction to Sound – presents the basic concepts needed by the HC to be able to grasp many of the subjects examined in the rest of the book. The basic concepts are easy to understand and are accessible to the non-specialist, keeping, at the same time, a high scientific level. This reviewer was really pleased to find that the term "intensity" was properly used and is not confounded with "sound level" as is unfortunately the case in many publications.

Standards and Regulations, the following chapter, deals with the maze of federal and provincial USA legislations. The author has managed to make the reading easy and interesting.

The next two chapters deal with audiometric tests and analysis of the audiogram. The author stresses that the audiogram, by itself, does not help conserve the hearing of the exposed person, but has to be used as a diagnostic and awareness tool, both by the management and the subject being tested. Different types of instruments as well as the measurement techniques are described.

The next chapter examines the subjects of noise measurements and controls. It describes the basic instrumentation and some measurement techniques. Again, this reviewer was pleased to see that the author does not recommend to the HC to purchase the sophisticated instruments. A series of useful tips orient the HC regarding how to measure, what to look for and what to do if controls are needed. The author manages quite well to orient the reader towards the different aspects that are needed, without giving a false impression that noise control solutions are easy to design and implement.

Hearing protectors, by far the most used means for reducing noise exposures is the subject of the next chapter. This reviewer liked that the chapter starts by dealing with workers acceptance. There is no way of avoiding this issue, so difficult to manage in any hearing conservation program. Without saying explicitly, the author implies that wearing hearing protectors is not comfortable and that not wanting to wear them is a natural reaction from workers. This has to be respected, while, at the same time, explaining the why, where, when and how protectors should be worn. The author explains with ease some controversial issues such as the NRR and the OSHA legislation.

Training, motivation, recordkeeping and program evaluation are the last two chapters of the book.

Each one of the chapters contains a list of abbreviations, a list of recommended readings and a list of related websites. Also is included a quiz on the main issues dealt with in the text. Answers are included in one of the appendices.

The quality of the presentation, the graphs, tables and photographs, is commendable. It all makes for easy and pleasurable reading.

The only drawback for the Canadian reader is the abundance of information on the USA legislation, something natural, in a book written for American readers. It would not be too difficult to include some information on the Canadian legislation and standards, that will expand the circle of potential readers.

In summary, this reviewer highly recommends this book for anyone involved in hearing conservation, whether or not he intends to be certified as a HC by the Council.

**Alberto Behar, P.Eng., C.I.H.**
**Toronto, ON, Canada**

## NEWS / INFORMATIONS

### CONFERENCES

*The following list of conferences was mainly provided by the Acoustical Society of America. 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 francine.desharnais@drdc-rddc.gc.ca*

**2003**

1-2 September: First Congress of the Alps-Adria Acoustics Association (AAAA), Portoroz, Slovenia. Contact: Jurij Preselj, Mechanical Engineering, University of Ljubljana, Askerceva 6, 1000 Ljubljana, Slovenia; Fax: +386 1 251 8567; Web: www.fs.uni-lj.si/sda

1-4 September: Eurospeech 2003, Geneva, Switzerland. Contact: SYMPORG SA, Avenue Krieg 7, 1208 Geneva, Switzerland; Fax: +41 22 839 8485; Web: www.symporg.ch/eurospeech2003

7-10 September: World Congress on Ultrasonics, Paris, France. Web: www.sfa.asso.fr/wcu2003

16-19 September: Autumn Meeting of the Acoustical Society of Japan, Nagoya, Japan. Fax: +81 3 5256 1022; Web: wwwsoc.nii.ac.jp/asj/index-e.html

18-19 September: Surface Acoustics 2003, Salford University, Manchester, UK. Web: www.ioa.org.uk/salford2003

23-25 September: 2nd International Symposium on Fan Noise, Senlis, France. Contact: CETIAT, B.P. 2042, 69603 Villeurbanne, France; Fax: +33 4 72 44 49 99; Web: www.fannoise2003.org

