

canadian acoustics

acoustique canadienne

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

SEPTEMBER 2009

Volume 37 -- Number 3

SEPTEMBRE 2009

Volume 37 -- Numéro 3

EDITORIAL / EDITORIAL	3
PROCEEDINGS OF THE ACOUSTICS WEEK IN CANADA 2009/ ACTES DE LA SEMAINE CANADIENNE D'ACOUSTIQUE 2009	
Table of Contents / Table des matières	4
Conference Calendar	11
Plenary Presentations	
<i>Architectural Acoustics: Some Historical Developments and Ongoing Issues - John Badley</i>	21
<i>I have Never Seen a Sound - Murray Schafer</i>	32
<i>Biomedical Ultrasound Imaging from 1 to 1000 MHz - Michael Kolios</i>	35
Architectural Acoustics	44, 72
Digital Signal Processing	56, 78, 90
Sound Quality	64
Aeroacoustics	86
Musical Acoustics	98
Auditory Scene Analysis	104
Hearing Conservation	106
Underwater Acoustics	110
Environmental Noise	122
Hearing Aids	126
Speech	138, 162, 188
Psychological Acoustics	148, 176
Rural Noise and Noise Control	154
Biomedical Acoustics	168
Vibrations	178, 198
Music Perception	182
Presentations without Summary Papers	208

PROCEEDINGS



ACOUSTICS WEEK IN CANADA
 SEMAINE CANADIENNE D'ACOUSTIQUE
 ACOUSTICS WEEK IN CANADA
 SEMAINE CANADIENNE D'ACOUSTIQUE
 ACOUSTICS WEEK IN CANADA
 SEMAINE CANADIENNE D'ACOUSTIQUE
 ACOUSTICS WEEK IN CANADA

COMPTES RENDUS

canadian acoustics

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

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

CANADIAN ACOUSTICS is published four times a year - in March, June, September and December. The deadline for submission of material is the first day of the month preceding the issue month. Copyright on articles is held by the author(s), who should be contacted regarding reproduction. Annual subscription: \$30 (student); \$70 (individual, institution); \$300 (sustaining - see back cover). Back issues (when available) may be obtained from the CAA Secretary - price \$10 including postage. Advertisement prices: \$600 (centre spread); \$300 (full page); \$175 (half page); \$125 (quarter page). Contact the Associate Editor (advertising) to place advertisements. Canadian Publication Mail Product Sales Agreement No. 0557188.

acoustique canadienne

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

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

ACOUSTIQUE CANADIENNE est publiée quatre fois par année - en mars, juin, septembre et décembre. La date de tombée pour la soumission de matériel est fixée au premier jour du mois précédant la publication d'un numéro donné. Les droits d'auteur d'un article appartiennent à (aux) auteur(s). Toute demande de reproduction doit leur être acheminée. Abonnement annuel: \$30 (étudiant); \$70 (individuel, société); \$300 (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: \$600 (page double); \$300 (page pleine); \$175 (demi page); \$125 (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.

EDITOR-IN-CHIEF / RÉDACTEUR EN CHEF

Ramani Ramakrishnan
Department of Architectural Science
Ryerson University
350 Victoria Street
Toronto, Ontario M5B 2K3
Tel: (416) 979-5000; Ext: 6508
Fax: (416) 979-5353
E-mail: rramakri@ryerson.ca

EDITOR / RÉDACTEUR

Chantai Laroche
Programme d'audiologie et d'orthophonie
École des sciences de la réadaptation
Université d'Ottawa
451, chemin Smyth, pièce 3062
Ottawa, Ontario K1H 8M5
Tél: (613) 562-5800 # 3066; Fax: (613) 562-5428
E-mail: claroche@uottawa.ca

ASSOCIATE EDITORS / REDACTEURS ASSOCIES

Advertising / Publicité

Jason Tsang
7 Parkwood Crescent
Ottawa, ONTARIO
K1B3J5
E-mail: jtsangeng@yahoo.ca

Canadian News / Informations

Jérémie Voix
Sonomax Hearing Healthcare Inc
8375 Mayrand Street
Montréal, QC, H4P 2E2, Canada
Tel: (514) 932-2674
Fax: (514) 932-4994
E-mail: jvoix@sonomax.com



Acoustics Week in Canada

Semaine canadienne d'acoustique

Niagara-on-the-Lake, 14 - 16 October 2009

Niagara-sur-le-Lac, 14-16 octobre 2009

Local Organizing Committee/ Comité d'organisation

Conference Chair/ Président: Ramani Ramakrishnan [rramakri@ryerson.ca]

Technical Chair/Directeur scientifique: Frank Russo [russo@ryerson.ca]

Technical Committee/Comité scientifique: Ben Dyson [bdyson@ryerson.ca]

Local Co-Chair/Directeur: Moustafa Osman [moustafa.osman@sympatico.ca]

Treasurer/Trésorier: Dalila Guisti [dalila@jadeacoustics.com]

Equipment Exhibition/Exposition technique: Rich Peppin [RPeppin@aol.com]

Registration/Inscription: Mandy Chan [machan@hgcengineering.com]

Registration/Inscription: Megan Munro [mmunro@hgcengineering.com]

Registration/Inscription: Payam Ezzatian [payam.ezzatian@gmail.com]

Website/Site Web: Payam Ashtiani [pashtiani@aercoustics.com]

Translations/Traductrice: Inna Petrennic [inna@echologics.com]

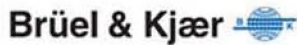
Sponsors



J A D E
ACOUSTICS



Exhibitors



EDITORIAL

Now that I am the chair of Acoustics Week in Canada 2009, I get to write the editorial after a few years' hiatus. It is my pleasure to invite you to this scenic spot of southern Ontario, Niagara-on-the-Lake, for three exciting days of acoustical communing. The organizing committee has put together an eclectic conference programme for your enjoyment.

Niagara-on-the-Lake is at the confluence of Niagara River and Lake Ontario with award winning vineyards as its backyard. It is half-hour ride away from the world famous Canadian Horseshoe Falls. It is also walking distance away from the Shaw Festival theatres. The conference will be held in the refurbished Pillar and Post Hotel.

The conference themes are architectural acoustics, psychological acoustics and bio-acoustics. One of the main sponsors of the conference is Ryerson University and fittingly, the Chair of the Department of Architectural Science, Dr. Kendra Schank Smith, will welcome the gathering on Wednesday. We will have an acoustical walk around Niagara-on-the-Lake by famous composer and acoustical ecologist, Mr. Murray Schafer. The banquet, on Thursday, will be highlighted with Emoticon, music sensed through vibrations.

The exciting technical program includes plenary lectures on architectural acoustics, psychological acoustics and bio-acoustics. It includes around 125 technical papers in over 20 subject areas in acoustics and vibration, from architectural to speech to underwater. This proceedings issue contains two-page summaries of many of them, along with the detailed conference calendar. The equipment exhibition on Thursday, 15 October, will have 12 exhibitors. The awards will be presented during the banquet. In short, the organizing committee will present 3-days of interesting and enjoyable acoustical activities. I, on behalf of our organizing committee, invite you to join us in Niagara-on-the-Lake in October, and look forward to welcoming you there.

Ramani Ramakrishnan
Conference Chair
Ryerson University

Comme le Président de la semaine canadienne d'acoustique, c'est un plaisir pour moi d'écrire cet éditorial après une pause de quelques années et de vous inviter à Niagara-on-the-Lake, un endroit pittoresque au sud de l'Ontario, pour y passer trois jours très intéressants et partager votre passion pour l'acoustique avec les autres. Le comité d'organisation a préparé un programme exceptionnel pour vous.

Niagara-on-the-Lake est situé au point de confluence de la rivière Niagara et le lac Ontario, et au coeur des vignobles primés. La ville est à trajet de demi-heure de Horseshoe – chutes canadiennes mondialement connues – et à courte distance à pied du théâtre du Festival Shaw. La conférence se tiendra à l'hôtel "Pillar & Post Inn" récemment renouvelé.

Les sessions de la conférence porteront sur acoustique architecturale, psycho-, et bioacoustique. Ryerson University est un des principaux sponsors de la conférence. Le mercredi, la Directrice du Département d'architecture Pr. Kendra Schank Smith ouvrira la conférence par un discours de bienvenue. Nous irons à la promenade dans la ville avec le célèbre compositeur et "le père d'écologie acoustique" M. Murray Schafer. Le point fort du banquet, que aura lieu mardi, sera Emoticon – un concert musical, où la musique est perçu par vibrations.

Le programme technique passionnant comprend les sessions plénières dans les domaines de l'acoustique architecturale, psycho- et bioacoustique. Environ 125 articles techniques sont soumis à plus de 20 sujets dans les domaines d'acoustique et vibration, y compris architecture, langage et hydroacoustique. Cette édition présente les résumés de deux pages de plusieurs de ces articles, aussi comme le programme détaillé de la conférence. Douze exposants participeront dans l'exposition d'équipement, que aura lieu jeudi, le 15 octobre. Les prix seront décernés pendant le banquet. Enfin, le comité d'organisation vous préparera trois jours d'activités intéressantes dans le domaine de l'acoustique. Au nom de notre comité d'organisation, j'espère avec plaisir de vous accueillir en octobre à Niagara-on-the-Lake.

Ramani Ramakrishnan
Président
Ryerson University

TABLE OF CONTENTS/TABLES DES MATIÈRES

Conference Poster	2
Editorial/Éditorial	3
Table of Contents/Tables des matières	4
Conference Calendar	11
Plenary I	
Architectural Acoustics: Some Historical Developments and Ongoing Issues John Bradley	21
I Have Never Seen a Sound Murray Schafer	32
Plenary III	
Biomedical Ultrasound Imaging from 1 to 1000 MHz Michael Kolios	35
Architectural Acoustics I	
SPMSoft and Speech Privacy Measurement in Open-Plan Offices John Bradley	44
NRC-IRC Computer Controlled Acoustic Measurement and Quality System Timothy Estabrooks	46
NRC-IRC Flanking Sound Transmission Facility Timothy Estabrooks, Frances King, Trevor Nightingale & Ivan Sabourin	48
Characterizing Flanking Transmission Paths in the NRC-IRC Flanking Facility Frances King, Stefan Schoenwald & Ivan Sabourin	50
The Effect of Resilient Channels on Ceiling Flanking Transmission Paths Frances King & Ivan Sabourin	52
Effects of Structural Load and Joist Type on Flanking Sound Transmission Ivan Sabourin, Berndt Zeitler & Frances King	54
Digital Signal Processing I	
Trends in cell phone Voice processing Chris Forrester	56
Wideband Echo Control Challenges Sylvain Angrignon, Chris Forrester & Malay Gupta	58
An Analysis of Loudspeaker Distortion in the Context of Acoustic Echo Cancellation Trevor Burton & Rafik Goubran	60
Subband Autoregressive Modelling For Speech Enhancement Brady Laska, Rafik Goubran & Miodrag Bolić	62

Sound Quality

- Loudness Prediction Model Comparison Using The Equal Loudness Contours
Jeremy Charbonneau, Colin Novak & Helen Ule 64
- Comparison of Non-Stationary Loudness Results to Equal Loudness Contours
Colin Novak, Jeremy Charbonneau & Helen Ule 66
- Prediction of Psychoacoustic Metrics from Suspension Induced Vibration and Road Induced Noise using Transfer Path and Psychoacoustic Analysis Techniques
Nebojsa Radic Colin Novak & Helen Ule 68
- Description of the Multiple Look Approach for Calculating Unsteady Loudness
Helen Ule & Colin Novak 70

Architectural Acoustics II

- Tool for Predicting Air-Borne and Structure-Borne Sound in Buildings
J. David Quirt & John K. Dickinson 72
- A Comparison of Three Methods for the In Situ Determination of Acoustic Absorption Coefficients
Scott Mallais & John Vanderkooy 74
- Low-frequency Room Analysis and Equalization using Properly-Modeled Source Terms in Finite-Difference Time-Domain Simulations
Ryan Matheson 76

Digital Signal Processing II

- Noise Suppression in Cellular Telephony
Malay Gupta, Chris Forrester & Sean Simmons 78
- The Cocktail Party Problem: Solutions and Applications
Karl Wiklund & Simon Haykin 80
- Region-Growing Permutation Alignment Approach in Frequency-Domain Convolutional Blind Source Separation
Lin Wang, Heping Ding & Fuliang Yin 82
- Review of Wideband Speech Noise Reduction Techniques
Malay Gupta, Chris Forrester & Sean Simmons 84

Aeroacoustics

- Screech Suppression of Supersonic Jet Noise
Ramani Ramakrishnan, Sergio Raimondo, Anant Grewal & Gary Elfstrom 86
- Trends in Aero-acoustic Wind Tunnel Testing
Gary M. Elfstrom 88

Digital Signal Processing III

- A Residual-Cepstrum Method of Pitch Estimation from Noisy Speech
Celia Shahnaz, Wei-Ping Zhu & M. Omair Ahmad 90
- Band-Adaptive Formant Frequency Estimation from Noisy Speech in Correlation Domain
Shaikh Anowarul Fattah, Wei-Ping Zhu, & M. Omair Ahmad 92

Digital Earplug for Brain Plasticity Research Marc Schönwiesner, Jérémie Voix & Philippe Pango	94
Acoustic Shock Signal Detection and Control Henry Luo	96
Musical Acoustics/Soundscapes	
“Sound Created Form” Ben Gaum & Ramani Ramakrishnan	98
The Sounding Museum - Cultural Soundscapes as a Tool in Museum Education Hein Schoer	100
Tapping to Pitched Auditory Feedback Tones: Effects of Pitch Contour and Interval Size on Intertap-Interval and Tap Force Paolo Ammirante, William F. Thompson & Frank A. Russo	102
Auditory Scene Analysis	
Temporal Dynamics of Brain Activity during Dichotic Listening Bernhard Ross	104
Hearing Conservation	
Hearing Protectors Standards Alberto Behar & Willy Wong	106
Monitoring Sound Pressure at the Eardrum for Hearing Conservation Anthony J. Brammer, Gongqiang Yu, Eric R. Bernstein, Donald R. Peterson, Martin G. Chemiack & Jennifer B. Tuft	108
Underwater Acoustics	
Acoustic Localization of an Autonomous Underwater Vehicle Nicos Pelavas, Garry J. Heard, Gordon R. Ebbeson, Carmen Lucas, Derek Clark	110
Bayesian Ocean Acoustic Source Tracking with Environmental Uncertainty Stan E. Dosso & Michael J. Wilmut	112
Bayesian Source Track Prediction in an Uncertain Ocean Environment Stan E. Dosso & Michael J. Wilmut	114
MRI Measurement of Dynamics of the Liquid and a Dissolved Gas during Acoustic Cavitation I.V.Mastikhin, B.Newling, N.Hetherington & S.Kristoffersen	116
Eigenfrequency Analysis of Fluid-filled Pipes N M Alam Chowdhury, Zaiyi Liao, Lian Zhao & Ramani Ramakrishnan	118
Modal Solutions of the Acoustic Wave Propagation through Fluid-filled Pipes N M Alam Chowdhury, Zaiyi Liao, Lian Zhao & Ramani Ramakrishnan	120
Environmental Noise	
Characteristics of Wind Turbine Noise Ramani Ramakrishnan	122
An Approach to Generating and Calibrating a Sound Contour Map Ryan Daneluzzi, Steve Bogensberger, Mike Dobrin, Colin Novak & Helen Ule	124

Hearing Aids

Music and Hearing Aids Marshall Chasin	126
Evaluation of Aided and Unaided Auditory Functions for Royal Canadian Mounted Police (RCMP) Members Véronique Vaillancourt, Chantal Laroche, Christian Giguère, Marc-André Beaulieu, & Jean-Pierre Legault	128
Automatic Volume Settings for Environment Sensitive Hearing Aids Rana Rahal, Wail Gueaieb, Christian Giguère & Hisham Othman	130
A New Feature Selection Method for Volume Control in Direct-Learning Hearing Aid Systems Jin Zhou, Hisham Othman, Hilmi Dajani, & Tyseer Aboulnasr	132
Subband Adaptive Filtering for Acoustic Feedback Compensation in Hearing Aids Brady Laska, Miodrag Bolić & Rafik Goubran	134
Binaural Objective Intelligibility Measurement and Hearing Aids Nicolas Ellaham, Christian Giguère & Wail Gueaieb	136

Speech I

Within-Category Variation in L2 English Vowel Learning Ron I. Thomson and Talia Isaacs	138
Vowel Imitation in Spontaneous Phonetic Imitation Molly Babel	140
Acoustic and Auditory Comparisons of Polish and Taiwan Mandarin Sibilants Chenhao Chiu	142
Talking while Chewing: Speaker Response to Natural Perturbation of Speech Connor Mayer, Bryan Gick & Elizabeth Ferch	144
Magnetic Resonance Imaging and Videofluoroscopy of Partial Glossectomees' Speech: Preliminary Results Janette Quintero, Tim Bressmann, Katalin Mady & Ambros Beer	146

Psychological Acoustics I

The Robustness of Infants' Early Word Representations Marieke van Heugten & Elizabeth K. Johnson	148
Controlling for Age Related Hearing Loss Can Eliminate Aging Differences in Lexical Competition: Evidence from Eye-tracking as an Online Measurement of Age and Noise Effects on Listening B.M Ben-David, C.G. Chambers, M. Daneman, M. K. Pichora-Fuller, E. Reingold & B. A. Schneider	150
The Environment for Auditory Research Bruce E. Amrein & Tomasz R. Letowski	152

Rural Noise and Noise Control

An Investigation of the Wind Induced Background Sound Levels in Rural Areas for Existing and Proposed Wind Farms Ian Bonsma & Megan Munro	154
Inadequacy of Wind Turbine Noise Regulations and their Application John P. Harrison	156
Source-Receptor Path Prioritization Considerations for the Specification of Silencers in Ventilation Systems Brian Chapnik and Alex Lorimer	158

Cooling Tower Noise Control for the Niagara Falls Casino
Thomas Paige 160

Speech II and Ecological Acoustics

Acoustic Correlates of Neutral versus Angry Affect in Real and Non-Word Sentences
Heather Flowers & Tim Bressmann 162

Underwater Acoustic Localization of Marine Mammals in an Uncertain Environment
Brendan Rideout, Stan E. Dosso & David Hannay 164

Andean ultrasonics: bioacoustics of two tropical montane katydids
Glenn Morris 166

Biomedical Acoustics

Measuring Scattering in Apoptotic Cancer Cells using Ultra High Frequency Acoustic Microscopy
Eric Strohm & Michael Kolios 168

A Comparison of Ultraosund-Based Imaging Modalities to Monitor Thermal and Mechanical Ultrasound Tissue Therapies
Arthur Worthington, Sankar Narasimhan, Jahan Tavakkoli & Michael C. Kolios 170

An Enhanced Numerical Model to Simulate Nonlinear Continuous Wave Ultrasound Field and Resulting Temperature Response in Tissue
Shahram Mashouf & Jahan Tavakkoli 172

Old and New Cochlear Maps
Reinhart Frosch 174

Psychological Acoustics II

Effect of Headphone Type on Listening Levels and Localization Abilities of Portable Audio Device Users in Quiet and Traffic Noise
I. Ibrahim, R. Malcolmson, M. B. Jennings & M. F. Cheesman 176

Vibrations I

Age and Neuropathies in Vibration Exposed Manual Workers
Martin G Cherniack, Anthony J Brammer, Ronnie Lundström, Timothy F. Morse, Greg Neely, Tohr Nilsson et al 178

Vibrotactile Perception as a Test for Numbness and Pain
Anthony J. Brammer, Paivi Sutinen, Sourish Das, Ilmari Pyykkö, Esko Toppila & Jukka Starck 180

Music Perception

Reliable Perception of Musical Sound Sources with Impoverished Spectral Envelopes
Michael Hall 182

Changes in Melodic Perception as a Function of Age and Lowest Harmonic Number
Huiwen Goy, D. Timothy Ives, Frank A. Russo, M. Kathleen Pichora-Fuller & Roy D. Patterson 184

Development and Validation of a Sensory-Substitution Technology for Music
Carmen Branje, Michael Maksimowski, Gabe Nespoli, Maria Karam, Deborah Fels & Frank Russo 186

Speech III

- The Sound of Stroop: Acoustic Effects in Stroop Interference
Boaz M. Ben-David, Pascal van Lieshout & Alex Nishta 188
- Establishing Normative Voice Characteristics of Younger and Older Adults
Jessica Banh, Konstantin Naumenko, Huiwen Goy, Pascal van Lieshout, David N. Fernandes & Kathy Pichora-Fuller 190
- Assessing the Intrinsic Relationship between Facial Motion and Acoustics in Patients with Parkinson's Disease
Luyao Ma, Huawei Colin Li, Akiko Amano-Kusumoto, Willy Wong & Pascal van Lieshout 192
- Men, Women and Lenition: Gender Differences in the Production of Intervocalic Voiced Stops in Mexican Spanish
Anna Limanni 194
- Separating Normal Variation in Movement Amplitudes from Gradient Speech Errors
Anneke Slis & Pascal van Lieshout 196

Vibrations II

- Transfer Path Analysis using Engine Radiated Sound and Mount Vibration
Nikolina Samardzic & Colin Novak 198
- An Accurate Coupled Structural-Acoustic Analytical Framework for the Prediction of Random and Flow-Induced Noise in Transport Vehicles: its Validation
Joana da Rocha, Afzal Suleman & Fernando Lau 200
- Experimental studies on the in-plane vibrations and sound radiation in an annular thick disk
Salem Bashmal, Rama Bhat & Subash Rakheja 202
- Modeling Electrostatic Field in MEMS Devices using Artificial Springs in Rayleigh-Ritz Method
Avinash K. Bhaskar, Muthukumaran Packirisamy & Rama Bhat 204
- Dynamic Behaviour of Micro-structure in Microfluidic Channel
MD Shakhawat Hossain, Muthukumaran Packirisamy & Subhash Rakheja 206
- Abstracts for Presentations without Summary Papers** 208

GOAL:

- ✓ REDUCE NOISE
- ✓ REDUCE WEIGHT
- ✓ REDUCE COST
- ✓ INCREASE CUSTOMER PROFITS



GOAL ACHIEVED.

Blachford specializes in testing, designing and producing quality acoustical materials.

Strategic, fast, flexible and loaded with talent, we bring our customers big-impact solutions. We're a responsive, reliable partner fully committed to continual improvement and well-known for our strengths in technology, innovation and rigorous process.

In the past three years we've saved our customers millions in both costs and in pounds of excess weight. And we've reduced product noise levels an average of 38%. That's something to celebrate.

Featuring state-of-the-art resources and expertise, Blachford is your ideal partner for producing noise control products while improving your bottom line.

Quality Management System
ISO 9001:2000 & ISO/TS 16949

Environmental Management System
Responsible Care (CDN) & ISO 14001 (US)

Laboratory Management System
ISO 17025 for ASTM C423 & SAE J1400

- Engineering
- Designing
- Testing
- Manufacturing

For details call

630.231.8300
or visit us at blachford.com

Blachford

WEDNESDAY 14TH OCTOBER 2009

8.00 – 8.10	AWC09 INTRODUCTORY REMARKS		
8.10 – 8.55	PLENARY 1 (XXX)		
8.55 – 9.40	PLENARY 2 (XXX)		
10.00 – 12.10	ARCHITECTURAL ACOUSTICS I	SOUND QUALITY	DIGITAL SIGNAL PROCESSING I
2.00 – 3.50	ARCHITECTURAL ACOUSTICS II	SINGING (AIRS)	DIGITAL SIGNAL PROCESSING II
4.10 – 6.00	AEROACOUSTICS	MUSICAL ACOUSTICS/ SOUNDSCAPES	DIGITAL SIGNAL PROCESSING III
6.10 – 7.00	ACOUSTIC WALK (MURRAY SCHAFFER)		

THURSDAY 15TH OCTOBER 2009

9.00 – 9.45	PLENARY 3 (XXX)		
10.00 – 12.20	UNDERWATER ACOUSTICS	HEARING CONVERSATION	AUDITORY SCENE ANALYSIS
2.00 – 4.10	ENVIRONMENTAL NOISE	HEARING AIDS	SPEECH I
4.30 – 6.30	RURAL NOISE & NOISE CONTROL	PSYCHOLOGICAL ACOUSTICS I	SPEECH II / ECOLOGICAL ACOUSTICS
6.00 – 7.30	CAA ANNUAL GENERAL MEETING		
7.30 – 9.30	AWC09 BANQUET		

FRIDAY 16TH OCTOBER 2009

9.00 – 9.45	PLENARY 4 (XXX)		
10.00 – 12.10	VIBRATIONS I	PSYCHOLOGICAL ACOUSTICS II	BIOMEDICAL ACOUSTICS
2.00 – 4.10	VIBRATIONS II	MUSIC PERCEPTION	SPEECH III

DAY ONE

WEDNESDAY 14TH OCTOBER 2009 (MORNING)

8.00 – 8.10	AWC09 INTRODUCTORY REMARKS		
8.10 – 8.55	PLENARY 1 (XXX)		
8.55 – 9.40	PLENARY 2 (XXX)		
9.40 – 10.00	COFFEE BREAK		
	ARCHITECTURAL ACOUSTICS I (Chair: John Bradley)	SOUND QUALITY (Chair: Colin Novak)	DIGITAL SIGNAL PROCESSING I (Chair: Chris Forrester)
	ROOM	ROOM	ROOM
10.00 – 10.20	John S. Bradley - SPMSoft and Speech Privacy Measurement in Open-Plan Offices	Jeremy Charbonneau, Colin Novak & Helen Ule - Loudness Prediction Model Comparison Using The Equal Loudness Contours	Chris Forrester - Trends in Cell Phone Voice Processing
10.20 – 10.40	Timothy Estabrooks - NRC-IRC Computer Controlled Acoustic Measurement and Quality System	Gary Newton Jr., Colin Novak, David Bogema & Helen Ule - Acoustic Imaging for BSR in Non-Quiet Environments using Spherical Beamforming	Sylvain Angrignon, Chris Forrester & Malay Gupta - Wideband Echo Control Challenges
10.40 – 11.00	Timothy Estabrooks, Frances King, Trevor Nightingale & Ivan Sabourin - NRC-IRC Flanking Sound Transmission Facility	Colin Novak, Jeremy Charbonneau & Helen Ule - Comparison of Non-Stationary Loudness Results to Equal Loudness Contours	Trevor Burton & Rafik Goubran - An Analysis of Loudspeaker Distortion in the Context of Acoustic Echo Cancellation
11.00 – 11.10	BREAK		
11.10 – 11.30	Frances King, Stefan Schoenwald & Ivan Sabourin - Characterizing Flanking Transmission Paths in the NRC-IRC Flanking Facility	Nebojsa Radic, Colin Novak & Helen Ule - Prediction of Psychoacoustic Metrics from Suspension Induced Vibration and Road Induced Noise Using Transfer Path and Psychoacoustic Analysis Techniques	Brady Laska, Rafik Goubran & Miodrag Bolić - Subband Autoregressive Modelling For Speech Enhancement
11.30 – 11.50	Frances King & Ivan Sabourin - The Effect of Resilient Channels on Ceiling Flanking Transmission Paths	Helen Ule & Colin Novak - Description of the Multiple Look Approach for Calculating Unsteady Loudness	Nazanin Pourmand & Vijay Parsa - Objective Assessment of Speech Enhancement Algorithms
11.50 – 12.10	Ivan Sabourin, Berndt Zeitler & Frances King - Effects of Structural Load and Joist Type on Flanking Sound Transmission		
12.10 – 2.00	LUNCH		

DAY ONE

WEDNESDAY 14TH OCTOBER 2009 (EARLY AFTERNOON)

	<i>ARCHITECTURAL ACOUSTICS II</i> (Chair: John Bradley)	<i>SINGING (AIRS)</i> (Chair: Annabel Cohen)	<i>DIGITAL SIGNAL PROCESSING II</i> (Chair: Chris Forrester)
	ROOM	ROOM	ROOM
2.00 – 2.20	J. David Quirt & John K. Dickinson - Tool for Predicting Air-Borne and Structure-Borne Sound in Buildings	Steven Brown & Peter Pfordresher - Imprecise Singing is Widespread	Malay Gupta, Chris Forrester & Sean Simmons - Noise Suppression in Cellular Telephony
2.20 – 2.40	Mandy Chan & Bill Gastmeier - Acoustical Performance Criteria and Treatment Protocols for Learning Spaces at a Large Institutional Teaching Facility	Lisa Chan & Frank A. Russo - The Role of Mimicry in Perception of Sung Emotion	Karl Wiklund & Simon Haykin - The Cocktail Party Problem: Solutions and Applications
2.40 – 3.00	Scott Mallais & John Vanderkooy - A Comparison of Three Methods for the In Situ Determination of Acoustic Absorption Coefficients	Rayna Friendly, Laurel Trainor & Steven Brown - Development of Singing: The Current State of Our Knowledge and an Outline of Critical Questions	Lin Wang, Heping Ding & Fuliang Yin - Region-Growing Permutation Alignment Approach in Frequency-Domain Convolutional Blind Source Separation
3.00 – 3.10	BREAK		
3.10 – 3.30	Jean-Philippe Migneron - Passive Noise Attenuation for Natural Ventilation System	Jeffery A. Jones, Colin S. Hawco & Dwayne Keough - Holding a Note: Sensorimotor Control during Singing	Phil Hetherington - Robust Multi-Channel Mixer & Spatial Post Filter
3.30 – 3.50	Ryan J. Matheson - Low-frequency Room Demerit Analysis and Equalization using Properly-modeled Source Terms in Finite-Difference Time-Domain Simulations	Annabel Cohen, Jenna Coady, Marsha Lannan, Emily Gallant & Vickie Armstrong - A Test Battery for Singing Suited for Lifespan and Longitudinal Studies	Malay Gupta, Chris Forrester & Sean Simmons - Review of Wideband Speech Noise Reduction Techniques
3.50 – 4.10	COFFEE BREAK		

DAY ONE

WEDNESDAY 14TH OCTOBER 2009 (LATE AFTERNOON)

	AEROACOUSTICS <i>(Chair: Rama Bhat)</i>	MUSICAL ACOUSTICS/SOUNDSCAPES <i>(Chair: Chris Waltham)</i>	DIGITAL SIGNAL PROCESSING III <i>(Chair: Chris Forrester)</i>
	ROOM	ROOM	ROOM
4.10 – 4.30	Ramani Ramakrishnan, Sergio Raimondo, Anant Grewal, & Gary Elfstrom - Screech Suppression of Supersonic Jet Noise	Shira Daltrop, Andrzej Kotlicki & Chris Waltham - The Vibroacoustic Behavior of Harp Soundboxes	Celia Shahnaz, Wei-Ping Zhu & M. Omair Ahmad - A Residual-Cepstrum Method of Pitch Estimation from Noisy Speech
4.30 – 4.50	Kareem Aly & Samir Ziada - Partially Spinning Behaviour of Flow-Excited Diametral Modes of Internal Shallow Cavities	Ben Gaum & Ramani Ramakrishnan - Sound Created Form	Shaikh Anowarul Fattah, Wei-Ping Zhu, & M. Omair Ahmad - Band-Adaptive Formant Frequency Estimation from Noisy Speech in Correlation Domain
4.50 – 5.10	Gary M. Elfstrom - Trends in Aero-acoustic Wind Tunnel Testing	Hein Schoer - The Sounding Museum: Cultural Soundscapes as a Toll in Museum Education	Marc Schönwiesner, Jérémie Voix & Philippe Pango - Digital Earplug for Brain Plasticity Research
5.10 – 5.30	Werner Richarz - Extracting Dual Spectra from Measured Jet Noise Data	Paolo Ammirante, William F. Thompson, & Frank A. Russo - Tapping to Pitched Auditory Feedback Tones: Effects of Pitch Contour and Interval Size on Intertap-Interval and Tap Force	Henry Luo - Acoustic Shock Detection and Control
6.10 – 7.00	ACOUSTIC WALK (MURRAY SCHAFFER)		

DAY TWO

THURSDAY 15TH OCTOBER 2009 (MORNING)

9.00 – 9.45	PLENARY 3 (XXX)		
9.45 – 10.00	COFFEE BREAK		
	UNDERWATER ACOUSTICS (Chair: Stan Dosso)	HEARING CONVERSATION (Chair: Alberto Behar)	AUDITORY SCENE ANALYSIS (Chair: Claude Alain)
	ROOM	ROOM	ROOM
10.00 – 10.20	Nicos Pelavas, Garry J. Heard, Gordon R. Ebbeson, Carmen Lucas, Derek Clark - Acoustic Localization of an Autonomous Underwater Vehicle	Gurjit Singh, Kathy Pichora-Fuller, Thomas Behrens & Tobias Neher - The Effects of Hearing Loss and Linear vs. Nonlinear Hearing Aid Fitting Algorithms on Auditory Spatial Attention	Stephen Arnott - Neural Generators Underlying Concurrent Sound Segregation
10.20 – 10.40	Stan E. Dosso & Michael J. Wilmut - Bayesian Ocean Acoustic Source Track With Environmental Uncertainty	Julie Ladiges, Vijay Parsa, Susan Scollie & Danille Glista - Electroacoustic Measures of Frequency Lowering Hearing Aids	Sandra Campeanu, Boaz Ben-David, Kelly Tremblay & Claude Alain - Neural Correlates of Rapid Speech Identification Learning
10.40 – 11.00	Stan E. Dosso & Michael J. Wilmut - Bayesian Source Track Prediction in an Uncertain Environmental Inversion	Alberto Behar & Willy Wong - Hearing Protectors Standards	Ben Dyson - Something or nothing? Electrophysiological Variations in Encoding and Retrieval during the Processing of Complex Auditory Scenes
11.00 – 11.20	COFFEE BREAK		
11.20 – 11.40	I.V.Mastikhin, B.Newling, N.Hetherington & S.Kristoffersen - MRI Measurement of Dynamics of the Liquid and a Dissolved Gas during Acoustic Cavitation	Anthony J. Brammer, Gongqiang Yu, Eric R. Bernstein, Donald R. Peterson, Martin G. Cherniack & Jennifer B. Tuft - Monitoring Sound Pressure at the Eardrum for Hearing Conservation	Nicholas Haywood, Valter Ciocca, Julie Chang - The Perceived Continuity of a Tone through Noise and Silence
11.40 – 12.00	N M Alam Chowdhury, Zaiyi Liao, Lian Zhao & Ramani Ramakrishnan - Eigenfrequency Analysis of Fluid-filled Pipes	Cheng Qian & Maximilian Touzel - Direct Measurement of Actual Listening Levels of Digital Music Players in the Field	Ada Leung - An Object-Based Account of Auditory Scene Analysis
12.00 – 12.20	N M Alam Chowdhury, Zaiyi Liao, Lian Zhao & Ramani Ramakrishnan - Modal Solutions of the Acoustic Wave Propagation Through Fluid-Filled Pipes		Bernhard Ross - Temporal Dynamics of Selective Attention During Dichotic Listening
12.20 – 2.00	LUNCH		

DAY TWO

THURSDAY 15TH OCTOBER 2009 (EARLY AFTERNOON)

	ENVIRONMENTAL NOISE <i>(Chair: Bill Gastmeier)</i>	HEARING AIDS <i>(Chair: Christian Giguere)</i>	SPEECH I <i>(Chair: Bryan Gick)</i>
	ROOM	ROOM	ROOM
2.00 – 2.20	Michael Medal - Windmachine Noise Diagnostics	Marshall Chasin - Music and Hearing Aids	Ron I. Thomson and Talia Isaacs - Within-Category Variation in L2 English Vowel Learning
2.20 – 2.40	Al Lightstone, Guangsheng (Sam) Du, Ian Matthew & Irshad Rizvi - Simultaneous Measurements of Sound Level and Wind Profile	Véronique Vaillancourt, Chantai Laroche, Christian Giguère, Marc-André Beaulieu, & Jean-Pierre Legault - Evaluation of Aided and Unaided Auditory Functions for Royal Canadian Mounted Police (RCMP) Members	Molly Babel - Vowel Imitation In Spontaneous Phonetic Imitation
2.40 – 3.00	Ramani Ramakrishnan - Characteristics of Wind Turbine Noise	Rana Rahal, Wail Gueaieb, Christian Giguère & Hisham Othman - Automatic Volume Settings for Environment Sensitive Hearing Aids	Chenhao Chiu - Acoustic and Auditory Comparisons of Polish and Taiwanese Mandarin Sibilants
3.00 – 3.10	BREAK		
3.10 – 3.30	Scott Penton, Kevin Carr, Craig Vatcher & Russ Lewis - Airport Noise Modelling for the New Quito International Airport	Jin Zhou, Hisham Othman, Hilmi Dajani, & Tyseer Aboulnasr - A New Feature Selection Method for Volume Control in Direct-Learning Hearing Aid Systems	Takashi Mitsuya, Ewan MacDonald & Kevin Munhall - Compensatory response to perturbed voice onset time feedback
3.30 – 3.50	Ryan Daneluzzi, Steve Bogensberger, Mike Dobrin, Colin Novak & Helen Ule - An Approach to Generating and Calibrating a Sound Contour Map	Brady Laska, Miodrag Bolić & Rafik Goubran - Subband Adaptive Filtering for Acoustic Feedback Compensation in Hearing Aids	Connor Mayer, Bryan Gick & Elizabeth Ferch - Talking While Chewing: Speaker Response To Natural Perturbation of Speech
3.50 – 4.10		Nicolas Ellaham, Christian Giguère & Wail Gueaieb - Binaural Objective Intelligibility Measurement and Hearing Aids	Janette Quintero, Tim Bressmann, Katalin Mady & Ambros Beer - Magnetic Resonance Imaging and Videofluoroscopy of Partial Glossectomees' Speech: Preliminary Results
4.10 – 4.30	COFFEE BREAK		

DAY TWO

THURSDAY 15TH OCTOBER 2009 (LATE AFTERNOON)

	RURAL NOISE & NOISE CONTROL <i>(Chair: Werner Richarz)</i>	PSYCHOLOGICAL ACOUSTICS I <i>(Chair: Kathy Pichora-Fuller)</i>	SPEECH II & ECOLOGICAL ACOUSTICS <i>(Chair: Bryan Gick / Glenn Morris)</i>
	ROOM	ROOM	ROOM
4.30 – 4.50	Ian Bonsma & Megan Munro - An Investigation of the Wind Induced Background Sound Levels in Rural Areas for Existing and Proposed Wind Farms	Marieke van Heugten & Elizabeth K. Johnson - The Robustness of Infants' Early Word Representations	Björn Lyxell, Malin Wass, Birgitta Sahlén, Tina Ibertsson & Lena Asker-Arnåson - Cognitive Skills and Reading Ability in Children with CI
4.50 – 5.10	John P Harrison - Inadequacy of Wind Turbine Noise Regulations and their Application	Nicole Folland, Laurel Trainor & Nicholas Smith - Segregation of Simultaneous Auditory Objects by Harmonic Mistuning in Infants	Heather Flowers & Tim Bressmann - Acoustic Correlates of Neutral versus Angry Affect in Real and Non-Word Sentences
5.10 – 5.30	Brian Chapnik & Alex Lorimer - Source-Receptor Path Prioritization Considerations for the Specification of Silencers in Ventilation Systems	Rayna Friendly & Laurel Trainor - Development of Voice Discrimination: Perceptual Narrowing through Experience in Infancy	Brendan Rideout, Stan E. Dosso & David Hannay - Underwater Acoustic Localization of Marine Mammals in an Uncertain Environment
5.30 – 5.50	Thomas Paige - Cooling-Tower Noise Control for the Niagara Falls Casino	B.M Ben-David, C.G. Chambers, M. Daneman, M. K. Pichora-Fuller, E. Reingold & B. A. Schneider - Controlling for Age Related Hearing Loss Can Eliminate Aging Differences in Lexical Competition: Evidence from Eye-Tracking as an Online Measurement of Age and Noise Effects on Listening	Glenn K. Morris - Andean Ultrasonics: Bioacoustics of Two Tropical Montane Katydid
5.50 – 6.10	Michael Bowie, Colin Novak & Helen Ule - Absorptive Properties of Magnesium Panels for Automotive Dash Panel Applications	Bruce E. Amrein & Tomasz R. Letowski - The Environment for Auditory Research	
6.00 – 7.30	CAA ANNUAL GENERAL MEETING		
6.30 – 7.30	CASH BAR		
7.30 – 9.30	AWC09 BANQUET		

DAY THREE

FRIDAY 16TH OCTOBER 2009 (MORNING)

9.00 – 9.45	PLENARY 4 (XXX)		
9.45 – 10.00	COFFEE BREAK		
	<i>VIBRATIONS I</i> (Chair: Tony Brammer)	<i>PSYCHOLOGICAL ACOUSTICS II</i> (Chair: Ben Dyson)	<i>BIOMEDICAL ACOUSTICS</i> (Chair: Jahan Tavakkoli)
	ROOM	ROOM	ROOM
10.00 – 10.20	Martin G Cherniack, Anthony J Brammer, Ronnie Lundström, Timothy F. Morse, Greg Neely, Tohr Nilsson et al- Age and Neuropathies in Vibration Exposed Manual Workers	I. Ibrahim, R. Malcolmson, M. B. Jennings & M. F. Cheesman- Effect of Headphone Type on Listening Levels and Localization Abilities of Portable Audio Device Users in Quiet and Traffic Noise	Omar Falou, Min Rui, Ahmed El-Kaffas, Carl J. Kumaradas & Michael C. Kolios- Modelling High Frequency Acoustic Backscatter from Biological Cells
10.20 – 10.40	Anthony J. Brammer, Paivi Sutinen, Sourish Das, Ilmari Pyykkö, Esko Toppila & Jukka Starck- Vibrotactile Perception as a Test for Numbness and Pain	Blake Butler & Laurel Trainor- Temporal Representation of Pitch in Auditory Cortex	Eric Strohm & Michael C. Kolios- Measuring Scattering in Apoptotic Cancer Cells using Ultra High Frequency Acoustic Microscopy
10.40 – 11.00	Alice Turcot, Richard Larocque, Serge André Girard, Marilène Courteau & Samuel Tétreault- Noise and Hand-Arm Vibrations: A Dangerous Match	Antoine Shahin- Neural Mechanism of Phonemic Restoration	Arthur Worthington, Sankar Narasimhan, Jahan Tavakkoli & Michael C. Kolios- A Comparison of Imaging Modalities to Monitor Thermal and Mechanical Ultrasound Tissue Therapies
11.00 – 11.10	BREAK		
11.10 – 11.30	Donald Peterson, Takafumi Asaki, Anthony Brammer & Martin D. Cherniack - Sound and Vibration Exposures from Polishers and Scalars in a Dental Hygienist Population	Benjamin Rich Zendel & Claude Alain - Musicians Experience Less Age-related Hearing Loss	Shahram Mashouf & Jahan Tavakkoli- An Enhanced Numerical Model to Simulate Nonlinear Continuous Wave Ultrasound Field
11.30 – 11.50	Tammy Eger, Jim Dickey & Michele Oliver- Whole-body Vibration Exposure: Understanding, Evaluating and Predicting Negative Health Risks	Kathleen Corrigan & Laurel Trainor - Measuring Auditory Attention Networks	Reinhart Frosch- Old and New Cochlear Maps
11.50 – 12.10		Michael Schutz & Michael Kubovy - Tone Envelope: A Cue for Audio-Visual Unity	
12.10 – 2.00	LUNCH (STUDENT AWARDS)		

DAY THREE

FRIDAY 16TH OCTOBER 2009 (EARLY AFTERNOON)

	VIBRATIONS II <i>(Chair: Tony Brammer)</i>	MUSIC PERCEPTION <i>(Chair: Frank Russo)</i>	SPEECH III <i>(Chair: Pascal van Lieshout)</i>
	ROOM	ROOM	ROOM
2.00 – 2.20	Nikolina Samardzic & Colin Novak - Transfer Path Analysis using Engine Radiated Sound and Mount Vibration	Roy Patterson, Etienne Gaudrain & Tom Walters . The Role of Acoustic Scale in the Perception of Musical Notes and Instruments (Part I of Invited Talk)	David Purcell & Kevin Munhall - Auditory Feedback and Maintenance of English Vowel Production
2.20 – 2.40	Joana da Rocha, Afzal Suleman & Fernando Lau - An Accurate Coupled Structural-Acoustic Analytical Framework for the Prediction of Random and Flow-Induced Noise in Transport Vehicles: Its Validation	Roy Patterson, Etienne Gaudrain & Tom Walters . The Role of Acoustic Scale in the Perception of Musical Notes and Instruments (Part II of Invited Talk)	Boaz M. Ben-David, Pascal van Lieshout & Alex Nishta - The Sound of Stroop: Acoustic Effects in Stroop Interference
2.40 – 3.00	Salem Bashmal, Rama Bhat & Subash Rakheja - Experimental Studies on the In-Plane Vibrations and Sound Radiation in an Annular Thick Disk	Michael D. Hall - Perception of Musical Sources with Impoverished Spectral Envelopes	Jessica Banh, Konstantin Naumenko, Huiwen Goy, Pascal van Lieshout, David N. Fernandes & Kathy Pichora-Fuller - Establishing Normative Voice Characteristics of Younger and Older Adults
3.00 – 3.10	BREAK		
3.10 – 3.30	Avinash K. Bhaskar, Muthukumaran Packirisamy & Rama Bhat - Modeling Electrostatic Field in MEMS Devices using Artificial Springs in Rayleigh-Ritz Method	Peter Pfordresher & Justin Couchman - Does Altered Auditory Feedback Sound like Feedback?	Luyao Ma, Huawei Colin Li, Akiko Amano-Kusumoto, Willy Wong & Pascal van Lieshout - Assessing the Intrinsic Relationship between Facial Motion and Acoustics in Patients with Parkinson's disease
3.30 – 3.50	MD Shakhawat Hossain, Muthukumaran Packirisamy & Subhash Rakheja - Dynamic Behavior of Micro Structure in Micro Fluidic Channel	Huiwen Goy, D. Timothy Ives, Frank A. Russo, M. Kathleen Pichora-Fuller & Roy D. Patterson - Changes in Melodic Perception as a Function of Age and Lowest Harmonic Component	Anna Limanni - Men, Women and Lenition: Gender Differences in the Production of Intervocalic Voiced Stops in Mexican Spanish
3.50 – 4.10		Carmen Branje, Michael Maksimowski, Gabe Nespoli, Maria Karam, Deborah Fels & Frank Russo - Development and Validation of a Sensory-Substitution Technology for Music	Anneke Slis & Pascal van Lieshout - Separating Normal Variation in Movement Amplitudes from Gradient Speech Errors

NOTICE

The Board of Directors is proposing a bylaw change for the date of the financial year-end from August 31st to June 30th to facilitate management of the Association, as indicated during the 2008 Annual General Meeting (AGM) in Vancouver on Oct. 7 and in the minutes of the AGM posted in the December issue of Canadian Acoustics. Final debate and voting on this issue will take place during the 2009 AGM in Niagara-on-the-Lake on Oct. 15th.