5-8 October: IEEE International Ultrasonics Symposium, Honolulu, HI. Contact: W.D. O'Brien, Jr., Bioacoustics Research Lab., Univ. of Illinois, Urbana, IL 61801-2991; Fax: 217-244-0105; Web: www.ieee-uffc.org

15-17 October: Acoustics Week in Canada, Edmonton, Alberta, Canada. Fax: +1 780 414 6376; Web: caa-aca.ca/edmonton-2003.html

15-17 October: 34th Spanish Congress on Acoustics, Bilbao, Spain. Contact: Sociedad Española de Acústica, Serrano 144, 28006 Madrid, Spain; Fax: +34 91 411 7651; Web: www.ia.csic.es/sea/index.html

30-31 October: Autumn Meeting of the Swiss Acoustical Society, Basel, 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

5-6 November: Institute of Acoustics (UK) Autumn Conference, Oxford, UK. Contact: Institute of Acoustics, 77A St. Peter's Street, St. Albans, Herdfordshire AL1 3BN, UK; Fax: +44 1727 850553; Web: www.ioa.org.uk

### CONFÉRENCES

*La liste de conférences ci-jointe a été offerte en majeure partie par l'Acoustical Society of America. 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 francine.desharnais@drdc-rddc.gc.ca*

**2003**

1-2 septembre: Premier congrès de l'Association d'acoustique Alpes-Adria (AAAA), Portoroz, Slovénie. Info: Jurij Preselj, Mechanical Engineering, University of Ljubljana, Askerceva 6, 1000 Ljubljana, Slovenia; Fax: +386 1 251 8567; Web: www.fs.uni-lj.si/sda

1-4 septembre: Eurospeech 2003, Genève, Suisse. Info: SYMPORG SA, Avenue Krieg 7, 1208 Geneva, Switzerland; Fax: +41 22 839 8485; Web: www.symporg.ch/eurospeech2003

7-10 septembre: Congrès mondial sur les ultra-sons, Paris, France. Web: www.sfa.asso.fr/wcu2003

16-19 septembre: Rencontre d'automne de la Société japonaise d'acoustique, Nagoya, Japon. Fax: +81 3 5256 1022; Web: wwwsoc.nii.ac.jp/asj/index-e.html

18-19 septembre: Acoustique de surface 2003, Salford University, Manchester, Royaume-Uni. Web: www.ioa.org.uk/salford2003

23-25 septembre: 2e Symposium international sur le bruit de ventilateur, Senlis, France. Info: CETIAT, B.P. 2042, 69603 Villeurbanne, France; Fax: +33 4 72 44 49 99; Web: www.fannoise2003.org

5-8 octobre: Symposium international IEEE sur les ultrasons, Honolulu, HI. Info: W.D. O'Brien, Jr., Bioacoustics Research Lab., Univ. of Illinois, Urbana, IL 61801-2991; Fax: 217-244-0105; Web: www.ieee-uffc.org

15-17 octobre: Semaine canadienne d'acoustique, Edmonton, Alberta, Canada. Fax: +1 780 414 6376; Web: caa-aca.ca/edmonton-2003.html

15-17 octobre: 34e Congrès espagnol d'acoustique, Bilbao, Espagne. Info: Sociedad Española de Acústica, Serrano 144, 28006 Madrid, Spain; Fax: +34 91 411 7651; Web: www.ia.csic.es/sea/index.html

30-31 octobre: Rencontre d'automne de la Société suisse d'acoustique, Bâle, 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

5-6 novembre: Conférence d'automne de l'Institut d'acoustique (Royaume-Uni), Oxford, UK. Info: Institute of Acoustics, 77A St. Peter's Street, St. Albans, Herdfordshire AL1 3BN, UK; Fax: +44 1727 850553; Web: www.ioa.org.uk