AVIS

The Comité de Direction désire modifier la fin de l'année financière du 30 juin au 31 août pour faciliter la gestion de l'Association. Cette proposition de changement aux statuts de l'association a été annoncée lors de l'Assemblée générale annuelle le 7 octobre 2008 à Vancouver et le procès-verbal de cette Assemblée a été publié dans l'édition de décembre de l'Acoustique Canadienne. Un débat et vote final sur cette question aura lieu lors de la prochaine Assemblée générale annuelle à Niagara-on-the-Lake, le 15 octobre 2009.

When "BUY" does not apply, give RENTAL a try!

At Scantek, Inc. we specialize in **Sound and Vibration Instrument Rental with *expert assistance***, and fully calibrated instruments for:

Applications

- Building acoustics
- Sound power measurement
- Community noise
- Building vibration
- Industrial noise
- Human body vibration
- Machine diagnostics
- Vibration measurement

Instruments analyzers •

- FFT and real-time*
- 1/3 and 1/1 octave bands*
- noise and vibration dosimeters •
- vibration meters •
- human body dose/vibration •
- A-weighted sound level meters •
- rangefinders •
- GPS •
- windscreens •
- wide range of microphones •
- and accelerometers

Scantek, Inc.

**Sound & Vibration Instrumentation
and Engineering**

www.scantekinc.com
info@scantekinc.com

800-224-3813

Architectural Acoustics: Some Historical Developments and Ongoing Issues

John S. Bradley

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

Abstract

This paper gives an overview of the history of the development of architectural acoustics and a sampling of the results of several more recent research studies. There are now standardised measures for assessing room acoustics character and there are criteria for room acoustics in terms of these measures. For rooms for speech, such criteria are not so well defined for some special groups of listeners such as the very young or old, hearing impaired listeners, and those listening in a second language. For rooms for musical performance, there is a growing understanding of the more important acoustical characteristics of concert halls, but still a need to improve tools for applying this knowledge to the design of these spaces.

RÉSUMÉ

Cet article donne un aperçu de l'historique du développement de l'acoustique architecturale ainsi qu'un échantillon des résultats des derniers projets de recherche. Il existe actuellement des mesures standardisées pour évaluer l'acoustique d'une pièce et des critères relatifs à ces mesures. Pour les pièces réservées à la communication orale, de tels critères ne sont pas très bien définis pour certains groupes d'auditeurs, tels que les très jeunes ou les gens âgés, les malentendants ou ceux qui écoutent un discours dans une langue qui n'est pas la leur. Dans le domaine musical, les caractéristiques acoustiques les plus importantes des salles de concerts sont de mieux en mieux connues, mais il faut améliorer les outils de façon à appliquer ces connaissances à la conception de ces espaces.

Introduction

Architectural acoustics has a long history that can be traced back to ancient Greece and Rome and the designs of early theatres. More recent quantitative research has progressed over the past 160 years with a number of re-discoveries of the results of earlier research. This paper gives an overview of some key points of this history and illustrates more recent progress with a sampling of various research studies. Although not intended to be comprehensive, a broad range of topics are included and a number of indications of remaining significant problems are mentioned.

Impulse Responses

Most important acoustical characteristics of rooms can be obtained from impulse response measurements in the rooms. Graphs of measured impulse responses are very helpful for explaining the important characteristics of sound in rooms and are introduced here so that they can be used in later explanations of room acoustics effects. Although impulse responses can be obtained by simply recording the response to an impulsive sound in a room (e.g. gun shots or balloon bursts), modern techniques allow us to calculate impulse responses more accurately from the measured response to various broad spectrum test signals such as sine sweeps or maximum length sequences sequences.

Fig. 1 shows an example of the first 300 ms of an impulse response measured in a large auditorium. Although there

are a few distinct and higher amplitude features near the start of the impulse response, much of the response seems to consist of somewhat random pressure fluctuations. In practice it is the result of the combination of a large number of reflections of the original sound. Fortunately it is not necessary to understand the effects of each detail in the response and these details would probably change for relatively small movements of the source or the receiver. What is important is the total sound energy associated with each of the 3 main features of the response: (a) the direct sound, (b) the early-arriving reflections, and (c) the later-arriving reflections (or reverberant decay).

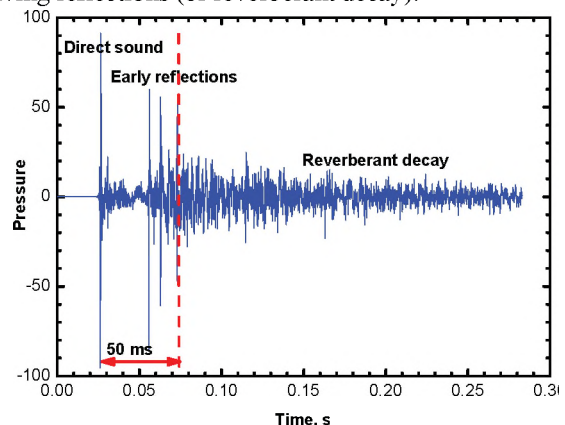


Fig. 1 Impulse response measured in a large auditorium plotted as pressure versus time.

Early-arriving reflections are important physically, because they usually include a large portion of the total

sound energy in the impulse response, and subjectively, because they are not heard as separate events but our hearing system integrates them with the direct sound. Thus more early-arriving reflections increase the effective level of the direct sound and increase the perceived clarity of sounds in rooms. Conversely, increased later-arriving sounds would decrease the clarity of sounds. The history of how we have come to understand these features will be discussed in the following section.

Fig. 2 shows a plot of the same impulse response but with the amplitude in decibels. Plotting the response in this way tends to lead to approximately linear sound decays after the initial reflections. We determine the reverberation time, that is the time for sound to decay 60 dB, from the slope of the decay of the later-arriving sound. While reverberation time is widely used as a simple room acoustics design parameter, considering the details of impulse responses, as illustrated in Fig. 1, gives a more complete understanding of room acoustics phenomena.

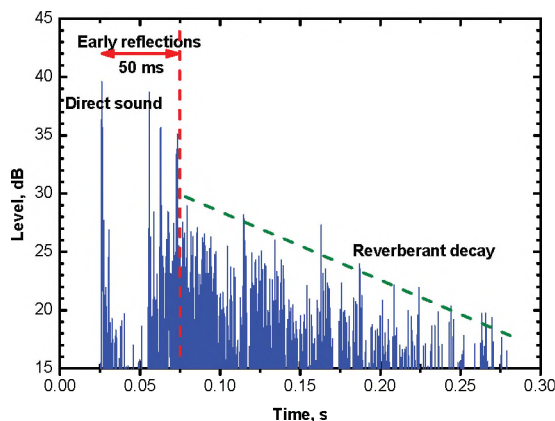


Fig. 2. Impulse response measured in a large auditorium plotted as sound level in dB versus time.

In Fig. 1 and 2 early-arriving is shown as the first 50 ms after the direct sound. This is appropriate for speech sounds but 80 ms is more appropriate for music.

Early History

The history of architectural acoustics is an interesting sequence of discovery and later re-discovery of various key details. Surprisingly, a quite complete understanding of the benefits of early-arriving reflections was reported in the 1850s. Joseph Henry, when head of the Smithsonian in Washington, reported the results of his experiments listening to the reflections of impulsive sounds outdoors [1]. By varying the relative arrival times of the direct sound and the reflection, he discovered that the reflection was not separately heard but enhanced the direct sound when it arrived within up to about 50 ms after the direct sound. He also appreciated that this was important in the design of rooms for speech. He designed a lecture theatre at the Smithsonian shaped to maximize the benefit of

early-arriving reflections or as he put it, “to husband every articulation of the voice”. This knowledge appears to have been lost for about 100 years before it was re-discovered by Hass [2-4].

Approximately 50 years after Joseph Henry’s work, Sabine determined experimentally the relationship between reverberation time and the combination of room volume and the total area of sound absorption in a room. Ever since then his reverberation time equation has been used as a simple method of predicting room acoustics characteristics.

Equation (1) shows the Sabine reverberation time equation, where T_{60} is reverberation time (s), V the room volume (m^3), and A the total sound absorption (m^2).

$$T_{60} = 0.161 V/A \quad (1)$$

While Sabine’s equation was an important step forward, it took acousticians another 50 years to realize that there was more to room acoustics than reverberation time, something that Joseph Henry had understood in the 1850s.

Haas’ research [2-4], established in a quantitative manner, the importance of early-arriving reflections. With the aid of electronics he was able to create early-arriving reflections with varied delay and amplitude. Fig. 3 shows the results of several researchers showing how early-arriving reflections can be as much as 10 dB higher in level than the direct sound and still be completely integrated with the direct sound, depending on the actual delay of the reflection relative to the arrival of the direct sound.

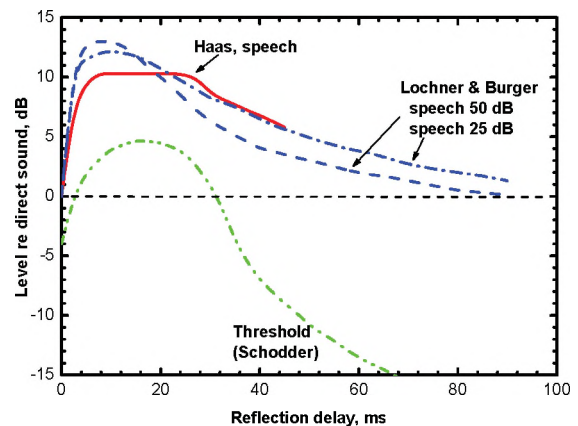


Fig. 3. Upper three curves show the limit of subjective integration of reflections with the direct sound by Haas [2-4] and Lochner and Burger [5].

As shown in Fig. 3 Lochner and Burger’s results [5] for speech sounds were very similar to those of Haas. Haas also found that the shape of the curves shown in Fig. 3 vary as the bandwidth of the test signal is varied. The Haas effect also includes the perception that the reflected sound arrives from the same direction as the first arriving sound,

which is usually the direct sound. Because of this other aspect of the effect of early-arriving reflections, the *Haas Effect* is also referred to as the *Precedence Effect*. This name was given by Wallach et al. [6] who published their results in the same year as Haas.

Although the papers formally reporting experimental results to demonstrate the Haas or Precedence effect didn't appear until 1949, there is evidence of a much earlier appreciation of these effects. For example, in 1935 Fay and Hall [7] reported using the Haas or Precedence effect to design a sound reinforcement system. By arranging for the sound from the loudspeaker to arrive shortly after the arrival of the direct sound from the talker they could preserve the realism of the sound appearing to come from the location of the actual talker. Gardener has reviewed the history of research related to the Haas effect [8] and cites a number of other earlier studies indicating some understanding of the concepts we now refer to as either the Haas effect or the Precedence effect.

After the publications of Haas and Wallach et al., others considered the effects of more than one reflection. For example, Seraphim [9] explored the subjective integration of a variable second reflection in the presence of a direct sound and a fixed reflection for speech sounds. Lochner and Burger explored the subjective integration of early reflections of speech sounds as a function of the delay and amplitude of the reflections, to develop a model for useful-to-detrimental sound ratios [5].

More recently Bradley et al. [10] published results to confirm the benefits of early reflections for the intelligibility of speech. As illustrated in Fig. 4, they also presented results demonstrating the practical importance of early-arriving reflections by showing that early-arriving reflections can increase the effective direct sound level by as much as 8 or 9 dB in rooms. The early-arriving reflections make it possible to understand speech in

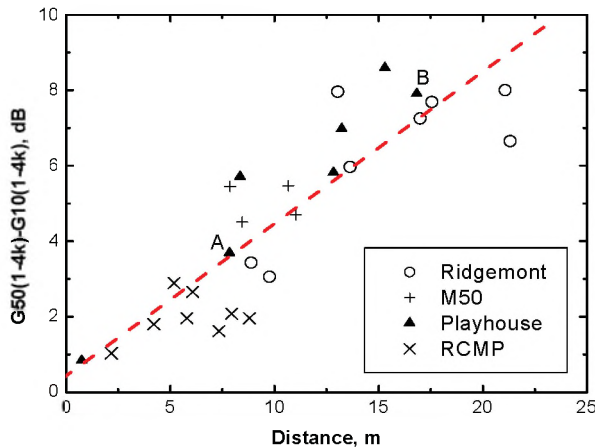


Fig. 4. The increasing contribution of early reflections to early sound strength versus source receiver distance [10].

situations where the direct sound is reduced such as when the talker is turned away from the listener or when the listener is simply farther from the source.

Room Acoustics Measures

The various studies of the subjective integration of early-arriving reflections led to an understanding that the balance between early- and late-arriving sound influenced the perceived clarity of sounds in rooms. Table 1 describes 4 such measures proposed over a 20-year interval. In addition to the mathematical definitions, the calculations are illustrated by pictograms with the shaded areas indicating the integrated parts.

Lochner and Burger took these measures one step further to create the useful-to-detrimental sound ratio (U) concept.

$$U = 10 \log \left\{ \frac{E_d + E_e}{E_l + E_n} \right\}, \text{ dB} \quad (2)$$

Here, E_d is the direct sound energy, E_e is the early-arriving sound energy, E_l is the later-arriving sound energy and E_n is the noise energy. This measure includes both the effects of signal-to-noise as well as an accurate rating of the clarity of the room acoustics and is an excellent predictor of the intelligibility of speech in rooms

Although this was published by Lochner and Burger as a

<p>Deutlichkeit (Thiele, 1953)</p> $D = \frac{\int_0^{50} p^2(t) \cdot dt}{\int_0^{\infty} p^2(t) \cdot dt}$	
<p>Running Liveness (Schultz, 1965)</p> $R = 10 \log \left(\frac{\int_0^{\infty} p^2(t) \cdot dt}{\int_0^{50} p^2(t) \cdot dt} \right)$	
<p>Clarity (Reichardt, 1975)</p> $C_{80} = 10 \log \left(\frac{\int_0^{80} p^2(t) \cdot dt}{\int_{80}^{\infty} p^2(t) \cdot dt} \right)$	
<p>Centre Time (Kürer, 1973)</p> $TS = \frac{\int_0^{\infty} t \cdot p^2(t) \cdot dt}{\int_0^{\infty} p^2(t) \cdot dt}$	

Table 1. Measures of clarity and the balance between reverberance and clarity [12-15].

new concept [5], it was very similar to the remarkable earlier work of Aigner and Strutt [11]. In addition to mentioning an integration limit of about 60 ms for early reflections, they proposed a measure of syllable articulation (intelligibility) Q , which was the same as equation (2) except it was a linear ratio without the '10log' part.

Although U ratios have been shown to be good predictors of the intelligibility of speech in rooms [16, 17], a standard format has never been proposed and it is not established how one should best combine information over the range of speech frequencies.

In the 1970s Houtgast and Steeneken [18] proposed the Speech Transmission Index (STI) as a measure of the combined effects of room acoustics and signal-to-noise on the intelligibility of speech in rooms. The measure is based on the assumption that intelligibility is related to the degradations of the amplitude modulation of speech by noise and reverberant sound. It uses modulation transfer functions to determine the expected speech intelligibility between points. The procedure calculates modulation reductions for the 98 combinations of 7 acoustical octave bands (125 to 8k Hz) and 14 1/3-octave bands for modulation frequencies from 0.63 to 12.5 Hz. Houtgast and Steeneken have made great efforts to fully develop STI and it is described in an IEC standard [19].

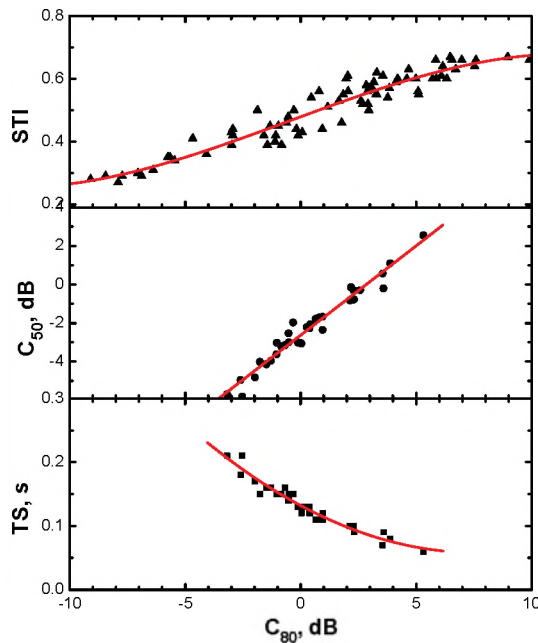


Fig. 5. Plot of measured STI, C_{50} and TS values versus C_{80} values. The STI values were calculated without the influence of signal-to-noise ratios so that room acoustics only measures can be compared [20].

Although STI seems completely different than U ratios, values of the two types of measures are highly correlated and are seem to provide the same information about

conditions for speech in rooms. Fig. 5 includes a plot of STI values (with the signal-to-noise component ignored) versus C_{80} values (an early-to-late arriving sound ratio with an 80 ms early time limit). The STI values as well as the C_{50} and TS values are all quite well related with the C_{80} values. That is, the room acoustic component of STI is equivalent to an early-to-late arriving sound ratio. The results in Fig. 6 similarly show that STI values (including the signal-to-noise component) are strongly correlated with $U_{50}(A)$ values (an A-weighted useful-to-ratio that included a 50 ms early time interval). Although the Useful-to-Detrimental Ratio and the Speech Transmission Index measures are based on quite different concepts, they seem to be similarly accurate predictors of the intelligibility of speech in rooms.

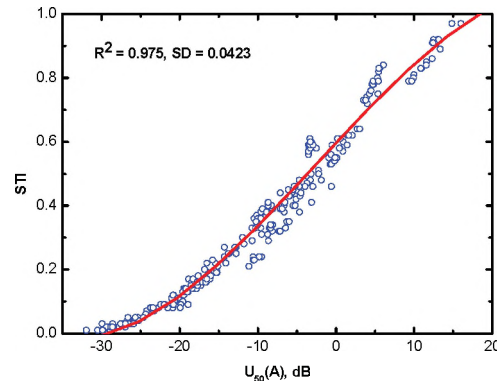


Fig. 6. Plot of STI values versus $U_{50}(A)$ values [20].

There are also measures of reverberance. The reverberation time, T_{60} , (s), is the time for sound to decay by 60 decibels after the source has stopped. It is usually measured from the slope of the sound decay between points -5 dB and -35 dB below the initial maximum. T_{60} is an important physical parameter, but Early Decay Times, EDT (s), are better related to the perceived degree of reverberance. EDT is calculated from the slope of the sound decay between points 0 dB and -10 dB below the initial maximum.

One of the most subjectively important characteristics of sound in rooms is the level of the sounds and in particular the contribution of the room to the level of the sound. This is measured by the sound strength, G dB, which is the level of the sound relative to the level of the same source at a distance of 10 m in a free field. It is an indicator of how the room affects the level of sounds. G values can also be calculated separately for the early-arriving or late-arriving sounds [22].

All of the room acoustics parameters are measured in octave bands and are described in ISO3382 measurement standard [21].

Classrooms and Younger Listeners

Although criteria exist to indicate ideal conditions for speech communication in rooms, there still are questions

as to appropriate criteria for some groups of listeners with special needs. These would include: younger children, second language listeners, older listeners and, of course, the hearing impaired. The case of conditions for younger children will be used to illustrate how criteria for this one special group have been studied and how their needs differ from the average non-impaired adult listener.

Many studies have shown that younger children are less able to understand simple speech in noise than adults. Although previous results all showed that younger children have greater difficulties, the relationships between intelligibility scores and signal-to-noise ratios were quite different among the studies and did not agree well with previous results in actual classrooms. The differences were thought to be due to the often unrealistic conditions of the tests that frequently used monaural headphone listening and usually made no attempt to duplicate binaural listening conditions in typical classrooms.

Fig. 7 shows the results of speech tests on grade 1, 3 and 6 students carried out while they were seated at their normal seats in their own classrooms [23]. The scores for the three age groups increase with increasing age as expected and agree with the limited previously available classroom data [17]. The results show that for very high signal-to-noise ratios all children could get scores of 98% or better. As signal-to-noise ratios decrease, scores decrease, and the differences among the 3 age groups grow. The scatter of the results about the mean trends also increases with decreasing signal-to-noise ratio.

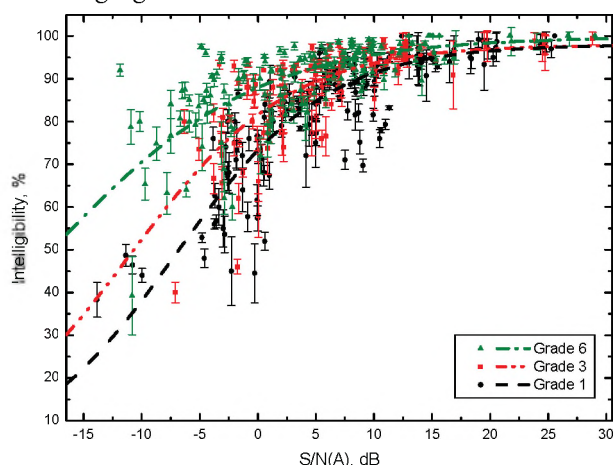


Fig. 7. Mean speech intelligibility scores versus A-weighted signal-to-noise ratio ($S/N(A)$) for grade 1, 3 and 6 students [23].

Although ideal conditions for speech communication are usually said to require a $S/N(A)$ of at least 15 dB, an analysis of these data concluded that grade 1 children needed an $S/N(A)$ of at least 20 dB, the grade 3 students at least 18 dB and the grade 6 students at least 15 dB. Using the measured speech levels of the teachers, it was possible to determine that the preferred maximum ambient noise

level of 35 dBA, recommended in ANSI S12.60, was easily justified.

Some question why 35 dBA is needed when children themselves often create higher levels of disturbing sounds than this. However, the children's sounds are influenced by the existing ambient noise of heating and ventilating systems. Due to an effect similar to the Lombard effect [24] they will tend to talk to their neighbours and make other sounds at a level that can be heard above the sounds of the existing ambient noises. It can be argued that we do not completely understand how room acoustics conditions and existing ambient noise interact with the behaviours of the teachers and students to create the combined interfering sound levels that exist in classrooms.

Classroom acoustics criteria usually also specify acceptable reverberation times. Frequently reverberation is assumed to be a negative factor for speech communication and therefore it should be as short as possible. Of course, this ignores room acoustics and the fact that reflected sounds enhance speech levels and can provide better conditions for speech communication. Joseph Henry understood this 160 years ago but many are not familiar with the importance of early-arriving reflections and why they are so important in rooms.

To explore optimum reverberation times for elementary school classrooms, children were tested using the same speech tests as in the previous study and by recording the tests in simulated conditions with a range of reverberation times [25]. Great care was taken to create conditions representative of classrooms. Binaural recordings of the speech test material were played back to individual students using headphones. Fig. 8 plots one of the results showing mean intelligibility scores versus reverberation time. In this example, sound fields were created in which the added reflections with increasing reverberation time were allowed to increase the overall speech levels as can occur in rooms.

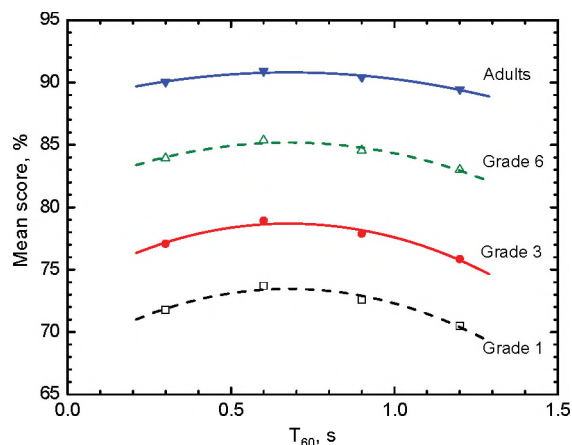


Fig. 8. Mean intelligibility scores versus reverberation time for 4 age groups [25].

The results in Fig. 8 show large difference between the age groups because of the different effects of signal-to-noise ratio with age. However, the variations with reverberation time are quite similar for all age groups. These results indicate that a range of reverberation times would lead to approximately the same near maximum intelligibility scores. The differences between the curves in Fig. 8 indicate that excessive noise is a more important problem than non-optimum room acoustics. Of course, other particular groups of listeners may be differently affected by reverberation time.

Rooms for Music: Bass Sounds

Most of the issues for rooms for speech are also important for rooms for musical performances. However, in rooms for music the acoustical issues are more complex and there are some additional concerns to consider. One of these is the strength of low frequency or bass sounds in rooms for music. Although not necessary for speech, the room should strongly reproduce bass sounds for most types of musical performances.

The classical approach for providing adequate strength of the bass sounds assumes that the strength of bass is related to low frequency reverberation times and that rooms for musical performances should have increased low frequency reverberation times. Typically the 125 Hz reverberation time is required to be 1.5 times the mid-frequency value. This sometimes leads to quite extreme efforts to reduce low frequency absorption in halls.

Listening tests were carried out in simulated conditions in an anechoic room to determine whether the level or the reverberation time of the low frequency sound determines perceived bass strength [26]. Subjects heard the same piece of music played for 8 different conditions consisting of the combinations of 4 levels of the early-arriving bass sounds and two low frequency reverberation times. Subjects rated the perceived bass strength on a 5-point scale. As shown in

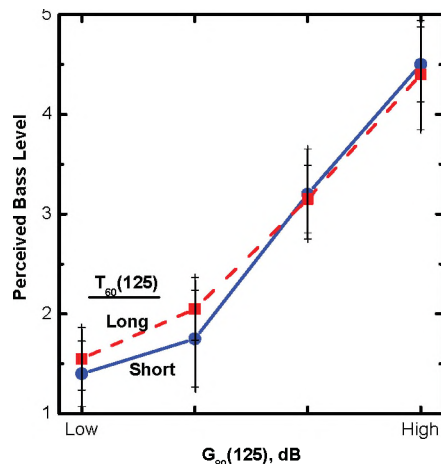


Fig. 9. Mean ratings of perceived bass level versus the strength of the early arriving bass sounds for two 125 Hz reverberation times (1.4 and 3.2 s) [26].

Fig. 9, the perceived strength of the bass sound was strongly and significantly affected by the level of the early arriving low frequency sound. However, the varied low frequency reverberation time (1.4 to 3.2 s) did not significantly influence the results. The perceived bass strength was determined by the level of the bass sounds and not by the reverberation time.

It is also important to understand the propagation of early-arriving low frequency sounds in concert halls. There is a phenomenon known as the seat dip effect [27, 28] that can strongly attenuate low frequency sound travelling at near grazing incidence to the audience surface. This attenuation is essentially due to an interference of the direct sound and reflections from the audience or the seating that cause large cancellations over a band of frequencies that is typically around 125 Hz. Fig. 10 plots measurements of this attenuation for varied angle of incidence. The frequency of the interference minimum shifts systematically higher in frequency as the angle of incidence increases. As indicated in Fig. 10, the maximum attenuation can be 10 dB or more. This is a much larger effect than can be obtained by decreasing the low frequency absorption in a hall. The usual design goal to increase 125 Hz reverberation times by 50% would only increase bass levels by about 2 dB.

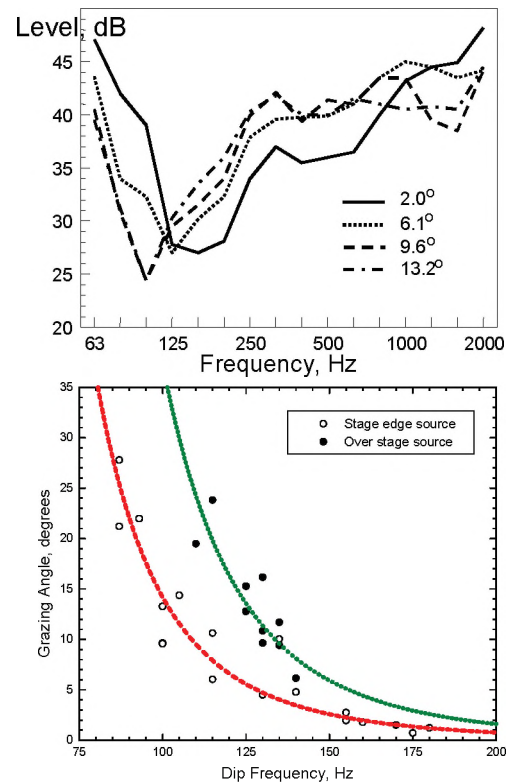


Fig. 10. (Upper) Spectrum of early arriving sound at positions above the audience seating for 4 different angles of incidence. (Lower) Frequency of the attenuation maximum for two different on-stage source positions and varied source height [29].

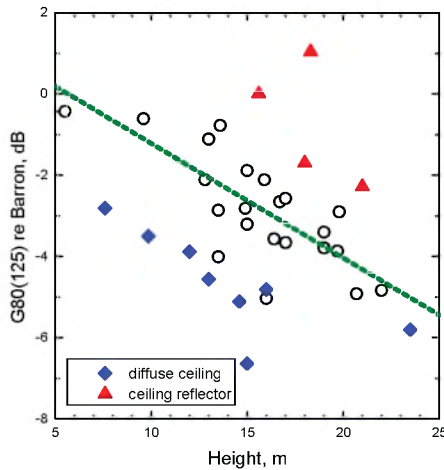


Fig. 11. Hall average 125 Hz octave band early-arriving sound strength ($G_{80}(125)$) versus ceiling height. $G_{80}(125)$ values were normalized relative to the predicted levels using Barron's theory to remove the effects of room volume and reverberation [30].

Although the problem is now better understood, a clear design approach to achieving strong bass sound is not available. One factor that does influence this bass attenuation is the reflecting properties of the ceiling. Fig. 11 plots the average strength of the early arriving 125 octave band sound levels versus ceiling height and identifies the ceilings as one of 3 types. Lower ceilings lead to increased low frequency sound levels. Diffusely reflecting ceilings reduce the strength of low frequency sound relative to other ceilings and halls with special ceiling reflectors can have increased levels of low frequency sound. Presumably stronger ceiling reflections can minimize the interference effects that lead to the seat dip attenuation. However, more research is needed to more fully understand the merits of various ceiling types in concert halls.

Rooms for Music: Spatial Impression

Another topic that must be considered in rooms for musical performance is spatial impression. Spatial impression was initially loosely described as the difference between more or less 'looking at the sound' as in outdoor situations and being immersed in the sound as in a good concert hall. Spatial impression was said to include apparent broadening of the sound source and a sense of being enveloped in the sound. Barron's pioneering work [31,32] showed that early-arriving lateral reflections led to perceived spatial impression. Fig. 12 shows Barron's graph indicating the combinations of the relative level of early arriving lateral reflections and the time delay after the arrival of the direct sound that led to the perception of spatial impression.

Barron's work also led to the definition of the lateral energy fraction (LF) as a measure of spatial impression.

LF is simply the linear ratio of the early-arriving lateral energy to the total early-arriving energy. The early time period is taken as the first 80 ms and the first 5 ms of the early lateral energy is ignored to avoid any influence of the direct sound.

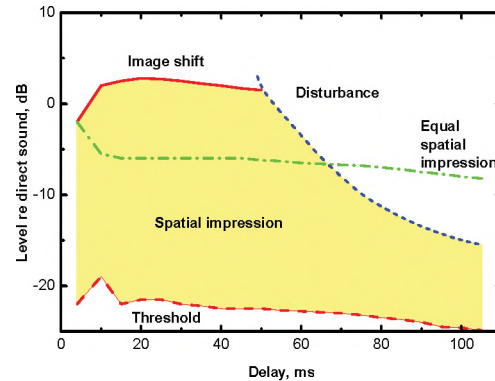


Fig. 12. The shaded area indicates combinations of the relative level of early-arriving lateral reflections and time delay after the arrival of the direct sound that led to perceived spatial impression [31].

Somewhat later Bradley and Souloudre tried to repeat Barron's experiments and found that they could not create a sense of envelopment with early-arriving lateral reflections. The early-arriving lateral reflections led to an impression of increasing apparent source width (ASW). After several experiments they demonstrated that a sense of listener envelopment (LEV) was created when later-arriving lateral reflections were present [33,34]. Fig. 13 includes the results of one experiment showing that increased reverberation time and increased late-arriving sound levels led to an increased sense of envelopment. Further experiments showed that the direction of arrival was important too.

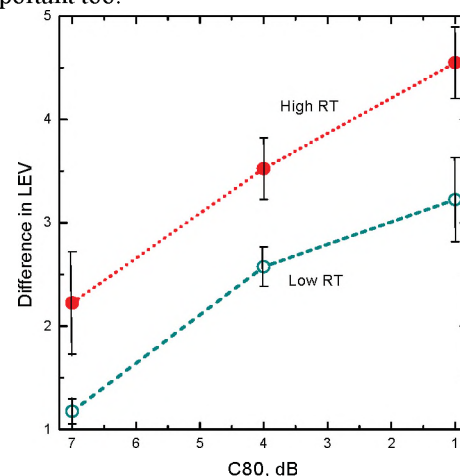


Fig. 13. LEV increases with reverberation time and the level of the late-arriving sound [33].

The strength of the later-arriving lateral energy was shown to be a good measure for predicting the amount of LEV

[34]. Fig 14. shows a plot of increasing LEV versus the strength of late-arriving lateral sound over the octave bands from 125 to 1000 Hz.

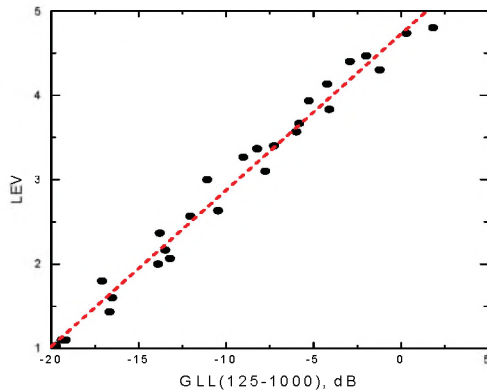


Fig. 14. LEV plotted versus the strength of the late-arriving sound energy calculated over the octave bands from 125 to 1000 Hz [34].

We now have a clearer understanding of spatial impression and how it has two components: ASW related to the strength of early-arriving lateral reflections and LEV related to the strength of later arriving lateral reflections.

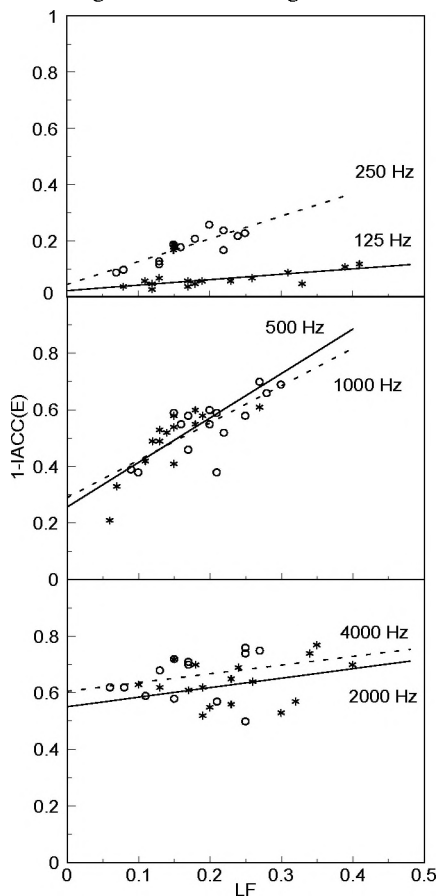


Fig. 15. Octave-band hall-average values of 1-IACC values plotted versus octave-band LF values [35].

While the LF measure proposed by Barron has been well accepted and used to evaluate concert halls, others have proposed using inter-aural cross correlations (IACC) to evaluate spatial impression. IACC values are usually calculated from impulse responses recorded at the two ears of an acoustical mannequin. When this is done for the early sound, that is the first 80 ms of the impulse response, the IACC values would be expected to relate to LF values and to be an indicator of the perceived ASW. Comparisons of hall-average values of these quantities in terms of octave band values are compared in Fig. 15. The 1-IACC values are quite significantly related to the LF values in all 6 octave bands. This strong correlation between the two measures suggests that they provide similar information about the average acoustical conditions of the halls. However in the lowest octave band (125 Hz), the variation in 1-IACC values is relatively small compared to the range of LF values. The same effect is seen in the 2k and 4k Hz octave bands.

When the individual seat measurements are compared in Fig. 16, the results are quite different. For the mid-frequency octave bands (500 and 1000 Hz) there is a reasonably good correlation between the two quantities.

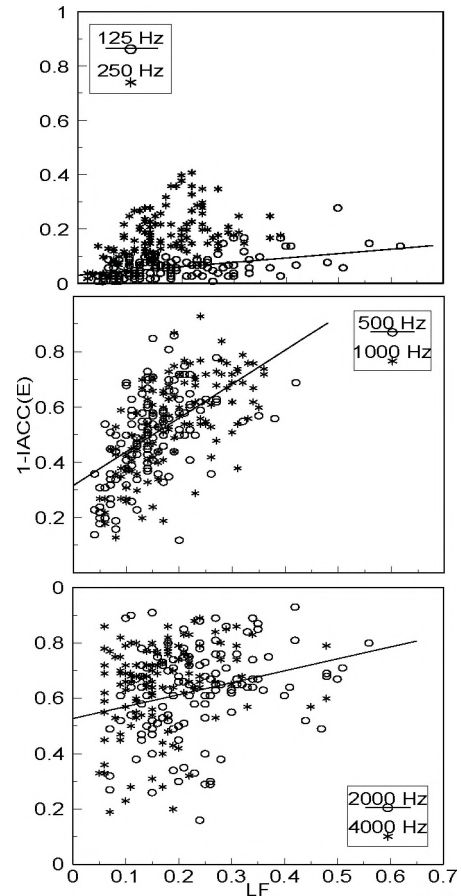


Fig. 16. Octave-band values of individual measurements of 1-IACC values plotted versus octave-band LF values [35].

However, for the lowest and highest two octave bands, the scatter is greater than the variation of the mean trend as indicated by the regression lines. There are other factors that cause greater variations in the I-IACC values that do not have the same effects on LF values. At lower frequencies the differences might be due to interference effects that influence I-IACC values but not the LF values which are a result of energy additions. The differences at higher frequencies might be due to scattered sound. However, neither explanation has been confirmed and it is not clear which measure best relates to the perceived changes in spatial impression.

Design and Evaluation of Rooms for Music

Room acoustics computer models have taken the acoustical design of new auditoria and modifications to existing halls a large step forward. The models usually are based on geometrical acoustics principles and trace rays (or some similar process) and add up calculated reflections on an energy basis. They often include ‘fudge’ factors to account for the effects of non-specular reflections and the scattering of sound. These have significant effects on the final results but the choice of appropriate values is greatly influenced by the experience of the user. Being based on geometrical acoustics they also do not include interference effects such as are important for understanding the low frequency seat dip attenuation. Many of the models can be used to create auralizations of sound in a planned hall, but of course although these can sound very impressive they are limited by the basic limitations of the models. These models are particularly useful for determining the basic geometry of a new hall and are, in most cases, a much more convenient replacement for scale models.

Although there are many issues that are not fully understood, there has been progress and there are better techniques for assessing conditions in existing rooms intended for musical performances. The ISO 3382 standard [21] describes a number of room acoustics quantities and measurement procedures that can be used to assess conditions in auditoria and concert halls. These are increasingly being used to document the acoustical conditions in halls and to help understand and resolve problems that are reported. All new halls and major renovations should be objectively evaluated using these measurements to ensure all parties know exactly what has been achieved. There is a long history of small but often expensive modifications to halls being recognized as ineffective after initial opinions to the contrary.

Figures 17 and 18 help to illustrate the use of measurements to understand conditions by comparing results from two very different halls. Both figures include plots of early-arriving and late-arriving strength values in the 1000 Hz octave versus source-to-receiver distance. Fig. 17 shows measurements from Boston Symphony Hall, a hall that is famous for its acoustical quality. Fig. 18 shows

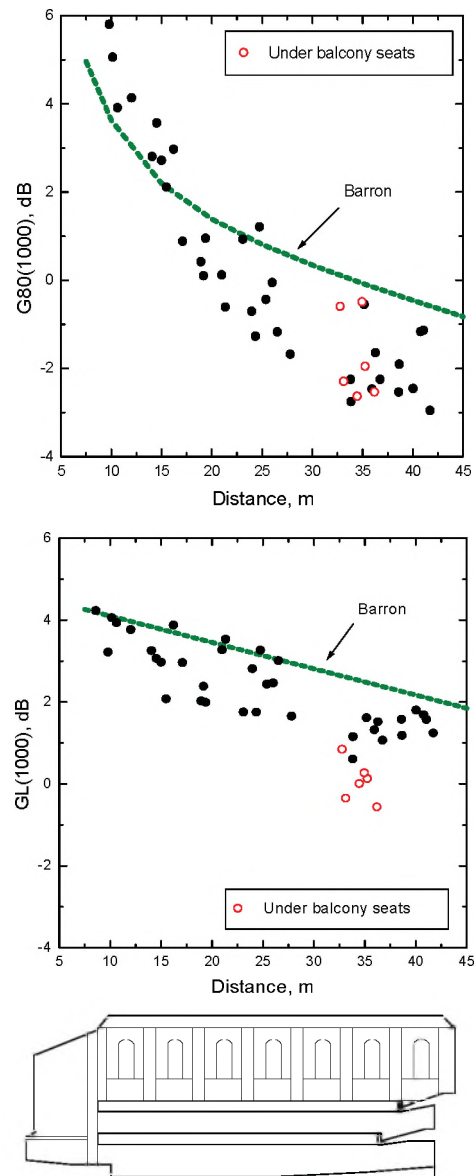


Fig. 17. Early-arriving 1000 Hz strength (upper) and later-arriving 1000 Hz strength (middle) versus source receive distance in Boston Symphony Hall. Dashed line shows values predicted using Barron’s theory [37].

results measured in the Orpheum in Vancouver before it was renovated a few years ago [36].

The upper part of Fig. 17. shows that early-arriving sound levels decrease in a regular manner with increasing distance and that seats located under the relatively small balconies have early-arriving sound levels similar to seats at the same distance in the balconies. The middle part of Fig. 17 shows a similar decrease in late-arriving sound levels with distance but the seats under the balcony do tend to have lower values than those in the balcony at the same

distance. This would suggest that in the under-balcony seats, it might sound a little less reverberant and a little less enveloping. For both early-arriving and later-arriving sounds the decrease in level with distance is a little more rapid than predicted by Barron's theory.

The results in Fig. 18 are quite different. The variation of early-arriving sound levels with distance show a large amount of scatter and especially for the seats in the balcony. This was determined to be due to various concave surfaces in the ceiling focusing early reflection energy to particular areas. However, there is again no evidence that under-balcony seats have systematically different early-

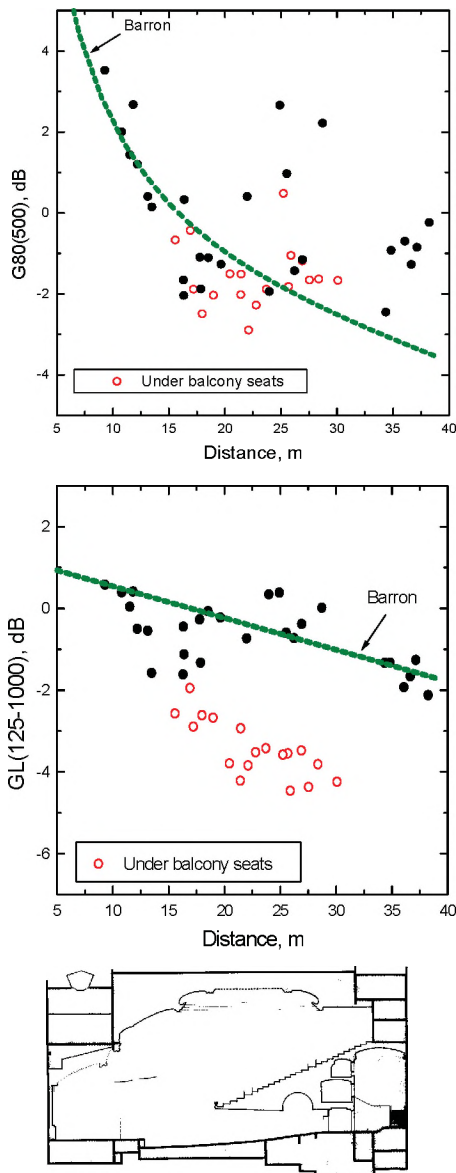


Fig. 18. Early-arriving 1000 Hz strength (upper) and later-arriving 1000 Hz strength (middle) versus source receive distance in the Orpheum in Vancouver. Dashed line shows values predicted using Barron's theory[37].

arriving levels than seats in the balcony at the same distance. The later-arriving sound levels show reasonably good agreement with Barron's theory for seats in the balcony but much lower levels at seats under the large balcony. The measurement results clearly identify the problems and helped the acoustical consultant (John O'Keefe, Aercoustics) to successfully renovate the hall [36].