7-9 November: Reproduced Sound, Oxford, UK. Contact: Institute of Acoustics, 77A St. Peter's Street, St. Albans, Herdfordshire AL1 3BN, UK; Fax: +44 1727 850553; Web: www.ioa.org.uk

10-14 November: 146th Meeting of the Acoustical Society of America, Austin, TX. 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 November: Workshop on detection and localization of marine mammals using passive acoustics, Dartmouth, NS, Canada. Contact: francine.desharnais@drdc-rddc.gc.ca

10-12 December: 3rd International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications, Firenze, Italy. Fax: +39 55 479 6767; Web: www.maveba.org

**2004**

17-19 March: Spring Meeting of the Acoustical Society of Japan, Atsugi, Japan. Fax: +81 3 5256 1022; Web: wwwsoc.nii.ca.jp/asj/index-e.html

22-25 March: Joint Congress of the French and German Acoustical Societies (SFA-DEGA), Strasbourg, France. Fax: +33 1 48 88 90 60; Web: www.sfa.asso.fr/cfa-daga2004

31 March – 3 April: International Symposium on Musical Acoustics (ISMA2004), Nara, Japan. Fax: +81 774 95 2647; Web: www2.crl.go.jp/jt/al32/isma2004

5-9 April: 18th International Congress on Acoustics (ICA2004), Kyoto, Japan. Web: ica2004.or.jp

11-13 April: International Symposium on Room Acoustics (ICA2004 Satellite Meeting), Hyogo, Japan. Fax: +81 78 803 6043; Web: rad04.iis.u-tokyo.ac.jp

17-21 May: International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004), Montreal, Canada. Web: www.icassp2004.com

24-28 May: 75th Anniversary Meeting (147th Meeting) of the Acoustical Society of America, New York, NY. 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

8-10 June: Joint Baltic-Nordic Acoustical Meeting, Mariehamm, Åland, Finland. Contact: Acoustical Society of Finland, Helsinki University of Technology, Laboratory of Acoustics and Signal Processing, P.O. Box 3000, 0215 TKK, Finland; Fax: +358 09 460 224; e-mail: asf@acoustics.hut.fi

5-8 July: 7th European Conference on Underwater Acoustics ECUA 2004, Delft, The Netherlands. Contact: Debbie Middendorp, D'Launch Communications, Forellendaal 141, 2553 JE The Hague, The Netherlands; Tel.: +31 70 3229900; Fax: +31 70 3229901; E-mail: middendorp@dlaunch.nl

7-9 novembre: Sons reproduits, Oxford, Royaume-Uni. Info: Institute of Acoustics, 77A St. Peter's Street, St. Albans, Herdfordshire AL1 3BN, UK; Fax: +44 1727 850553; Web: www.ioa.org.uk

10-14 novembre: 146e rencontre de l'Acoustical Society of America, Austin, TX. 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 novembre: Atelier sur la détection et la localisation de mammifères marins par l'acoustique passive, Dartmouth, NE, Canada. Info: francine.desharnais@drdc-rddc.gc.ca

10-12 décembre: 3e Atelier international sur les modèles et analyse d'émissions vocales avec applications bio-médicales, Florence, Italie. Fax: +39 55 479 6767; Web: www.maveba.org

**2004**

17-19 mars: Rencontre de printemps de la Société japonaise d'acoustique, Atsugi, Japon. Fax: +81 3 5256 1022; Web: wwwsoc.nii.ca.jp/asj/index-e.html

22-25 mars: Congrès combiné des Sociétés française et allemande d'acoustique (SFA-DEGA), Strasbourg, France. Fax: +33 1 48 88 90 60; Web: www.sfa.asso.fr/cfa-daga2004

31 mars – 3 avril: Symposium international sur l'acoustique musicale (ISMA2004), Nara, Japon. Fax: +81 774 95 2647; Web: www2.crl.go.jp/jt/al32/isma2004

5-9 avril: 18e Congrès international sur l'acoustique (ICA2004), Kyoto, Japon. Web: ica2004.or.jp