Conclusions

The development of architectural acoustics has many occasions of discovery and later-rediscovery, but progress is being made. There are many unresolved problems and with adequate resources substantial progress could be readily achieved.

References

- [1] Henry, J., "On Acoustics Applied to Public Buildings", Scientific Writings of Joseph Henry, Smithsonian Institution, Washington, Part II, 403-421 (1856).
- [2] Haas, H., "Th Influence of a Single Echo on the Audibility of Speech", Library Commun. No. 363, Dept. Sci. and Indust. Res., Bldg. Res. Sta., Watford, England (1949).
- [3] Haas, H., "Über den Einfluss eines Einfachechos auf die Hörsamkeit von Sprache", *Acustica* 1, 49 (1951).
- [4] Haas, H., "The Influence of a Single Echo on the Audibility of Speech", *J. Aud. Eng. Soc.* 20 (2) 146-159 (1972).
- [5] Lochner, J.P.A. and Burger, J.F., "The Influence of Reflections on Auditorium Acoustics", *J. Sound Vibr.* 1 (4) 426-454 (1964).
- [6] Wallach, H., Newman, E.B. and Rosenzweig, M.R., "The Precedence Effect in Sound Localization", *J. Am. J. Psychol.* 52, 315-336 (1949).
- [7] Fay, R.D. and Hall, W.M., "Historical Notes on the Haas Effect", *J. Acoust. Soc. Am.* 28, 131-132 (1956).
- [8] Gardener, M.B. "Historical Background of the Haas and/or Precedence Effect", *J. Acoust. Soc. Am.* 43 (6) 1243-1248 (1968).
- [9] H.P. Seraphim, "Über die Wahrnehmbarkeit mehrerer Ruckwürfe von Sprachschall", *Acustica* 11, 80-91 (1961).
- [10] Bradley, J.S., Sato, H. and Picard, M., "On the Importance of Early Reflections for Speech in Rooms", *J. Acoust. Soc. Am.* 113 (6) 3233-3244 (2003).
- [11] F. Aigner and M.J.O. Strutt, "On the Physiological Effect of Several Sources of Sound on the Ear and its Consequences in Architectural Acoustics", *J. Acoust. Soc. Am.*, 6, 155-159 (1935).
- [12] Thiele, R., "Richtungsverteilung und Zeitfolge der Schallruckwürfe in Raumen", *Acustica*, 3, 291-302 (1953).
- [13] Schultz, T.J., "Acoustics of the Concert Hall", *IEEE Spectrum*, 56-67 (1965).

- [14] W. Reichardt, O. Abdel Alim, and W. Schmidt, "Definition and Basis of Making an Objective Evaluation to Distinguish Between Useful and Non-useful Clarity Defining Musical Performances", *Acustica*, 32, 126-137 (1975).
- [15] Kürer, R., "Zur Gewinnung von Einzahlkriterien bei Impulsmessungen in der Raumakustik", *Acustica* 21, (6) 370-372 (1969).
- [16] Bradley J.S., Predictors of Speech Intelligibility in Rooms, *J. Acoust. Soc. Am.*, 80 (3) 837-845 (1986).
- [17] Bradley J.S. "Speech Intelligibility Studies in Classrooms", *J. Acoust. Soc. Am.*, 80 (3) 846-854 (1986).
- [18] T. Houtgast and J.M. Steeneken, "The Modulation Transfer Function in Room Acoustics as a Predictor of Speech Intelligibility", *Acustica*, 28, 66-73 (1973).
- [19] IEC 60268-16, Ed. 3, 2003, "Sound System Equipment – Part 16, Objective Rating of Speech Intelligibility by Speech Transmission Index", International Electrotechnical Commission, Geneva.
- [20] Bradley, J.S., "Relationships Among Measures of Speech Intelligibility in Rooms", *J. Aud. Eng. Soc.*, 46 (5) 396-405 (1998).
- [21] ISO3382, "Acoustics—Measurement of the Reverberation Time of Rooms with Reference to Other Acoustical Parameters," International Organisation for Standardisation, Geneva, Switzerland (1998).
- [22] Bradley, J.S., "Using ISO3382 Measures, and Their Extensions, to Evaluate Acoustical Conditions in Concert Halls", *Acoustical Science and Technology*, 26 (2) 170-178 (2005).
- [23] Bradley, J.S. and Sato, H., "The Intelligibility of Speech in Elementary School Classrooms", *J. Acoust. Soc. Am.*, 123 (4) 2078-2086 (2008).
- [24] Lane H., Tranel B., "The Lombard Sign and the Role of Hearing in Speech", *J. Speech Hear. Res.* 14, 677-709 (1971).
- [25] Yang, W. and Bradley, J.S., "Effects of Room Acoustics on the Intelligibility of Speech in Classrooms for Young Children", *J. Acoust. Soc. Am.* 125 (2) 923-933, 2009.
- [26] Bradley, J.S., Soulodre, G.A., and Norcross, S., "Factors Influencing the Perception of Bass", *J. Acoust. Soc. Am.*, Vol. 101, No. 5, Pt. 2, p.3135, (1997).
- [27] T.J. Schultz and B.G. Watters, "Perception of Music Heard via Interfering Paths", *J. Acoust. Soc. Am.* 36 (5) 897-902 (1964).
- [28] G.M. Sessler and J.E. West, "Sound Transmission Over Theater Seats", *J. Acoust. Soc. Am.* 36 1725-1732 (1964).
- [29] Bradley J.S., "Some Further Investigations of the Seat Dip Effect", *J. Acoust. Soc. Am.* 90, 324-333 (1991).
- [30] Bradley, J.S. "The Sound Field for Listeners in Concert Halls and Auditoria", *Computational Architectural Acoustics*, Editor J.J. Sendra WIT, Press, UK. (1999).
- [31] Barron, M., "The Subjective Effects of First Reflections in Concert Halls: The Need for Lateral Reflections", *J. Sound Vibr.*, 15, 475-494 (1971).
- [32] Barron, M. and Marshall, A.H., "Spatial Impression due to Early Lateral Reflections in Concert Halls: The Derivation of a Physical Measure", *J. Sound Vibr.*, 77, 211-232 (1981).
- [33] Bradley, J.S., and Soulodre, G.A., "The Influence of Late Arriving Energy on Spatial Impression", *J. Acoust. Soc. Am.*, Vol. 97, No. 4, pp. 2263-2271, April, 1995.
- [34] Bradley, J.S. and Soulodre, G.A., "Objective Measures of Listener Envelopment", *J. Acoust. Soc. Am.*, Vol. 98, No. 5, pp. 2590-2597, (1995).
- [35] Bradley, J.S., Comparison of Concert Hall Measurements of Spatial Impression, *J. Acoust. Soc. Am.*, Vol. 96, No. 6, pp. 3525-3535, (1994).
- [36] O'Keefe, J. and Bradley, J.S., "Acoustical Renovations of The Orpheum, Vancouver: I. Measurements Prior to Renovations", *Canadian Acoustics*, Vol. 28, No. 1. 21-33, (2000).
- [37] Barron, M., and Lee, L.J., "Energy Relations in Concert Auditoriums. I", *ASA*, 84 (2). 618-628 (1988).
- [37] Barron, M., and Lee, L.J., "Energy Relations in Concert Auditoriums. I", *ASA*, 84 (2). 618-628 (1988).

I HAVE NEVER SEEN A SOUND

R. Murray Schafer

Indian River, Ontario, Canada

Now I wish to speak of sounds.

The world is full of sounds.

I cannot speak of them all.

I shall speak of sounds that matter.

To speak of sounds, I make sounds.

I create - an original act which I performed the moment I emerged on this earth.

Creation is blind. Creation is soundful.

“In the beginning God created the heaven and earth” – with his mouth.

God named the universe, thinking aloud.

The Egyptian gods came into being when Atum, the creator, named them.

Mithra came into being out of vowels and consonants.

The terrible gods came into being out of thunder.

The fruitful gods came into being out of water.

The magic gods came into being out of laughter.

The mystic gods came into being out of distant echoes.

All creation is original. Every sound is new.

No sound can be repeated exactly. Not even your own name. Every time it is pronounced it will be different. And a sound heard once is not the same as a sound heard twice, nor is a sound heard before the same as a sound heard after.

Every sound commits suicide and never returns. Musicians know this. No musical phrase can be repeated exactly the same way twice.

Sounds cannot be known the way sights can be known, seeing is analytical and reflective. It places things side by side and compares them (scenes, slides, diagrams, figures...). This is why Aristotle preferred sight as ‘the principal source of knowledge.’

Sights are knowable. Sights are nouns.

Sounding is active and generative. Sounds are verbs. Like all creation, sound is incomparable. Thus there can be no science of sound, only sensations...intuitions...mysteries....

In the Western world, and for some time, the eye has been the referent for all sensory experience. Visual metaphors and scaling systems have dominated. Interesting fictions have been invented for weighing or measuring sounds: alphabets, music scripts, sonograms. But everybody knows you can't weigh a whisper or count the voices in a choir or measure a child's laughter.

It is probably going too far to say that in aural culture, science, especially physics and mathematics and their dependants (statistics, physiology, empirical psychology, drafting, demography, banking, etc. – the list is long) would disappear. It is probably enough to say that in purely aural cultures they don't appear.

Have I got off the track?

I was saying that everything in the world was created by sound and analysed by vision. God spoke first and saw that it was good second.

What happens if it isn't good? Then God destroys with sound. Noise kills. War. The Flood. The Apocalypse.

Noise cancels. It turns language into a polyglot; the case at Babel. When the noise of the world became so great that it disturbed “even the inner parts of the gods,” they released the Flood (Epic of Gilgamesh).

Some say the noise of the apocalypse will be of ear-splitting intensity (Mohammed in the Qur'an or John of Patmos in Revelation). Others maintain “the world will not end with a bang but a whimper.” In any case, it will sound, because all traumatic events maintain sound as their expressive medium: war, violence, love, madness. Disease alone is silent and yields to analysis.

Come with me now and sit in the grandstand of life. The seats are free and the entertainment is continuous.

The world orchestra is always playing; we hear it inside and outside, from near and far.

We have no earlids.

We are condemned to listen.

I hear with my little ear...

Most of the sounds I hear are attached to things. I use sounds as clues to identify these things. When they are hidden, sounds will reveal them. I hear through the forest, around the corner, over the hill.

Sound gets to places where sight cannot.

Sound plunges below the surface.

Sound penetrates to the heart of things.

When I disregard the things to which sounds are attached, the phenomenal world disappears. I become blind. I am swept away sensuously by the vast music of the universe.

Everything in this world has its sound – even silent objects. We get to know silent objects by striking them. The ice is thin, the box is empty, the wall is hollow.

Here is a paradox: two things touch but only one sound is produced. A ball hits a wall, a drumstick strikes a drum, a bow scrapes a string. Two objects - one sound.

Another case of 1 plus 1 equals 1.

Nor is it possible to join sounds without them changing character. Zeno's paradox: "If a bushel of corn turned out upon the floor makes a noise, each grain and each part of each grain must make a noise likewise, but, in fact, it is not so."

In acoustics, sums equal differences.

Sounds tell me about spaces, whether small or large, narrow or broad, indoor or outdoor. Echoes and reverberation inform me about surfaces and obstructions. With practice I can begin to hear "acoustic shadows," just like the blind.

Auditory space is very different from visual space. We are always at the edge of visual space, looking in with the eye. But we are always at the centre of auditory space, listening out with the ear.

Thus, visual awareness is not the same as aural awareness. Visual awareness faces forward. Aural awareness is centred.

I am always at the heart of the sounding universe.

With its many tongues it speaks to me.

With the tongues of gods it speaks to me.

You cannot control or shape the acoustic universe. Rather the reverse. This is why aural societies are considered unprogres-

sive; they don't see straight ahead.

If I wish to order the world I must become "visionary." Then I close my ears and create fences, property lines, straight roads, walls. All the major themes of science and mathematics as developed in the Western world are silent (the space-time continuum of relativity, the atomic structure of matter, the wave-corpuscular theory of light) and the instruments developed for their study, the telescope and the microscope, the equation the graph and above all, number, are silent likewise.

Statistics deals with a world of quantities that is presumed to be silent.

Philosophy deals with a phenomenal world that is presumed to be silent.

Economics deals with a material world that is assumed to be silent.

Even religion deals with a God who has become silent.

Western music is also conceived out of silence. For two thousand years it has been maturing behind walls.

Walls drove a wedge between music and the soundscape. The two fell apart and became independent.

Music within; pandemonium (i.e. devilry) without.

But everything that is ignored returns. The vehement obscurity of the soundscape pushes back to confront us as noise pollution.

As an articulated problem, noise belongs exclusively to Western societies. It is the discord between visual and acoustic space. Acoustic space remains askew because it can't be owned. It becomes disenfranchised – a sonic sewer. Today we view the world without listening to it, from behind glassed-in buildings.

In an aural society all sounds matter, even when they are only casually overheard.

In Ontario, the signal to stop tapping the maple trees is when the spring frogs are heard; then the ice has melted, the sap is darker, the syrup inferior.

Another example: a man walks across the snow. You know the temperature from the sound of his footsteps. This is a different way of perceiving the environment; one in which the sensorium is undivided; one which recognizes that all information tracks are interconnected.

Some sounds are so unique that once heard they will never be forgotten: a wolf's howl a loon's call, a steam locomotive,

a machine gun.

In an aural society sounds like this can be brought forward and mimicked in song and speech as easily as a visual society can draw a picture or a map.

The visual society is always amazed at the aural retentiveness of people who have not yet passed through the visual phase. The Qur'an the Kalevala and the Iliad were once memorized.

Remember that.

Visual man has instruments to help him retain visual memories (paintings, books, photographs). What is the device for retaining aural memories?

Repetition.

Repetition is the memory medium for sound.

Repetition is the means by which sounds are retained and explained.

Repetition is the means by which the history of the world is affirmed.

Repetition never analyses; it merely insists.

Repetition makes the listener participate in the statement not by comprehending it but by knowing it.

"It is written, but I say unto you...." And I will say it again and again and again, because Hearing is Believing.

As the grip of the visual-analytical world weakens and is replaced by intuition and sensation, we will begin to discover again the true tuning of the world and the exquisite harmony of all its voices.

We will find the centre.

Then the whole body will become an ear and all sounds will come to you, the known and the unknown, the sweet, the sad and the urgent.

When my body lies white and blue in bed at night, then all sounds come to me of their own accord, unhurried, strangely blended, the light-toned and the slow grinding of mountains. Then hearing is most alert...and there is singing before me... as I pass beyond "to the land that loveth silence."

TAPPING just got easier!

The rugged brand new Norsonic N-277 Tapping Machine is ideal for making structureborne impact noise tests for floor/ceiling combination in the field and in the laboratory. This third-generation unit meets all international and US standards.

- Impact sound transmission testing according to ISO140 part VI, VII and VIII, ASTM E-492 and ASTM E-1007.
- Remote operation from hand switch or PC; Mains or battery operation.
- Low weight 10 kg (22 lb) incl. battery and wireless remote option.
- Built in self check of hammer fall speed, and tapping sequence for automatic calibration of major components.
- Retractable feet and compact size provide easy transportation and storage.

Scantek, Inc.
Sound & Vibration Instrumentation
and Engineering

www.scantekinc.com
info@scantekinc.com
800-224-3813



BIOMEDICAL ULTRASOUND IMAGING FROM 1 TO 1000 MHz

MICHAEL KOLIOS

Department of Physics

Ryerson University, Toronto, ON, Canada

INTRODUCTION

Ultrasound refers to mechanical waves whose vibrations are above the upper limit of the human audible range (above 20 kHz). Upon recognition that ultrasound can be used in a manner similar to radar in the 1950s (in a pulse echo mode), development of ultrasound as an imaging modality followed^{1,2}. In the several decades following the introduction of ultrasound, it has become one of the most popular forms of imaging of the human anatomy in the world³. Close to 25% of all imaging procedures are based on ultrasound. The main reason is that it uses non-ionizing radiation to produce images of very good spatial resolution (in the millimeter range), is portable, relatively inexpensive and it can be used to measure function (e.g. blood flow). Most people are familiar with ultrasound imaging of the fetus, which provides detail about fetal anatomy and function. When the fetus is imaged the frequency range used is between 1 to 10 MHz, and the corresponding wavelengths are sub-millimeter. Ultrasound however is also known to also be operator dependent (in other words the operator of the ultrasound instrument can have a great effect on the quality of the image) and has, in comparison with other imaging modalities, relatively poor inherent soft tissue contrast. Therefore, many of the developments in the field of ultrasound imaging are focused on technological improvements to improve the signal to noise of the instrumentation or in the use of new techniques to increase the soft tissue contrast, such as microbubble contrast agents⁴⁻⁶, elastography^{7, 8}, applied radiation force imaging (ARFI)^{9,10} and shear wave imaging^{11,12}. Moreover, attempts are made to make these methods quantitative, so as to avoid the operator dependence. One way to achieve this is by performing a quantitative analysis of the backscatter data prior to image formation. In this paper our attempts to study ultrasonic scattering at different length scales is presented. To achieve this, imaging at frequencies ranging from 1 MHz to 1000 MHz is performed.

ULTRASOUND IMAGING

Contrast in an ultrasound image is typically based on the strength of the backscattered ultrasound. Soft tissues adjacent to bone or air cavities strongly scatter ultrasound due to the large change in mechanical properties encountered by the pressure wave. In order to achieve good image quality and resolve small objects (such as tumors), focused transducers are used to achieve good spatial resolution in the lateral direction of the Imaging device. This focusing could be geometric (a focused transducer) or electronic (a phased array)². The lateral resolution of an ultrasound imaging device depends on the frequency used

and the geometrical characteristics of the focusing and is given by¹³:

$$R_{lateral} = \bar{\lambda} f_{number}$$

where $\bar{\lambda}$ is the average wavelength of the ultrasound (since usually a pulse is used for imaging) and f_{number} is the ratio of the focal length of the of the transducer to its diameter. While the equation for the lateral resolution shares many similarities with similar concepts in optics, the axial resolution is dictated by the ultrasound pulse length in time. Wide-band transducers (i.e. spanning a large frequency range or equivalently short pulses) are used to achieve greater axial resolution. The equation that describes the axial resolution of the transducer is given by:

$$R_{axial} = \frac{1}{2} \frac{c}{B}$$

where c is the speed of sound in the propagating medium and B is the bandwidth of the transducer. Increasing the transducer bandwidth increases the ultrasound spatial resolution. Therefore, higher frequencies are typically used to image smaller structures of interest to provide better axial and lateral resolution. The tradeoff, as in optics, is that the depth of field of the focused transducer decreases with greater geometrical focusing. This results in large variations in the intensity of the backscattered ultrasound in images (which is greater in the focal zone) that are a result of this focusing. The DOF can be approximated by¹³:

$$DOF = 7.0 \bar{\lambda} (f_{number})^2$$

The advent of high frequency ultrasound instrumentation has made ultrasound popular beyond the clinical regime. A field in which ultrasound has grown in the last decade is that of small animal imaging (often referred to as pre-clinical imaging). For this application, typically 20-60 MHz ultrasound is used to image structures on the order of tens to hundred of microns. Typical spatial resolutions achieved are in the order of 200-30 μm lateral resolutions and 60 – 30 μm axial. This field is also termed ultrasound biomicroscopy and a comprehensive review article has been written by Foster *et al.*¹³ Exquisite images of the mouse fetus have been acquired using such ultrasound imagers, clearly outlining the mouse embryo anatomy (as well as function through measurements of flow). A Canadian company based in Toronto Ontario (VisualSonics Inc. www.visualsonics.com) is currently the market pioneer and leader in this field. The increase in spatial resolution however comes at a price: that of penetration depth. For

frequencies between 20-60MHz, the imaging penetration depth is of the order of a few centimeters. At 60MHz and for a system with a dynamic range of 80dB, a tissue penetration depth of up to 6mm is expected.

In the acoustic microscopy regime the frequencies used are beyond this range (> 100 MHz). The penetration depth is limited to millimeters (100 MHz) or microns (at GHz frequencies) and therefore the applications are limited to surface measurements (such as probing tooth enamel)¹⁴, the examination of thinly prepared specimens^{15, 16} or the study of single cells^{17, 18}. The same imaging considerations as stated above apply for acoustic microscopy. In this frequency range, axial and lateral resolutions of 1 μm and 3 μm can be achieved, approaching the limits of optical microscopy. One of the main advantages of acoustic compared to optical microscopy is that cells do not need to be stained and depth information can be obtained which cannot be obtained with conventional optical techniques.

ULTRASOUND SCATTERING AND TISSUE CHARACTERIZATION

While the increase in spatial resolution with frequency is straight forward and well understood, more subtle are the effects of the physics of scattering at higher frequencies, especially as the wavelength of the ultrasound approaches the size of cells, one of the most basic organizational units of tissues. Our laboratory is interested in scattering from tumors which typically have a high concentration of cells. A good understanding of the physics of scattering is required to better exploit the contrast mechanisms that ultrasound imaging can provide.

A difference in the acoustic properties of the target tissue (or the tissue structures that scatters the sound) and the acoustic properties of the background medium create the scattering of the propagating wave - which is the basis of the ultrasound signal detected (typically the backscatter). The ability to differentiate tissues based on their scattering is related to the contrast resolution of the imaging system and is an important parameter when considering various applications^{2, 19}. Large contrast is achieved for example between the bone and soft tissues, as well as soft tissues and fluid filled cavities (such as cysts). When the wavelength of the ultrasound is much smaller than the structure that scatters the ultrasound, then the scattering strength can be found by using the well known equations of reflection and refraction. Assuming scattering from a plane surface between two fluid-like media with the incident wave normal to the surface, the reflected power in dB can be found:

$$dB = 10 \log_{10} \frac{(Z_1 - Z_2)^2}{(Z_1 + Z_2)^2}$$

where Z_1 is the characteristic impedance of the first medium and Z_2 is the characteristic impedance of the second medium.

When the wavelength is of the order of the scatterer size, then the full equations of a traveling pressure wave incident on a structure have to be solved²⁰. For high frequency ultrasound applications the ultrasound wavelengths are 20-80 μm and therefore approach the size of the typical mammalian cell. The cell shape can be approximated as spherical. The mathematical solution to a plane wave incident on a sphere exists and is given as an infinite summation. The solution is slightly simplified for fluid spheres²¹ (that do not support shear waves) compared to elastic spheres²² (that do support shear waves). Approximations however exist when the scattering structure is much smaller than the wavelength (Rayleigh scattering) or the scattering is weak (the Born approximation, an approximation that is used frequently in biomedical ultrasound)^{23, 24}. These solutions typically provide the scattered pressure field as a function of the frequency of the monochromatic incident wave. An example of the solution of a pressure wave incident on a polystyrene microsphere is shown in Figure 1. This solution, even though for an elastic sphere (and therefore not fluid-like as most tissues), illustrates many important concepts in understanding scattering from biological tissues. An important parameter is the product of the wavenumber k and the scatterer size a . For low values of ka , the scattering power rapidly increases with the size of the object (or equivalently increasing frequency). This is the region of Rayleigh scattering. A transition occurs slightly above $ka = 1$, and the backscattering power does not increase substantially after this. As the frequency of the ultrasound increases above this (or equivalently the diameter of the object increases), one can identify patterns that correspond to surface modes of the sphere²⁰. Resonant scattering theory, which can solve more complex cases than a plane wave incident on a homogenous isotropic sphere, has been developed for target identification in acoustic scattering problems of greater complexity²⁵.

We have been able to accurately reproduce these predicted values by using high-frequency ultrasound²⁶. An example is shown in figure 2. In the experiment presented a 43 μm polystyrene microsphere was suspended from a micropipette. Confocally aligned with an optical microscope was a commercial high-frequency ultrasound imager (the VEVO 770 from VisualSonics Inc., Toronto, Ontario). Scattering from the bead was recorded with a 55 MHz transducer. The red trace in the figure (left axis) corresponds to a RF (radiofrequency) line from the scattering of the bead. The blue line (right axis) corresponds to the power spectrum derived from the RF line. The complex resonances are evident both from the image and from the frequency response. Using acoustic microscopy we have recently been able to reproduce these scattering patterns using smaller diameter beads at frequencies between 100 and 1,000 MHz. For these experiments, the SASAM 1000 (Kibero GmbH, Saarbrücken, Germany) was used^{27, 28}. Figure 3(A) shows C-scan and B-scan images that were generated using 200 MHz and 400 MHz ultrasound transducers from the backscatter of the polystyrene microspheres of sizes ranging

from 2 to 10 μm . As would be expected, the C-scan image generated increases in size with the increased bead diameter. The B-scan images also show the complex resonant patterns which were also seen at the 20 to 60 MHz frequency range (since the ka values are similar). Figure 3(B) shows the power spectra calculated from these backscatter patterns for different combinations of bead diameter and ultrasound frequency. Reasonable agreement is seen in our first attempts when comparing experimental data with theoretical predictions in this frequency range.

The solutions indicate that the scattered pressure variations are sensitive to the size of the scattering structures (as well as their physical attributes). This suggests the possibility of scatterer size identification in biological tissues by the examination of the frequency dependence of the scattering patterns. This has been studied extensively in the past for the clinical ultrasound frequency range²⁹⁻³¹ and is now being extended to higher frequency applications^{32, 33}. To examine these patterns, one must have access to the raw radiofrequency data that were used to produce the ultrasound images (an example is the red trace of figure 2). These data are typically stored as voltage as a function of time (and related to space through the speed of sound) that can then be examined as to their frequency content by their Fourier transform. Our group has extensively studied the frequency dependence of scattering as a means to identify scattering structures with an emphasis on monitoring the effects of cancer treatment and how cell scattering changes when the cells respond to treatment^{26, 34-36}. It is hypothesized that since large changes in cellular structure and composition occur during cell death, the backscatter patterns would reflect these changes.

One example of how the contrast resolution changes as the frequency of the interrogating ultrasound pulse changes is that of blood scattering. In the low frequency ultrasound imaging range, blood vessels appear as black cylinders of low echogenicity as the scattering strength from the red blood cells ($ka \ll 1$) is very low compared to the tissue surrounding the blood vessels. However, at high frequencies ($ka \approx 1$) the scattering from the red blood cells is comparable to that of the surrounding tissue; in some cases the signal from the blood is stronger than that of the surrounding tissues. Another example is shown in figure 4. In this example, two beads are imaged with a 200 MHz transducer, one of 5 μm diameter ($ka \approx 2$) and another one of 2 μm diameter ($ka \approx 0.8$). Both beads are within the field of view of the transducer scan area. Even though the beads are of comparable sizes and have identical acoustic properties, only the 5 μm bead can be visualized with the 200 MHz transducer. The 2 μm bead cannot be detected due to its reduced scattering (roughly 25dB less, figure 1). It is hypothesized that when biological cells are imaged, an increase in the imaging frequency will introduce similar effects. When the combination of structure size and frequency of interrogation approaches $ka \approx 1$, this will result in greater contributions from those scattering structures to the overall backscatter of the biological sample.

The examples mentioned in this paragraph illustrate this point. Therefore, increasing the ultrasound frequency not only improves image resolution but changes the population of scattering structures that contribute to the backscattered signal and therefore form the final image.

Due to the wave nature of ultrasound, another feature of the images produced by ultrasound is speckle. Several excellent reviews of the nature of speckle in imaging are available³⁷⁻³⁹. Speckle arises when there are many unresolved scatterers of similar scattering strength per resolution volume of the imaging device. In the context of ultrasonic imaging, speckle is often referred to as "speckle noise". This is because speckle can interfere with the delineation of boundaries between two media due to its "grainy" appearance. However, speckle is deterministic and for a given scattering structure the speckle pattern will remain the same. Therefore, speckle is not noise in the traditional sense. In fact, for the high frequency imaging of cells in which individual cells cannot be resolved, the entire content of the image is speckle⁴⁰. This is seen in figure 5. Moreover, when one examines the scattering from tumors in mice, a similar speckle pattern is detected^{41, 42}. Therefore the speckle pattern can relay information that is related to the scattering structures interrogated. In figure 5 a compact aggregate of acute myeloid leukemia (AML) cells is imaged. These cells have been centrifuged so as to form a tightly packed aggregate that emulates certain types of tumors in which the cell number density is very high. This aggregate was imaged at multiple frequencies: 6, 20, 40 and 200 MHz using devices that permit access to the RF data. The images generated, irrespective of the frequency used, produce a speckle pattern. This is another illustration of how the nature of the scattering structures changes as the frequency increases: the presence of speckle indicates that there are several unresolved scatterers in the resolution volume of each of the transducers used. The data presented in figure 5 suggest that at the lower frequencies, due to the large ultrasound wavelengths compared to the size of the cells (the cells are 10 μm in diameter, and the ultrasound wavelength is approximately 250 μm at 6MHz) the cell aggregate is more like a homogeneous medium and an impedance mismatch that occurs at the surface causes the specular-like reflection. However, at 40MHz, the wavelength is 37 μm which is approaching the physical size of the cell. Therefore, the surface reflection is more like diffuse scattering, and the scattering is similar in strength to that seen for further depths within the cell aggregate. This also illustrates why high frequency ultrasound is required to provide adequate sensitivity to distinguish features on the order of the cell size: it is clear that at 6MHz predominantly interfaces that enclose these cells can be distinguished rather than the signal from the cells.

CONCLUDING REMARKS

Ultrasound imaging is one of the most used imaging methods worldwide accounting for almost a quarter of all imaging procedures worldwide. In this presentation we show how the physics of scattering can influence the images

that are generated when different frequencies are used to image biological samples. Increasing the frequency of the interrogating ultrasound does not only improve spatial resolution, as would be expected from first principles, but it also changes the nature of the dominant ultrasonic scattering structures: as $ka \approx 1$ then the new structures that meet this criterion start to contribute more to the backscattered ultrasound signal detected by the transducers. Several examples presented here illustrate this principle. Our laboratory is equipped with instrumentation that allows ultrasonic interrogation with frequencies starting from 1 MHz to 1000 MHz which allows the experimental investigation of these relationships. As novel forms of imaging are developed that deal with tracers and their interactions with predominantly cellular targets (the field of molecular medicine) it is anticipated that a solid understanding of these interactions is very important for the interpretation of the ultrasound images.

ACKNOWLEDGMENTS

This research was undertaken, in part, thanks to funding from the Canada Research Chairs Program, the Canada Foundation for Innovation, the Ontario Ministry of Research and Innovation, Ryerson University and the Canadian Institutes of Health Research. The author would like to acknowledge Patrick Kennedy, Omar Falou, Eric Strohm, Dr. Min Rui, Ralph Baddour and Arthur Worthington for much of the work presented in these studies. The author would like to acknowledge the scientific input of Drs. Gregory J. Czarnota, Eric Weiss and J. Carl Kumaradas.

REFERENCES

1. R. S. C. Cobbold, *Foundations of biomedical ultrasound* (Oxford University Press, Toronto, 2007).
2. T. L. Szabo, *Diagnostic ultrasound imaging: inside out* (Elsevier/Academic Press, Boston, 2004).
3. F. Forsberg, "Ultrasonic biomedical technology: marketing versus clinical reality," *Ultrasonics* **42**(1), 17-27 (2003).
4. B. B. Goldberg, J. B. Liu, and F. Forsberg, "Ultrasound contrast agents: a review," *Ultrasound Med Biol* **20**(4), 319-333 (1994).
5. F. Forsberg, D. A. Merton, J. B. Liu, L. Needleman, and B. B. Goldberg, "Clinical applications of ultrasound contrast agents," *Ultrasonics* **36**(1-5), 695-701 (1998).
6. B. A. Kaufmann and J. R. Lindner, "Molecular imaging with targeted contrast ultrasound," *Curr Opin Biotechnol* **18**(1), 11-16 (2007).
7. J. Ophir, S. K. Alam, B. Garra, F. Kallel, E. Konofagou, T. Krouskop, and T. Varghese, "Elastography: ultrasonic estimation and imaging of the elastic properties of tissues," *Proc Inst Mech Eng H* **213**(3), 203-233 (1999).
8. B. S. Garra, "Imaging and estimation of tissue elasticity by ultrasound," *Ultrasound Q* **23**(4), 255-268 (2007).
9. J. J. Dahl, D. M. Dumont, J. D. Allen, E. M. Miller, and G. E. Trahey, "Acoustic radiation force impulse imaging for noninvasive characterization of carotid artery atherosclerotic plaques: a feasibility study," *Ultrasound Med Biol* **35**(5), 707-716 (2009).
10. L. Zhai, M. L. Palmeri, R. R. Bouchard, R. W. Nightingale, and K. R. Nightingale, "An integrated indenter-ARFI imaging system for tissue stiffness quantification," *Ultrason Imaging* **30**(2), 95-111 (2008).
11. S. Chen, M. W. Urban, C. Pislaru, R. Kinnick, Y. Zheng, A. Yao, and J. F. Greenleaf, "Shearwave dispersion ultrasound vibrometry (SDUV) for measuring tissue elasticity and viscosity," *IEEE Trans Ultrason Ferroelectr Freq Control* **56**(1), 55-62 (2009).
12. M. Muller, J. L. Gennisson, T. Deffieux, M. Tanter, and M. Fink, "Quantitative viscoelasticity mapping of human liver using supersonic shear imaging: preliminary in vivo feasibility study," *Ultrasound Med Biol* **35**(2), 219-229 (2009).
13. F. S. Foster, C. J. Pavlin, K. A. Harasiewicz, D. A. Christopher, and D. H. Turnbull, "Advances in ultrasound biomicroscopy," *Ultrasound Med Biol* **26**(1), 1-27 (2000).
14. S. D. Peck, J. M. Rowe, and G. A. Briggs, "Studies on sound and carious enamel with the quantitative acoustic microscope," *J Dent Res* **68**(2), 107-112 (1989).
15. K. Raum, "Microelastic imaging of bone," *IEEE Trans Ultrason Ferroelectr Freq Control* **55**(7), 1417-1431 (2008).
16. Y. Saijo, C. S. Jorgensen, P. Mondek, V. Sefranek, and W. Paaske, "Acoustic inhomogeneity of carotid arterial plaques determined by GHz frequency range acoustic microscopy," *Ultrasound Med Biol* **28**(7), 933-937 (2002).
17. J. Bereiter-Hahn, "Mechanics of crawling cells," *Med Eng Phys* **27**(9), 743-753 (2005).
18. G. A. Briggs, J. Wang, and R. Gundle, "Quantitative acoustic microscopy of individual living human cells," *J Microsc* **172**(Pt 1), 3-12 (1993).
19. J. F. Greenleaf and C. M. Sehgal, *Biologic System Evaluation with Ultrasound* (Springer-Verlag New York, Inc., 1992), p. 127.
20. O. Falou, J. C. Kumaradas, and M. C. Kolios, "Finite-element Modeling of Elastic Surface Modes and Scattering from Spherical Objects," in *Proceedings of the COMSOL Conference 2007*, (COMSOL, Boston, 2007).
21. V. C. Anderson, "Sound scattering from a fluid sphere," *Journal of the Acoustical Society of America* **22**, 426-431 (1950).
22. J. J. Faran, "Sound Scattering by Solid Cylinders and Spheres," *Journal of the Acoustical Society of America* **23**(4), 405-418 (1951).
23. P. M. Morse and K. U. Ingard, *Theoretical acoustics* (McGraw-Hill, New York, 1968).

24. K. S. Ratan and K. S. Subodh, "Validity of a modified Born approximation for a pulsed plane wave in acoustic scattering problems," *Physics in Medicine and Biology* **12**, 2823 (2005).
25. L. Flax, L. R. Dragonette, and H. Uberall, "Theory of elastic resonance excitation by sound scattering," *The Journal of the Acoustical Society of America* **63**(3), 723-731 (1978).
26. R. E. Baddour, M. D. Sherar, J. W. Hunt, G. J. Czarnota, and M. C. Kolios, "High-frequency ultrasound scattering from microspheres and single cells," *J Acoust Soc Am* **117**(2), 934-943 (2005).
27. S. Brand, E. C. Weiss, R. M. Lemor, and M. C. Kolios, "High Frequency Ultrasound Tissue Characterization and Acoustic Microscopy of Intracellular Changes," *Ultrasound Med Biol* **34**(9), 1396-1407 (2008).
28. R. M. Lemor, E. C. Weiss, G. Pilarczyk, and P. V. Zinin, "Mechanical properties of single cells - measurement possibilities using time-resolved scanning acoustic microscopy," presented at the IEEE International Ultrasonics Symposium, 2004.
29. F. L. Lizzi, M. Astor, E. J. Feleppa, M. Shao, and A. Kalisz, "Statistical framework for ultrasonic spectral parameter imaging," *Ultrasound Med Biol* **23**(9), 1371-1382 (1997).
30. F. L. Lizzi, S. K. Alam, S. Mikaelian, P. Lee, and E. J. Feleppa, "On the statistics of ultrasonic spectral parameters," *Ultrasound Med Biol* **32**(11), 1671-1685 (2006).
31. F. L. Lizzi, M. Greenebaum, E. J. Feleppa, M. Elbaum, and D. J. Coleman, "Theoretical framework for spectrum analysis in ultrasonic tissue characterization," *J Acoust Soc Am* **73**(4), 1366-1373 (1983).
32. M. L. Oelze and J. F. Zachary, "Examination of cancer in mouse models using high-frequency quantitative ultrasound," *Ultrasound Med Biol* **32**(11), 1639-1648 (2006).
33. M. L. Oelze, W. D. O'Brien, Jr., J. P. Blue, and J. F. Zachary, "Differentiation and characterization of rat mammary fibroadenomas and 4T1 mouse carcinomas using quantitative ultrasound imaging," *IEEE Trans Med Imaging* **23**(6), 764-771 (2004).
34. M. C. Kolios, G. J. Czarnota, M. Lee, J. W. Hunt, and M. D. Sherar, "Ultrasonic spectral parameter characterization of apoptosis," *Ultrasound Med Biol* **28**(5), 589-597 (2002).
35. O. Falou, R. E. Baddour, G. Nathanael, G. J. Czarnota, J. C. Kumaradas, and M. C. Kolios, "A study of high frequency ultrasound scattering from non-nucleated biological specimens," *J Acoust Soc Am* **124**(5), EL278-283 (2008).
36. O. Falou, J. C. Kumaradas, and M. C. Kolios, "Finite-element modelling of acoustic wave scattering from fluid, rigid and elastic objects," *Canadian Acoustics - Acoustique Canadienne* **33**(3), 84-85 (2005).
37. J. W. Goodman, "Some fundamental properties of speckle," *J. Opt. Soc. Am.* **66**(11), 1145-1150 (1976).
38. R. F. Wagner, M. F. Insana, and S. W. Smith, "Fundamental correlation lengths of coherent speckle in medical ultrasonic images," *IEEE Trans Ultrason Ferroelectr Freq Control* **35**(1), 34-44 (1988).
39. S. W. Smith, G. E. Trahey, S. M. Hubbard, and R. F. Wagner, "Properties of acoustical speckle in the presence of phase aberration. Part II: Correlation lengths," *Ultrasound Imaging* **10**(1), 29-51 (1988).
40. A. S. Tunis, G. J. Czarnota, A. Giles, M. D. Sherar, J. W. Hunt, and M. C. Kolios, "Monitoring Structural Changes in Cells with High Frequency Ultrasound Signal Statistics," *Ultrasound in Medicine & Biology* **31**(8), 1041-1049 (2005).
41. R. Vlad, S. Brand, A. Giles, M. C. Kolios, and G. J. Czarnota, "Quantitative Ultrasound Characterization of Responses to Radiotherapy in Cancer Mouse Models," *Clinical Cancer Research* **15**(6), in press (2009).
42. R. M. Vlad, N. M. Alajez, A. Giles, M. C. Kolios, and G. J. Czarnota, "Quantitative ultrasound characterization of cancer radiotherapy effects in vitro," *Int J Radiat Oncol Biol Phys* **72**(4), 1236-1243 (2008).

FIGURES

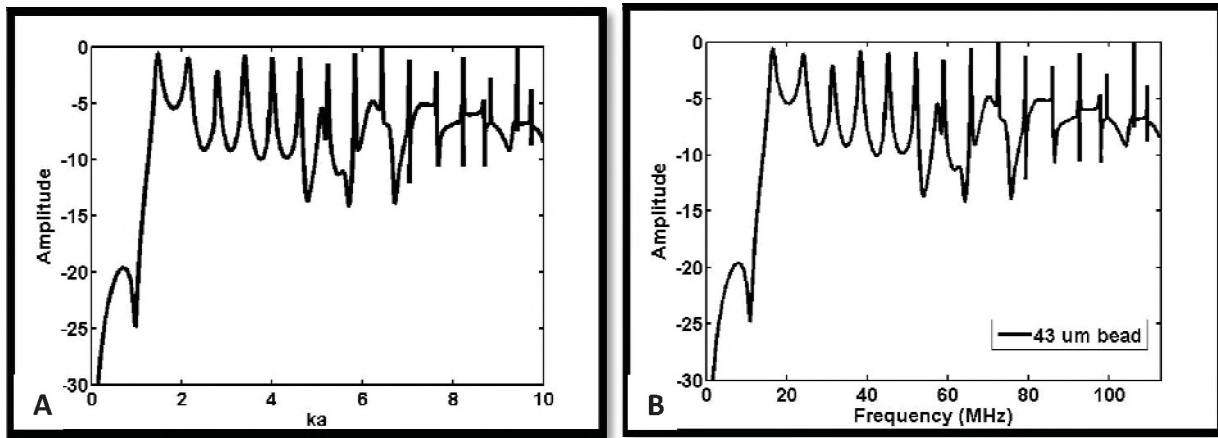


Figure 1: The Faran solution to scattering of a plane wave incident on an elastic, isotropic and homogeneous sphere. A) Normalized backscatter power (to the maximum value) is plotted as a function of the wave number times the sphere diameter (ka). B) Normalized backscatter power plotted as a function of ultrasound frequency for a given scatterer size ($a = 43 \mu\text{m}$) for the same parameters in (A).

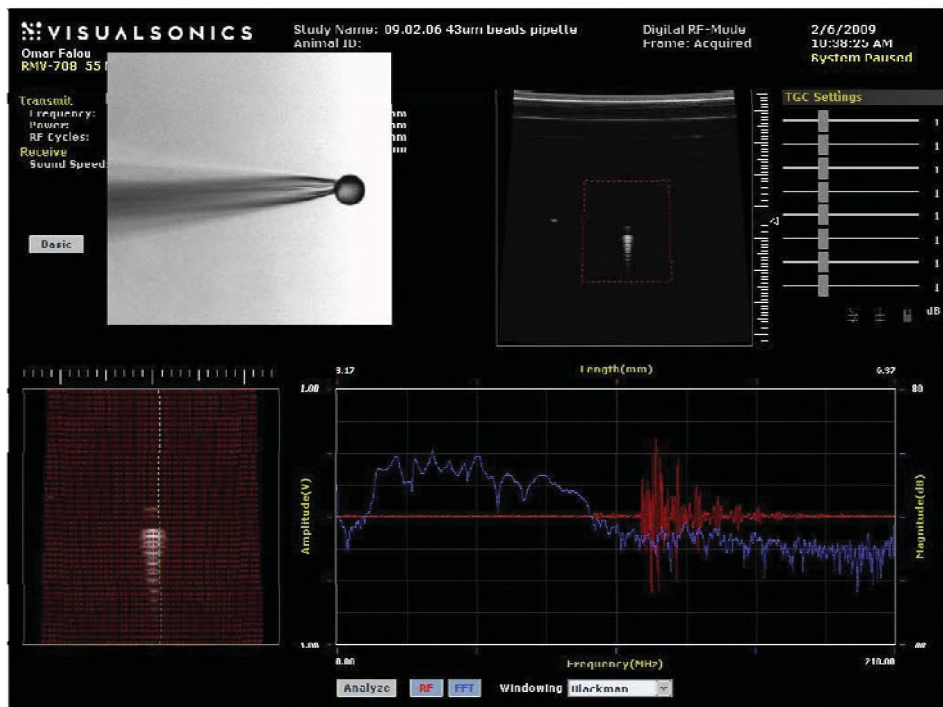


Figure 2: Single bead scattering measured using a high-frequency ultrasound system (the VEVO 770 using a 55MHz transducer in this experiment). In the top left corner of the image an optical microscopy image of a 43 μm polystyrene microsphere that is held in position by a micropipette and confocally aligned with the transducer). To the right of this picture the B-mode ultrasound image of the bead (seen closer up in the bottom left). The resonant pattern is clearly seen. In the bottom right of the image the red line represents a single radiofrequency trace of the scattered ultrasound and the blue line represents power spectrum derived from this a line. The resonant patterns of figure 1(B) can be seen in the power spectrum.

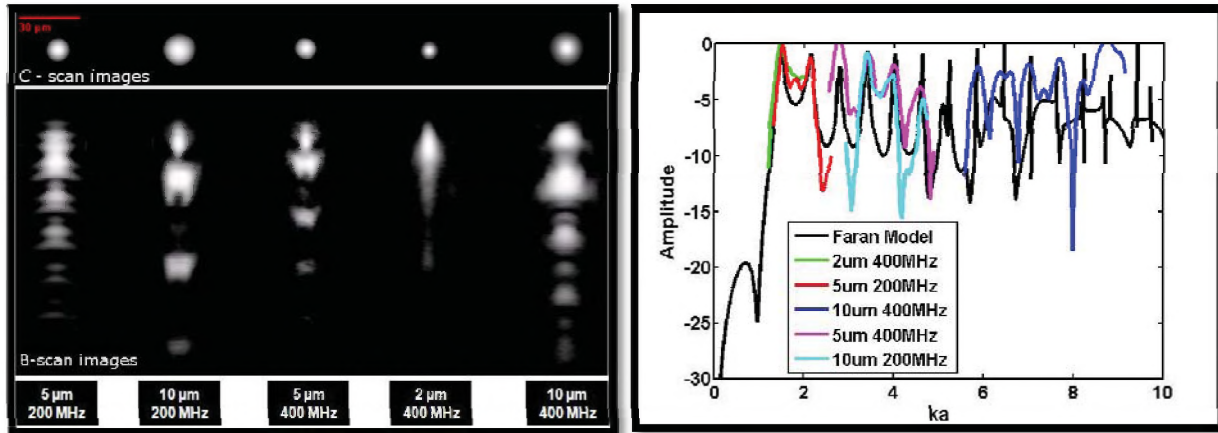


Figure 3: Bead scattering using acoustic microscopy. A) C-scan (top) and B-scan images shown for 5 combinations of bead size (2-10 μm) and interrogation frequency (200 or 400MHz). C-scans illustrate full-width half maximums of spheres of similar size to actual bead dimensions, and B-scans with characteristic resonant patterns. B) Power spectra of the various combinations of frequencies and transducers superimposed on the Faran solution for scattering, showing good agreement at locations of the peaks in the backscatter power spectrum.

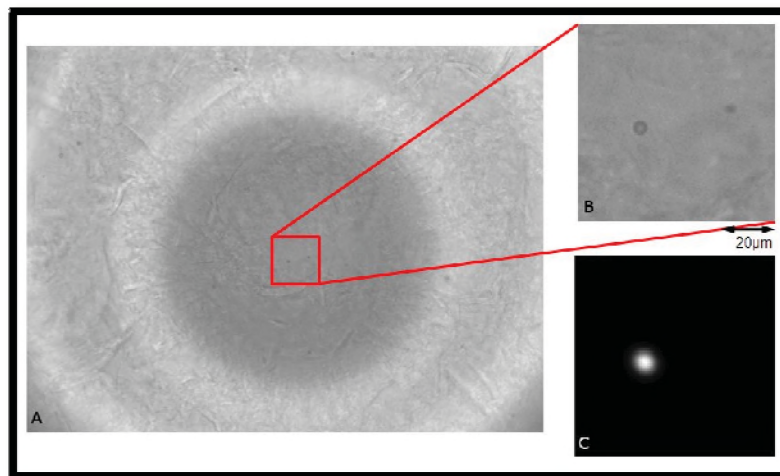


Figure 4: Demonstration of the effect of ultrasound contrast resolution. The above image was taken using the Kibero SASAM 1000 acoustic microscope that confocally aligns and optical microscope with the acoustic transducer (200MHz used in this example). A) Optical microscopy image of the specimen imaged with the acoustic microscope. The circle in this image represents the transducer cavity as seen from the optical microscope underneath. The region enclosed in the red square was scanned with the acoustic microscope. B) a zoom of the region scanned shows two polystyrene microspheres in the region, one of 5 μm diameter (to the left) and one of 2 μm diameter (to the right). C) The acoustic microscopy image clearly shows the 5 μm bead but detects no signal from the area the 2 μm bead is located. This could be understood by inspection of figure 1: the scattering from the smaller bead is much weaker.

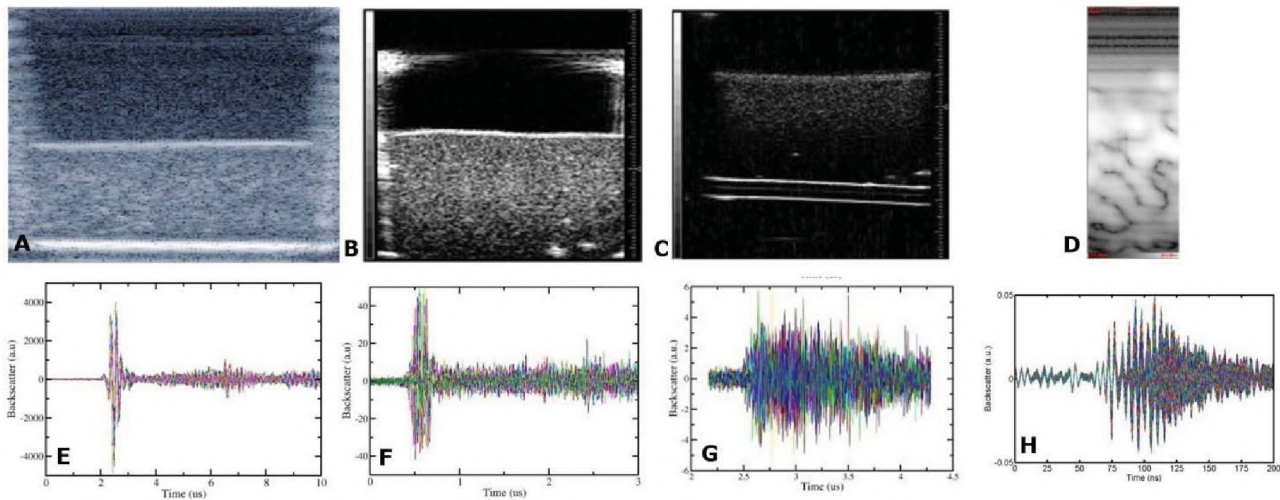


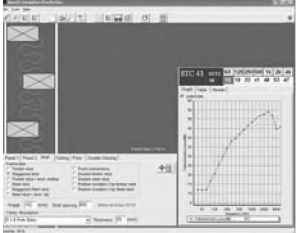



Figure 5: B scan images (top row) and a superimposed collection of RF lines that formed the images (bottom row) from a compact aggregate of acute myeloid leukemia cells. The aggregate has a height ranging from 1-10mm. The data were collected at 6MHz (with an Ultrasonix RP instrument, panels A,E), 20MHz (B,F with a VisualSonics VEVO 770 instrument, 8mmX8mm), 40MHz (C,G with a VisualSonics VEVO 770 instrument, 8mmX8mm) and 200MHz (D,H with a Kibero SASAM 1000 instrument). Image scale varies in each panel. Assuming the cell as the dominant scattering source, this represents a large range in ka . In all cases a speckle pattern is detected. At the lower frequencies, scattering from the aggregate surface dominates while at higher frequencies scattering from the cells dominates.

 <h1>INSUL</h1>	 <h1>SoundPLAN</h1>
 <p>INSUL is an easy to use software tool for predicting airborne sound insulation of simple or complex partitions consisting of various materials & structural systems, floors and glazing and impact sound insulation of concrete floors. It can be used to quickly evaluate new materials and systems or to investigate the effects of changes to existing designs. The INSUL v6.3 release incl. <u>Rain Noise Assessments</u>, <u>Triple Glazing</u> and <u>Outdoor to Indoor Noise</u> calculations.</p> <p>>>> August 2009 - INSUL v6.3 Release <<<</p> <p>Trial Version: www.navcon.com/insul.htm</p> <p>Navcon Engineering Network Phone: 714-441-3488 Email: forschner@navcon.com</p>	 <p>SoundPLAN is a graphics oriented noise prediction program used for noise planning, noise assessment & the development of noise mitigation measures. The database and management structure allows for a quick & easy generation of variants for small & complex noise models (i.e., Road & Railroad Projects, Industrial Plants, Quarry & Mines Operation, Power Plants, Amusement Parks, Wind Farms, Manufacturing Buildings/Rooms & Enclosures).</p> <p>SoundPLAN is based upon 30⁺ standards such as ISO 9613, Concawe, Nord2000, FHWA RD 77-108, TNM™2.5, FTA/FRA, VDI 3760. It generates traceable result tables and professional looking maps visualizing the input & output data.</p> <p>>>> September 2009 - SoundPLAN v7 Release <<<</p> <p>For more information go to www.navcon.com/soundplan.htm. Occasional users please ask for SoundPLAN essential www.navcon.com/soundplan_essential.htm</p>

In a Class of its Own

The unmistakable look of Hand-held Analyzer Type 2270 can overshadow a number of discrete yet significant distinctions which make this powerful instrument the complete toolbox for sound and vibration professionals. These include:

- Integrated digital camera
- Two-channel measurement capability
- Integrated LAN and USB interfaces for fast data transfer to PC and remote control and monitoring of Type 2270
- Environmental protection IP 44

Versatile in the Extreme

Type 2270 also boasts a wide range of application software modules that can be licensed separately so you get what you need when you need it.

Currently available measurement software includes:

- Sound Level Meter application
- Real-time frequency analysis
- Logging (noise level profiling)
- Sound and vibration recording
- Building acoustics
- Tonal assessment

Type 2270 meets the demands of today's wide-ranging sound and vibration measurement tasks with the accuracy and reliability associated with Brüel & Kjær instrumentation.

To experience the ease-of-use of Type 2270, just go to www.bksv.com and view the on-line video demonstrations.

For more information please contact your local Brüel & Kjær representative



BN 0281-11

HEADQUARTERS: DK-2850 Nærum · Denmark · Telephone: +4545800500
Fax: +4545801405 · www.bksv.com · info@bksv.com

Australia (+61)29889-8888 · Austria (+43)18657400 · Brazil (+55)115188-8166
Canada (+1)514695-8225 · China (+86)1068029906 · Czech Republic (+420)267021100
Finland (+358)9-755950 · France (+33)169907100 · Germany(+49)42117870
Hong Kong (+852)25487486 · Hungary (+36)12158305 · Ireland (+353)18037600
Italy (+39)025768061 · Japan (+81)337798671 · Republic of Korea (+82)234730605
Netherlands (+31)318 55 9290 · Norway (+47)66771155 · Poland (+48)228167556
Portugal (+351)214711453 · Singapore (+65)3774512 · Slovak Republic (+421)254430701
Spain (+34)916590820 · Sweden (+46)84498600 · Switzerland (+41)18807035
Taiwan (+886)227139303 · United Kingdom (+44)1438739000 · USA (+1)8003322040

Local representatives and service organisations worldwide

Hand-held Analyzer *Type 2270*

Brüel & Kjær 

SPMSOFT AND SPEECH PRIVACY MEASUREMENT IN OPEN-PLAN OFFICES

John S. Bradley

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

Introduction

A successful open-plan office design requires ‘acceptable’ or ‘normal’ privacy corresponding to $AI \leq 0.15$ and an ambient noise level of close to 45 dBA¹⁻³. This can only be achieved by careful attention to all important design details. If any key design parameters are ignored, $AI \leq 0.15$ will not be reached, and the office will not be a success. To ensure that changes systematically improve speech privacy, it is helpful to have in situ objective measurements to give unbiased assessments of conditions. Although this has been difficult to do using conventional measurements, the new SPMSoft program makes possible convenient in situ assessments.

New Speech Privacy Measurement Software

SPMSoft is new measurement software intended to measure privacy in open-plan offices during occupied conditions without unduly disturbing the occupants. Sound attenuations between locations in an office are determined from impulse response measurements. Daytime ambient noise levels can also be measured and speech privacy measures calculated in situ at each measurement location. SPMSoft calculates values of the Articulation Index (AI)⁴, the Speech Intelligibility Index (SII)⁵, the Speech Transmission Index (STI)⁶ and an A-weighted signal-to-noise ratio (SNR(A)). The user sees immediately how close the measured conditions are to the goals of $AI \leq 0.15$ and an ambient noise level of 45 dBA.

The program displays the envelope of the measured impulse response and the spectrum of the measured ambient noise as illustrated in Fig. 1. These displays give the user the key information required to diagnose the cause of acoustical problems. The two most important paths are usually the initial ceiling reflection and the initial screen diffracted path (whereby speech sound diffracts over the separating partial height panel). After the user enters some dimensions, the program indicates the components of the impulse response that are due to the sound arriving by the initial ceiling reflection path and via the initial panel diffraction path (red and blue boxes respectively in the lower graph on Fig. 1). The relative heights of these two components of the impulse response relate to the relative importance of increasing either the panel height or the ceiling absorption for improving speech privacy of the measured condition.

An ambient noise level of close to 45 dBA is an optimum compromise between being too loud and disturbing and too quiet and not masking unwanted speech sounds¹⁻³. It is typical of ventilation system noise levels found in many offices. Privacy measures can be recalculated with a more ideal ambient noise spectrum, than that measured, to explore the importance of changes to ambient noise levels. From the impulse response display and ambient noise level effects, the user gets an immediate indication of the relative importance of improvements to the key parameters: panel height, ceiling absorption and ambient noise levels, for improving speech privacy.

Software Operation

Fig. 2 illustrates a typical measurement setup and a block diagram of the external hardware required. A portable computer is connected to an external sound card. The program outputs sine-sweep test signals via the sound card and power amplifier to a small loudspeaker located in one workstation. A microphone is typically located in the adjacent workstation and connected to the computer via the external sound card.

The user simply proceeds, from top to bottom, through the 8 measurement buttons shown in Fig. 1. New measurement information is entered using the first button and the microphone is calibrated using the second button. The third button starts the measurement of ambient noise levels for a user-specified duration. The results are stored as 1/3-octave band levels but displayed on the upper part of the main screen as octave band levels.

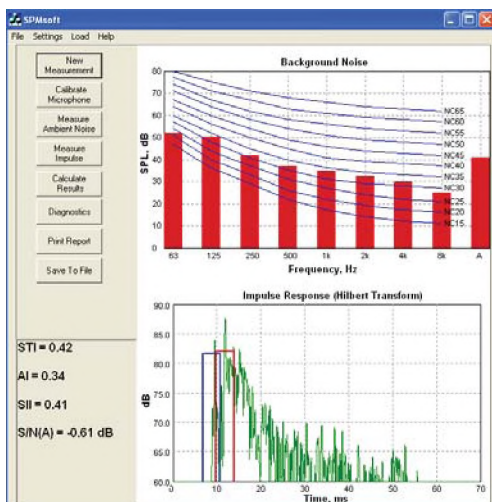


Fig. 1 SPMSoft main screen: measured ambient noise spectrum (upper) and impulse response envelope with the blue and red boxes identifying the initial diffracted path sound and the initial ceiling reflection sound (lower).

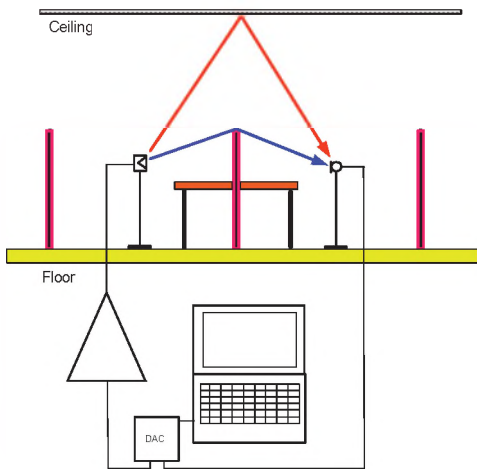


Fig. 2. Typical measurement setup showing a section through a pair of adjacent workstations.

The fourth button starts the impulse response measurement, which is displayed as the impulse response envelope in the lower part of the main screen in Fig 1. By using up to 99 repeats and synchronously averaging the responses, good results can be obtained without disturbing the occupants. Pressing the fifth button calculates the speech privacy indices AI, SII, STI and SNR(A).

The user can enter the dimensions of the office by using the 6th button to allow the program to determine the expected arrival times of the initial ceiling reflection (red box, Fig. 1) and the screen diffraction pulse (blue box, Fig. 1). The last 2 buttons make possible printing of results and saving them to files. The relative importance of the identified two sound paths (red and blue boxes on Fig. 1) and the measured ambient noise levels give the user immediate feedback as to the most critical factors for improving speech privacy.

Measurement Examples

SPMSoft is available for download⁷ along with a users manual⁸ and reports describing its use⁹ as well as the results of open-plan office case studies using SPMSoft¹⁰. Systematic measurements in real offices have demonstrated the effects of: ceiling absorption, panel height, window reflections, furniture placement, ambient noise levels and excessive reflected sound^{9,10}. Because of the speed of measurements, more detailed studies are possible such as calculating contours of speech privacy within workstations [11]

The measurement studies demonstrated that more absorptive ceilings can decrease AI values by 0.1 to 0.2 and increase the area around the source within which acceptable privacy is achieved by a factor of 8. Higher workstation panels were shown to decrease AI values by as much as 0.15.

Sound reflections from adjacent windows were shown to increase AI values by 0.1 compared to other workstations

without window reflections. The reflections from windows could be identified on the measured impulse response envelopes.

Surprisingly the furniture arrangements in workstations were also found to systematically effect privacy. Where there were tables or desks near the centre of the workstations, AI values were increased by 0.1 to 0.15 due to the added reflection energy from these flat horizontal surfaces.

Conclusions

The new SPMSoft program makes it possible to systematically identify and solve acoustical problems in open-plan offices and to better improve acoustical conditions in the offices. This complements existing acoustical design software for open-plan offices¹².

Acknowledgement

This work was supported by Public Works and Government Services Canada.

References

- [1] Bradley, J.S., "Acoustical Design for Open-Plan Offices", CTU #63, 2004. http://irc.nrc-cnrc.gc.ca/pubs/ctus/63_e.html, http://irc.nrc-cnrc.gc.ca/pubs/ctus/63_f.html
- [2] Bradley, J.S., "The Acoustical Design of Conventional Open Plan Offices", Can. Acoust. 31(2) 23-30 (2003).
- [3] Bradley, J.S., "A Renewed Look at Open Office Acoustical Design", Paper N1034, Proceedings Inter Noise 2003, Seogwipo, Korea, August 25-28, 2003.
- [4] ANSI S3.5-1969.
- [5] ANSI S3.5-1997.
- [6] IEC 60268-16, (2003).
- [7] SPMSoft download: http://irc.nrc-cnrc.gc.ca/ie/acoustics/open/SPMSoft/index_e.html
- [8] Bradley, J.S. and Estabrooks, T., "SPMSoft (Speech Privacy Measurement Software) User's Manual, Software user's manual, IRC publication, RR-266, August 2008. <http://irc.nrc-cnrc.gc.ca/pubs/rr/rr266/>
- [9] Bradley, J.S. and Gover, B.N., "Development and Evaluation of Speech Privacy Measurement System", IRC Research Report, RR-262, August 2008. <http://irc.nrc-cnrc.gc.ca/pubs/rr/rr262/>
- [10] Bradley, J.S. and Gover, B.N., "Open-Plan Office Speech Privacy Case Studies", IRC Research Report, RR-263, August 2008. <http://irc.nrc-cnrc.gc.ca/pubs/rr/rr263/>
- [11] Bradley, J.S., "Assessing speech privacy in open-plan offices", Proceed. 19th ICA (2007), Paper RBA-10-007 .
- [12] COPE-Calc Open-plan office Acoustical design software, http://irc.nrc-cnrc.gc.ca/ie/cope/07_e.html

NRC-IRC COMPUTER CONTROLLED ACOUSTIC MEASUREMENT AND QUALITY SYSTEM

Timothy Estabrooks

National Research Council, 1200 Montreal Road, Ottawa, ON, K1A 0R6 timothy.estabrooks@nrc-cnrc.gc.ca

1. INTRODUCTION

Computer measurement systems are becoming the norm for laboratory measurement of transmission loss (ASTM E90), impact sound insulation (ASTM E492), and sound absorption (ASTM C423). Such measurement systems are capable of archiving and quickly comparing current test data with those of previous measurements. This paper describes the architecture, development, and functionality of acoustics software developed at the National Research Council Canada, which is used to organize data from test specimens. The software moves beyond simple data storage and incorporates methods to formalize data verification and data approval for quality control. This is the first in a suite of five papers.

2. DATA ORGANIZATION

Test management software is used to organize specimens and test results. In very large data sets organization is critical to functionality, therefore, the software was designed to allow the user to access various data results quickly and efficiently as well as provide a quick visual reference to users. Using the software it is possible to review any test result, or the raw data from which that result was derived. All values calculated, or measured, are stored in an Oracle database while the raw unprocessed data are stored on a secure backed-up network disk system. This system enables quick and easy access to all data.

The system design includes a relational database with the capability of linking the contract, specimens and test data together. This relational database allows data to be cross-referenced quickly and efficiently. Figure 1 shows the interface that is used to access tests that have been performed. Once a test is selected, it is possible to review the data for that specimen and compare it to data from previous tests. Since research requires use of non-standard tests, the measurement system was designed so that data from other tests could be associated with the standard test results. This allows the user to spot data trends and perform basic data analysis before exporting the results for detailed analysis. This organization forms the foundation for recording all test data and using the simple tree structure it is possible to see all tests performed on a given specimen.

In Figure 1, the tree in the left of the window depicts how the contract, specimens and test are related. The icon indicates if the test has been approved, rejected or is awaiting approval. The graph in the lower right depicts measured data. Once a test has been approved all computed data are written to the database ensuring as standards are revised and

calculation requirements change, the archived data will remain true to the version of the test method used.

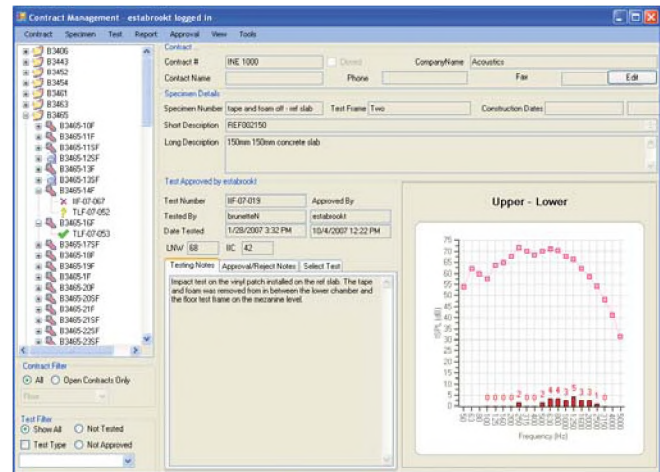


Figure 1: Main window of the measurement system.

3. DATA ENTRY

Defining the specimen to be tested is important for future use of the test data. The system allows the user to enter specimen specifications and to record many aspects of the specimen to be built and tested. For example, the type of material, dimensions, mass, as well as manufacturer, are recorded for all materials used to construct the specimen. For sheet materials the placement, orientation and fastener type, length and spacing, are also recorded. The specifications are used to generate and print detailed instructions that a carpenter can use to build the actual specimen to be tested. The system also records who built the specimen, when it was built, and who performed inspections at various stages of construction. During construction the carpenter records the actual measured mass and dimensions of the materials used. From these actuals it is possible to determine the total mass of the specimen.

A sophisticated algorithm is used to check the mass and other metrics of each element. If the data does not match known values within a certain margin of uncertainty the system flags a possible error due to user entry, measurement of the physical quantity, or a product that deviates from the establish norm.

4. SYSTEM CHECKS AND COMPUTER CONTROLLED MEASUREMENTS

A daily system check is conducted which involves the calibration of the measurement system and a speaker check.

A calibrator is placed on each measurement channel and the channel sensitivity is recorded and then compared to historical records. Once the system is calibrated, the speaker check is conducted by comparing each speaker measurement levels with a stored baseline for that speaker.

Once the specimen has been defined, the system calibrated and system checks have been performed, a test may be conducted. Once the user selects the test to be conducted, the computer measurement system uses the information to set up and perform all operations needed to conduct the test. Some tests require user interaction and the user will be directed at the appropriate point during the test to perform the user task, such as moving the light impact hammer box to its next position.

The system records all acoustical signals to file, after which the signal can be processed multiple times allowing the user to extract as much detail from the measurement file as is needed. For example, a signal recorded from a sound source can be processed to determine levels in both octave and one-third octave bandwidths. The signal could also be exported as an audio file to be used for subjective testing or to troubleshoot a problem with a measurement.

5. QUALITY CONTROL

The software moves beyond simple data storage and incorporates methods to formalize data verification and data approval for quality control. The quality control method includes both automated checks, and human review and approvals.

Once specimen test information and measured data are recorded in the database, a quality control audit is performed before the test data is used in any research project or report. The system provides a systematic method to vet both the measured data and the user entered data. A number of automatic checks are performed to ensure the proper data has been entered and measured data is not corrupted. Figure 2 shows the quick-check list provided to assist the user in the quality assurance process.

The quick-check guides the user through the vetting process. First, the software verifies that the measured data meets the requirements of the test method used. Examples of such requirements might include minimum surface area, temperature and humidity variances. The system also performs statistical analysis on the data collected such as the spatial variation of the sound field, and background levels.

The quick-check also uses previous records in the database to verify that the measured data is within historical norms. For example, the system can verify normal background levels and if the levels are higher than normal it will flag a potential problem with the measurement.

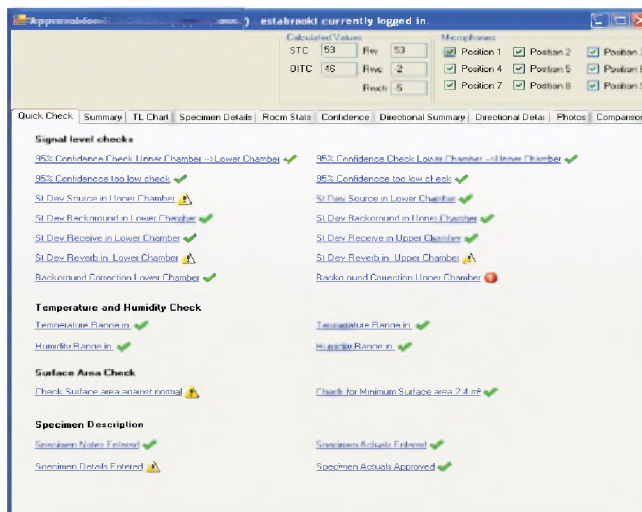


Figure 2: Quick check tab of data approval window

A graphical indication format is also provided which allows the user to quickly identify items that need closer scrutiny. A green checkmark indicates the data is acceptable, while a yellow caution symbol indicates a deviation from comparable historical data. A red exclamation point indicates requirements of the test standard are not satisfied.

The software enables review of each piece of measured data and identifies similar specimens that can quickly be compared to the test being vetted. This gives the user the ability to spot problems that may not be due to the test data, but rather due to the specimen under test. An example would be a wall with resilient channels that have screws shorting the gypsum board directly to the framing. While the measured data would not indicate a problem; comparison of the current test with historical data would highlight a discrepancy. Upon completion of the data vetting process the user must then approve or reject the test.

6. DATA MINING

The system also incorporates sophisticated data mining capability. Using the data miner, it is possible to view historical test data filtered by various criteria. For example, the user could perform a search of all floors that have both a minimum STC and IIC, and further narrow the results by specifying construction details such as joist type, spacing, etc. This enables for quick comparison of data for quality assurance and retrieval of results for more detailed analysis.

7. CONCLUSION

The NRC-IRC created Computer Controlled Acoustic Measurement System allows standard acoustics tests to be performed with less user interaction while increasing the data quality. A rigorous quality control system gives the user confidence in the data, while enabling later reprocessing of the measured signals. The data mining capabilities allow access to all test data and the ability to create cross-references between related test sets.

NRC-IRC FLANKING SOUND TRANSMISSION FACILITY

Timothy Estabrooks, Frances King, Trevor Nightingale, and Ivan Sabourin

National Research Council, 1200 Montreal Road, Ottawa, ON, K1A 0R6 timothy.estabrooks@nrc-cnrc.gc.ca

1. INTRODUCTION

Walls and floors evaluated using ASTM laboratory test methods are built in structural isolation to evaluate direct transmission. However, in real buildings, junctions between walls, floors, enable structure borne vibration to be transmitted around the nominally separating element. This ‘flanking’ transmission reduces apparent the sound isolation effectiveness of walls and floors in real world settings but is not accounted for in the ASTM laboratory methods. National Research Council persons (T. Estabrooks, B. Fitzpatrick, R. Halliwell, F. King, D. MacMillan, T. Nightingale, and D. Quirt) recently designed and built a flanking facility to investigate how flanking paths affect the sound insulation performance of building systems (coupled wall and floor elements). This paper describes some of the key features of the facility that is unique to the world. This is the second paper in a suite of five that examine flanking.

2. FACILITY DESIGN

The research facility has 8 rooms, or more specifically, 8 walls, 4 floors and 6 junctions, enabling evaluation flanking of bearing and non-bearing wall/floor junctions, as well as wall/wall paths with a single specimen. To ensure the facility itself does not contribute to vibration transmission, fixed surfaces are massive and isolated from one another using resilient mountings and each surface has its own independent structural framing. Each room has a volume that differs by about 10% from adjacent rooms, to avoid modal matching, and follows the recommendation of ISO 140.

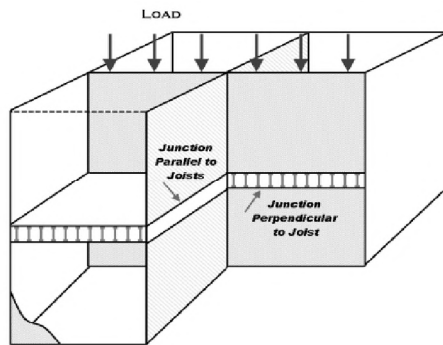


Figure 1: Cut-away sketch showing the bearing walls of the specimen to which a static load can be applied.

The entire facility is run by a computer system¹, which controls noise sources, robot movements and signal capture. In each room a pressure microphone is precisely moved to pre-determined positions by a computer-controlled robot having 4 degrees of freedom. The analogue signal from the microphones in all 8 rooms is sampled simultaneously,

digitized, and saved to disc for later processing. This system enables measurements to be run after business-hours when the ambient environment is quietest, and virtually eliminates uncertainty due to measurement repeatability. This is the second generation of flanking facility built at the NRC. It has the enhancement of being able to simulate the static load using 6 hydraulic cylinders which distribute 4500lbs on each of the East and West loading beams (*Figure 1*) - the equivalent of a two story building. The effect of loading on flanking transmission in wood frame construction is discussed in a companion paper².

3. COMMISSIONING

There were three phases in commissioning the facility. First was to establish microphone locations. Second was to establish consistency of results from the first generation flanking facility. Third was to establish flanking limits.

Selection of microphone positions – Microphone positions must be selected such that their mean value approximates that of a very large population that samples the entire room volume. To accomplish this each room was divided into a grid of 5 x 5 x 5, and the 125 positions measured in each third octave band between 50 and 5000 Hz. All points met ASTM standards, namely, they were farther than 0.5 m from a room surface and 1 m from a noise source. The average of all 125 positions was then calculated. One position was then selected within the centre zone of the grid, and an algorithm was written to randomly select the remaining 8 microphone positions. None of the resulting 9 microphone positions were permitted to be no closer than 0.5m of each other. This ensured that the microphones were spread throughout the volume of the room. After this process was completed for all eight rooms, measurements were conducted at the selected locations and then compared to the average of the 125 positions. Positions were deemed acceptable when the mean value of the 9 positions (plus/minus one standard deviation) equaled that of the larger set. Physical checks were conducted to ensure the automated system properly positioned the microphones. The robots were cycled though all 9-microphone positions many times, and positions confirmed by a measuring tape.

Consistency of Results – The first construction built in the new facility is shown in *Figure 2* and was one that had been thoroughly evaluated in the first generation facility. Full details of the construction details can be found elsewhere³, which consisted of floors with 3/4” OSB, wood-I joists spaced 16” O.C., 6” batt insulation, resilient channels spaced 16” O.C., and two layers of 5/8” Type C gypsum board. Wall consisted of 2x4 wood studs 16” o.c. batt insulation,

resilient channels 16" o.c., and three layers of 5/8" Type C gypsum board. The joists used in the first generation facility were slightly different than those used in the new facility and are listed in Table 1.

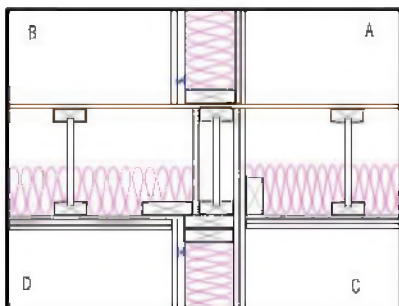


Figure 2: Sketch showing the specimen used in commissioning.

So any differences between results from the two facilities will then be the sum of two systematic effects and one random effect. Systematic effects are different joist flanges, and differences in room sizes, microphone positions, and measurement system. The random effect is the uncertainty introduced by rebuilding a (complex) specimen.

Table 1. Joist Properties.

Joist	Description	Web
#1	Solid Wood 2" x 10"	N/A
#2	Laminated veneer lumber 1-1/2"X1-1/2" flange	3/8" thick OSB
#3	Spruce-pine-fur 2-1/2"X1-1/2" flange	3/8" thick OSB

Figure 3 shows that despite differences in the joists, Apparent Transmission Loss (ATL) results from the two facilities are remarkably similar, especially considering the added uncertainty of completely rebuilding the complicated construction. Agreement between the apparent airborne and impact sound insulation between other room pairs was similar.

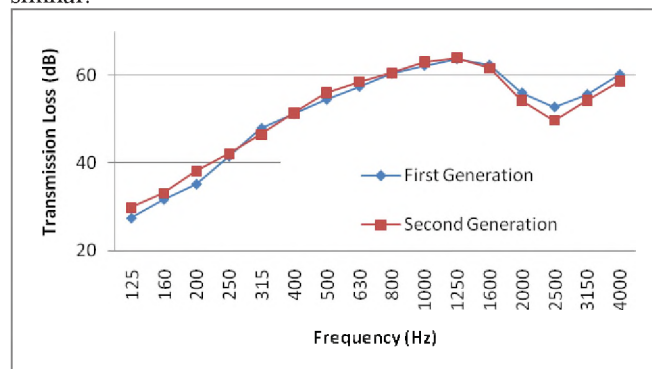


Figure 3: ATL for rooms separated by the floor-ceiling assembly in the first generation facility and the second generation facility.

Facility Flanking Limits – For the facility to be effective as a research facility to assess the various transmission paths in framed construction, transmission involving the facility (facility flanking) must be considerably less than that of the

specimen under test. Establishing flanking limits is difficult and time consuming, and basically follows this approach. A specimen is installed whose sound insulation is systematically improved by adding some treatment (typically a topping for floors, and additional layers of the gypsum board for walls) and at each stage the change in sound insulation is compared to expected changes. When commissioning the new facility it was possible to install some of the “high sound insulation” assemblies from the first generation facility and compare results. In reality for any sound transmission facility there is no unique flanking limit because facility flanking changes with the type and construction details of the specimen installed.

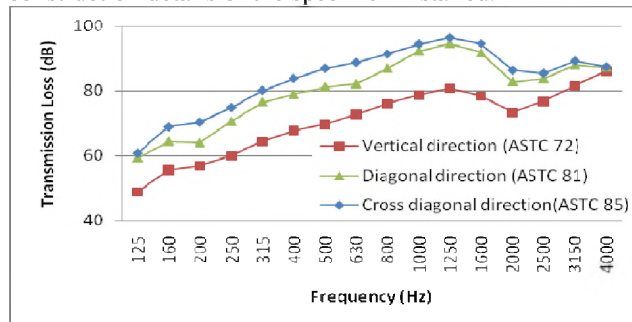


Figure 4: ATL between various room pairs indicating very sound insulation can be measured before the facility will affect results.

4. ESTIMATING SPECIFIC PATHS

The apparent sound insulation is the sum of all transmission paths between two rooms, and is what determines the subjective response of occupants. However, it is often very useful to obtain estimates of the sound insulation of particular paths in order to determine the most effective treatment to improve the apparent sound insulation. The NRC-IRC facility employs the ISO 10848 procedure with a minor variant³.

Systematically, all wall surfaces are shielded, and for each shielding condition the airborne and impact sound insulation measured. The next paper in this suite, shows that by recognizing which transmission paths are active for each shielding condition it is possible to create a set of simultaneous equations, which can be solved for the various paths. Obviously, the effectiveness of this method is strongly dependent on the quality of the measurements – they must be highly repeatable – hence the need for computed controlled robotic measurement system.

REFERENCES

- ¹ Timothy Estabrooks, “NRC-IRC Computer Controlled Acoustic Measurement and Quality System,” *Canadian Acoustics*, V37 No.3, 2009
- ² Ivan Sabourin, Berndt Zeitler, and Frances King, “Effects of Structural Load and Joist Type on Flanking Sound Transmission” *Canadian Acoustics*, V37 No.3, 2009.
- ³ “Flanking Transmission in Multi-Family Dwellings: Phase IV”, NRC-IRC-RR-218, 01 March 2006.

CHARACTERIZING FLANKING TRANSMISSION PATHS IN THE NRC-IRC FLANKING FACILITY

Frances King, Stefan Schoenwald, and Ivan Sabourin

Institute for Research in Construction, National Research Council, Ottawa, Ontario, Canada Frances.King@nrc-cnrc.gc.ca

1. INTRODUCTION

This paper is the third of a series of papers that provide an overview of work conducted by NRC in the past years on flanking transmission. In the previous papers the measurement system and the new NRC-IRC Flanking Facility are described^{1,2}. The ISO 10848 method for measuring flanking paths is time-consuming and not feasible for a facility of this complexity. This paper describes the methodology used in the Flanking Facility to systematically measure airborne sound transmission through the individual flanking paths for the 6 junctions of the specimen.

2. METHODOLOGY

Between two adjacent rooms there is 1 direct path through the partition and a total of 12 flanking sound transmission paths (3 for each of the 4 junctions, shown in Table 1 for the horizontal side-by-side case) and even diagonal room pairs that are connected by only one junction have four flanking paths. The room arrangement in the Flanking Facility allows evaluation of 8 horizontal, 4 vertical, 8 diagonal and 8 cross diagonal room pairs. The relative importance of these paths will depend on the properties of the floor, ceiling, party wall and side walls and of the junctions. Hence, extensive research studies are carried out in the Facility to characterize the sound transmission loss (TL) of each flanking path for different specimens^{3,4,5}.

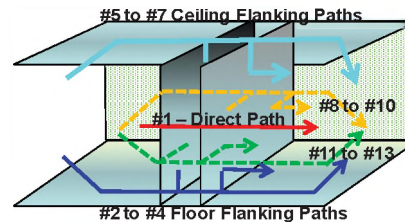
ISO 10848 suggests to measure all flanking paths one-by-one by shielding the surfaces of all other building elements that are not part of the considered path, but are either excited in the source room or radiate sound into the receive room. In the NRC-IRC flanking facility, 16 mm gypsum board on 90 mm glass wool is put in front of the building elements as shielding with no rigid connection to the test specimen. However, applying shielding to horizontal surfaces, the floor or ceiling, is not always feasible. In the NRC-IRC facility, some flanking paths are characterized with a slightly modified approach by extracting single path data from measurements with different shielding conditions.

For example, the paths listed in Table 2 are measured between two lower horizontal rooms of the Flanking Facility using both the ISO standard and the modified approaches as described below. The measurement results

presented in Figure 1 are for the specimen that was used for the commissioning of the Facility².

Table 1: Direct and flanking transmission paths between two horizontal side-by-side rooms.

Path #	Direct and Flanking Transmission Paths	Airborne
#1	Party Wall-Party Wall	Direct
#2	Floor-Floor	Flanking
#3	Floor-Party Wall	Flanking
#4	Party Wall-Floor	Flanking
#5	Ceiling-Ceiling	Flanking
#6	Ceiling-Party Wall	Flanking
#7	Party Wall-Ceiling	Flanking
#8	Side Wall-Side Wall (1)	Flanking
#9	Side Wall-Party Wall (1)	Flanking
#10	Party Wall-Side Wall (1)	Flanking
#11	Side Wall-Side Wall (2)	Flanking
#12	Side Wall-Party Wall (2)	Flanking
#13	Party Wall-Side Wall (2)	Flanking



The Flanking-TL of the *ceiling-ceiling* (#5) path in Figure 1 is measured according to ISO 10848. For the measurement, only the party wall and one pair of the side walls that belongs to the specimen have to be shielded in both rooms since measures have been taken to suppress sound transmission between the rooms through the permanent shell of the Facility¹ which forms one pair of side walls and the floor of the considered rooms.

The remaining paths in Table 2 are obtained with the modified approach. The shielding is removed from the party wall in the receiving room and the TL due to the *ceiling-wall* (#6) and the *ceiling-ceiling* (#5) path is measured. The TL of the *ceiling-wall* (#6) path shown in Figure 1 is obtained by subtracting the TL of the *ceiling-ceiling* (#5) path from the measured data. Similarly, the TL due to the *wall-ceiling* (#7) path and the *ceiling-ceiling* (#5) path is obtained by moving the shielding from the party wall in the source room to the receiving room. The TL of the *wall-ceiling* (#7) (Figure 1) is obtained by subtracting the TL of the *ceiling-ceiling* (#5) path from the measured data.

Finally, all shielding is removed from the party walls and the TL due to all four paths of Table 2 is measured. Since the TLs of three of the paths are known, the remaining direct path (#1) can be extracted.

For the side wall paths listed in Table 1 (#8 to #10 or #11 to #13), they can be characterized by shielding all the party walls and changing the shielding condition of the side wall systematically. The floor flanking paths (#3 to #5) can be characterized in the upper rooms of the Facility. The flanking paths of the vertical and diagonal room pairs can also be estimated following the same methodology.

Table 2: Direct and floor-ceiling paths in the Flanking Facility.

Path #	Direct and Flanking Transmission Paths	Airborne
#1	Party Wall-Party Wall	Direct
#5	Ceiling-Ceiling	Flanking
#6	Ceiling-Party Wall	Flanking
#7	Party Wall-Ceiling	Flanking

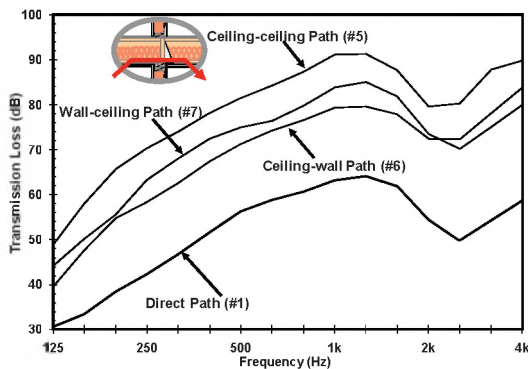
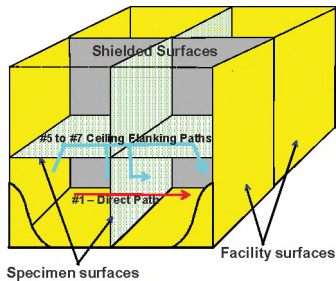


Figure 1: TL of ceiling-ceiling, wall-ceiling, ceiling-wall and direct path between two rooms separated by a party wall.

3. LIMITS OF MEASUREMENT METHOD

Although this paper shows that the applied method works fine, its limitations are discussed in this section. In Figure 2, the TL of the direct, the *ceiling-wall* and *ceiling-ceiling* path is presented for an extended frequency from 63 Hz to 4 kHz. Below 125 Hz, the TL of all three paths is very low and converges around the Apparent-TL. Hence, the applied shielding is not effective at low frequencies - most sound is transmitted directly through the shielded party wall - and the TL of all flanking paths is underestimated. Similar to the limitation of TL due to the mass-spring-mass resonance of double leaf walls (leaves are the masses and the spring is the air in the cavity), the TL of the shielded wall (now a system with 4 masses coupled by 3 springs) is limited by resonances. The shielding could be improved by increasing the mass of the applied gypsum board, which generally makes the shielding method even more impractical. Thus, previous study⁶ has shown that it is reasonable to fit ‘tails’

with a 6 dB increase per octave band to the measured flanking TL below 315 Hz as shown in Figure 2.

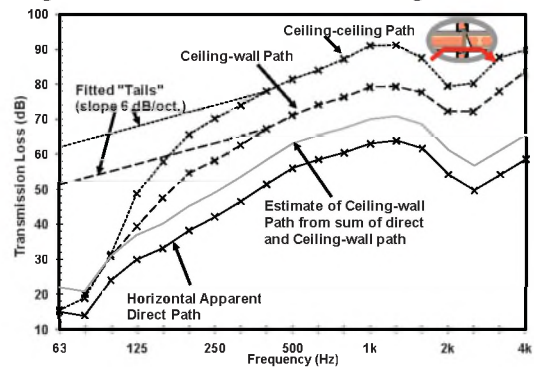


Figure 2: Shielding limits at low frequencies – flanking TL with fitted “tails”; Conservative estimate of Ceiling-wall path due to small measured differences (grey)

Another limitation that affects the extraction of flanking paths is discussed in the following. In the example in Figure 1 an ideal case is presented with rather big differences in measured TL because the flanking path with the highest TL was measured separately. But in some cases the measured differences are small, and sometimes less than the repeatability uncertainty of measurements despite the high precision measurement system¹. In such cases, only a conservative estimate can be defined as shown below. It is assumed that in this example the TL of the direct path with smallest TL is measured separately. The difference of this TL to the apparent TL due to transmission by the direct path and by any other ceiling paths that is measured next is small. If it is less than the measurement uncertainty of 1 dB then it could certainly not be related to a change of the shielding condition. In the extraction of the path data, the difference must be assumed to be 1 dB. This gives a conservative estimate for TL of the second path that is only 7 dB greater than the TL of the direct path. In most cases this estimate grossly underestimates the flanking TL as shown by the grey line in Figure 2.