11-13 avril: Symposium international sur l'acoustique des salles (Rencontre satellite de ICA2004), Hyogo, Japon. Fax: +81 78 803 6043; Web: rad04.iis.u-tokyo.ac.jp

17-21 mai: Conférence internationale sur l'acoustique, la parole, et le traitement de signal (ICASSP 2004), Montréal, Canada. Web: www.icassp2004.com

24-28 mai: 75e rencontre anniversaire (147e rencontre) de l'Acoustical Society of America, New York, NY. 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

8-10 juin: Rencontre acoustique jointe Baltique-Nordique, Mariehamm, Åland, Finlande. Info: Acoustical Society of Finland, Helsinki University of Technology, Laboratory of Acoustics and Signal Processing, P.O. Box 3000, 0215 TKK, Finland; Fax: +358 09 460 224; courriel: asf@acoustics.hut.fi

5-8 juillet: 7e Conférence européenne sur l'acoustique sous-marine ECUA 2004, Delft, Pays-Bas. Info: Debbie Middendorp, D'Launch Communications, Forellendaal 141, 2553 JE The Hague, The Netherlands; Tél.: +31 70 3229900; Fax: +31 70 3229901; Courriel: middendorp@dlaunch.nl

11-16 July: 12th International Symposium on Acoustic Remote Sensing (ISARS), Cambridge, UK. Contact: S. Bradley, School of Acoustics and Electronic Engineering, Brindley Building, Room 301, University of Salford, Salford M5 4WT, UK; Fax: +44 161 295 3815; Web: www.isars.org.uk

3-7 August: 8th International Conference of Music Perception and Cognition, Evanston, IL. Contact: School of Music, Northwestern Univ., Evanston, IL 60201; Web: www.icmpc.org/conferences.html

22-25 August: Inter-noise 2004, Prague, Czech Republic. Web: www.internoise2004.cz

23-27 August: 2004 IEEE International Ultrasonics, Ferroelectrics, and Frequency Control 50th Anniversary Conference, Montreal, Canada. Contact: R. Garvey, Datum, 34 Tozer Road, Beverly, MA 01915-5510; Fax: +1 978 927 4099; Web: www.ieee-uffc.org/index2-asp

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

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

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

**2005**

7-10 August: Inter-Noise, Rio de Janeiro, Brazil. Details to be announced later.

28 August – 2 September: Forum Acusticum Budapest 2005, Budapest, Hungary. E-mail: sea@fresno.csic.es

11-16 juillet: 12e Symposium international sur la télédetection acoustique (ISARS), Cambridge, Royaume-Uni. Info: S. Bradley, School of Acoustics and Electronic Engineering, Brindley Building, Room 301, University of Salford, Salford M5 4WT, UK; Fax: +44 161 295 3815; Web: www.isars.org.uk

3-7 août: 8e Conférence internationale sur la perception et la cognition de la musique, Evanston, IL. Info: School of Music, Northwestern Univ., Evanston, IL 60201; Web: www.icmpc.org/conferences.html

22-25 août: Inter-noise 2004, Prague, République tchèque. Web: www.internoise2004.cz

23-27 août: 50e Conférence anniversaire internationale IEEE 2004 sur les ultra-sons, la ferroélectricité et la régulation par la fréquence, Montréal, Canada. Info: R. Garvey, Datum, 34 Tozer Road, Beverly, MA 01915-5510; Fax: +1 978 927 4099; Web: www.ieee-uffc.org/index2-asp

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

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

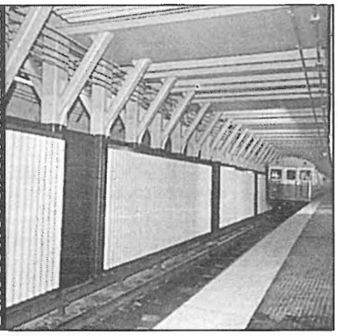
15-19 novembre: 148e rencontre de l'Acoustical Society of America, San Diego, CA. Info: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tél.: 516-576-2360; Fax: 516-576-2377; Courriel: asa@aip.org; Web: asa.aip.org