Like every measurement method, the one applied in the NRC-IRC Flanking Facility has its limitations. Hence, thorough planning of the tests and care in the data analysis is required.

REFERENCES

- ¹Estabrooks T., “NRC-IRC Computer Controlled Acoustic Measurement and Quality System”, *Canadian Acoustics*, V37, No3, 2009
- ²Estabrooks T. et al., “NRC-IRC flanking sound transmission facility”, *Canadian Acoustics*, V37, No3, 2009
- ³King F., Sabourin I., “The effect of resilient channels on ceiling flanking transmission paths”, *Canadian Acoustics*, V37, No3, 2009
- ⁴Sabourin I., et al., “Effects of structural load and joist type on flanking sound transmission”, *Canadian Acoustics*, V37, No3, 2009.
- ⁵Nightingale T.R.T. et al., “A hierarchy of flanking transmission paths in lightweight wood frame construction”, *Internoise 2009*
- ⁶Nightingale, T.R.T. et al., “Flanking Transmission in Multi-Family Dwellings: Phase IV”, *Research Report 218, NRC-IRC, 2006*

THE EFFECT OF RESILIENT CHANNELS ON CEILING FLANKING TRANSMISSION PATHS

Frances King and Ivan Sabourin

Institute for Research in Construction, National Research Council, Ottawa, Ontario, Canada Frances.King@nrc-cnrc.gc.ca

1. INTRODUCTION

This paper is the fourth of a series of papers that provide an overview of work conducted by NRC in the past years on flanking transmission. NRC-IRC has carried out extensive studies to characterize flanking transmission through various building components and assemblies in the Flanking Facility¹. One of the studies is to examine the effect of resilient channels on ceiling flanking paths. The floor-ceiling assembly in apartment buildings separates two different dwellings and the gypsum board ceiling is usually mounted on resilient channels to meet the requirements of the building code. However for row housing, the floor-ceiling assembly separates two rooms of the same dwelling and typically the floor is not sound or fire rated. In this case, the gypsum board ceiling is often directly attached to the floor joists and ceiling flanking transmission paths between the two side-by-side rooms shown in Figure 1 can reduce the apparent STC significantly. This paper describes the improvement due to the use of resilient channels.

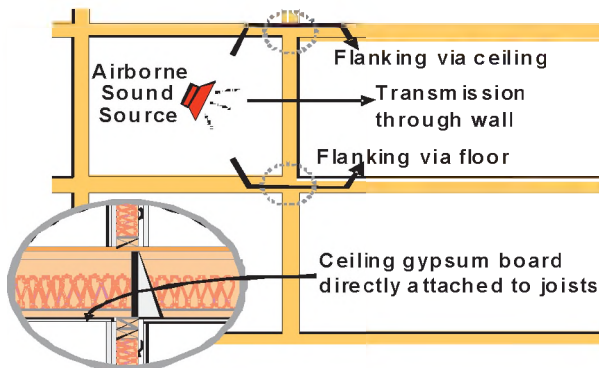


Figure 1: Types of flanking transmission paths in row housing.

2. MEASUREMENTS AND RESULTS

The apartment type floor-ceiling assembly was built in the NRC-IRC Flanking Facility for the commissioning study¹. For the row house floor ceiling assembly, the cavity insulation, the two layers of 16 mm gypsum board and the resilient channels were removed and replaced by a single layer of 13 mm gypsum board directly attached to the wood I-joists. The direct and flanking transmission paths were evaluated for both types of ceilings using airborne and impact sound sources. This paper focuses on airborne sound

transmission. The methodology used for characterizing individual flanking paths is described in the third paper². The airborne measurements were carried out simultaneously in the flanking facility for three different types of junction: 1.) floor joists parallel to the wall, 2.) floor joists perpendicular to the wall and continuous, and 3.) floor joists perpendicular to the wall but discontinuous at the junction. Figure 2 shows the Transmission Loss (TL) of the horizontal *ceiling-ceiling* path for both types of ceilings for the three types of junctions. Figure 3 shows the TL of the diagonal *ceiling-floor* path. Because the method of shielding a surface does not adequately suppress low frequency transmission, the estimates for paths in the low frequencies tend to be very conservative². In this paper, the raw data are shown and typically a tail is fitted to the measured estimate. Since the ceiling flanking paths are comparable for the three junctions, TLs are averaged and the differences between the two types of ceiling are shown in Figure 4. By attaching the ceiling with resilient channels to the floor joists, TL of the apartment type of ceiling is much bigger than the row house type for all paths. The vertical direct path increases by about 15 dB, the horizontal *ceiling-ceiling* path by about 20 dB and the diagonal *ceiling-floor* path by about 10 dB. No significant difference is found on the *floor-floor* path.

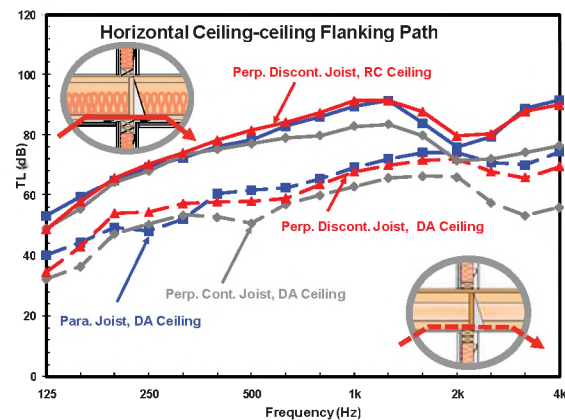


Figure 2: Flanking-TL of horizontal *ceiling-ceiling* path for apartment-type ceiling and row house-type ceiling and 3 different junctions.

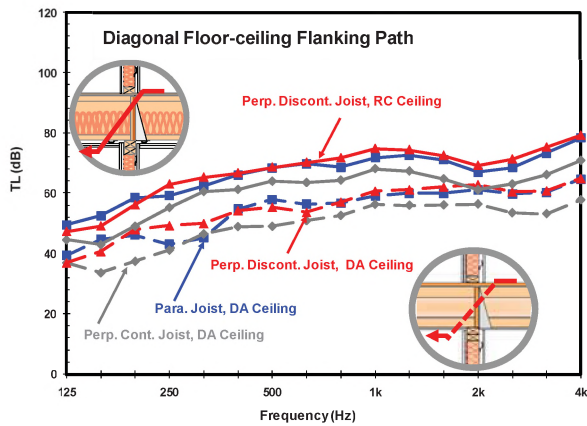


Figure 3: Flanking-TL of diagonal floor-ceiling paths for apartment-type ceiling and row house-type ceiling with 3 different junction constructions.

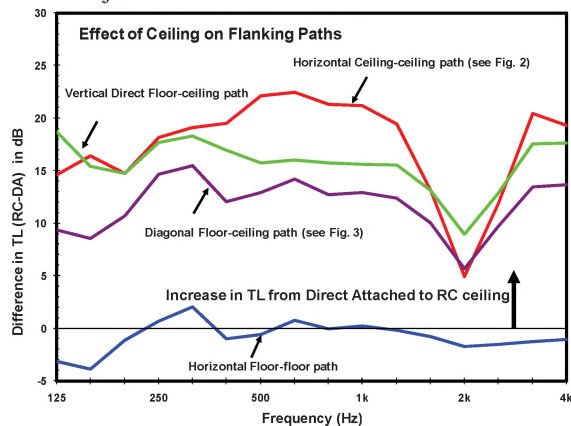


Figure 4: Differences in transmission loss between apartment and row house type ceilings for direct and flanking transmission paths.

3. DISCUSSION

There are three ceiling flanking paths as discussed in another companion paper. The *ceiling-wall* paths can be extracted from TL measurements of different shielding conditions². For a single stud wall with a directly attached ‘leaf’ on one side, the Apparent Sound Transmission Class (ASTC) between two side-by-side rooms depends on the amount of noise transmitted through the partition wall and the flanking paths. The *ceiling-ceiling* and *ceiling-wall* flanking paths can limit the ASTC when the ceiling and wall are directly attached as shown in Figure 5. The diagonal ASTC is limited by the *floor-wall* and the *floor-ceiling* flanking paths and comparable to the horizontal ASTC as shown in Figure 6. In this scenario with both the ceiling and wall directly attached and no additional topping on the floor, a lower room in a row housing can be exposed to flanking noise coming from the neighbour’s upper and lower rooms. Currently, the building code does not have any requirements on flanking transmission.

A recent paper shows that when a dominant path is treated, the hierarchy changes and other path(s) become more

important³. As the knowledge of dominant flanking paths for different building assemblies increases, it is possible to develop design strategies and provide guidance to practitioners for designing better and cost effective sound isolation buildings.

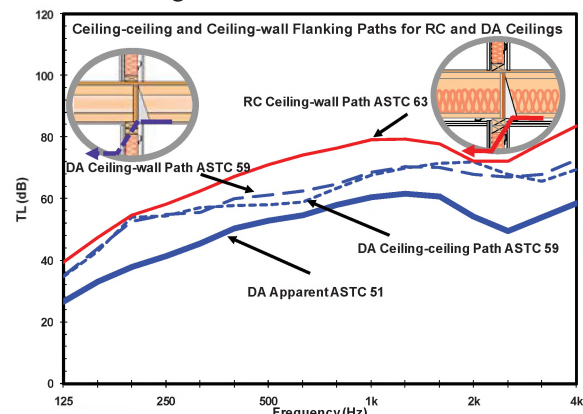


Figure 5: Horizontal Apparent-STC between two side-by-side dwellings can be limited by the *ceiling-ceiling* and *ceiling-wall* transmission paths.

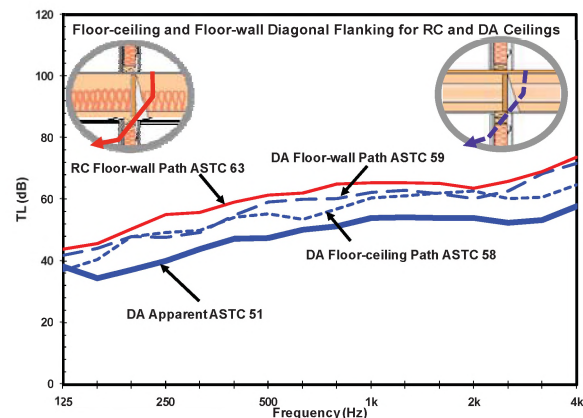


Figure 6: Diagonal Apparent-STC between an upper and a lower room is limited by the *floor-ceiling* and *floor-wall* transmission paths.

REFERENCES

- Estabrooks T., et al., “NRC-IRC flanking sound transmission facility”, *Canadian Acoustics*, V37, No.3, 2009
- King F., et al., “Characterizing flanking transmission paths in the NRC-IRC flanking facility”, *Canadian Acoustics*, V37, No.3, 2009.
- Nightingale T., et al., “A hierarchy of flanking transmission paths in lightweight wood frame construction”, *InterNoise 2009*, Ottawa.
- Nightingale, T.R.T., et al. “Flanking Transmission in Multi-Family Dwellings: Phase IV”, *NRC-IRC, RR-218*, pp. 425, March 01, 2006.

EFFECTS OF STRUCTURAL LOAD AND JOIST TYPE ON FLANKING SOUND TRANSMISSION

Ivan Sabourin, Berndt Zeitler, and Frances King

National Research Council Canada, 1200 Montreal Road, Ottawa, ON, K1A 0R6, Canada – ivan.sabourin@nrc-cnrc.gc.ca

1. INTRODUCTION

It is well known that flanking sound transmission in lightweight wood frame construction is dependent on details around the junctions such as joist orientation and joist continuity¹. In this paper (the fifth paper of a five part suite) two further parameters, namely structural load on the load bearing wall and joist type, are presented. They are both relevant for the building Codes and guides². It is important to assess if laboratory measurements under- or over-predict the sound transmission in a real building and if construction elements such as joists can be treated as generic or not.

Measurements to assess the effect of these parameters were conducted at two NRC flanking facilities: the first generation four-room (two above two room), and second generation eight-room (four above four room) flanking facilities³. The later facility is unique in the world and has the capability of measuring flanking sound transmission of loaded partitions. The construction assemblies used in these studies are of common North American style wood frame construction³. The appropriate surfaces of the rooms are shielded to limit the sound transmission to the paths of interest⁴.

2. EFFECT OF STRUCTURAL LOAD

This section on the effect of loading on flanking sound transmission is separated in two parts. The first part, an initial loading study, deals with the reproducibility of load and general tendency of the effect of loading. The second part deals with the effect of loading relative to joist orientation and continuity.

In the initial study the flanking sound transmission was investigated between two horizontally separated rooms, with discontinuous joists perpendicular to the separating wall; all sidewalls were shielded. Different loads were applied to the separating wall simulating no stories above, 0.5 stories above, and 0.7 stories above the rooms of interest. The apparent transmission loss (ATL) for the sum of these paths increases due to adding load to the party wall and tends to increase with frequency as shown in Figure 1. The effect of load stagnates at an extra load of 0.5 stories. This is a very important finding because it means that the same load can be used to assess the effect of load on buildings of different height, be they 2 stories or 4 stories tall. The second important finding is that the effect of load is reversible. After removing the load, the ATL returns to its original state. The same is true when reapplying the load. The very small variations seen are most likely due to the long duration between tests which was spread out over a period of weeks.

The second part of this loading study compares the effect of load relative to joist orientation and continuity.

Single sound transmission paths were isolated to understand the effect of loading. It is to be noted that the partition wall is only loaded for the junction cases where the joists are perpendicular to the partition wall whereas it is the side walls which are loaded for the parallel case. Figure 2 shows the effect of loading for the horizontal floor-floor and ceiling-ceiling paths. Loading has a similar effect for both horizontal paths through the same junction. In general, loading improves both perpendicular cases while it causes a slight negative effect on the parallel paths. Very similar results were observed for the change of ISPL through floor-floor paths.

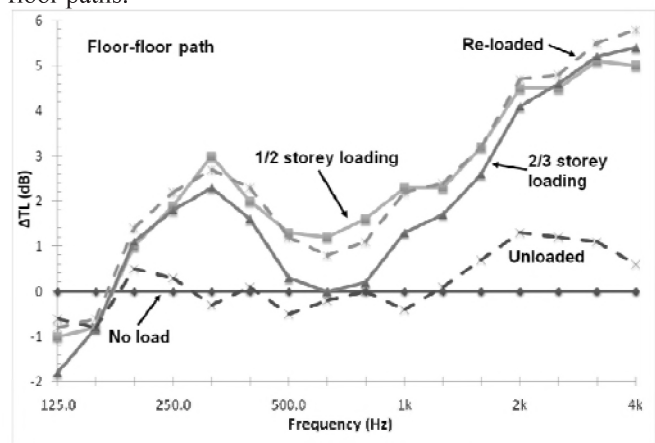


Figure 1: Change of ATL of horizontal room pairs due to loading, unloading, and reloading partition wall.

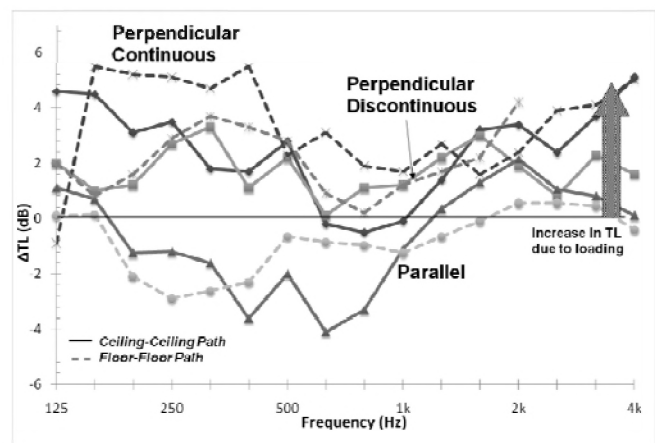


Figure 2: Effect of loading on ATL for different junction types for horizontal floor-floor and ceiling-ceiling flanking paths.

The fact that loading increases attenuation of sound transmission in cases where the joists are perpendicular and decreases attenuation slightly in parallel joist cases, can be explained by a recent study⁵ using a laser vibrometer to capture the structural intensity. From this study, it has been shown that much of the flanking sound power is transported

along the head- and soleplates of the wall into the receiver room. By loading the partition wall the contact between the floor and the soleplate of the wall becomes more rigid. Thereby more sound power from the floor is transmitted into and along the soleplate, and travels parallel to the load bearing junction instead of perpendicular to the junction through the floor causing the perpendicular paths to be attenuated more and the parallel path to be attenuated less.

Table 1: Summary of the effect of loading

Junction type	Direct Paths -vertical -horizontal	Horizontal Flanking Paths -floor-floor -ceiling-ceiling	Diagonal Flanking Paths -floor-ceiling
Perpendicular continuous	No significant change	Improvement (approx. 3 dB)	Improvement (approx. 2 dB)
Perpendicular discontinuous		Improvement (approx. 3-4 dB)	Improvement (approx. 1-2 dB)
Parallel discontinuous		No change or slight worsening	No change or slight worsening

3. EFFECT OF JOIST TYPE

The flanking sound transmission of assemblies constructed using three different types of joists is investigated here. The first comparison is between the ATL of specimens built with 2x10 lumber (Joist #1) and wood I-joists (Joist #2) measured at NRC’s first generation flanking facility. The second comparison is between ATL using two different types of wood I-joists, both with OSB web. Joist #2 has a smaller laminated veneer lumber flange where as Joist #3 has a larger spruce-pine flange³. The latter assemblies to compare the effect of wood I-joist were constructed in different NRC flanking facilities; hence other parameters such as different room size, junction length, might have an influence on the results. Changing the joist type could affect many components of the transmission path – the power injected by the source, the structural attenuation, the junction attenuation, and radiation to the receiver room.

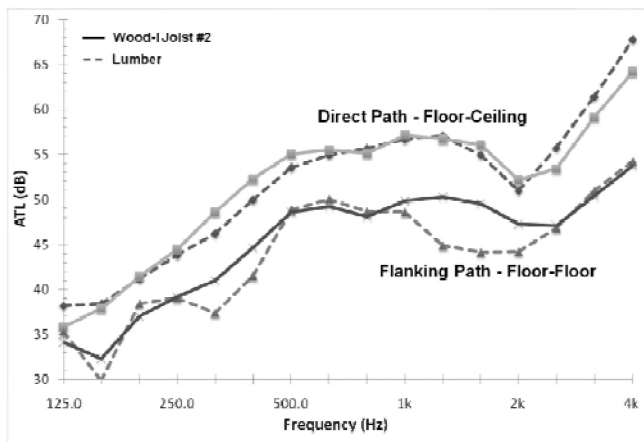


Figure 3: ATL difference for floor-floor paths due to joist types: 2x10 lumber vs. wood I-joists in same facility

A comparison of ATL between lumber and wood I-joists is shown in Figure 3 for the direct vertical and horizontal floor-floor cases. The joist type which performs better varies over the whole frequency range for the direct vertical path. For the floor-floor path however, the Wood-I

joist #2 performs better throughout most of the frequency range. This order of performance changes for impact sound pressure levels though¹.

As expected, the largest difference due to different joist type is observed for floor-floor flanking transmission path. It is however difficult separate the effect due to the joist types and the room properties.

The comparison of Joist#2 and #3 can be seen in Figure 4, where the horizontal floor-floor, the direct vertical, diagonal floor-ceiling path are displayed for the parallel junction case. The differences for all paths are of similar magnitude. At high frequencies the floor-floor path difference is probably a reproducibility issue due to varied contact area between the floor and the head- and soleplate.

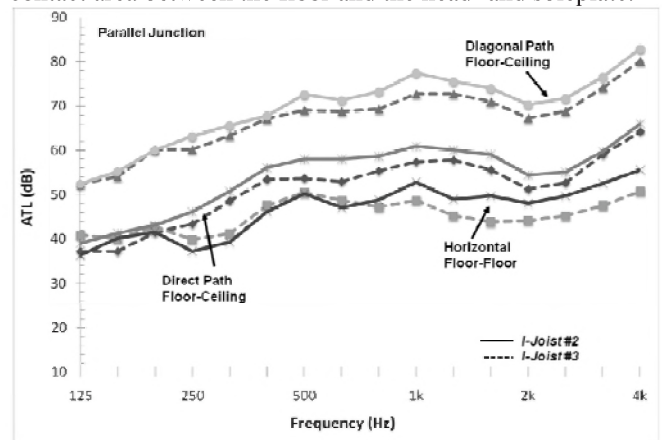


Figure 4: ATL for parallel floor-floor paths due to joist types – wood I-joists #1 vs. wood I-joist #2 in different facilities

4. CONCLUSION

Changes to the flanking sound transmission due to loading are small and might not be of any importance if the direct paths are dominant. Current estimates without loading are mostly conservative except for parallel cases. The load correction factor for relevant paths depends on the joist orientation and continuity.

Early data suggests that joists can be treated as generic, because the effect of the joist type is considered small relative to the influence of the joist orientation and continuity.

REFERENCES

- ¹ T.R.T. Nightingale, et al., “Flanking transmission in multi-family dwellings: phase IV”, NRC-IRC RR-218, 2006
- ² J.D. Qurt, Nightingale, T.R.T.; King, F., “Guide for sound insulation in wood frame construction”, NRC-IRC, RR-219, 2006
- ³ T. Estabrooks, T. Nightingale, I. Sabourin, “NRC-IRC flanking sound transmission facility”, *Canadian Acoustics*, V37, No3, 2009
- ⁴ F. King, S. Schoenwald, I. Sabourin, “Characterizing flanking transmission paths in the NRC-IRC Flanking Facility”, *Canadian Acoustics*, V37, No3, 2009
- ⁵ S. Schoenwald, et al., “Investigation of flanking sound transmission in lightweight building structures using a scanning laser vibrometer”, Proceedings of Euronoise 2009, Edinburgh

TRENDS IN CELL PHONE VOICE PROCESSING

Chris Forrester

Research in Motion, 295 Phillip St., Waterloo Ontario, Canada, N2L 3W8 cforrester@rim.com

1. INTRODUCTION

Cell phone voice signal processing technology has come a long way since the first digital cellular phone call in 1991 on a second generation (2G) wireless network. For those early digital handsets, DSP processors lacked sufficient horse power to implement much voice processing beyond the speech codec itself. Since then, the speed of DSP's has increased by more than an order of magnitude and modern cell phones have added multiple processor cores to support multi-media in addition to voice. So given this increase in capability, what trends are emerging in cell phone audio design and what impact does this have on voice processing?

2. DISCUSSION

In 1989, Texas Instruments introduced the highest performing fixed point DSP in the industry, operating at 28 MIPS [1]. This DSP became very popular for use in cell phones. The main priority of the DSP is always to implement the radio signal processing needed to support mobile communications. By necessity, this means that voice signal processing is allotted only a small fraction of the MIPS budget. So for early 2G cell phones, there was very little time budgeted for voice processing.

This lack of available processing power made life difficult for the audio pioneers of digital cellular. The acoustic design of early 2G phones had to meet GSM voice-band phone specifications without the benefit of DSP voice equalization. This forced the use of analog circuitry to provide any filtering needed to meet voice-band specifications. As well, there was little DSP processing power available to perform echo or noise control functions that are common today. As a result, it was not uncommon to hear echo of your voice when talking to a person on a digital cell phone. Received speech volume levels needed to be kept low to control echo which led to 'can't hear' complaints by cell phone users. Without background noise reduction, users had to make calls in moderate background noise conditions or suffer complaints from the called party. But of course it did not take long for this to change.

By the mid to late 1990s, DSP processing power had increased sufficiently to provide more enhanced voice signal processing, targeted at addressing the early mobile voice quality issues. To address the echo problems, it became the standard to provide a basic time-domain echo canceller

based on Normalized Least-Mean-Squares (NMLS) with some form of basic gating or gain switching function to control residual echo [2]. As much as 30dB of echo return loss enhancement is possible under typical conditions. The standard for noise cancellation became some form of sub-band or full-band gain reduction based on SNR (speech-signal-to-noise) calculations [3] - as much as 15-20dB of stationary noise reduction is achievable with this technology without harming speech quality too much. Other more basic features such as transducer equalization and dynamic range signal processing to limit high-level signals and boost low-level signals became common as well [4].

Of course, DSP processing speeds continue to increase and current platforms are more than 50 times more powerful than the early 2G platforms. The increase in processing power has enabled DSP audio engineers to use more sophisticated voice processing to provide features and solve problems that were only dreamed of as recent as five year ago.

Since the release of the first wireless phone there has been a continued evolution towards smaller, thinner and lighter phones. This trend to shrink the cell phone has had a huge impact on the acoustic design and has driven a need for ever greater voice processing features and performance.

As devices shrink, microphones move farther away from the mouth which reduces the signal-to-noise (SNR) at the microphone. The SNR may drop by as much as 10dB or more depending on the design of the phone. So the background noise reduction must be improved by this amount to maintain performance. At the same time, cellular service providers are demanding improvements in non-stationary noise reduction.

A trend that has emerged recently is to provide multiple microphones to address the need for improved background noise reduction. There has been a lot of interest lately in the area of blind-source separation (BSS) and beam forming signal processing techniques for application in cell phones [5]. With these techniques it is possible to provide improved noise reduction, especially to address non-stationary noise where one microphone is insufficient. It is expected that future devices will add more microphones with more complex signal processing to achieve even better voice quality performance in non-stationary noise environments.

With multiple microphones it is possible to achieve significant noise reduction to the point that the far end user finds it unnaturally quiet. Additionally, it is not uncommon for noise reduction algorithms to severely change the quality of the noise, so that it becomes very unnatural to the far end listener. The ideal is to achieve a reduction in the transmission level of noise to some optimal level, but leave the character of the noise fully intact for a more natural experience. So some attention now is being given to these problems and it is expected that future noise reduction systems will provide good suppression without harming the quality of the noise.

The first cellular speakerphone devices used loudspeakers as large as 34mm diameter. To save space in modern cell phones, it is common to use speakers that are rectangular with dimensions on the order of 14x20mm – and the trend is towards even smaller speakers. Due to physical limitations, the impact on audio performance is a reduction in maximum acoustic output and an increase in distortion. As well, in some speakerphone designs, the loudspeaker is placed literally next to the microphone due to space limitations, which leads to a very high echo signal. So signal processing engineers are being challenged to provide improved dynamic range compression to increase output and minimize distortion, and better echo control algorithms to handle the increased distortion at high volume levels.

A modern approach for providing residual echo suppression is to use sub-band analysis or spectral matching techniques to suppress the bands only with high echo levels – this can provide a more virtual full-duplex experience [6]. As well, researchers continue to work on non-linear filtering to address echo components from distortion added by the loudspeaker and plastics vibrations. A common approach is to use a Volterra filter to model the loudspeaker distortion [7]. Often it is the third harmonic of the input signal which is the most significant distortion component which leads many to use third order filters. It is expected that these techniques will enhance the echo return loss at high volume levels by 6dB or more.

One of the biggest cellular voice quality enhancements expected to take place in the near future is the rollout of wideband telephony. Wideband doubles the voice bandwidth from about 3.5 kHz to 7 kHz and promises to provide a more natural voice experience with increased user satisfaction. Existing narrowband voice processing will need to be upgraded to support the extended bandwidth. So this will require modifications to echo cancellation, noise suppression and all other blocks. Generally it will require as much as twice the processing power to support wideband, but this will not be a problem for modern DSP's.

While wideband voice will provide a significant leap forward for voice quality, the Third Generation Partnership Project (3GPP) Technical Specification Groups are already

working on enhanced mobile audio specifications for future networks. This includes such features as 'superwideband' which doubles wideband to 14 kHz, stereo telephony and further extensions to multi-channel audio such as 5.1. These capabilities will allow other audio program material from presentations or teleconferences to be more faithfully transmitted with voice.

Finally, a recent area of research focus is non-intrusive, single-ended voice quality monitoring and assessment. These techniques analyze the incoming speech using various algorithms which attempt to relate the quality of the speech to some subjective equivalent. This capability may be used to dynamically adapt the phone or network performance to improve voice quality, or to simply report the voice quality experience to the cellular operator for diagnostic purposes. Different approaches exist in the literature based on the ITU G.114 E-Model [8] and P.563 [9] methods. More work is needed to find simple but reliable solutions which are suitable for mobile devices which have limited resources if this capability is to be implemented however.

3. CONCLUSIONS

With the advancements in digital signal processing hardware, it is clear that voice processing engineers have more opportunity than ever to have a positive impact on voice quality. The ever shrinking cell phone is forcing voice processing engineers to use more sophisticated techniques to deliver more loudness yet provide better echo, noise and dynamic range control. However, the rollout of wideband and the continued evolution in voice processing will no doubt mean that one day soon wireless voice quality will exceed landline performance. Voice signal processing engineers will play a very important role in achieving this milestone in wireless voice.

REFERENCES

- [1] Texas Instruments, www.ti.com.
- [2] Hänslér, E. and Schmidt, G. Acoustic Echo and Noise Control: A Practical Approach. John Wiley & Sons, 2004.
- [3] Loizou, P., Speech enhancement: Theory and Practice, CRC Press, Boca Raton, FL, 2007.
- [4] Zölzer, U., Digital Audio Signal Processing, Wiley & Sons, 1997
- [5] Makino, S. et al, Blind Speech Separation. Springer, 2007.
- [6] Hoshuyama, O., Nonlinear Echo Suppression Technology Enabling Quality Handsfree Talk for Compact Equipment. NEC Journal Vol. 2 No.2/2007.
- [7] Guerin et al, Nonlinear Acoustic Echo Cancellation Based on Volterra Filters. IEEE Transactions on Speech and Audio Processing, Vol 11, No. 6, 2003.
- [8] ITU-T, 2005, G.107, 2005, E-Model, a computational model for use in transmission planning.
- [9] ITU-T, 2004, P.563, 2004, Single-ended method for objective speech quality assessment in narrow-band telephony applications.

WIDEBAND ECHO CONTROL CHALLENGES

Sylvain Angrignon, Chris Forrester, and Malay Gupta

Research in Motion, 295 Phillip St., Waterloo Ontario, Canada, N2L 3W8 sangrignon@rim.com

1. INTRODUCTION

The introduction of wideband speech coding in telecommunication results in a significant improvement in the overall user experience. The frequency components from 50 to 200 Hz add to the naturalness and the components from 3400 to 7000 Hz add to the intelligibility of the speech. However, this increase in bandwidth introduces many challenges in speech processing and in particular echo canceller design.

Many of the traditional echo canceller algorithms work well on narrow-band signals but may not be suitable for wideband applications. Furthermore, most of the echo canceller performance standards are based on level calculation and do not discriminate between frequency bands. Some of the factors which contribute to the challenge include the obvious increase in the adaptive filter length and thus reduction in convergence speed as well as a significant increase in the required memory and CPU usage. Other less tangible factors include user expectations and possible increase in the perception of residual echoes. All of these factors need to be considered when choosing an algorithm for wideband echo cancellation.

This paper presents an overview of some of the possible candidates for echo cancellation algorithms in wideband telecommunication.

2. TYPES OF ECHO

There are two types of echo sources; electrical and acoustic. Electrical echo is caused by the hybrid (2 to 4 wire converter) found in the public switched telephone network (PSTN). The electrical echo impulse response is typically very short (< 8 ms) and the echo return loss (ERL) provided by the PSTN hybrid is relatively high (> 10 dB). In general, the electrical echo impulse response remains stable during a call and can easily be handled by an electrical (or line) echo canceller (LEC).

Acoustic echo is present when the signal from the loudspeaker is picked up by the microphone. This echo is more noticeable when a device is used in speaker mode. In this case, an acoustic echo canceller (AEC) must be used to remove this acoustic echo. Contrary to hybrid echo, the acoustic echo impulse response can vary significantly during a call. The ERL can also change during the call so

the AEC must be able to adapt to the changing acoustic environment.

3. ACOUSTIC ECHO CANCELLATION METHODS

Acoustic echo cancellation algorithms use adaptive filters [1]-[4] to remove echo. Figure 1 shows a block diagram of a typical acoustic echo canceller algorithm. Echo is picked up by the microphone from a direct path and also from reflections on walls and other objects. For this reason, acoustic echo impulse response can be very high (> 100 ms) and represents a major challenge to AEC designers.

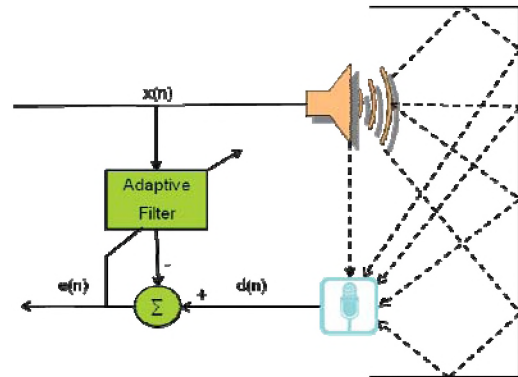


Figure 1. Acoustic Echo Canceller block diagram

The AEC input signal $d(n)$ contains the near-end speech as well as the acoustic echo. A double-talk detector (DTD) is needed to detect the presence of near-end speech and freeze the adaptation process in order to prevent the adaptive filter from diverging. The acoustic echo ERL can be very low (~ 0 dB) which produces a challenge for the DTD. A traditional level-based DTD is often not very effective for AEC.

There are many different types of algorithms used to implement the adaptive filter used by the AEC. In general, the AEC algorithms can be categorized as time-domain or frequency-domain.

4. TIME-DOMAIN AEC ALGORITHMS

In a time-domain adaptive filter AEC, an input signal $x(n)$ is filtered by an adaptive filter. The output from this filtering operation is subtracted from a desired signal $d(n)$ to produce the error signal $e(n)$. This error is used to update the filter coefficients. The filter update is typically done using the Least Mean Squares (LMS) algorithm.

5. FREQUENCY-DOMAIN AEC ALGORITHMS

Traditional time-domain adaptive filter algorithms adapt filter coefficients on every input sample. This adaptation process requires significant amount of DSP resources. In order to reduce complexity, some frequency-domain approaches process the input signal in “blocks” of samples.

5.1 BFDAF Algorithm

The Block Frequency Domain Adaptive Filter (BFDAF) algorithm diagram is shown in figure 2. In this algorithm, the input samples are accumulated in blocks of N samples. The block of samples is then converted to the frequency domain using the Fast Fourier Transform (FFT). Filtering is accomplished by a multiplication in the frequency-domain which is equivalent to performing a convolution in the time-domain. The filter coefficient adaptation can be done directly in the frequency-domain using a LMS based process.

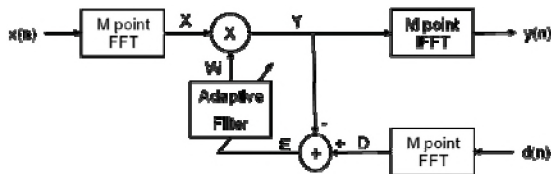


Figure 2. BFDAF AEC Block Diagram

There are two major drawbacks in the BFDAF algorithms. The first drawback is the delay introduced by the block processing. This delay can be significant as it is dictated by the length of the filter. The other significant drawback in the BFDAF algorithm is the speed of convergence since the adaptive filter coefficients are only updated once every N samples. To improve on the convergence speed and to reduce the processing delay, the multi-delay block frequency-domain adaptive filter algorithm [2] (MDF) was developed.

5.2 MDF Algorithm

The MDF adaptive filter algorithm diagram is shown in figure 3. In this algorithm, the input signal is accumulated into $m \cdot N$ blocks of samples. The size of the adaptive filter is controlled by the number of blocks (m) and the block length (N). This algorithm improves the speed of convergence significantly over the traditional BFDAF algorithm.

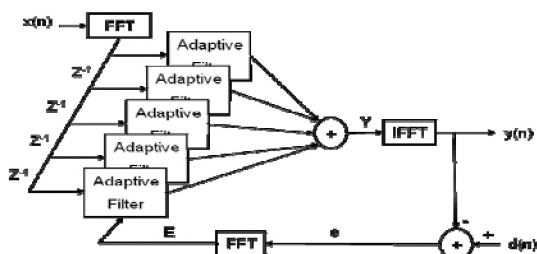


Figure 3. MDF AEC Block Diagram

Furthermore, the processing delay introduced by the MDF algorithm is smaller since the block length (N) is smaller than the BFDAF algorithm. The performance of the MDF algorithm is described in [2].

6. CHALLENGES

Acoustic echo cancellation in cellular telephony presents many challenges to DSP engineers. One of the challenges is the nature of the acoustic echo itself. The length of the adaptive filter (number of taps) required for the AEC can be very high, especially in wideband mode where the sampling frequency is 16 kHz.

Another important challenge that faces AEC designers is the environment in which the algorithm must operate. Mobile devices are often used in very noisy environment and the AEC algorithm must be robust to high levels of background noise. This also puts additional constraints on the design of the double-talk detector (DTD). This challenge is even more present in wideband telephony as more noise will be present along with the near-end speech.

The introduction of wideband devices on the market represents another important challenge for DSP algorithms developers. The first impact of wideband is the obvious increase in DSP resources (MIPS/memory) required by going from 8 kHz to 16 kHz sampling rate. The other impact of wideband is the user expectation. Residual echo artifacts will be more noticeable in a wideband cellular call. Wideband cellular users may also be less tolerant on echo artifacts as they are paying more for devices and will expect better voice quality.

7. CONCLUSIONS

This paper describes some basic acoustic echo canceller algorithms and challenges. It was shown that the MDF algorithm could be a good candidate to perform AEC in the context of wideband telephony. Furthermore, the MDF algorithm can be implemented to work for both narrowband and wideband mode by controlling the block length (N) at run-time. This paper aims to give algorithm developers an overview of challenges involved in the design of acoustic echo cancellation algorithms for cellular telephony.

REFERENCES

- [1] M. M. Sondhi, “An adaptive echo canceler,” Bell System Technical Journal, vol. 46, no. 3, pp. 497–511, 1967.
- [2] J.-S. Soo and K. Pang, “Multidelay block frequency domain adaptive filter,” IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 38, no. 2, pp. 373–376, 1990.
- [3] S. Haykin, “Adaptive Filter Theory”, Prentice-Hall, Englewood Cliffs, NJ, USA, 1996.
- [4] E. Hansler and G. Schmidt, “Acoustic Echo and Noise Control A Practical Approach”, John Wiley and Sons, 2004.

AN ANALYSIS OF LOUDSPEAKER DISTORTION IN THE CONTEXT OF ACOUSTIC ECHO CANCELLATION

Trevor Burton¹ and Rafik Goubran¹

¹Department of Systems and Computer Engineering, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, Canada, K1S 5B6

1. INTRODUCTION

Acoustic echo is inherent in all hands-free communication systems and can corrupt the conversation between parties if it is not sufficiently suppressed. The conventional approach to removing echo is with a digital echo canceller (EC), which is typically implemented as a linear adaptive filter. A block diagram of a simplified echo cancellation system is shown in Figure 1. Here the reference signal, $x(n)$, is the far-end signal that is played through the loudspeaker and the response signal, $d(n)$, is the signal captured by the microphone. The microphone signal is comprised of the echo, $y(n)$, local noise, $\eta(n)$, and local talker, $v(n)$, signals. The echo signal is formed from the direct loudspeaker to microphone signal along with the reflections from the walls and objects within the acoustic environment. The EC determines an approximation of the loudspeaker-enclosure-microphone-system (LEMS) transfer function via an adaptive filtering algorithm, and produces an echo signal estimate, $\hat{y}(n)$, that is subtracted from $d(n)$ to cancel the unwanted $y(n)$. The resulting error signal, $e(n)$, is then transmitted back to the far-end of the communication system. It should be noted that the EC is only adapted under quiet local talker conditions (i.e. $v(n)=0$). For practical acoustic echo cancellation (AEC), a doubletalk detector is used to determine if a local talker is active or not and controls adaptation of the EC accordingly.

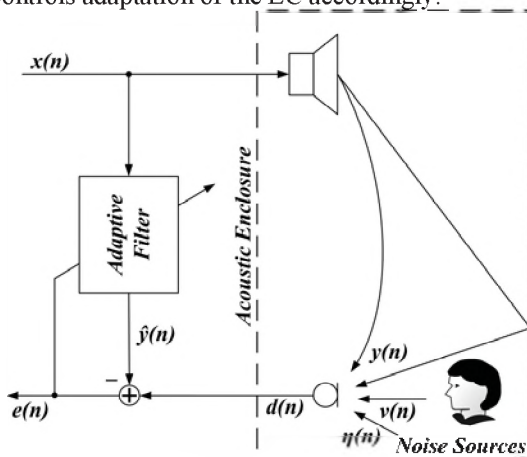


Figure 1 – Simplified echo cancellation system.

The performance of an EC is limited by many factors including undermodeling of the echo path, time variations and background noise within the hands-free environment, and by echo path nonlinearities [1]. Given that the EC accurately models the linear portion of the echo path and that the hands-free environment is stationary with an acceptable amount of background noise, the main factor

limiting its performance becomes the echo path nonlinearities. These nonlinearities include loudspeaker distortion, vibrations within the hands-free device, and amplifier saturation [1], [2]. Assuming that the amplifier in the hands-free device is not overdriven, the main source of echo path nonlinearity can become loudspeaker distortion which is effectively modeled by Volterra series expansions [3]. Thus, a nonlinear EC is required to prevent loudspeaker distortion from degrading the quality of hands-free communication between parties.

This work presents an analysis of loudspeaker distortion based on experimental measurements obtained from several hands-free systems under various operating conditions. The results of the analysis reveal trends in the frequency domain nature of the loudspeaker distortion, which provides insight into designing computationally efficient nonlinear echo cancellers.

2. LOUDSPEAKER MODELING IN AEC

To improve AEC performance compared to a linear EC, adaptive Volterra filters can be used to model the linear portion of the unknown acoustic system along with the nonlinear loudspeaker distortion. The general discrete time P -th order finite memory Volterra series expansion of an input signal, $x(n)$, is given by [3]:

$$y(n) = h_0 + \sum_{p=1}^P \left[\sum_{m_1=0}^{L_p-1} \sum_{m_2=m_1}^{L_p-1} \cdots \sum_{m_p=m_{p-1}}^{L_p-1} h_p(m_1, m_2, \dots, m_p) \right] \times x(n-m_1)x(n-m_2)\cdots x(n-m_p) \quad (1)$$

where h_0 is a constant term and $h_p(m_1, \dots, m_p)$ are the p -th order Volterra kernels or impulse responses that characterize the nonlinear system with memory lengths L_p . The biggest drawback to using adaptive Volterra filters is their very large complexity requirement, which can be prohibitive for practical AEC especially as the order of the Volterra series increases. Thus, reduced complexity nonlinear AEC structures based on adaptive Volterra filters are highly desirable, particularly with the movement towards wideband telephony systems. As wideband systems operate at a 16 kHz sampling rate, twice as many taps are required to model the same LEMS as a standard 8 kHz narrow band system.

3. EXPERIMENTAL RESULTS

In this section a harmonic distortion analysis is presented for a telephone set with a 2.5 inch loudspeaker and for a cellular telephone with a miniaturized loudspeaker,

that were both operated under high volume conditions in hands-free mode. The distortion data was collected by recording the microphone signal, under maximum preamplifier gain, at a 48 kHz sampling rate for various test tones between 50 Hz and 7 kHz for the telephone set, and between 300 Hz and 7 kHz for the cellular telephone. Each test tone was applied to the telephone set loudspeaker at a root-mean-square (RMS) voltage of 1.4 and to the cellular telephone loudspeaker at 1.1. Ten harmonics were included in the distortion calculation or as many that could fit up to the Nyquist rate for fundamental frequencies higher than 2.4 kHz.

As shown in Figure 2 the loudspeaker of the telephone set exhibits significant second harmonic distortion of up to 52% for low fundamental frequencies between 50 and 300 Hz. At frequencies beyond this range the second harmonic distortion decreases rapidly to much smaller values. Significant third harmonic distortion of up to 28% occurs as well for frequencies between 50 and 150 Hz, before decaying towards zero at higher frequencies. Fourth and higher harmonic distortion of up to 25% occurs for frequencies between 50 and 100 Hz and around 400 Hz.

The loudspeaker of the cellular telephone reaches up to 25% third harmonic distortion at low fundamental frequencies between 300 and 500 Hz. At higher frequencies the third harmonic distortion tends to be quite small with the exception of around 800 Hz and 3 kHz. Considerable second harmonic distortion of up to 55% occurs for fundamental frequencies near 1 and 1.5 kHz, and between 3 and 4 kHz. Only a small amount of fourth and higher harmonic distortion of up to 12% is displayed between 300 and 500 Hz and around 1.5 kHz.

Based on these harmonic distortion results, nonlinear AEC structures with adaptive Volterra filters up to the third order and possibly higher would be required to effectively compensate for the resulting nonlinear echoes under these high signal level conditions. As hands-free devices tend to be operated under high volume conditions, it is important that the underlying EC models and removes the resulting distortion to ensure a high quality of communication. With the increased signal bandwidth of wideband hands-free systems, the need for nonlinear ECs will become more essential.

One approach to reduced complexity nonlinear ECs based on Volterra filters is with subband adaptive filtering structures. In this case complexity is reduced by performing the adaptive filtering operations in multiple frequency bands at a reduced sampling rate, where shorter adaptive filter lengths are employed. Furthermore, a subband structure allows the frequency domain nature of loudspeaker distortion to be exploited. As shown in Figures 2 and 3 the loudspeaker distortion of typical hands-free devices is confined mainly to specific frequency regions, thus applying

adaptive Volterra filtering in only the correspondingly affected subbands would help to further reduce complexity.

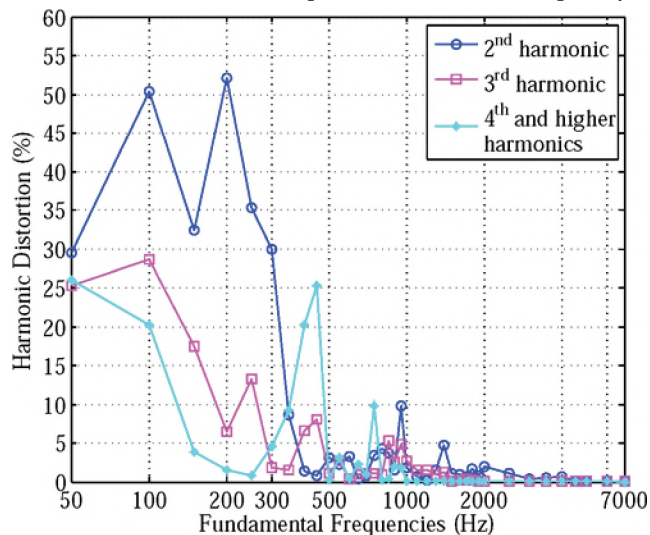


Figure 2 – Harmonic distortion for the telephone set.

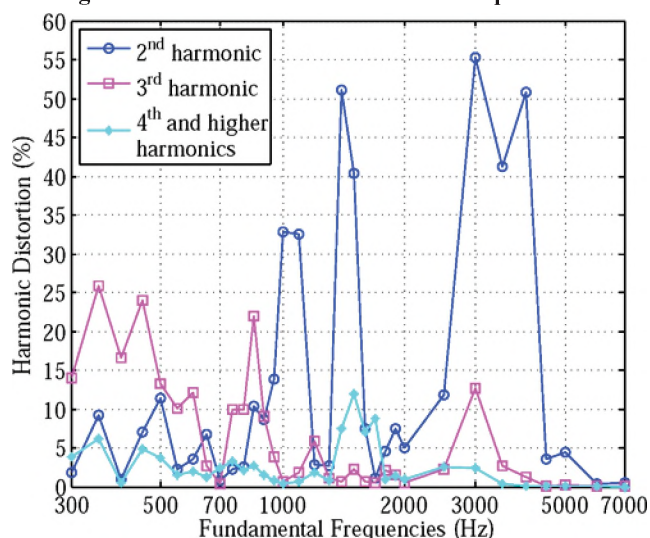


Figure 3 – Harmonic distortion for the cellular telephone.

4. CONCLUSIONS

This paper presented a distortion analysis for the loudspeakers of typical hands-free devices under high volume conditions. The results revealed that the distortion is confined to specific frequency regions, which provides insight into designing reduced complexity nonlinear ECs based on subband adaptive Volterra filters.

REFERENCES

- [1] A. N. Birkett and R. Goubran, "Nonlinear loudspeaker compensation for hands free acoustic echo cancellation," *IEE Electronics Letters*, vol. 32, no. 12, pp. 1063–1064, Jun. 1996.
- [2] A. Stenger, L. Trautmann, and R. Rabenstein, "Nonlinear acoustic echo cancellation with 2nd order adaptive Volterra filters," in Proc. IEEE ICASSP, vol. 2, Mar. 1999, pp. 877–880.
- [3] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*. New York, NY: Wiley, 2000.

SUBBAND AUTOREGRESSIVE MODELLING FOR SPEECH ENHANCEMENT

Brady Laska¹, Rafik Goubran¹, and Miodrag Bolic²

¹Dept. of Systems and Computer Engineering, Carleton University, Ottawa, Canada, {laska.goubran}@sce.carleton.ca

²University of Ottawa School of Information Technology and Engineering, Ottawa, Canada mbolic@site.uottawa.ca

1. INTRODUCTION

The speech waveform can be efficiently represented as an autoregressive (AR) process. Speech AR modelling is referred to as linear predictive coding (LPC) because the current speech signal sample $x(n)$ is represented as a linear combination of previous samples:

$$x(n) = \sum_{k=1}^p a(k)x(n-k) + u(n) = \mathbf{a}^T \mathbf{x}(n) + u(n), \quad (1)$$

where $\mathbf{a}^T = [a(1), \dots, a(p)]$ is the vector of autoregressive model coefficients, $\mathbf{x}^T(n) = [x(n-1), \dots, x(n-p)]$ is the signal vector, p is the model order and $u(n)$ is the excitation sequence. The use of AR models for speech signals has a physiological justification as $u(n)$ corresponds to the excitation from the lungs and the filter defined by the AR model coefficients corresponds to the all-pole vocal tract filter. The roots of the AR coefficient polynomial define the resonances of the filter which produce the characteristic formant peaks in the speech spectrum.

The parametric form of the AR model provides an efficient and low-variance representation of the speech signal spectrum. This allows for substantial compression gains in speech communications, and can also be applied to speech signal enhancement. Additive background noises - such as building ventilation system or in-car road noise - reduce speech intelligibility for human listeners and degrade the performance of automated speech and voice recognition systems. Speech enhancement algorithms attempt to remove the additive noise without distorting the desired speech signal. If the AR parameters of the clean speech signal are known, and the excitation and measurement signals are white Gaussian noise, Equation (1) can be arranged into a linear state-space form. This allows the Kalman filter equations to be used to obtain the minimum mean-square error (MMSE) optimal estimate of the clean speech waveform (Paliwal and Basu, 1987). By enforcing an AR model structure, Kalman filter speech enhancement provides high quality enhanced speech with natural sounding residual noise.

2. SUBBAND AR MODELLING

The rise of digital networks has enabled the emergence of systems transmitting wideband (16 kHz sampling rate)

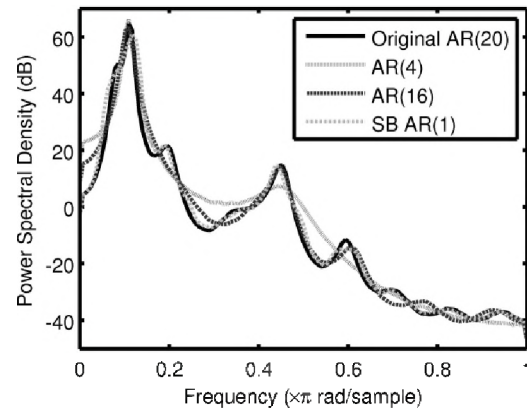


Fig. 1. Spectra obtained from fullband and subband AR modelling of an AR (20) process.

speech. While wideband speech improves the listener experience, it creates problems for AR speech modelling. The computational complexity of systems using AR speech models grows rapidly as the model order increases. Also, high-order models require more data to reliably estimate the parameters, but speech signals are only stationary over a short time. Low-order models are therefore desirable, however very high-order models are required to capture the pitch structure (Puder, 2006), and the deep troughs and steep slopes of the wideband speech spectrum.

An alternative to using a single high-order AR model is to use a filterbank to decompose the speech signal, and to model each subband channel with a very low-order AR model. Since the model parameters and energy level of each band are determined independently, subband AR models need not exhibit the same spectral smoothness of fullband models, and may permit better modelling of steep spectral slopes and troughs. Furthermore, as the processing of each subband signal is carried out at a decimated (time-reduced) rate, the number of computations per unit time may be decreased. It has been demonstrated (Rao and Pearlman, 1996) that with ideal filterbanks, subband AR models can achieve lower modelling error. Here we investigate the performance of realizable filterbanks.

To compare the performance of low-order fullband and subband AR modelling of complex signals, an AR(20) process was generated by passing white noise through an all-pole filter measured from a segment of voiced speech.

Fullband AR(4) and AR(16) models were fit to the signals, and an AR(1) model was fit to each band of a 16-band cosine modulated filterbank, designed using the approach in (Lin and Vaidyanathan, 1998). The resulting spectral estimates are shown in Fig. 1. The AR(4) curve shows that when low-order fullband AR modeling is used to estimate a higher order process, the smooth fitting of a curve between the poles causes the signal energy between the poles to be significantly over-estimated. While the AR(16) model is an improvement over the AR(4), the subband estimate still provides the closest fit. The root mean squared error between the true estimated signal spectra for the AR(4), AR(16) and subband AR(1) models are 6.56 dB, 4.16 dB and 3.32 dB respectively.

3. SUBBAND KALMAN FILTERING

In the context of Kalman filter speech enhancement, the over-estimation of the spectral troughs by the low-order fullband AR models leads to increased residual noise, as the noise between spectral peaks is treated as speech. If the residual noise is sufficiently far from a formant peak, it will not be masked by the formant and will be perceptually noticeable. This problem is more prominent in wideband speech where the spectral dynamic range within a speech segment is higher, and there can be high and low frequency energy in the same frame. The better spectral modelling provided by a subband AR model may therefore be exploited in Kalman filter speech enhancement to achieve better noise reduction, and has been shown to offer good sound quality (Puder, 2006). As an added benefit, the subband signal decomposition enables frequency-dependent processing strategies. Speech formant peaks are concentrated in the low-frequency regions, while the high frequency regions are spectrally flatter; using different model orders in different bands could provide a further complexity reduction without degrading the model estimate.

Simulations were carried out to compare fullband and subband Kalman filter enhancement of a speech signal degraded by white Gaussian noise at an overall SNR of approximately 10 dB. Fullband AR(8) and AR(16) configurations were tested as were two subband configurations: one using AR(1) models in all bands, and another using AR(1) models in the lower 8 bands and AR(0) models in the upper 8. Table 1 presents the signal to noise ratio (SNR) results and Fig. 2 presents the spectra of a voiced speech segment obtained from the enhanced output. The improved modelling of the spectral troughs allows the subband configurations to achieve higher noise suppression, especially in the higher frequencies.

4. DISCUSSION

Wideband speech signal spectra can possess multiple peaks and deep troughs within the same short time segment. This diverse spectral character motivates the use of frequency-dependent processing strategies such as

subband AR modelling. Kalman filter speech enhancement was used to demonstrate the potential benefits of this approach. In addition to offering better modelling performance at a lower complexity than the same order fullband model, subband systems allow the designer to vary the model parameters with frequency. A heterogeneous processing strategy using different model orders in high and low bands was shown to offer comparable noise reduction at a reduced complexity.

Table 1: Segmental and overall SNR scores.

Configuration	Segmental SNR	Overall SNR
Noisy Signal	0.96 dB	10.32 dB
Fullband AR(8)	5.21 dB	15.27 dB
Fullband AR(16)	5.28 dB	15.33 dB
Subband AR(1)	7.48 dB	18.25 dB
Subband AR(1/0)	7.46 dB	18.21 dB

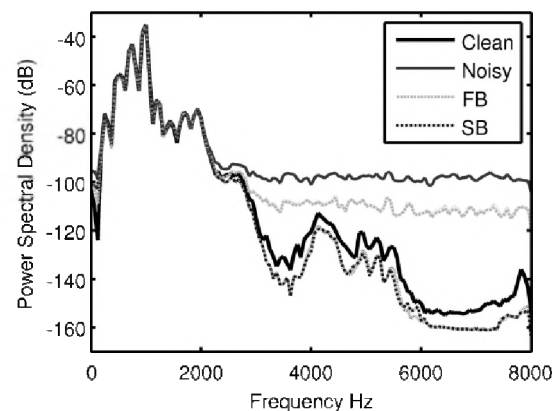


Fig. 2. Spectra of voiced speech segment from clean, noisy and fullband and subband Kalman filter enhanced signals.

REFERENCES

- Lin, Y. and Vaidyanathan, P. (1998). A Kaiser window approach for the design of prototype filters of cosine modulated filter banks, *IEEE Signal Proc. Letters*, vol. 5. no. 6. pp. 132-134.
- Paliwal, K., Basu, A. (1987). A speech enhancement method based on Kalman filtering. *Proc. IEEE ICASSP*, pp. 177-180.
- Puder, H., (2006). Noise reduction with Kalman filters for hands-free car phones based on parametric spectral speech and noise estimates. In: *Topics in Acoustic Echo and Noise Control*, Springer.
- Rao, S., Pearlman, W., (1996). Analysis of linear prediction, coding, and spectral estimation from subbands. *IEEE Trans. Info. Theory* 42 (4), pp. 1160-1178.

ACKNOWLEDGEMENTS

This work was funded in part by OGS, NSERC and Siemens AG.

LOUDNESS PREDICTION MODEL COMPARISON USING THE EQUAL LOUDNESS CONTOURS

Jeremy Charbonneau¹, Colin Novak², and Helen Ule³

¹⁻³Dept. of Mechanical Engineering, University of Windsor, Windsor, Ontario, Canada, N9B 3P4,
Email: charbo6@uwindsor.ca¹, novak1@uwindsor.ca², ule@uwindsor.ca³

1. INTRODUCTION

With the increase in demand for customer satisfaction, psychoacoustic metrics are playing a larger role in product development. For instance, the loudness level of a particular device could greatly influence a potential buyer's decision. Loudness can be described as the perceived intensity level of a noise source as compared to a reference level. Once calculated, the application of loudness can be broadened to calculate a variety of other perception models including roughness, annoyance and modulation. By understanding how these different characteristics of sound effect human perception, it is possible to apply the findings to a variety of related fields such as product health and safety.

When describing loudness levels, the ISO 226 standard for equal loudness contours is one of the most descriptive and well documented resources available. Updated in 2003, this standardized document represents a collaborative study describing the strong and weak points of loudness recognition in human hearing.¹ A plot of frequency versus sound pressure level indicates pure tones varying in amplitude and frequency as being equally loud across the hearing spectrum. Since the initial ISO 226 document produced in 1961 several loudness models have appeared to predict the loudness level of stationary noise sources.

This paper focuses on two of the few standardized methods for calculating steady-state loudness that are available. The first is a German standard, DIN 45631, based on an approach developed by Zwicker. The second metric studied is the American ANSI S3.4:2005 standard based on the work of Glasberg and Moore.²⁻³ The ISO 532B method for predicting loudness was also included in the initial investigation. However, due to the poor performance of the model it was neglected from this paper due to space restrictions.

The German organization Deutsches Institut für Normung (DIN) adopted a stationary loudness metric commonly referred to as the Modified Zwicker Method. Originally based on the 1975 ISO 532B document, the DIN 45631 approach calculates the loudness levels in a similar manner with minor modifications to the procedure. As an English version of this German document was not available at the time of the study, specific differences were not able to be identified by the authors of this study. It is known, however, that this modified version does improve upon the

loudness level calculations for the lower frequency ranges; a problematic area in the original ISO 532B document.

The second commonly used loudness metric is an American standard which has origins to work produced by S.S. Stevens. The ANSI S3.4:2007 standard was recently updated to better approximate human perception and is the only known model to match and account for the latest updates made to the ISO equal loudness contour document.

2. PROCEDURE

To compare the models under investigation, the ISO 226:2003 equal loudness curves were selected as reference values. This set of contours was selected due to their distinctive shapes across the frequency spectrum and general acceptance within the scientific community. As the equations contained within the ISO 226 document allow a high resolution plot to be generated, a comparison set of data was produced with each 3rd octave from 20 Hz to 12.5 kHz in the frequency spectrum. Equal loudness contour lines were then located from 20 to 90 phons representing the applicable range of the standard. To locate the respective curves of equal loudness pure tones could now be recorded and analyzed for each prediction model.

Loudness levels using the German standard DIN 45631 were determined directly using the commercially available Brüel & Kjær PULSE LabShop software. To control the signal quality, pure tones were generated using a B&K 3560 B-Frame and fed back directly into an input channel on the acquisition system. The raw data was recorded this way in order to minimize any external influences which might affect the quality of the signal and results. Pure tones were then generated by varying the voltage output at each center frequency until the target loudness level could be located, recorded and saved.

A standalone executable file from the University of Cambridge Auditory Perception website was used to derive the ANSI S3.4:2007 contours.⁴ The program named LOUD2006a accompanies the standard and was updated to correspond to the 2003 revisions to the ISO 226 contours. Within the program itself, pure binaural tones can be specified at any sound pressure level within the range of human perception. The program then takes into account the recording method specified (free field, binaural, ect.), and reports back the corresponding loudness level. For the

application used in this study, data was collected for in-harmonic tones recorded binaurally over the frequency spectrum using a free field application. Using this procedure the corresponding equal loudness contours were located and plotted as the DIN standard.

3. ANALYSIS

Once the equal loudness contours were located for each of the loudness models under study, a thorough comparison of the models could be completed between them with reference to the ISO 226 contours. As the ISO standard only applies between 20 phons and 90 phons, other contours were extrapolated using the given equations to help with the comparisons (dotted lines). In order to graphically show the differences, each model will be compared separately as follows.

As shown in **Figure 1**, the DIN 45631 tends to follow the reference curves quite well, particularly in the mid frequency ranges. However, both the lower end and higher end frequencies drop off when compared to the current ISO 226 standard. This result was expected as when this model was accepted, the current equal loudness contours were developed in 1987, having a much shallower slope. At the time, the DIN 45631 standard performed exceptionally, matching the equal loudness contours almost entirely. Due to space restrictions this plot was not included.

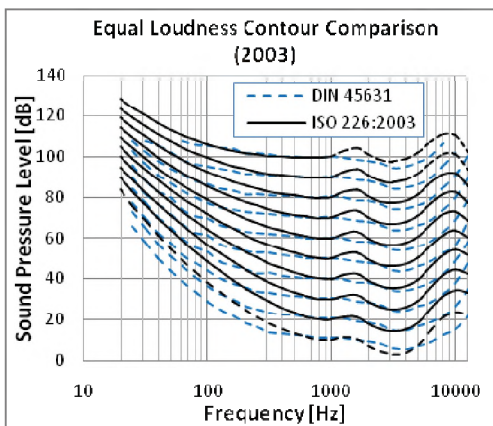


Figure 1: DIN 45631 Equal Loudness Contours compared to ISO 226:2003

As the ANSI S3.4:2007 model was updated to the recent changes of the ISO document, the contours tend to follow the reference curves more consistently. As seen in **Figure 2**, the ANSI S3.4:2007 was able to follow the trends of the present standard with minimal drop-offs. Above 1 kHz, the contours drop slightly but still reasonably follow the characteristics of the ISO 226:2003 standardized curves. As with the DIN standard, this model is unable to completely describe the hump in the curves between 1000 Hz and 2000 Hz. A reason for this may be that this particular characteristic was not as exaggerated in the 1987 version of

the contours. Like the DIN method, this standard was initially fit to the equal loudness contours of ISO 226:1987. When compared to the older contours, the upper frequencies for the ANSI standard match well to the standardized curves, particularly below 80 phons. Although the higher frequencies are slightly shifted from the ISO 226:2003 curves, the ANSI derived equal loudness contours overall perform the best out of the two calculation procedures.

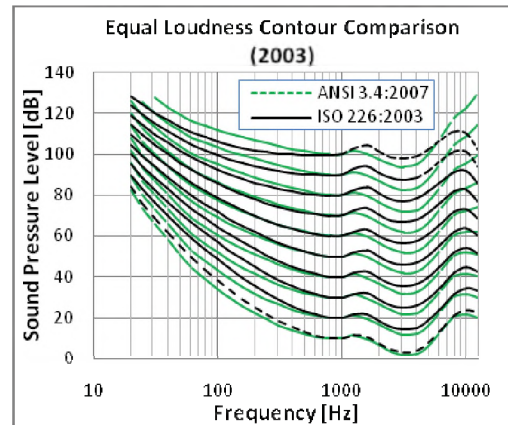


Figure 2: ANSI S3.4:2007 equal loudness contours compared against ISO 226:2003

4. CONCLUSION

For each of the stationary loudness metrics under study, equal loudness contours were produced and compared thoroughly. Each model was compared against the set of reference contours as described in the ISO 226:2003 document. From this comparison, it was apparent that the ANSI S3.4:2007 stationary loudness prediction model correlated the best with the updated ISO experimental data, outperforming the other models in all areas across the applicable frequency spectrum. As the ANSI best approximates the reference contours, it would therefore be recommended for use when determining the loudness of stationary noise sources. The DIN 45631 model for predicting loudness does an accurate job of estimating the older, obsolete equal loudness contours of ISO 226:1987. However, based on the updated data this model should not be used for very high frequency tones.

5. REFERENCES

- ¹ ISO226:2003. "Acoustics -- Normal equal-loudness-level contours," Stage: 90.93 (2008-12-22), *International Organization for Standardization*, Standard, Geneva, Switzerland. (2008).
- ² DIN 45631. "Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum. Verfahren nach E. Zwicker." (1990).
- ³ ANSI S3.4-2007. "American National Standard Procedure for the Computation of Loudness of Steady Sound," *American National Standards Institute*, (Acoustical Society of America, 2007).
- ⁴ B.R. Glasberg and B.C.J Moore. "Loudness model calculated according to ANSI S3.4-2007," *Auditory Perception Group University of Cambridge*, accessed April 09 2009, (2006), <http://hearing.psychol.cam.ac.uk/Demos/demos.html>.

COMPARISON OF NON-STATIONARY LOUDNESS RESULTS TO EQUAL LOUDNESS CONTOURS

Colin Novak, Jeremy Charbonneau, Helen Ule

Dept. of Mechanical Engineering, University of Windsor, 401 Sunset, Ontario, Canada, N9B 3P4

1. INTRODUCTION

The increased interest and use of sound quality metrics has led to the refinement of existing calculation techniques as well as the development of new metrics. The most fundamental psychoacoustic metric is the loudness model for stationary sounds for which over the past 30 years several internationally accepted models have been introduced and standardized including ISO 532B and the similar German DIN 45631. Another is the ANSI 3.4:2005 Glasberg and Moore based loudness metric. For the most part, these approaches have demonstrated good correlation to human perception for steady sounds. However, many sounds of interest do not fall into the category of steady state which necessitates the development of loudness calculation methods that are capable of characterizing the loudness of unsteady sounds.

More recent studies have developed alternative loudness models for use with non-stationary sounds. Given the inherent complexity of evaluating the abilities of these models to correlate to human perception for unsteady sounds, a first comparison should be done using steady sounds as an unsteady model should work equally well with such sounds. This study used a methodical and scientific approach to evaluate the performance of two such models, the proposed German DIN 45631/A1 and the Glasberg and Moore unsteady model, to accurately predict reference steady loudness values and compare these to the ISO 226 equal-loudness contours. The goal of this study was to evaluate any differences between the DIN 45631/A1 and Glasberg and Moore approaches.

2. CALCULATING NON-STATIONARY LOUDNESS

For the calculation of unsteady loudness, two relatively common methods exist. The proposed German DIN 45631/A1 method is loosely based and an extension of what is commonly referred to as the Zwicker approach. The method for calculating stationary loudness developed by Eberhart Zwicker (1) provides the foundation for the standardized loudness calculation specified in ISO 532 (1975). This was procedure was later improved at the low frequency end of the spectrum in the DIN 45631 standard for stationary noise. Eventually this standard was expanded

upon and now provides the basis for the draft DIN 45631/A1 for calculating unsteady loudness.

A second approach has been developed by Moore and Glasberg (2,3) which uses a method which they called long term integration. For this method a waveform is used as the input into the transfer function of the outer and middle ear followed by the calculation of the short term spectrum via six parallel FFTs. The excitation pattern is then calculated followed by its transformation to a specific loudness pattern. Finally the area under the specific loudness pattern is determined for the overall loudness.

3. APPROACH

Both of the above described methods are very different approach but purport to achieve the same end result of outputting a value for non-stationary loudness which correlates with human perception. For this to be an accurate statement, a first test used in this study was to evaluate the non-stationary approach using stationary sounds. For this, pure tone sinusoidal signals were used as the stationary input which served to further simplify things. From this, the resulting loudness values can be compared to the equal loudness curve specified by ISO 223:2003 (4).

The method used in this study was similar to that used in a previous investigation (5). The pure tones inputs which were used in the above referenced stationary study were inserted into each of the time varying loudness models. This also gave the advantage that the resulting loudness values could be directly compared to the results of the stationary loudness investigation.

The DIN model took in the signals directly as part of the commercially available Bruel & Kjaer Sound Quality software. For the Glasberg and Moore (G&M) model, a program available on their website was used. For this, 100 dB full scale sinusoids were recorded and scaled down to match the total sound pressure levels found in the stationary loudness test (ie 1000 Hz @ 80dB was derived by scaling down the 100 dB sinusoid by subtracting 20 dB). The lower levels used a 50 dB full scale sinusoid in order to prevent the loss of important information in the scaling process.

It was then assumed that any deviations between the models could be extended back to the equal loudness contours for comparisons.

4. RESULTS

Given in Figure 1 are the differences in loudness of the DIN unsteady results compared to the DIN steady calculation for the steady pure tone inputs across the frequency and level range. Also given is the same only for the G&M (indicated as the ANSI TVL Model) unsteady approach compared to the comparable ANSI S3.4:2007 steady method. For a perfect match, the resulting “difference” lines would be straight and horizontal.

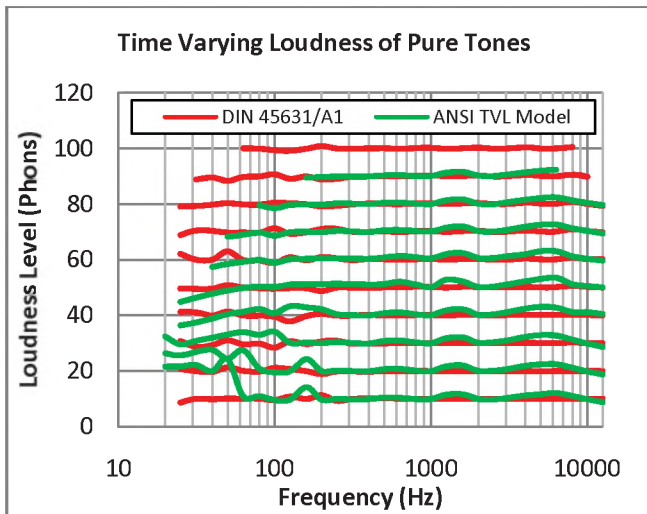


Figure 1: Loudness Difference between DIN (red) and G&M (green) Unsteady and Steady calculations for Pure Tone Input Signals.

Inspection of Figure 1 shows good correlation for both unsteady models when compared to their respective steady models at frequencies above 100 Hz with the DIN being moderately better.

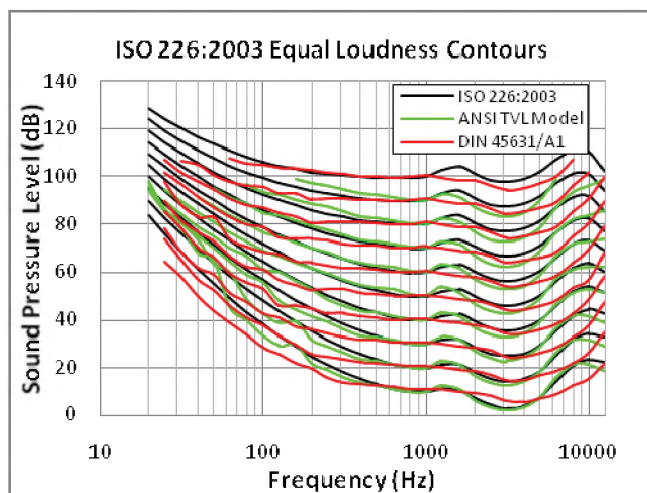


Figure 2: DIN and G&M unsteady Loudness plotted against the ISO 226:2003 Equal Loudness Curves.

Illustrated in Figure 2 are the calculated DIN 45631/A1 and Glasberg & Moore unsteady loudness results plotted against the ISO 332:2003 equal loudness curves. While the purpose of Figure 1 was to illustrate how well the unsteady calculations of steady tones correlated with the steady calculation methods, the results shown in Figure 2 provides insight on how well the unsteady results compare to standardized human perception.

Again, it can be seen in Figure 2 that good correlation between both methods and the equal loudness contours is realized. Both models performed less ideally below 100 Hz, particularly at lower levels. It should also be noted that the G&M model followed the equal loudness curves in the 1 kHz to 5 kHz range. This is particularly important as this is a very important frequency range for sound quality analysis. It is due to this that one may conclude that the G&M approach would be better for cases where human perception is important, as usually is the case for psychoacoustic studies.

5. CONCLUSION

This study set out to evaluate the performance of two models for the calculation of loudness for non-stationary sounds. These models were the proposed German DIN 45631/A1 and the Glasberg and Moore unsteady model. The results of these models were compared to the results using the same input signals into the respective stationary calculation models as well as to the ISO 226:2003 equal-loudness contours. The goal was to quantitatively evaluate any differences between the two approaches. Using a signal generator, a full spectrum of pure tone input signals were fed into each of the loudness models over a large range of sound pressure levels. Using this input matrix, plots were presented to facilitate the comparison. On a common plot, minor discrepancies to the equal-loudness contours are identified with the Glasberg and Moore model showing slightly better performance in some frequency ranges. Both models did not perform well at frequencies below 100 Hz.

6. REFERENCES

1. Zwicker, E., Erratum: “Procedure for calculating loudness of temporally variable sounds”. *Journal of the Acoustical Society of America*, 63(1), 283.
2. Moore, B. (2003). Temporal integration and context effects in hearing. *Journal of Phonetics*, 31, 563-574.
3. Glasberg, B., Moore, B., (2002). A model of loudness applicable to time-varying sounds. *Journal of the Audio Engineering Society*, 50(5), 331-342.
4. ISO226:2003. “Acoustics -- Normal equal-loudness-level contours,” Stage: 90.93 (2008-12-22), *International Organization for Standardization*, Standard, Geneva, Switzerland. (2008).
5. Charbonneau, J., Novak, C., Ule, H., (2009). Comparison of Loudness Calculation Procedure Results to Equal Loudness Contours. *Inter-Noise 2009*.

PREDICTION OF PSYCHOACOUSTIC METRICS FROM SUSPENSION INDUCED VIBRATION AND ROAD INDUCED NOISE USING TRANSFER PATH AND PSYCHOACOUSTIC ANALYSIS TECHNIQUES

Nebojsa Radic, Colin Novak¹, and Helen Ule

ABC Air Management Systems Inc., 51 Rexdale Blvd., Toronto, Ontario, Canada, M9W 1P1

¹Dept. of Mechanical Engineering, University of Windsor, 401 Sunset, Ontario, Canada, N9B 3P4 novak1@uwindsor.ca

1. INTRODUCTION

In terms of noise generation, the automobile is simply a set of different systems that when excited at specific frequencies will eventually lead to the creation of noise. This statement was of course also true in the early days of automobiles, however, it was always taken to be a secondary issue that was simply accepted since more important factors had to be addressed. Both technology improvements and legislative advancements have since led to the evolution of the modern automobiles and the development of new performance targets, including noise.

Today, automakers invest significant time and money in research and development associated with the reduction of vehicle noise pollution. Since automakers are also more aware of the importance of the perception of noise, vibration and harshness (NVH) emissions, there is also an increased focus on the sound quality of vehicle cabin noise. Consumers now also demand safer and more comfortable vehicles, especially given the significant increased use of cellular phones, entertainment and interactive voice controls in vehicles. As part of this, it is recognized that the investigation and performing of psychoacoustics analysis of vehicle cabin acoustics is essential in the improvement of today's vehicles.

A significant source of unwanted cabin noise is the result of road-induced excitation of the vehicle suspension which propagates into the vehicle cabin. Keeping this in mind, the main objective of this study was to evaluate the transmission paths of the excitation energy into the vehicle cabin from road induced noise and vibration using frequency response functions. Finally, the objective was to establish a correlation between the noise and vibration measurements taken outside of the vehicle to noise and psychoacoustic observations inside the vehicle. These included Loudness, Roughness and Fluctuation Strength as well as the A-weighted sound pressure level.

2. EXPERIMENTAL DETAILS

For this investigation, acoustic pressure measurements were taken inside a 2004 Chevrolet Epica cabin at the driver's left ear location using conventional microphones as well as at

the passenger's ears position with a binaural head for the evaluation of the resulting sound quality. A microphone was also located outside the car near the front wheel. Acquisition recording time was 15 seconds per run. The testing was done at the Brüel & Kjær Application Research Center (ARC) in Canton, Michigan, USA within a semi-anechoic room equipped with a 4WD dynamometer.

For vibration acquisition, one accelerometer positioned on the wheel hub and was stationary for the entire experiment. A second accelerometer was moved between tests, initially located on the lower A-arm and later moved to the top of the McPherson strut. Six Brüel & Kjær Type 2635 and Type 2626 amplifiers were used to condition the two 3-axis accelerometers.

Data acquisition was performed during motored and driven conditions. For both cases, the following operation conditions were considered:

- Idling/Ambient
- Steady speeds (20, 40, 50, 60 & 80 km/hr)
- Acceleration run-up from 0 to 80 km/hr

3. Data Analysis

The ten acquired time signals were used for the determination of the frequency response functions (FRF) and coherence between the pressure and acceleration excitations from outside of the vehicle and the sound pressure obtained inside the cabin with the intent to predict several psychoacoustic metrics. Estimation of the relationship between the input and the output of the system was performed using Matlab. For this process, each of the ten input signals needed to be correlated to each of the three output signals. Two different trials were also considered at six different steady speeds. Additionally, two different driving conditions were considered which gave frequency response functions and coherence functions for correlation between the sound pressure from outside and sound pressure inside the vehicle cabin.

For the first case, in order to generate the frequency response and coherence functions between any input and output signal, several steps were required. First, the two signals of interest were selected and the auto-spectrum and

cross-spectrum of each of them are found using narrow band spectrum analysis. The FFT window type was then selected as well as the window overlap and the number of FFT lines. To reduce computational time, the frequency band was specified in this case to be limited to 1000 Hz. The cross-spectrums were calculated in the similar matter making sure that all of the selected parameters match the ones used for auto-spectrum calculations to be able to compute the Transfer Functions, FRFs and Coherence values. A Matlab function was also used to estimate the transfer function of the system with input A and output B using Welch's averaged periodogram method. Coherence was also estimated this way.

4. RESULTS

Testing was done at six different steady speeds. These were all performed with the conditions of the vehicle being driven and motored. As this provided hundreds of different FRFs and coherence functions, discussions are limited to a few examples to illustrate the main points.

From the results of the FRF calculation procedure, it was found that both the microphone and accelerometer stimuli can be used to obtain pressure inside the car. Using this information, the psychoacoustic metrics of interest were then predicted. The following Figures 1 through 4 show the agreement between the original and calculated levels at the passenger's right ear for all metrics for the motored car. The condition of the 60 km/hr motored vehicle is taken arbitrary simply to illustrate the point of the typical results. For this speed, the predicted mean levels of A-weighted sound level are within 2 dB with respect to the original data. This represents a variation that cannot be distinguished by human perception. Roughness and fluctuation strength showed excellent agreement as well as shown in the figures.

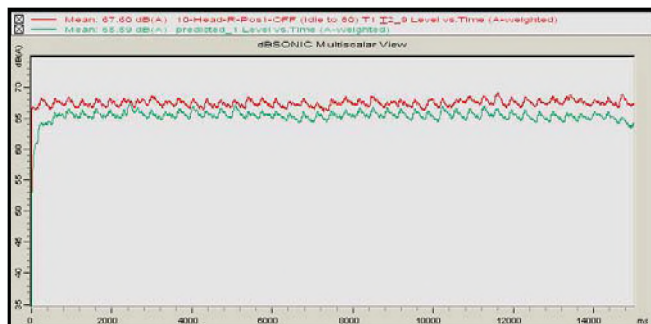


Figure 19: Original (Red) and Predicted (Green) A-weighted SPL for Motored Car at 60 km/hr.

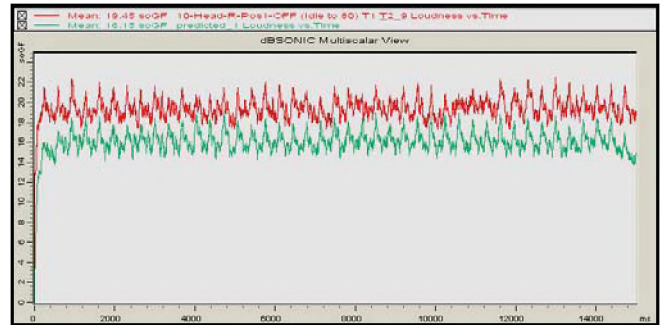


Figure 20: Original (Red) and Predicted (Green) Loudness for Motored Car at 60 km/hr.

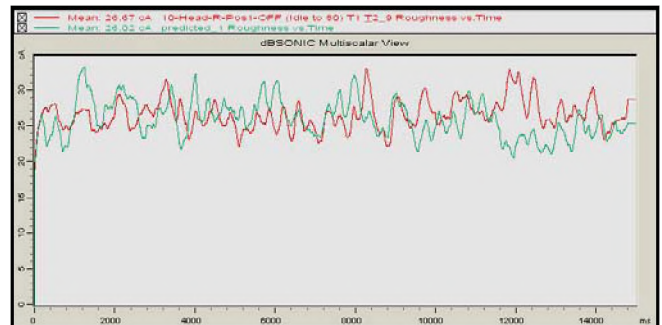


Figure 21: Illustration of Original (Red) and Predicted (Green) Roughness for Motored Car at 60 km/hr.

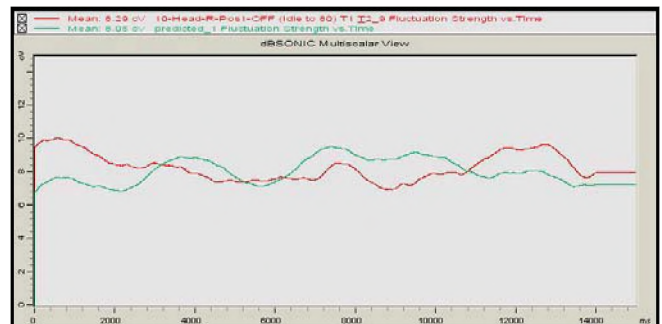


Figure 22: Original (Red) and Predicted (Green) Fluctuation Strength for Motored Car at 60 km/hr.

5. CONCLUSIONS

An attempt was made to establish a correlation between the noise and vibration measurements from the outside of the vehicle to the noise and psychoacoustic observations inside the vehicle. This was proven to be possible with some inherent limitations. Direct prediction of the sound quality metrics inside the vehicle from both acceleration and sound pressure observed only outside of the vehicle cabin did not show compatible results to the measured data. However, it was proven possible to predict the sound pressure for which the psychoacoustic metrics could be calculated indirectly.

DESCRIPTION OF THE MULTIPLE LOOK APPROACH FOR CALCULATING UNSTEADY LOUDNESS

Helen Ule¹, and Colin Novak²

¹Dept. of Mechanical Engineering, University of Windsor, 401 Sunset, Ontario, Canada, N9B 3P4 ule@uwindsor.ca

²Dept. of Mechanical Engineering, University of Windsor, 401 Sunset, Ontario, Canada, N9B 3P4 novak1@uwindsor.ca

1. INTRODUCTION

Human perception of the quality of sound of a product or noise source is important because it is tied to human comfort. Therefore, it is important for design engineers to design various industrial products related to automotive, consumer electronics, computer, etc. so that sound quality is enhanced. The ongoing work presented in this paper is focused on the development and improvement of the specific psychoacoustic metric of loudness such that it can better correlate with real human perception. This metric characterizes the strength attribute of the human auditory response in terms of the perception of sounds from quiet to loud and has become a very important sound quality metric for the acoustic evaluation of product noise. However, no calculation method for determining the loudness of sounds which vary with time has been successfully correlated with all experiments involving human perception. The fundamental goal of this research is to develop an approach which will calculate the time varying loudness using the multiple-look model which is thought to be able to better process human perception characteristics not identified in other proposed models.

2. CALCULATING UNSTEADY LOUDNESS

For the calculation of unsteady loudness, two methods exist. The first is referred to as a Zwicker base approach since it is an extension of the stationary calculation method developed by Eberhart Zwicker (1) and provides the foundation for the loudness calculation method given by ISO 532 (1975). This method has been expanded upon and provides the basis for the draft DIN 45631/A1 for calculating unsteady loudness. A second approach has been developed by Moore and Glasberg (2) which uses a method which they called long term integration.

For the Zwicker based approach, the loudness is based on the distribution of the specific loudness along the critical band scale only here they are treated as time dependant values. For each of the critical bands, the effect of temporal masking is accounted for by applying corrections based on post-masking data developed by Zwicker. It has been found

that in the presence of impulsive signals, the apparent loudness decreases with the duration of the impulse. This is accounted for in the Zwicker approach. Finally, a network summing all specific loudness values along the critical bands is applied in order to get the overall loudness value.

Glasberg and Moore's temporal integration approach involves combining information over time which they say is to improve detection or discrimination over the Zwicker approach and is done in 5 steps. First, the sound is sampled at a rate of 32 kHz and put through filter representing the function of the outer and inner ear. Next, the excitation pattern, or specific loudness, is calculated from a fast fourier transform of the filter signal. The specific loudness is then integrated to get the instantaneous loudness. Finally it is averaged to get the short term and long term time varying loudness.

3. JUSTIFICATION OF MULTIPLE LOOK APPROACH

The classical approach to the calculation of time varying loudness is temporal integration, often called temporal summation, which involves the combining of information over time to improve detection and is often thought of as a simple accumulation process such as energy integration (2). However, recent studies (3,4,5) suggest there is a better approach to the calculation of time varying loudness which is to consider the combination of audible information from multiple independent looks (2).

Viemeister and Wakefield (4) studied the effect of signal duration on detection and discrimination thresholds of sound pressure levels. From experiments, they have shown that pulse tones with temporal separations greater than 10 ms are processed independently which subsequently draw a conclusion that the classical view of long-term integration is not an acceptable method for the characterization of signals containing gaps which are often prevalent in a non-stationary noise signal. Instead, the results are consistent with the idea that the input signal is sampled at a fairly high rate and each sample, or "look" is stored in memory for selective access and processing. It is this idea that provides the foundation for the multiple-look model as a more accurate perceptual representation, particularly in the presence of temporal gaps.

Moore (2) has performed various experiments which demonstrate that, "Information extracted from one part of a sound may influence the evaluation and interpretation of information extracted from another part, at a different time." These results not only support the conclusions of Viemeister and Wakefield, but also do so through more rigorous experimentation. The outcome of this work suggests that the temporal integration which occurs for unsteady sounds cannot be characterized by a basic summation or accumulation process, but instead supports the approach of the multiple-look concept.

The work done by Pedersen (16) on the various stages of analyzing time varying sounds is also of importance. Pedersen was able to demonstrate that different types of temporal processing are involved in hearing, each of which are dependent on the specific task given to the listener. The result of this work reinforces the reasoning as to why long-term integration techniques do not correlate with the experiments of Viemeister and Wakefield. More importantly, this work yields further justification to seek alternative approaches such as the multiple-look model proposed in this research plan.

While algorithms and computer programs have been proposed for time-varying loudness, a version which utilizes the multiple-look concept has not been realized. It is anticipated that the creation of such an algorithm would provide the ability to determine the loudness of an unsteady signal which would include the ability to discriminate varying thresholds due to short-term and burst temporal effects.

4. METHODOLOGY OF MULTIPLE LOOK APPROACH

To overcome the shortcomings of the present unsteady loudness calculation techniques, the multiple look method is based on a short time constant process of 2 to 3 ms where decisions are based upon selected "looks" and can in principle account for the temporal resolution for modulation detection, gap detection, etc. Another important point of the multiple-look model is that it allows for intelligent processing of the "looks". As these are stored in memory, they can be selectively used for further processing and decision making.

While it is believed that the multiple look approach will overcome the issues that the other methods have, the present research is focused on the application for gap detection. The following is the proposed approach:

The time signal is first acquired in short step durations of 2 ms for which an FFT is applied for each stepped duration, as well as for longer, or multiple step durations, to account for low frequency components of the sound. Next the sampled signal has been put through the response filters representing the inner and outer ear functions. From here, the specific

loudness of the signal can be calculated for each of the small duration segments, or "looks", which are sampled at a high rate. The program then stores the vector of "looks" in temporary memory and is put through standardized computations and comparisons. This is analogous to how non-varying sounds are compared to values in "look-up" tables for the calculation of stationary loudness in the present ISO 532B method, only here, it is done for extremely short duration signals in very quick succession. The treatment of the "looks" is dependent on the task, and for this study, gap detection is targeted. As such, using the 2 ms loudness, it must be decided if a gap is present by looking at three successive samples for immediate drops in level. If no gap is detected, the process returns to the beginning of the signal and repeats for up to 100 ms and apply long term integration unless a gap is detected before 100 ms. If a gap is found to exist, the program modifies the threshold detection for loudness of the 3 sampled steps of data. Finally with this long term integration is applied to the eventual 100 ms of signal.