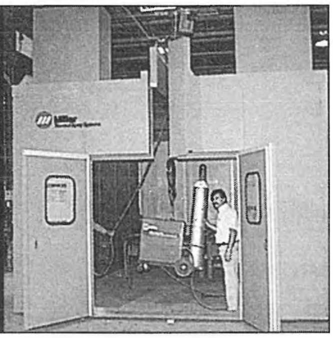
**2005**

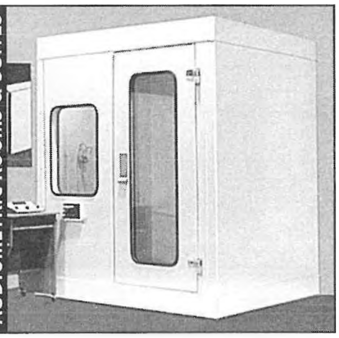
7-10 août: Inter-Noise, Rio de Janeiro, Brésil. Information à suivre.

28 août – 2 septembre: Forum Acusticum Budapest 2005, Budapest, Hongrie. Courriel: sea@fresno.csic.es


ENGINE TEST FACILITY

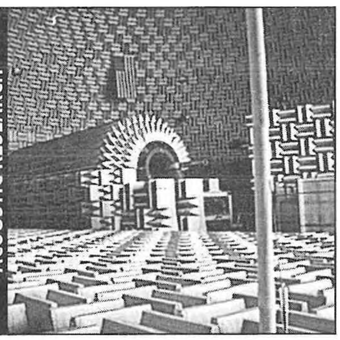

ECKOUSTIC FUNCTIONAL PANELS

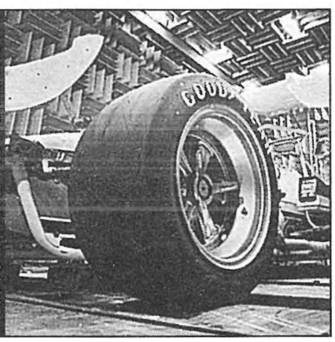

O.E.M. ENCLOSURE


AUDIOMETRIC ROOMS & SUITES

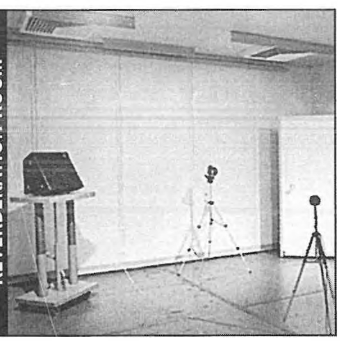



ARCHITECTURAL ACOUSTICS


ACOUSTIC RESEARCH



AUTOMOTIVE TEST CHAMBER


REVERBERATION ROOM

## 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 Canadienneet 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* __________

City/Province __________ Postal Code __________ Country __________

**Areas of Interest:** (Please mark 3 maximum)

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. Noiand Lewis - (650) 595-8588
acopac@acopacific.com - Belmont, CA

**Acoustec Inc.**
Dr. J.G. Migneron - (418) 682-2331
information@acoustec.qc.ca - Montréal, 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. Biaise Gosselin - (514) 840-3000ext5134
gosselin.biaise@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@aibn.on.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

**Nelson Industries Inc.**
Mr. Ray Cheng - (608) 873-4370
- Stoughton, WI

**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. Warren Tait - (905) 612-4000
info@silex.com - Mississauga, ON

**SNC/Lavalin Environment Inc.**
Mr. J-L. Allard - (514) 651-6710
allaj@snc-lavalin.com - Longueuil, QC

**Soft dB Inc.**
M. André L'Espérance - (418) 686-0993
Info@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