5. CONCLUSION

This paper described an alternative approach to calculating time varying loudness using a multiple look model. This approach involves the dissection of a time signal into small duration segments which are sampled at a very high rate. These multiple looks are then processed independently and are thought to be able to better process the perception characteristics, specifically gap detection, not identified in the temporal integration model traditionally used. While this work is still in progress, the algorithm being developed is hoped to proven to result in the calculation of loudness with better correlation to attributes of human perception.

6. REFERENCES

1. Zwicker, E., Erratum: "Procedure for calculating loudness of temporally variable sounds". *Journal of the Acoustical Society of America*, 63(1), 283.
2. Moore, B. (2003). Temporal integration and context effects in hearing. *Journal of Phonetics*, 31, 563-574.
3. Glasberg, B., Moore, B., (2002). A model of loudness applicable to time-varying sounds. *Journal of the Audio Engineering Society*, 50(5), 331-342.
4. Viemeister, N., Wakefield, G. (1991). Temporal integration and multiple looks. *Journal of the Acoustical Society of America*, 90(2), 858-865.
5. Pedersen, B. (2006). *Auditory Temporal Resolution and Integration – Stages of Analyzing Time-Varying Sounds*. Doctoral dissertation, Department of Acoustics, Aalborg University.
6. Pedersen, B., Ellermeier, W. (2005). Temporal and spectral interaction in loudness perception. *Journal of the Acoustical Society of America*, 117(4), 2397-2407.

TOOL FOR PREDICTING TRANSMISSION OF AIR-BORNE AND STRUCTURE-BORNE SOUND IN BUILDINGS

J. David Quirt, and John K. Dickinson

National Research Council, Institute for Research in Construction, Ottawa, K1A 0R6, Canada

1. INTRODUCTION

In recent years, the science and engineering for controlling sound transmission in buildings have shifted from a focus on individual assemblies such as walls or floors, to a focus on performance of the complete system. A calculation framework based on extensive experimental studies of lightweight wood-framed constructions has been developed, and a first design guide¹ was published in 2006. This paper presents an overview of the design context and the basic features of a new internet-based acoustical application intended to make the design process more effective.

For decades, North American building codes, have considered only the rating for the assembly separating adjacent dwellings—Sound Transmission Class (STC) for airborne sources—as if sound were transmitted only through the obvious separating assembly. In reality, the problem is more complex. An airborne sound source excites all the surfaces in the source space causing them to vibrate in response. Some of this vibration is transmitted across the surfaces abutting the separating assembly, through the junctions where these surfaces join the separating assembly, and into surfaces of the adjoining space, where part of the transferred energy is radiated as sound. Occupants of the adjacent space actually hear the combination of sound due to direct transmission through the separating assembly and any leaks, plus sound due to structure-borne flanking transmission involving all the other elements coupled to the separating assembly. For design or regulation, the terminology to describe the overall sound transmission including all paths is well established; the ASTM descriptor for system performance including flanking is Apparent Sound Transmission Class (ASTC). While measuring the ASTC in a building is quite straightforward, predicting the ASTC due to the set of transmission paths in a building is quite complex, and requires data on structure-borne transmission that is only gradually becoming available.

A model for air-borne and structure-borne transmission must account for all five factors indicated in Figure 1. Framed assemblies are anisotropic and highly damped – the vibration field exhibits a strong gradient that is different in the directions parallel and perpendicular to the joists. In general, this vibration field is a poor approximation of a diffuse field, which limits the applicability of simple SEA models. Not only are vibration levels strongly attenuated across the surface of the structural assembly, but also some added surface layers (such as concrete floor toppings)

change the attenuation across the structural assembly, with different changes in the three orthogonal directions pertinent to direct and flanking transmission.

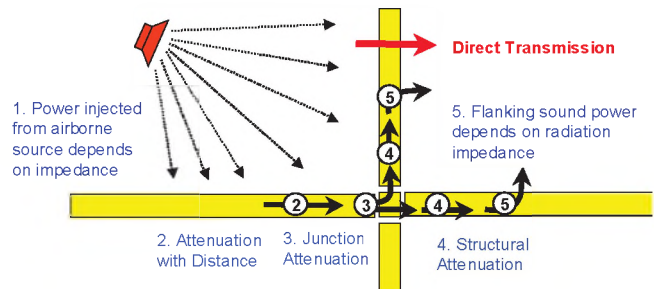


Figure 1: Five factors that affect flanking transmission, with an airborne source for the paths involving the floor surface in the source room. Similar factors apply for all other paths.

2. THE DESIGN GUIDE

A simplified guide for design of wood-framed buildings was developed¹, using a tabular approach to present alternative choices for all the surfaces likely to be significant to the overall sound transmission between adjacent spaces. It presents tables of ASTC and AIIIC ratings for sound transmission between units that are side-by-side, or one above the other, for common constructions. These were based on results from a multi-year experimental study.

Separating wall:	Basic Wall (STC 52)	Better Wall (STC 57)	
Gypsum board on side walls:	Direct or resilient	Direct resilient	
Floor Topping:	(Apparent-STC)		
Basic OSB subfloor	43	43	43
Second layer of 19-mm OSB	48	50	50
•	•	•	•
•	•	•	•
38 mm gypsum concrete + resilient mat on subfloor	51	53	55

Figure 2: Simplified example from Guide¹ for adjacent one-level apartments. Sidewalls abutting the separating wall also transmit sound, but resilient channels supporting the gypsum board ceiling block transmission via the ceiling/ceiling path.

The tabular approach discussed above does show the effect of changes to the surfaces controlling sound transmission—both the separating assembly and the key flanking surfaces

(hence indicating potential places to make improvements), and it also provides ASTC estimates for designers. Because tables are readily presented in conventional technical documents, distribution to builders and their generalist designers was effective. But there are obvious limitations:

- Each table (like Figure 2 above) applies to one specific combination of wall and floor constructions; therefore, many tables were required.
- Tabular form does not readily support comparison of different designs, or show relative strength of direct and flanking transmission paths in each case (to highlight which surfaces limit performance).
- A table can present only a few variants on possible elements such floor toppings, or floor coverings, or gypsum board type and attachment on flanking surfaces.

The obvious means to display more choices for each of the component materials—and to facilitate a more detailed analytic approach—is to implement the calculation framework in software, linked to a database of sound transmission data for each path, for the matrix of construction options that have been experimentally characterized. Implementing this as an internet-based tool (See Figure 3) should facilitate distribution, version control, and periodic updating and expansion of the database. A prototype screen image is shown in Figure 4, to illustrate the potential of such tools to provide acoustical performance estimates in an accessible form.

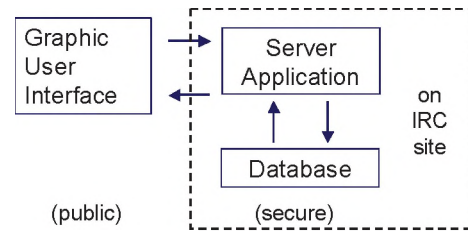


Figure 3: Conceptual structure for internet-based application

3. SUMMARY AND REFERENCES

This paper provides a terse overview of how experimental characterization of the direct and flanking sound transmission paths in wood-framed construction leads to a manageable set of path transmission terms (which depend on the specific construction details). By combining an interactive interface with a calculation framework, a new internet-based tool presents predicted energy transmitted via all paths in an intuitive and user-friendly form that supports informed design decisions.

We acknowledge collaboration with NRC for this extension of the flanking guide by FPI Forintek, Owens Corning, and Lauzon Floors, and also long term support for this activity by CMHC, Marriott International, Trus Joist, and USG.

1. J.D. Quirt, T.R.T. Nightingale, F. King, Guide for Sound Insulation in Wood Frame Construction, **RR219**, NRC-IRC Canada, (2006)

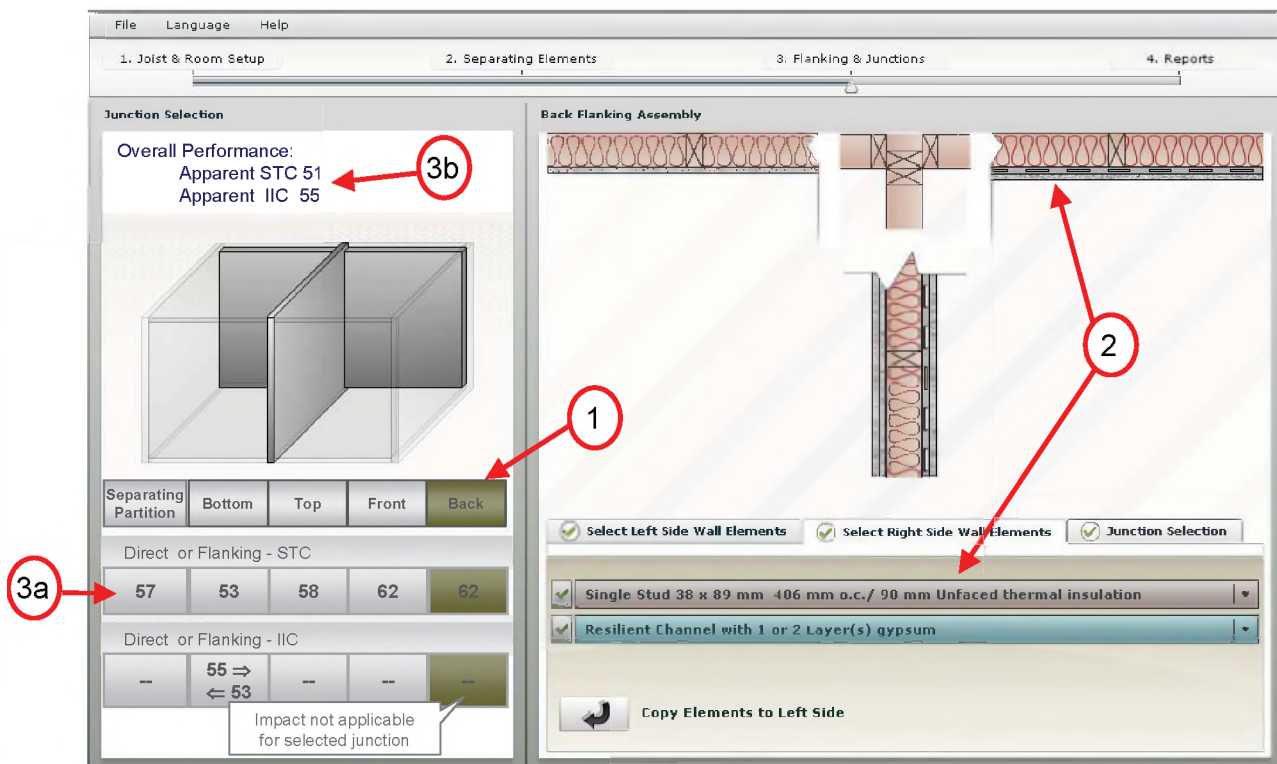


Figure 4: Example of user interface that displays sound transmission estimates for flanking and direct paths, to guide design decisions. Parts of the interface include: (1) buttons to select between the separating assembly or each of the four flanking junctions at its edges, (2) drop down menus to select details of framing and other components affecting transmission via the selected junction, (3a) calculated sound transmission ratings for each set of paths, (3b) calculated overall sound insulation estimate.

A COMPARISON OF THREE METHODS FOR THE IN SITU DETERMINATION OF ACOUSTIC ABSORPTION COEFFICIENTS

Scott Mallais, and John Vanderkooy

Audio Research Group, Dept. of Physics and Astronomy, University of Waterloo, 200 University Ave. West, Waterloo, Ontario, Canada, N2L 3G1 smallais@uwaterloo.ca jv@uwaterloo.ca

1. INTRODUCTION

This paper compares three methods used in order to determine the absorption coefficient of a surface *in situ*. The first two methods are well known, using a reflection method [1] and a subtraction technique [2] to separate the incident and reflected sound waves of the impulse response. The third method [3] utilizes two measurements, the first at a surface under study and the second at a rigid surface at the same location in order to calculate the absorption coefficient.

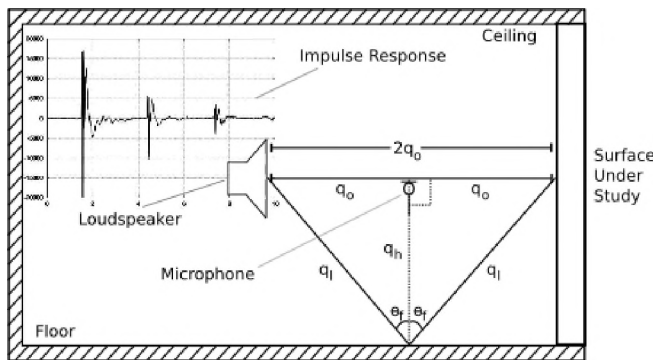


Fig. 1. Configuration of the reflection method. A microphone is placed halfway between a loudspeaker and a study surface. The inset is a typical impulse response showing the incident and two reflected signals.

2. IN SITU MEASUREMENT METHODS

2.1 Reflection Method

The reflection method is described in [1,4] and shown in Figure 1. The incident and reflected sound waves, p_I and p_R , are modeled as spherical waves. This results in the following absorption coefficient [4]

$$\alpha = 1 - \rho = 1 - \left| \frac{3p_R}{p_I} \right|^2 \quad (1)$$

where α and ρ are respectively the absorption and reflection coefficients for sound intensity. This development assumes that the centre of radiation of sound is at the baffle of the loudspeaker. In general the sound is radiating from the acoustic centre, δ , which is important at low frequencies.

2.2 Subtraction Technique

The subtraction technique is described in [2]. Two pressure measurements are determined as shown in Figure 2, a) p_I , and b) p_{IR} . The absorption coefficient may then be calculated by

$$\alpha = 1 - \left| \frac{p_{IR} - p_I}{p_I} \right|^2 \quad (2)$$

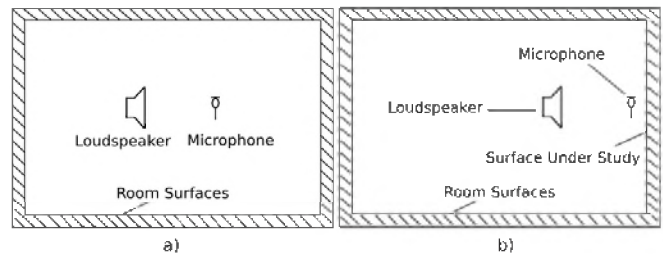


Fig. 2. Configuration of the subtraction technique. a) Measurement away from room surfaces. b) Measurement at the surface under study.

2.3 Surface Pressure Method

The surface pressure method was introduced in [3]. The *in situ* implementation of this technique is discussed in [5,6]. A rigid surface is approximated by mounting a thin steel sheet against the surface under study. Two measurements are taken, one directly on the surface under study, p_S , and the second similar measurement with an intervening reflective sheet, p_R . The configuration is similar to Figure 2 b). Approximations are discussed in [5] which lead to the absorption coefficient of the surface,

$$\alpha = 1 - \left| 2 \frac{p_S}{p_R} - 1 \right|^2 \quad (3)$$

3. RESULTS AND DISCUSSION

Three surfaces were measured: a resonant surface (wood panelled wall), an absorptive surface (office divider) and a rigid surface (glazed block wall). The results are shown in Figures 3 and 4.

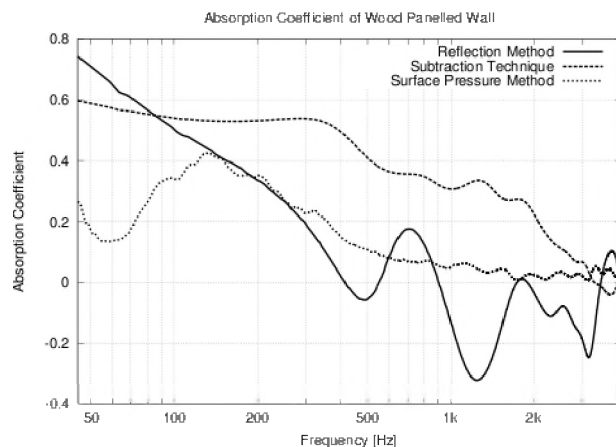


Fig. 3. Absorption coefficient of wood panelled wall. This wall should have little absorption besides a mass-air-mass resonance in the low frequency region (100-200 Hz).

In general, both figures show that each method yields different results. The reflection method and the subtraction technique have worse resolution than the surface pressure method. The former methods use approximately 2-3 ms of data whereas the last method uses 64 ms of data.

In particular, the data in Figure 3 should show a mass-air-mass resonance in the 100-200 Hz region. This is clearly seen using the surface pressure method, however this is not seen in the data for the other methods, since they do not have sufficient frequency resolution. Large differences and oscillations are apparent for the reflection method and subtraction technique. At low frequencies, this is caused by the truncation of the impulse response. At high frequencies the reflection method may suffer from diffraction effects caused by the loudspeaker. Note that the spherical wave assumption in the reflection method will not be valid at high frequencies. The subtraction technique is difficult since small variations in loudspeaker-microphone distance can result in an incomplete subtraction. Diffraction from the edges of the sheet may influence the surface pressure method.

Interestingly, the absorption coefficients for the office divider are much closer than for the wood panelled wall (see Figure 4). The results diverge at low frequencies due to frequency resolution limitations. Only the reflection method and surface pressure method were used on the glazed block wall. Results from the surface pressure method indicate that this surface should have zero absorption over the whole frequency range. However, the reflection method yields results that are negative over most of the frequency range of interest. The reflected pressure is greater than the incident pressure in this case (after multiplication by three, see (1)). Diffraction effects from the loudspeaker and the supporting structure are also suspected in these measurements. More details on the surface pressure method and other *in situ* methods is given in [6].

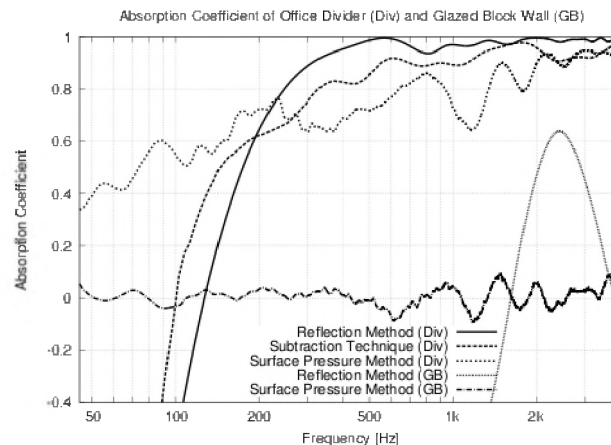


Fig. 4. Absorption coefficient of an office divider (three upper curves) and a glazed block wall (two lower curves). The office divider should have very high absorption at high frequencies and a reduction in absorption as frequency is decreased. The glazed block wall should have little absorption at all frequencies. This is confirmed by the measurement results.

In conclusion, we have compared three different methods that may be applied *in situ* to determine the acoustic absorption coefficient. The first two methods are quick to implement, however require the separation of incident and reflected sound waves in the time domain. The surface pressure method requires a thin sheet to be placed in front of the surface under study in order to create a rigid boundary. No windowing is necessary, and therefore the limitations on the frequency resolution are not the same as in the former measurement methods. A calibration measurement could aid the accuracy of this method.

REFERENCES

- [1] Garai, M. *Measurement of the sound-absorption coefficient in situ: The reflection method using periodic pseudo-random sequences of maximum length*. Applied Acoustics, vol 39, 119-139, 1993.
- [2] Mommertz, E. *Angle-dependent in situ measurements of reflection coefficients using a subtraction technique*. Applied Acoustics, vol 46, 251-263, 1995.
- [3] Ingård, U. and Bolt, R. H. *A free field method of measuring the absorption coefficient of acoustic materials*. JASA, vol 23, 509-516, 1951.
- [4] Mallais, S. *In situ determination of acoustic absorption coefficients*. AES 125th Convention, San Francisco, USA, 2008.
- [5] Mallais, S. and Vanderkooy, J. *In situ measurements of acoustic absorption coefficients using the surface pressure method*. AES 127th Convention, New York, USA, 2009.
- [6] Mallais, S. *In situ measurements of acoustic properties of surfaces*. Master's thesis, University of Waterloo, 2009.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the Natural Sciences and Engineering Research Council of Canada (NSERC) for research funding. Discussions with other members of the Audio Research Group: Ryan Matheson and Professor Stanley Lipshitz, have been very helpful.

LOW-FREQUENCY ROOM DEMERIT ANALYSIS AND EQUALIZATION USING PROPERLY-MODELED SOURCE TERMS IN FINITE-DIFFERENCE TIME-DOMAIN SIMULATIONS

Ryan J. Matheson

¹Dept. of Physics & Astronomy, University of Waterloo, 200 University Ave. West., Ontario, Canada, N2L 3G1
ryan.j.matheson@gmail.com

1. FDTD Source Term Derivation

In deriving the Finite Difference Time Domain Method (or FDTD Method) for Acoustics we require two essential equations. The first is the Linear Inviscid Force Equation, or Newtons' Equation,

$$\nabla p = -\rho \frac{\partial \mathbf{u}}{\partial t} \quad (1)$$

The second is the Equation of Continuity,

$$\frac{1}{\rho c^2} \frac{\partial p}{\partial t} = \frac{q(\mathbf{r}, t)}{\rho} - \nabla \cdot \mathbf{u} \quad (2)$$

where p is the sound pressure, \mathbf{u} is the particle velocity vector, ρ is the density of the medium and $q(\mathbf{r}, t)$ is the function that defines the rate of creation of fluid. Normally this equation is quoted without the function q , but if one carefully follows the FDTD derivation method laid out in [1] then it can be shown that at a point the equation for pressure in the FDTD scheme can be given by,

$$p(x, T) = p(x, T - k) + k \cdot c^2 \cdot q\left(x, T - \frac{k}{2}\right) - \frac{c^2 \cdot \rho \cdot k}{h} \left[u_x\left(x + \frac{h}{2}, T - \frac{k}{2}\right) - u_x\left(x - \frac{h}{2}, T - \frac{k}{2}\right) \right] \quad (3)$$

Equation (3)¹ contains the function q which is the rate of creation of fluid in the system, which has units of $\text{kg} \cdot \text{m}^{-3} \cdot \text{s}^{-1}$. But we want a relation that contains the volume velocity function Q , which has units of $\text{m}^3 \cdot \text{s}^{-1}$. Therefore we can relate q and Q by multiplying q by a volume and dividing it by the density of the medium. Thus

$$Q = \frac{\Delta V}{\rho} \cdot q \quad \text{or} \quad q = \frac{\rho}{\Delta V} \cdot Q \quad (4)$$

Here $\Delta V = \Delta x \cdot \Delta y \cdot \Delta z = h^3$. Substituting (4) into (3) yields,

$$p(x, T) = p(x, T - k) + \frac{k \cdot c^2 \cdot \rho}{\Delta V} Q\left(x, T - \frac{k}{2}\right) - \frac{c^2 \cdot \rho \cdot k}{h} \left[u_x\left(x + \frac{h}{2}, T - \frac{k}{2}\right) - u_x\left(x - \frac{h}{2}, T - \frac{k}{2}\right) \right] \quad (5)$$

We require the use of Q in this equation because conventional cone loudspeakers produce volume velocity, and at low frequencies are approximately point sources. This is exactly what equation (5) is, a pressure point source that is driven by a volume velocity.

It should be noted that the pressure at a distance r away from a volume velocity point source is given by

$$p(r) = \frac{\rho \cdot A(t)}{4 \pi \cdot r}$$

where $A(t)$ is the volume acceleration, or $dQ(t)/dt$. This result is observed in the simulations.

2. Cancellation Methods

2.1 End-Wall Speaker Placement

One of the easiest ways to equalize low frequencies in a room is to adjust the speaker placement such that fundamental modes of vibration in the room are not excited. For example, the corner of a room is a good way to excite all modes of vibration in a room. A possible placement scheme for a stereo setup in a room with dimensions width(W) x length(L) x height(H) would be to place the two speakers at ($W/4, L, H/2$) and ($3W/4, L, H/2$). At very low frequencies, in reasonable sized listening rooms in which L is the largest dimension, this speaker placement has the effect of creating a plane wave along the length of the room[2].

2.2 Rear Cancellation

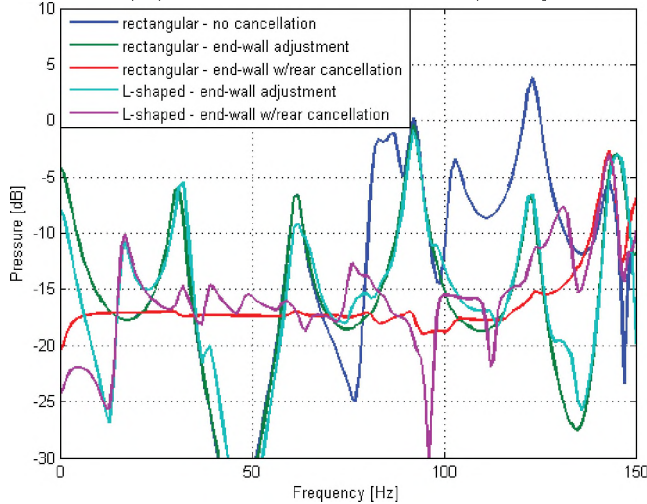
An active method of low frequency room equalization is the rear cancellation method. If we take our example low-frequency speaker placement in the last sub-section, ($W/4, L, H/2$) and ($3W/4, L, H/2$), then what we would want to do is place an identical pair of speakers at the opposite end of the room at ($W/4, 0, H/2$) and ($3W/4, 0, H/2$). These rear low-frequency speakers would then be fed an inverted and delayed version of the signal that would be sent to the front speakers.

2.3 Cancellation Results

As can be observed from the figure below, we see that the end-wall adjustment helps equalize things a bit because we are no longer exciting the lowest modes that correspond to the W dimension. This does not affect the lowest frequency modes that are attributed to the L dimension though, which causes the largest low frequency resonances. The L -dimensional modes are taken care of by the rear cancellation method rather well as we observe that we have removed all of the modes below roughly 130Hz; where the troublesome low frequency resonances occur.

¹ Equations (3) and (5) are the 1-D case.

Comparison of Responses for Cancellation Methods in Rectangular(4.2x5.6x2.4m) and L-shaped(5.0x2.8x2.4m subroom added onto main room) Listening Rooms



2.4 Cancellation of L-shaped rooms

Simulating and equalizing rectangular listening rooms is relatively easy but most rooms are not perfectly closed-off rectangles. Common rooms have doorways, windows or hallways. Outside doorways and windows are openings that open up to the larger outside world so these will not add any additional resonance modes to the room. On the other hand, closed-off hallways add an extra dimension to the room that creates coupled modes of vibration that, typically, adds an additional low frequency resonance mode. This additional mode can be seen in the above figure as the lowest mode for the L-shaped room plots, and it is not affected by our cancellation techniques.

3. Room Demerit

By visually comparing the room responses in the figure it is easy to tell which responses are more desirable. But all we can say from the plot is that the curves with cancellation look better than the ones without because they look smoother. This is a qualitative measure. But we all know from high school science that quantitative measures are more defensible. Thus we present two possible ways to calculate a 'Room Demerit' value which are variations of the demerit found in [3].

$$D1 = \frac{1}{N} \sum_{n=1}^N \left[\left| \frac{P_{unsmoothed}[n] - P_{smoothed}[n]}{P_{smoothed}[n]} \right| \right] \quad (6)$$

$$D2 = \frac{1}{N} \sum_{n=1}^N \left[\frac{(P_{unsmoothed}[n] - P_{smoothed}[n])^2}{P_{smoothed}[n]^2} \right] \quad (7)$$

In these calculations each mic uses its own smoothed result. Another approach might be to use the same smoothed result for all mics.

In most cases where smoothing is involved people like to use octave smoothing. This is done because at low frequencies it can be argued, based on how people hear in

the bass region, that a constant bandwidth smoothing is more appropriate, a bandwidth of 60Hz over the range from 0 to 100Hz was chosen that produced the results in the following table.

Table of Results from Room Demerit Calculation of 25 mics.

	rectangular - no cancellation	rectangular - end-wall adjustment	rectangular - w/rear cancel	L-shaped-end-wall	L-shaped-w/rear cancel
Mean(D1)	1.0437	1.1750	.0772	1.0314	.5610
Mean(D2)	2.5454	3.7271	.0133	2.5206	.8731
STD(D1)	.1060	.0495	.0106	.0465	.0427
STD(D2)	.9943	.5385	.0023	.4161	.1782
SDOM(D1)	.0106	.0049	.0011	.0046	.0043
SDOM(D2)	.0994	.0539	.0002	.0416	.0178

When performing any sort of room equalization technique it is best to compare results over several mic positions. History has shown that equalizing one area very easily un-equalize another. Thus the demerit data is collected over 25 microphones covering a 4m² area in our simulated room, and the mean demerit value is computed from these mics. We also know that from statistics that when using the mean to assess a distribution it is also helpful to calculate the Standard Deviation(STD) and the Standard Deviation of the Mean(SDOM). The STD and SDOM provide a measure of how consistent the distribution is about the mean value and is of interest to us because when we equalize the room we not only want a flat response, but we also want to ensure that the response is consistent across the listening room, so that anyone sitting anywhere in our room will receive the same performance as everyone else, thus attempting to change our 'acoustic sweet spot' into an 'acoustically sweet area'.

REFERENCES

- [1] S. Krarup Olesen, "Low Frequency Room Simulation using Finite Difference Equations", presented at the 102nd AES Convention, 1997 March 22-25, Munich, Germany. Paper 4422
- [2] A. Celestinos and S. Birkedal Nielsen, "Low Frequency Sound Field Enhancement System for Rectangular Rooms using Multiple Low Frequency Loudspeakers", presented at the 120th AES Convention, 2006 May 20-23, Paris, France. Paper 6688
- [3] J. Vanderkooy, "Multi-Source Room Equalization: Reducing Room Resonances", presented at the 123rd AES Convention, 2007 October 5-8, New York, NY, USA. Paper 7262

NOISE SUPPRESSION IN CELLULAR TELEPHONY

Malay Gupta, Chris Forrester, and Sean Simmons

Research in Motion, 295 Phillip St., Waterloo Ontario, Canada, N2L 3W8 mgupta@rim.com

1. INTRODUCTION

Noise suppression in cellular telephony is one of the several measures (ITU-T P.330) taken to improve the speech quality and intelligibility at the far end. With cellular/mobile phones increasingly becoming outdoor devices, noise suppression has become an integral part of the audio signal processing chain used in cellular telephony today. The term speech enhancement, which in principle has a broader scope, has been used synonymously with noise suppression and has been an active area of research since the late 1960's. Suppressing various types of background acoustic noise without adversely impacting the speech quality makes this problem particularly challenging.

Examples of noise environments that a typical cell phone has to encounter include: home, office, car, street, café etc. A noise suppression algorithm is desired to enhance speech quality/intelligibility in these environments and is expected to operate reliably in a broad range of stationary/non-stationary signal-to-noise-ratio (SNR) environments ranging from -3 dB to 30 dB (ITU-T G.330). With worldwide mobile sales beating almost any optimistic forecast, a good noise suppression algorithm is one of the key factors to keep voice quality high and meet customer requirements.

2. NOISE SUPPRESSION METHODS

The goal of noise suppression system is to aid speech encoder by presenting it a signal with reduced background noise (BGN) level in comparison to the original noisy signal. According to the number of microphones used in the noise suppression system, we can broadly group noise suppression techniques into two groups. Group 1 utilizes a single-microphone and Group 2 utilizes an array of microphones. In the next two subsections we briefly discuss these techniques and refer to some of the important references in the present context

2.1 Single Microphone Noise Suppression

Most of the techniques proposed in this category function in the frequency-domain and use the short-time analysis-synthesis technique. These techniques apply a frequency-dependent gain function to the spectral components of the noisy speech to attenuate the components with greater noise content. This gain function could be a simple thresholding function or could be a much more sophisticated perceptually motivated spectral weighting

function. These methods collectively belong to the class of spectral subtraction or short-time spectral amplitude (STSA) estimation methods [1]-[2].

With the changing BGN environment in cell phone applications, the most critical component of these algorithms is the estimation of BGN power spectrum. There are two distinct set of approaches towards BGN spectrum estimation. The first approach makes use of an explicit voice activity detector (VAD) to update noise spectrum. However, VAD based techniques are difficult to tune and become unreliable in low-SNR and non-stationary noise conditions. The other approach is based on minimum statistics and probabilistic update [3] of noise spectrum has been proposed for improved speech quality.

Further improvements in single microphone noise suppression were reported in 1995 [4] with the introduction of subspace processing in noise suppression applications. These methods decompose the noisy speech subspace in to two subspaces, one corresponding to the noise-plus-speech subspace and one corresponding to the noise subspace. Noise suppression is performed by removing the noise subspace and estimating the signal from the remaining subspace. Extensions of subspace algorithms to non-stationary and non-Gaussian noise cases can be found in [5].

2.2 Multi-Microphone Noise Suppression

Significant progress in digital signal processor (DSP) technology has enabled multi-microphones based noise suppressors in cell phones. The main advantage of having multiple microphones is their ability to utilize spatial dimension in addition to temporal dimension to effectively suppress the BGN. The popularity of multi-microphones in cellular telephony appears to be growing with several dual-microphone noise suppression solutions already available in the marketplace.

A common goal of these techniques is to optimally combine the multi-sensor data to achieve improved noise suppression. The procedure is more commonly known as the beamforming. Beamforming has been a well studied area of signal processing for over three decades [6]; in particular, the delay-and-sum (DS) beamformer and the adaptive beamformer (ABF) are the most commonly used techniques suitable for BGN suppression purposes. ABF generally performs better than the DS, but requires *a priori*

information about the *look direction*, and *speech absence* interval for ABF coefficient update.

Recently there has been a lot of interest in microphone array processing algorithms that do not require such *a priori* information. These approaches, under the assumption of linear speech/noise mixing are also known as blind source separation (BSS) and have been used for speech separation applications [7]. These approaches are based on one of the following criteria, 1.) Statistical independence of speech and noise, 2.) Signal diversities present in speech, e.g., time-frequency diversity, 3.) Temporal characteristics of speech, and 4.) Non-stationarity of speech signal. The techniques employing independence criteria are particularly popular and are collectively known as the Independent Component Analysis (ICA) techniques.

3. CHALLENGES

Single microphone solutions are easy to understand and low in computational complexity. However, most single-microphone solutions proposed report musical noise phenomenon associated with the processed speech. Musical noise manifests itself as randomly occurring pure tones in the processed speech and is very unpleasant to listen to. None of the solutions proposed so far claim complete elimination of musical noise in low SNR and non-stationary noise scenarios, despite the huge amount of research already been devoted to the problem. Hence effective noise suppression with a single microphone is still a challenge.

Multi-microphone systems have the potential of better noise suppression. In theory, a large array with many microphones is required for effective noise suppression either by beamforming or BSS methods. In beamforming for a fixed number of microphones, a larger array results in narrower mainlobe beamwidth, and for a fixed array length, more microphones result in lower sidelobe levels in the beampattern. Similarly, with BSS the system must have at least as many microphones as the number of sound sources for effect separation of the desired speech source from noise sources. For a fixed array geometry and fixed number of microphones, the dependency of beampattern on the signal frequency further complicates the design of such systems. Another limitation of these techniques is the heavy dependency on the application environment and the nature of the noise field. There have been recent performance studies of dual microphone systems under different microphone configurations and varying noise scenarios including directional and diffuse noise conditions in [8]-[9]. It remains, however, to study the amount of gain that can be obtained by (on-line) beamforming or BSS methods with small number of microphones under test conditions/signals specified by ETSI in ETSI EG 202 396-1 v1.2.3 (2009-03).

3.1 Performance evaluation

Assessing speech quality and intelligibility of the speech processed by the noise suppression system in the presence

of different types of noise environments is an important but difficult task. A standardized set of test cases is one of the most critical aspects in noise suppression algorithm evaluation. Telecommunication standardization bodies like ITU/ETSI/TIA all have recommended test cases for evaluating the noise suppressor performance and its impact on speech quality. The noise suppression performance is generally measured in terms of ambient noise rejection (ANR) value as described in ETSI TS 126 132 v4.1.0 (2001-09). Tests for algorithm impact on loudness ratings can be found in ITU-T P.79. Processed voice quality subjective and objective guidelines can be found in ITU-T P.800/P.835 and P.862 respectively. Recommendations for acceptable algorithmic delays, convergence time, active speech level changes etc, are described in 3GPP TS 26.077 V4.0.0 (2001-03) and ITU-T G.114. Artificial test signals are also described in ITU-T P.50 in order to reduce the variability associated with real speech. Note that most of the above mentioned test cases involving noise sources assume noise source stationarity. The development of test cases suitable for non-stationary noise sources is still underway and new cases are constantly being added (ITU-T P.835 Amendment 1: New Appendix III) to existing test cases.

4. CONCLUSIONS

This paper describes the current state-of-the-art of noise suppression algorithms suitable for cellular telephony. An overview of operating principles behind both single-microphone and multi-microphone systems was provided. Associated challenges along with performance guidelines were discussed. This paper aims to give algorithm developers an overview of challenges involved in the design of noise suppression algorithms for cellular telephony.

REFERENCES

- [1] Boll S. F., Suppression of acoustic noise in speech using spectral subtraction. IEEE Trans. Acoustics, Speech, Signal Processing, vol. 2, 113-120, 1979.
- [2] Ephraim Y. et al., Recent advancements in speech enhancement. The Electrical Engg. Handbook, CRC Press, 2004.
- [3] Cohen I., Noise spectrum estimation in adverse environments: Improved minima controlled recursive averaging. IEEE Trans. Speech, Audio Processing, Vol. 11, 466-475, 2003.
- [4] Ephraim Y., et al., A signal subspace approach for speech enhancement. IEEE Trans. Speech, Audio Processing, Vol. 4, 251-266, 1995.
- [5] Loizou, P., Speech enhancement: Theory and Practice, CRC Press, Boca Raton, FL, 2007.
- [6] Krim H., et al., Two decades of array signal processing research. IEEE signal processing magazine, vol. 13, no. 4, 1996.
- [7] Makino S., et al., Blind Speech Separation, Springer, 2007.
- [8] Kocinski J., Speech Intelligibility improvement using convolutive blind source separation assisted by denoising algorithms., Speech Communication, vol. 50, no. 1, Jan. 2008.
- [9] Chen J., et al., Performance evaluation of adaptive dual microphone systems. Speech Communication, *In Press*.

The Cocktail Party Problem: Solutions and Applications

Karl Wiklund¹ and Simon Haykin¹

¹Dept. of Electrical and Computer Engineering, McMaster University, Ontario, Canada, L8S 4K1

1. INTRODUCTION

If a genuine, real-time solution to the cocktail party problem[1] could be found, it would have many potential applications. These uses would include for example: interference suppression for hearing aids, improving environmental awareness for wearers of active noise control headsets, as well as improving the audio quality of telephone conversations. Surveying the literature however, one finds that while there has been much work done on the subject of cocktail party processing and computational auditory scene analysis[2], essentially none of it has dealt with the problems of real-time and embedded processing[3]. This is an important challenge given that the previously mentioned applications must run in real-time while using minimal computational resources.

2. Computational Auditory Scene Analysis (CASA)

As the name implies, Computational Auditory Scene Analysis, is the machine-based counterpart to Bregman's Auditory Scene Analysis[4]. While there are many different variants of the computational process, they all share the same essential feature: given some discrete time-frequency decomposition, assign an individual unit $s(t,f)$ a gain $m(t,f)$ such that the desired target signal is preserved, and the unwanted interference is suppressed.

The classification procedure is ultimately based on simple approximations of the four basic cues outlined by Bregman: the interaural time-difference (ITD), interaural intensity difference (IID), pitch and temporal onset. All four of these cues have simple signal processing analogues. The ITD and pitch, for example, can be computed via the cross-correlation and auto-correlation functions respectively, while IID and onset are based on comparisons involving the signal's power envelope. These basic computations are all well-established in the CASA literature.

3. Data Fusion in CASA Systems

Data fusion in CASA systems is not a trivial matter given the highly variable nature of the acoustic environment. This variability exists from moment-to-moment as a particular environment changes, as well as on a room-to-room basis, given that any CASA unit must be able to operate in a wide variety of different general environments. Attempts at producing statistically optimal fusion rules therefore have not been successful given the grave difficulties in actually forming the real-world probabilities for the range of scenarios that can exist. This is further complicated by the computational complexity of such estimation algorithms,

which severely limits their practicality in real-time and embedded systems.

3.1 Hierarchical Cue Fusion

Consideration must therefore be given to the behaviour of these cues in realistic environments, and their respective robustness to noise and reverberation.

We based our approach to cue fusion on the following two principles:

1)The most acoustically robust cues are the most important in terms of grouping. Less robust cues should be used in a supplementary role in order to constrain the association of the primary cues.

2)The variability of the cue distributions means that the interpretation of the cues must be in terms of the mean and variance over several channels, and not in terms of individual time-frequency units.

In practice, these principles can be realized given the knowledge that both the speech onset period and periods of approximately constant pitch[5] are relatively robust to noise and reverberation when compared to both the ITD and IID. These last two cues, although less robust, supply the spatial information necessary for speech separation, and must therefore be aggregated in order to resolve a stream's identity.

3.2 Fuzzy Logic Data Fusion

Owing to the previously mentioned limitations of probabilistic methods, as well as the rule-based nature of Bregman's outline, the use of fuzzy logic for data fusion is a natural choice. Cue fusion therefore, can be described in terms of a series of *IF-THEN* rules that make use of somewhat vague definitions. For example, the combination of the ITD and IID can indicate the presence of a target speaker at some pre-defined spatial location. In other words, *IF* "most" of the ITDs *AND* "most" of the IIDs indicate the presence of a target, *THEN* a target is "likely" to be present. The time-frequency mask for CASA segregation can be readily formed using the truth-values of the applicable fuzzy rules, thus forming a "softmask" approach to segregation.

3. Control and Adaptation

The reliability of the auditory cues, and as a consequence, the reliability of the fusion mechanisms, depends on the acoustic environment in ways that are difficult to quantify.

On the whole however, it can be said that increasing levels of noise and reverberation reduce the quality of the filtered signal. Two combat these different sources of interference, two adaptation methods were incorporated into the FCPP.

3.1 Recursive Smoothing

In the FCPP, this scheme takes the form of the double-sided single-pole recursion[6] shown in equation (2):

$$\hat{\rho}(t, j) = \beta(t) \cdot \rho(t, j) + (1 - \beta(t)) \cdot \hat{\rho}(t, j - 1) \quad (2)$$

where $\rho(t, j)$ is the truth value of the fuzzy rule relevant to the t th time-step, and j th frequency bin. The time-varying smoothing parameter is $\beta(t)$, and the smoothed gain estimate is $\hat{\rho}(t, j)$.

The smoothing parameter depends on the presence or absence of an onset period. As these segments are relatively free of reverberation, their associated spatial cues are more reliable, and the relevant frames should therefore be given greater emphasis in terms of the level of certainty (gain) attached to them. This is accomplished by letting the smoothing value take on two different values as shown below

$$\beta(t) = \begin{cases} 0.3 & \text{if } onset = TRUE \\ 0.1 & \text{if } onset = FALSE \end{cases} \quad (3)$$

3.2 The Gain Floor

The second aspect of the control problem performs the task of controlling the level of musical noise by adapting to changing levels of background noise. This problem was addressed not by smoothing, but by selectively adding in the unprocessed background noise. Specifically, the final gain calculation for the controller is expressed as

$$g(t, j) = \hat{\rho}(t, j) + \neg\hat{\rho}(t, j) \cdot FLOOR \quad (4)$$

where $g(t, j)$ is the gain for the j th frequency bin at time t , $\hat{\rho}(t, j)$ is the smoothed gain estimate from (2), $\neg\hat{\rho}(t, j)$ is its complement, and the value of $FLOOR$ is dependent on the currently estimated SNR.

4. Post-processing Using Spectral Subtraction

The cue fusion and estimation routines that have been described so far are unfortunately ambiguous with respect to noise sources located behind the listener. That is, the directional cues are unable to distinguish between sources that are in front of, or behind, the listener. In order to overcome this problem, the use of an additional pair of directional microphones is proposed. A very simple variation of the standard spectral subtraction algorithm can then be used to distinguish between frames dominated by the target, situated in front of the listener, and frames

dominated by interfering sources located somewhere behind the listener.

5. Results

In the experiment shown below, a target signal in a reverberant room was positioned straight ahead of a KEMAR dummy is embedded in a scenario with three competing talkers positioned at azimuths of 67° , 180° , and 270° , with the respective time positions of the signals being randomized.

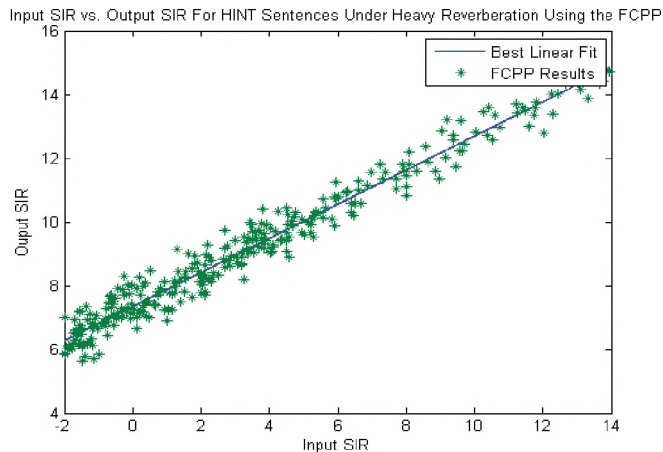


Figure 5. Output segmental SIR for a given input SIR in a reverberant room.

REFERENCES

- [1] Rong, Dong. Perceptual Binaural Speech Enhancement in Noisy Environments. M.A.Sc thesis, McMaster University, 2004.
- [2] Wang, Deliang and Brown, Guy J. "Fundamentals of Computational Auditory Scene Analysis" Ed. Deliang Wang and Guy J. Brown. Computational Auditory Scene Analysis. Piscataway, NJ: IEEE Press, 2006. 1-44.
- [3] Wang, Deliang. Computational Auditory Scene Analysis and its Application to Hearing Aids. IHCON 2008. August 13-17, 2008. Lake Tahoe, CA.
- [4] Bregman, Albert. Auditory Scene Analysis: The Perceptual Organization of Sound. Cambridge, MA: MIT Press, 1990.
- [5] Brown, Guy J. and Palomaki, Kalle J. "Reverberation". Ed. Deliang Wang and Guy J. Brown. Computational Auditory Scene Analysis. Piscataway, NJ: IEEE Press, 2006. 209-250.
- [6] Diethorn, Eric J. "Subband Noise Reduction Methods for Speech Enhancement". Ed. Yiteng Han and Jacob Benesty. Audio Signal Processing For Next Generation Multimedia Communication Systems. Norwell, MA: Kluwer, 2004. 91-118.

REGION-GROWING PERMUTATION ALIGNMENT APPROACH IN FREQUENCY-DOMAIN CONVOLUTIVE BLIND SOURCE SEPARATION

Lin Wang^{1,2}, Heping Ding¹, and Fuliang Yin²

¹ Institute for Microstructural Sciences, National Research Council Canada, lin.wang@nrc-cnrc.gc.ca

² School of Electronic and Information Engineering, Dalian University of Technology, P. R. China, 116023

1. INTRODUCTION

Blind source separation (BSS) is a technique for recovering the original source signals from observed signals with the mixing process unknown. One well recognized BSS application is the separation of audio sources that have been mixed and recorded by a microphone array in a real room environment, i.e. a cocktail party problem. The challenge of the problem is that the mixing process is convolutive, where observations are the combinations of the unknown filtered versions of the sources. Frequency-domain BSS methods are promising for convolutive BSS due to its fast convergence and small computation load^[1].

In frequency-domain BSS, the convolutive separation problem is converted to an instantaneous Separation problem in each frequency bin by short-time Fourier transform (STFT). Generally, satisfactory instantaneous separation may be achieved within all frequency bins, but combining them to recover the original sources is a challenge because of the unknown permutations associated with individual frequency bins. The permutation ambiguities should be looked after properly so that the separated frequency components from the same source are grouped together.

To solve the permutation problem in frequency-domain blind source separation, this paper proposes a new alignment method based on an inter-frequency dependence measure: the powers of separated signals. Bin-wise permutation alignment is applied first across all frequency bins, using the correlation of separated signal powers; then the full frequency band is partitioned into small regions based on the bin-wise permutation alignment result; and finally, region-wise permutation alignment is performed in a region-growing-like manner. Experiment results verify the effectiveness of the proposed method.

2. METHOD

2.1 Frequency-domain BSS

The convolutive separation problem can be converted to instantaneous separation in each frequency bin via short-time Fourier transform (STFT). In the frequency domain, it is possible to separate each frequency bin independently using complex-value instantaneous BSS algorithms such as Infomax^[2], which are considered as quite mature. However, there is a scaling and permutation ambiguity in each bin, where the order and scale of the separated signals are

unknown. It is necessary to correct the scaling and permutation ambiguity before transforming the signals back to the time domain. The permutation correction is a challenging problem. The permutation at each bin should be aligned so that the separated components originating from the same source are grouped together. The scaling ambiguity can be relatively easily resolved by using the Minimal Distortion Principle^[3]. Finally the unmixing network $W(n)$ is obtained by inverse discrete Fourier transforming (IDFT) and the estimated source $y(n)$ is obtained by filtering $x(n)$ through $W(n)$. The workflow of the frequency-domain BSS is shown in Fig. 1.

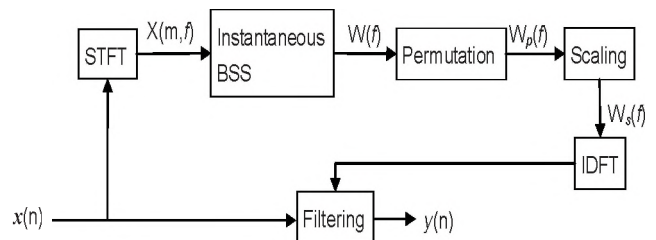


Fig. 1 Workflow of frequency-domain blind source separation

2.2 Proposed permutation alignment scheme

2.2.1 Inter-frequency dependence measure

The inter-frequency dependences of speech sources can be exploited to align the permutations across all frequency bins. An inter-frequency dependence measure proposed in [4], the correlation coefficient of separated signal power ratios, exhibits a clearer inter-frequency dependence among all frequencies. Its definition is

$$v_i(f) = \frac{\text{power of the } i\text{'th separated signal at } f}{\text{total power of all the separated signal at } f} \quad (1)$$

The power ratio represents the dominance of the i 'th separated signal in the observations at frequency f . Being in the range $[0, 1]$, (1) is close to 1 when the i 'th separated signal is dominant, and close to 0 when others are dominant. The power ratio measure can clearly exhibit the signal activity due to the sparseness of speech signals.

2.2.2 Permutation alignment scheme

The correlation of bin-wise signal power ratio tends to be high when the two components belong to the same source, but such dependence is not always evident. Generally the misalignment at an isolated frequency bin may spread to other frequencies easily and cause a big misalignment

beyond it. A region-growing-like permutation scheme is proposed which can survive this misalignment spread problem. The scheme is described in 4 steps as follows.

Step 1. Calculate the power ratio (1) for all frequency bins and all separated signals

Step 2. Correct permutation bin by bin, so that the power ratio time sequences at each frequency bin has the highest correlation coefficient with the previous bin.

Step 3. Partition the frequency band into K regions, where the frequency bins with high correlations belong to one region.

Step 4. Select a region with the largest number of elements as a seed; merge with its neighboring regions on both sides in a region-growing style.

Based on the description above, the workflow of the proposed scheme is shown in Fig. 2.

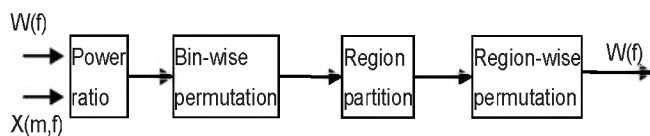


Fig. 2 Workflow of proposed permutation alignment method

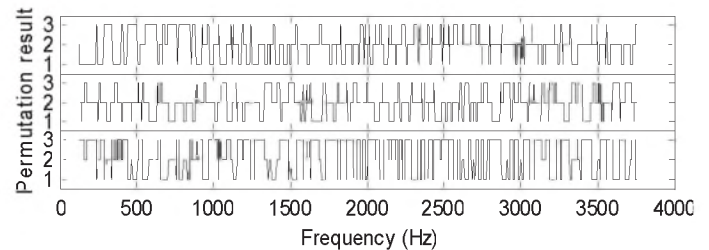
3. RESULTS

In this experiment the proposed algorithm is applied to the problem with three microphones and three sources in a simulated environment. The room size is $7\text{m} \times 5\text{m} \times 3\text{m}$, the simulated room reverberation time is $RT60 = 130$ ms, where $RT60$ is the time required for the sound level to decrease by 60 dB. To demonstrate the permutation alignment performance, we show the permutation result at three stages in Fig. 3: (a) before permutation alignment, (b) after bin-wise permutation alignment, and (c) after region-wise permutation alignment. The permutation result is calculated using the method proposed in [5], supposing the mixing filters are known. It can be seen from Fig. 3 that the permutation ambiguity is very severe before permutation alignment; the ambiguity is mitigated after bin-wise permutation alignment but large misalignments still occur; and the ambiguity is almost eliminated after region-wise permutation alignment except some misalignments on some isolated frequency bins. Such misalignments do not spread to nearby frequencies. Finally, the average output signal-to-interference ratio is 14.2 dB, a 16.8 dB improvement over the average input signal-to-interference ratio of -2.6 dB.

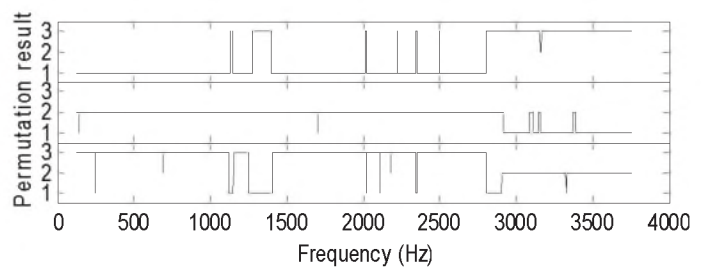
4. CONCLUSION

Studying frequency-domain convolutive blind source separation, this paper proposes a new permutation alignment method based on the inter-frequency dependence of the

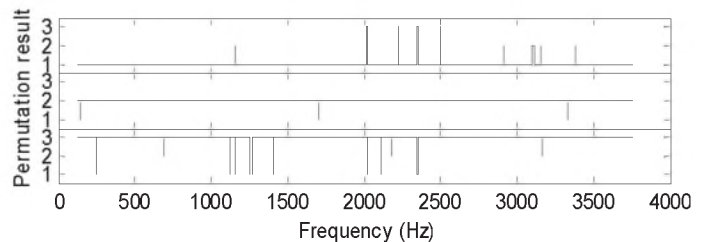
correlation of separated signal powers. With a region-growing-like permutation alignment style, the proposed method minimizes the spreading of misalignment at isolated frequency bins to others. Good separation results are observed in simulation.



(a) Before permutation alignment



(b) After bin-wise permutation alignment



(c) After region-wise permutation alignment

Fig. 3 Permutation result of the proposed method

REFERENCES

- [1] Sawada, H. (2007). Frequency-domain blind source separation, in *Blind Speech Separation*, 47-78, Springer.
- [2] Douglas, S.C. (2007). Scaled natural gradient algorithms for instantaneous and convolutive blind source separation, *ICASSP 2007*, vol. 2, 637-640.
- [3] Matsuoka, K. (2001). Minimal distortion principle for blind source separation, *2001 International Workshop on Independent Component*, 722-727.
- [4] Sawada, H. (2007). Measuring dependence of bin-wise separated signals for permutation alignment in frequency-domain BSS, *2007 IEEE International Symposium on Circuits and Systems*, 3247-3250.
- [5] Ikram, M.Z. (2000). Exploring permutation inconsistency in blind separation of speech signals in a reverberant environment, *ICASSP 2000*, vol. 2, 1041-1044.

ACKNOWLEDGEMENT

This work was supported by NRC-MOE Research and Post doctoral Fellowship Program from Ministry of Education of China and National Research Council Canada.

REVIEW OF WIDEBAND SPEECH NOISE REDUCTION TECHNIQUES

Malay Gupta, Chris Forrester, and Sean Simmons

Research in Motion, 295 Phillip St., Waterloo Ontario, Canada, N2L 3W8 mgupta@rim.com

1. INTRODUCTION

Wideband speech has been touted as the next great leap in telephone voice quality – analogous some would say to the difference between AM and FM radio. Wideband telephony extends the transmitted speech bandwidth from 3.4 kHz to 7.0 kHz. Wideband speech potentially offers improved speech quality and intelligibility and a more life-like voice experience. Realizing the potential of high quality speech communications, VoIP-based enterprise systems have been first to deploy wideband speech in a significant way in desktop phones within the office environments. Of course the public telephone networks are still narrowband, but with the advent of high-powered digital signal processors in mobile phones, increased network bandwidth and the standardization of wideband speech codecs, the deployment of wideband voice into the wireless public network appears to be on the horizon.

One challenge facing wireless phone designers is to realize the potential of wideband speech in noisy environments. With the possibility of enterprise mobile phones replacing the desktop phones in the near future [1], the removal of background acoustic noise has become very important in order to fully realize the benefits offered by wideband speech codecs. Narrowband systems constrain the low frequency bandwidth to 200-300 Hz and hence inherently filter out significant low frequency noise. Wideband systems have the potential to transmit low frequency noise down to 50-100Hz. This additional one to two octaves of low frequency noise can significantly degrade the speech quality unless great care is taken in suppressing it. This paper outlines the challenges involved and presents possible wideband noise suppression approaches.

1.1 Bandwidth of transmission

A bandwidth of 200 Hz to 3400 Hz has traditionally been considered sufficient to transmit the information contained in human speech over the telephone networks. However, this constrained bandwidth actually transmits only 20 percent of the frequencies that are present in the human speech [2]. This constrained bandwidth therefore results in a low quality/intelligibility speech, causing confusion and fatigue between the near-end and far-end users.

Increased bandwidth has been shown to be a critical factor determining the speech intelligibility and naturalness [2]. According to a Siemens 2003 press release, *wideband transmission can reduce speech ambiguities by as much as*

90 percent, increasing conversational intelligibility and reducing listener fatigue. Polycom reported in its 2006 technology white paper that *for single syllables, 3.3 KHz yields an accuracy of only 75 percent, as opposed to over 95 percent with 7 KHz bandwidth* [2]. However, this increased bandwidth also opens up a new challenge for noise suppression algorithms as now more noise will also be subsequently transmitted along with the near-end speech.

1.2 Wideband speech structure

The transmitted narrowband speech is mainly voiced speech mostly consisting of vowel sounds. On the other hand, a wideband signal in addition to vowel sounds has a significant amount of consonant energy. Consonants are unvoiced high frequency speech components contributing mainly towards the speech intelligibility. Consonants are often low in intensity when compared to the voiced low frequency content of the signal and have a noise-like structure in them. Therefore, consonant sounds are actually more prone to noise distortions. The increased susceptibility of wideband systems for noise distortions call for careful development of noise reduction algorithms for wideband.

2. NOISE REDUCTION IN WIDEBAND

For many decades, noise reduction algorithm developers have focused on the development and optimization of algorithms suitable only for the narrowband systems. Because of this huge inertia, very little information is available in the literature about noise reduction techniques suitable for the wideband frequency range. Extending a narrowband noise reduction algorithm to wideband is a non-trivial task, and requires a deep understanding of differences between the narrowband and wideband signals and systems. As such, algorithms that seem to work for narrowband systems are not guaranteed to work well in wideband systems even when accompanied with an increased computational complexity.

A majority of narrowband speech enhancement algorithms process the noisy speech in frequency domain and are based on short-time analysis-synthesis technique and follow the principles of spectral subtraction and short-time spectral amplitude (STSA) estimation [3]. In these techniques, the incoming noisy signal is processed in successive frames by computing the FFT of the windowed frames, and then estimating the clean speech spectrum from the noisy speech spectrum by using various techniques [3]. Potentially, the

underlying principles behind short-time analysis-synthesis are applicable to wideband systems with certain modifications. In order to develop noise reduction algorithms for wideband systems, algorithm developers have two choices, first is to modify the existing narrowband noise reduction algorithms for wideband applications, and the second choice is to account for structural differences between narrowband and wideband signals and device different algorithmic strategy for wideband signals.

2.1 Extending narrowband algorithms to wideband

The transition from narrowband to wideband involves the doubling of sampling frequency from 8 KHz to 16 KHz. For a fixed number of frequency bins, this results in a loss of spectral resolution in wideband as compared to narrowband systems. Hence in order to keep the spectral resolution same as narrowband, the length of FFTs in wideband processing must be doubled. One obvious impact of this change is increased computational complexity; however, that is not the only impact. In order to keep the algorithm behavior same as narrowband, various other parameters must also be carefully examined. These parameters may include attack and decay constants commonly used for speech activity based gain computations, smoothing parameters for periodograms etc., just to name a few [4]. The frequency-domain changes might also impact the time-domain evolution of the algorithm. For example, frequency-domain parameter changes might adversely impact the convergence time of the algorithm. Hence a proper compromise needs to be made between time-domain and frequency-domain performances when modifying narrowband algorithms for use in wideband systems.

2.2 New algorithms for wideband

New algorithms specifically tailored for wideband applications generally demand different processing strategies for different parts of the wideband signals. For example, a typical solution would consist of a *split-band* approach by dividing the signal spectrum into a low-frequency part and a high-frequency part as has been proposed in [5] for automatic speech recognition (ASR) applications. Different processing strategies could be devised for the low-frequency part and high-frequency part in terms of FFT size, spectral gain computations, or in terms of parameters/thresholds that are needed in the algorithm.

Structural/spectral richness of wideband signals could also be exploited to develop multi-state speech models, in which states represent the presence or absence of fricative, stop, vowel, glide, and nasal speech sounds [6]. An elaborate speech enhancement procedure would then consist of first identifying the state and then applying a state specific filter to the noisy speech frame. Because of the presence of more speech states, perceptual models may also have a greater impact in noise reduction process. Recent algorithms incorporating perceptual models can be found in [7]-[10].

2.3 Processed speech expectations

User expectations of a much better speech quality compared to narrowband raise the bar for algorithm developers in terms of preserving the speech quality in noisy conditions. It is more important than ever not to introduce any spectral distortion in the processed speech. It has also been noted that the quality of background noise transmission in a wideband system impacts overall speech quality differently than a narrowband system [11]. Hence preserving background noise characteristics also appears to be an important task in a wideband system. Fully characterizing the parameters that affect the overall speech quality in a wideband system is an ongoing process, with a specialist task force 294 (STF-294) of ETSI specifically working on wideband speech applications.

3. CONCLUSIONS

The standardization of wideband speech codec has made the deployment of wideband speech a reality in public telephone domain. In this paper, we have reviewed the current state-of-the-art of noise reduction algorithms suitable for wideband applications. We presented possible strategies for the development of noise reduction algorithms for wideband applications. We discussed few modifications that are needed for the application of narrowband noise reduction algorithms to wideband. Also due to the structural richness of wideband signals some new algorithms tailored for wideband applications have been mentioned. Lastly, we have outlined speech quality expectations resulting from narrowband to wideband transition.

REFERENCES

- [1] Gartner Press Release (IT section), Feb. 2009.
- [2] Rodman J., The effect of bandwidth on speech intelligibility. Polycom white paper, Sept. 2006.
- [3] Loizou, P., Speech enhancement: Theory and Practice. CRC Press, Boca Raton, FL, 2007.
- [4] Beaugeant C., et al., Challenges of 16 KHz in acoustic pre- and post-processing for terminals. IEEE Communications Magazine, pp. 98-104, May 2006.
- [5] Macho D., et al., On the use of wideband signal for noise robust ASR. Proc. IEEE ICASSP, vol. 2, pp. 109-112, April 2003.
- [6] Drucker H., Speech processing in a high ambient noise environment. IEEE Trans. Audio Electroacoust., vol. AU-16m, pp. 165-168, June 1968.
- [7] Thiemann J., et al., Noise suppression using a perceptual model for wideband speech signals. Proc. Biennial Symp. Commun., Kingston, ON, pp. 516-519, June 2002.
- [8] Jelinek M., et al., Noise reduction method for wideband speech coding. Proc. EUSIPCO., pp. 1959-1962, 2004.
- [9] Kato M., et al., A wideband noise suppressor for the AMR wideband speech codec. Proc. IEEE ICASSP, pp. 156-158. 2002.
- [10] C. Beaugeant et al., Noise reduction and preprocessing for the AMR-WB. ETSI STQ Wksp. Wideband Speech Signal Processing, June 2005.
- [11] Gierlich H.W., Wideband speech communication—The quality parameters as perceived by the user. ForumAcusticum, 2005.

SCREECH SUPPRESSION OF SUPERSONIC JET NOISE

Ramani Ramakrishnan^{1,2}, Sergio Raimondo², Anant Grewal³ and Gary Elfstrom²

¹Department of Architectural Science, Ryerson University, Toronto, ON, Canada rramakri@ryerson.ca

²Aiolos Engineering, Toronto, ON, Canada

³Institute for Aerospace Research, NRC, Ottawa, On, Canada

1. INTRODUCTION

Screech tones are a common by-product of supersonic jets. The screech tones are seen to get amplified due to the effect of solid surfaces. The amplification is clearly seen in V/STOL aircrafts and in particular during landing. Different techniques have been applied to reduce and/or eliminate the screech tones altogether. A detailed test programme was conducted to evaluate the noise generating potential of supersonic jets. Screech tone removal was one subset of the test programme conducted at the at the National Research Council Canada's Acoustic Reverberant Chamber in Ottawa. The results of the measurement programme are presented in this paper. In particular, the results of two techniques utilized for screech removal will be the focus of the presentation.

2. BACKGROUND

Ganesh Raman¹, in a review paper, described the mechanism of the shock wave periodicity-acoustic feedback loop with the nozzle lip as the main source of screeches. The intensity of the screech seems to be related to the stabilization of the above feedback loop by the impinging plate. Complete details of supersonic screech are also found in the references listed in Raman¹.

A large test programme was conducted at the National Research Council of Canada's Acoustic Reverberant Chamber in Ottawa in early 2009 to understand the noise output of supersonic jets. The results of the programme were summarized and presented at the Inter-Noise 2009 conference². The results of the test programme that pertain to jet screech removal are discussed below.

3. SCREECH REMOVAL

Different techniques have been applied to remove the screech from supersonic jet noise. Khan et. al.³ used a hemispherical reflector to remove screech. Another mechanism for screech removal is the use of small tabs (rectangular or chevron shaped) attached to the nozzle at the exit⁴. The application of the tabs, called vortex generators, and the relevant details of the tabs can be found in Reference 4. Choi et al.⁵ have demonstrated the application of pulsed micro-jets to remove screech tones.

Reflectors and tabs will be used for the current investigation of screech tone removal.

4. THE EXPERIMENT

A jet rig, as shown in Figure 1, was installed in a 540 m³

reverberation chamber. Different converging nozzles of sizes varying from 1" to 3.5" were used and the supply pressure ranged from 10 psi to 100 psi. The diffused sound pressure levels inside the chamber were measured using eight microphones.

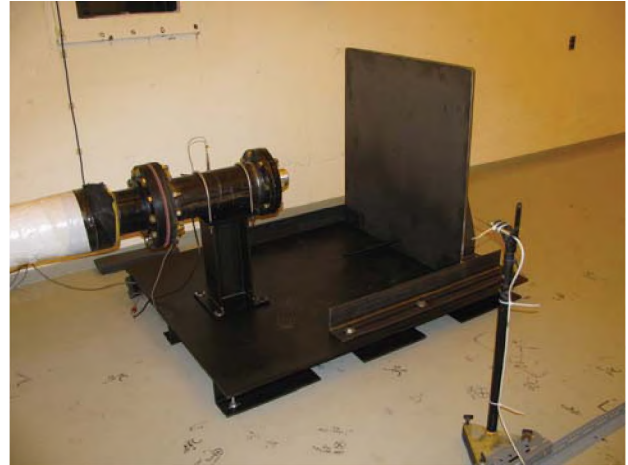


Figure 1. Jet rig to produce supersonic jet noise.

Various reflectors whose sizes were proportional to the nozzle diameters were used in the current test programme. In addition, rectangular or chevron shaped tabs, either one or three, were applied as the second mechanism for screech tone removal.

5. RESULTS AND DISCUSSION

A sample set of results from more than 50 runs are presented in this section. The effect of using spherical reflectors and/or tabs is also discussed in this section. The use of a spherical reflector is shown in Figure 2 below.



Figure 2. Jet Nozzle with bell mouth.

As mentioned earlier, different size reflectors were used for each nozzle size. The results for a 2" nozzle with four different reflector sizes are shown in Figures 3 and 4. The supply pressure for these tests was 35 psi.

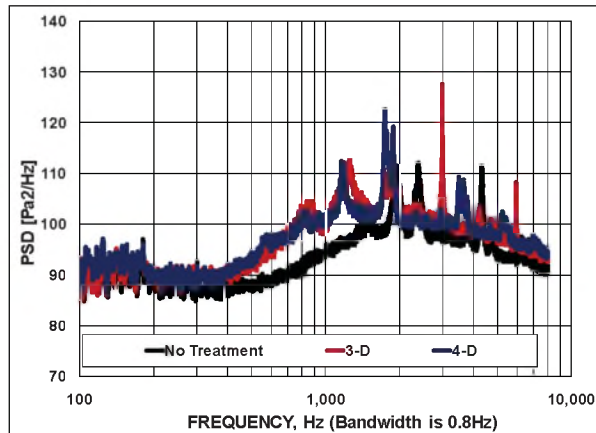


Figure 3. The effect of spherical reflectors.

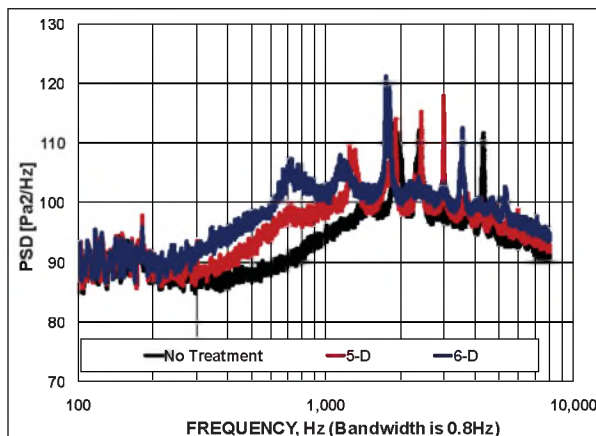


Figure 4. The effect of spherical reflectors.

The results of Figures 3 and 4 clearly show that the spherical reflectors are not useful as screech tone removal. In some instances, the reflectors actually produce screech tones as the reflectors are seen to stabilize the shock-wave patterns¹.

The nozzle with three chevron shaped tabs is shown in Figure 5.



Figure 5. Jet Nozzle with three chevrons.

The results for a 2" nozzle with one or three chevrons are shown in Figure 6. The Single tab was effective in attenuating the dominant tone by about 10 dB, but the tone is still present. The three chevrons are seen to remove the screech tones completely. Similar results were seen for other nozzles and across all supply pressures.

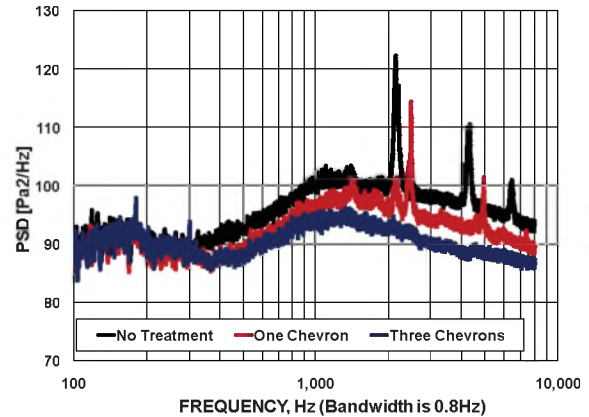


Figure 6. The effect of chevrons.

6. CONCLUSIONS

The impact of screech tones from V/STOL jet during landing can be considerable. Two methods were applied to remove the screech tones. The spherical reflectors were seen to be not useful mechanism of tone removal. The tabs, rectangular or chevron shaped, were successful in removing the screech tones completely.

REFERENCES

1. Ganesh Raman, "Advances in understanding supersonic jet screech." *AIAA Paper* 98-0279, 55 pages (1998).
2. R. Ramakrishnan, S. Raimondo, A. Grewal and G. Elfstorm, "High frequency noise generation by impinging jets." Proceedings of Inter-Noise 2009, Ottawa, Ontario. Institute of Noise Control Engineering, Poughkeepsie, NY, USA, August 2009
3. T. I. Khan, K. Seto, Z. Xu and H. Ohta, "An approach to noise reduction of a supersonic jet with a spherical reflector," *Acoust. Sci. & Tech.* **25**(2) 136-143 (2004).
4. M. Samimy, K. B. M. Q. Zaman and M. F. Reeder, "Effect of tabs on the flow and noise field of an axisymmetric jet," *AIAA Journal*, 31(4), 609-619 (1993).
5. J. J. Choi, A.M. Annaswamy, H. Lou and F. S. Alvi, "Active control of supersonic impingement tones using steady and pulsed microjets," *Exp. Fluids* DOI 10.1007/s00348-006-0189-7 (2006).

TRENDS IN AERO-ACOUSTIC WIND TUNNEL TESTING

Gary M. Elfstrom

Aiolos Engineering Corp, 2150 Islington Ave., Suite 100, ON, Canada, M9P 3V4 gme@aiolos.com

1. INTRODUCTION

Wind tunnels are an important tool in the aero-acoustic development of ground and air vehicles: for investigation of noise sources and characterization of interior and exterior noise environment associated with external flow aerodynamics. Recent trends in new aero-acoustic wind tunnels, upgraded aerodynamic wind tunnels, and climatic wind tunnels are presented in terms of innovation to serve as aero-acoustic tools for the vehicle designer.

2. THE NEED

Prototype ground vehicles, especially automobiles, require large wind tunnels for the development of quiet vehicles, and for comparative testing against competitor's models. High speed vehicles such as trains and aircraft are tested in smaller wind tunnels using scale models, primarily due to capital and operating costs of large high speed wind tunnels.

Hucho (1998) mentions three basic wind noise source types: leak noise (high frequency, about 4kHz), cavity noise (low frequency, about 40Hz), and wind rush noise (mid frequency, greater than 500Hz). Car designers are steadily pushing towards lower aerodynamically driven noise and so tools are needed to systematically identify and deal with all noise sources in a controlled environment.

3. THE TOOLS – WIND TUNNELS

The normal goal for good simulation in a wind tunnel is to have a “signal to noise” ratio of about 10dB between vehicle noise and background noise, across the full frequency spectrum of interest for a given vehicle. Lesser ratios, e.g. 5dB have been routinely used in cases where it is not feasible to achieve any better but care should be taken when comparing results from different wind tunnels.

The other important aspect of wind tunnel test capability is the cut-off frequency, f_c , of the test section. This should be as low as possible to capture the frequencies of interest. In this regard, scale model aircraft testing require much higher frequencies than full scale automotive vehicles.

3.1 Aero-Acoustic Wind Tunnels

The most recent publication on aero-acoustic wind tunnel capabilities is given by Duell (2002). The present paper can be considered an extension of that report in that it presents new aero-acoustic and related wind tunnel developments. Three new facilities are discussed here, each with its own unique requirements and challenges.

Agency for Defence Development (ADD) 3mx2.25m low speed aero-acoustic aircraft wind tunnel in Korea – Elfstrom (2007) – see Figure 1. This test section is a closed/open-jet, $f_c = 200\text{Hz}$ using wedges; circuit acoustic treatment on turning vanes and some of the airline walls.

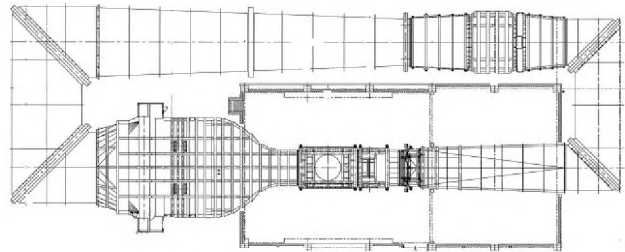


Figure 1. ADD wind tunnel

GIE S2A 24m² automotive aero-acoustic wind tunnel in France – Waudby-Smith et al (2004) – see Figure 2. The test section is a semi-open, $f_c = 80\text{Hz}$ using flat panels; circuit acoustic treatment on turning vanes and some walls.

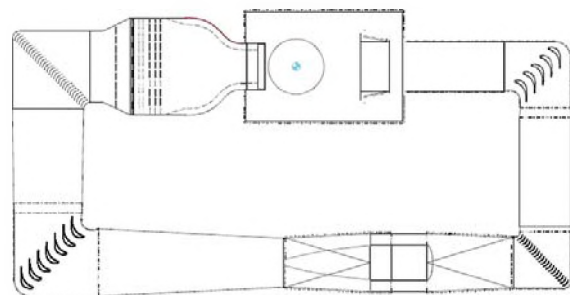


Figure 2. GIE wind tunnel

General Motors Aerodynamic Lab (GMAL) 56.2m² in USA – Yeh et al (2008) – see Figure 3. The test section is a closed-wall, $f_c = 50\text{Hz}$ using foam; circuit acoustic treatment on turning vanes. This facility is unique, not just because it is an upgrade of an aerodynamic wind tunnel.

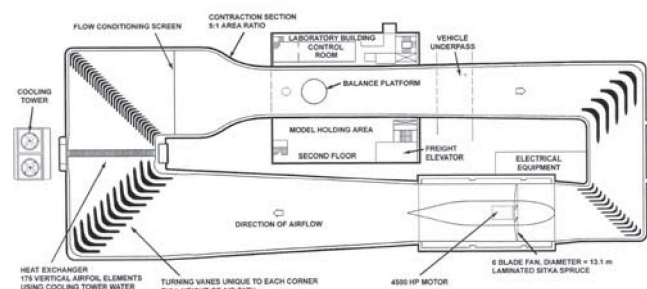


Figure 3. GMAL wind tunnel

Figures 4 and 5 show the in-flow and out-of-flow noise measured in the ADD, GIE, and GMAL facilities lie within the envelope of contemporary wind tunnels given by Duell et al (2002).

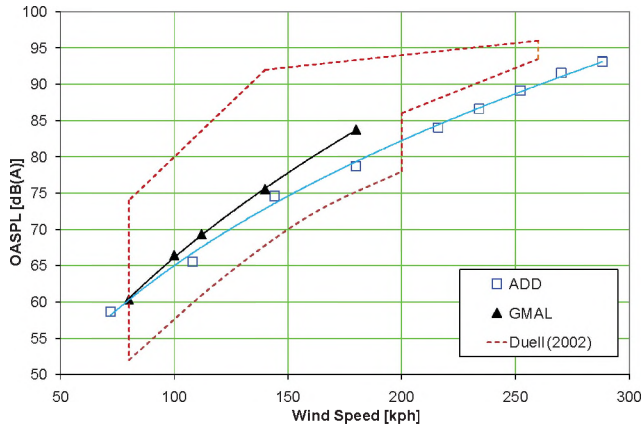


Figure 4. In-flow noise data

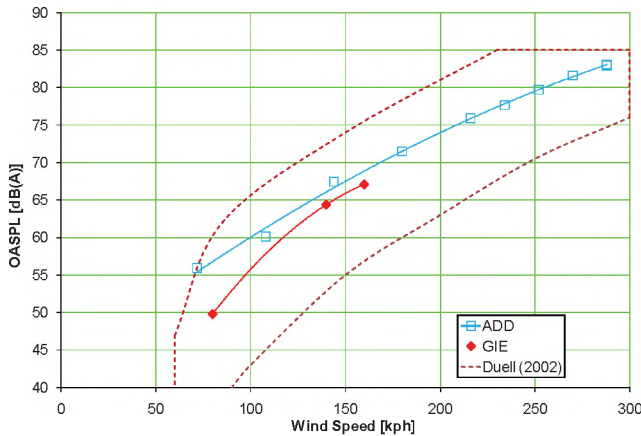


Figure 5. Out-of-flow noise data

The lower limit to the Duell (2002) envelope in Figure 5 is defined by a railroad wind tunnel, RTRI (1997). Being so low in OASPL, this at first seems out of character with its contemporaries. However, when examined in the context of circuit efficiency – see Figure 6 – it becomes clear that low noise may come with a penalty of high installed fan power.

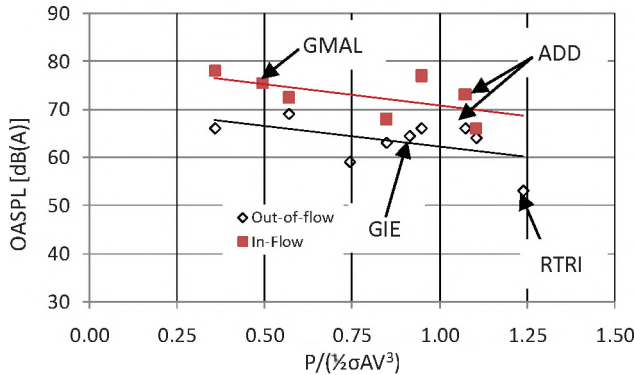


Figure 6. OASPL at 140kph vs non-dimensional fan power

The RTRI circuit is the least efficient of all due to high loss baffle sets included in each cross leg. In this context, the GMAL case deserves special mention because even with the additional losses incurred in the upgrade, it is still an extremely efficient circuit.

3.2 Climatic Wind Tunnels (CWT)

Noise sources other than external airflows, such as drive-train and exhaust system, do not require ultra low background noise levels and so recently there has been a trend towards outfitting CWT's with a modest amount of acoustic treatment. Figure 7 shows data from the Ford UK Engineering Test Laboratory CWT No.1.

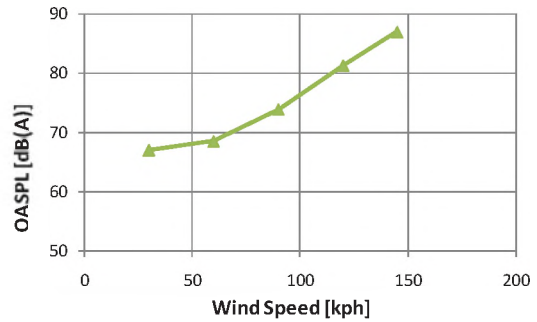


Figure 7. Climatic Wind Tunnel out-of-flow noise data

Figure 7 shows the background noise level is low enough that an experienced test engineer will be able detect any misfires, combustion instability, etc. and then to advise whether further driveability testing should be conducted.

REFERENCES

Duell, E., Walter, J., Arnette, S., and Yen, J. (2002) Recent Advances in Large Scale Aeroacoustic Wind Tunnels, AIAA Paper 2002-2503.
 Elfstrom, G., (2007) History of Test Facility Design Expertise at Aiolos Engineering Corporation, AIAA Paper 2007-0149.
 Hucho, W-H, *Aerodynamics of Road Vehicles (4th Edition)*, SAE, 1998, pp 343-355.
 Künstner, R., Potthoff, J., and Essers, U. (1995) The Aero-Acoustic Wind Tunnel of Stuttgart University, SAE Paper 950625.
 RTRI (1997) RTRI's Large Scale Low Noise Wind Tunnel, Informational Brochure, Railway Technical Research Institute.
 Waudby-Smith, P., Bender, T., Vigneron, R. (2004) The GIE S2A Full Scale Aero-acoustic Wind Tunnel, SAE Paper 2004-01-8080.
 Yeh, Y. P., Schenkel, F., Meinert, F., and Niemic, J. (2008). Calculation of Wind Tunnel Circuit Losses and Speed With Acoustic Foams, SAE Paper 2008-01-1203.

ACKNOWLEDGEMENTS

The author wishes to thank Aiolos colleagues David Van Every and Peter Waudby-Smith for assistance in preparing data for this paper; also to Jim Zurich of General Motors and Ken Huntington of Ford for useful comments and permission to publish the preliminary data on the aero-acoustic upgrade of GMAL and CWT data, respectively.

A RESIDUAL-CEPSTRUM METHOD OF PITCH ESTIMATION FROM NOISY SPEECH

Celia Shahnaz, Wei-Ping Zhu, and M. Omair Ahmad

Centre for Signal Processing and Communications
Dept. of Electrical and Computer Engineering, Concordia University,
Montreal, Quebec, Canada H3G 1M8

1. INTRODUCTION

Pitch is an important speech parameter in speaker recognition, speech synthesis, coding, and articulation training for the deaf. The pitch estimation is to determine the fundamental frequency (F_0) or period (T_0) of a vocal cord vibration causing periodicity in the speech signal. This task becomes very difficult when the speech observations are heavily corrupted by noise. Most of the methods proposed in the literature are capable of estimating pitch from clean speech [1]. As noise obscures the periodic structure of speech, many existing methods fail to provide accurate pitch estimates under noisy conditions.

In this paper, residual and cepstral representations of speech are utilized for pitch estimation in a noisy environment. For a voiced speech, the major excitation of the vocal tract within a pitch period occurs at the instant of glottal closure (GC). It is possible to determine the pitch period by careful analysis of the speech signal with the help of GC instants. Some characteristics of the GC instants can be better observed in a residual signal (RS) of speech in comparison to the speech signal itself. However, it is difficult to use the RS directly for pitch estimation because of its bipolar fluctuations around the GC instants. In order to overcome this limitation, we derive a Hilbert envelope (HE) of the RS, which presents a unipolar nature at the GC instants. Under a severe noisy condition, the time difference of successive peaks of the HE of the RS may not provide an accurate estimate of the true pitch period. Hence, with a view to overcome the adverse effect of noise on the HE of the RS, we propose a discrete Fourier transform (DFT) based power cepstrum (DFTPC) of the HE that exhibits a more prominent pitch-peak even in a heavily degraded condition in comparison to that demonstrated by the conventional cepstrum of the noisy speech. Simulation results testify that the global maximization of the DFTPC yields an accurate pitch estimate compared to the state-of-the-art methods in an intricate noisy scenario for a wide range of speakers.

2. PROPOSED METHOD

2.1 Pre-processing

Each windowed noisy frame of the observed noisy speech is low-pass filtered to remove very high-frequency contents. Such a windowed filtered noisy speech frame is given by

$$y(n) = x(n) + v(n) \quad (1)$$

where, $x(n)$ and $v(n)$ represent the windowed and low-pass filtered version of clean speech and uncorrelated additive noise, respectively. Such a pre-processing assumes to retain 4-5 formants, which facilitates the extraction of vocal-tract system parameters required for the RS generation

2.2 Pitch Estimation

One approach to derive the information about the GC instants for the extraction of pitch of speech signal is the Linear Prediction (LP) analysis. In order to remove the vocal-tract information from the process of pitch estimation, if we perform an inverse-filtering of the noise-corrupted speech $y(n)$ in a frame, where $y(n)$ is let to pass through an inverse vocal-tract system filter $\hat{A}(z)$ given by

$$\hat{A}(z) = 1 + \sum_{k=1}^p \hat{a}_k z^{-k} \quad (2)$$

the output of the inverse vocal-tract system filter is referred as the error or residual signal(RS):

$$e(n) = T^{-1}[\hat{A}(z)Y(z)] = y(n) + \sum_{k=1}^p \hat{a}_k y(n-k) \quad (3)$$

with T^{-1} representing the inverse operator of a z transform T , where $T[y(n)] = Y(z)$. In (2) and (3), p is the order of linear prediction and \hat{a}_k are the vocal-tract system parameters to be identified prior to inverse filtering. They are obtained using the LP analysis based on the autocorrelation function (ACF) $\phi_y(m)$ of $y(n)$ as

$$\phi_y(m) = -\sum_{k=1}^p \hat{a}_k \phi_y(m-k), \quad 0 < m \leq p \quad (4)$$

where m is the discrete lag variable and note that $\phi_y(m)$ obeys a recursive relation that relates the $\phi_y(m)$ values to the \hat{a}_k parameters. For an N -sample frame of $y(n)$, the ACF $\phi_y(m)$ in (4) can be estimated as,

$$\phi_y(m) = \frac{1}{N} \sum_{n=0}^{N-1-|m|} y(n)y(n+|m|), \quad m = 0, \pm 1, \dots, \pm M, \quad M < N \quad (5)$$

Since in the presence of noise, lower lags of $\phi_y(m)$ are generally become more corrupted than that of the higher lags, a few lower lags of $\phi_y(m)$ are avoided in the computation of the \hat{a}_k parameters. Utilizing the ACF coefficients $\phi_y(p+1), \dots, \phi_y(p+S)$ in (4) yields a set of linear equations, which can be represented in the following matrix form

$$\begin{bmatrix} \phi_y(p) & \phi_y(p-1) & \dots & \phi_y(1) \\ \phi_y(p+1) & \phi_y(p) & \dots & \phi_y(2) \\ \vdots & \vdots & & \vdots \\ \phi_y(p+S-1) & \dots & \dots & \phi_y(S) \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \vdots \\ \hat{a}_p \end{bmatrix} = - \begin{bmatrix} \phi_y(p+1) \\ \phi_y(p+2) \\ \vdots \\ \phi_y(p+S) \end{bmatrix} \quad (6)$$

where S governs the number of equations to be used in (6). The \hat{a}_k parameters can easily be obtained from the least-squares solution of (6) to generate the RS according to (3). It is found difficult to use the RS directly for the detection of the GC instants due to the occurrence of peaks of either polarity around the GC instants. Furthermore, in a noisy condition, the RS could be significantly different from the excitation signal due to the inaccurate estimates of \hat{a}_k parameters. However, this ambiguity can be reduced by computing the Hilbert envelope (HE) of the RS as

$$E(n) = \sqrt{e^2(n) + e_h^2(n)} \quad (7)$$

where $e_h(n)$ is the Hilbert transform of the RS $e(n)$. The HE of $e(n)$ is a unipolar positive function. The correlation among the samples of the HE of $e(n)$ is high compared to the corresponding samples in the $e(n)$. However, in a severe noisy condition, by detecting the peaks at the GC instants in the HE of $e(n)$ and taking the time difference of its successive peaks, we may not obtain the true pitch period. Hence, with a view to overcome the undesirable effect of noise on the HE of the $e(n)$, we propose a discrete Fourier transform based power cepstrum (DFTPC) of the HE as given by

$$c(n) = \left(T^{-1} \left[\log \left| T \left[E(n) \right]^2 \right| \right] \right)^2 \quad (8)$$

In (8), as usual natural logarithm is used and T^{-1} represents the inverse operator of the discrete Fourier transform T . The DFTPC of the HE is more effective in that it emphasizes the true pitch-peak even in a heavily degraded condition in comparison to that depicted by the conventional cepstrum of $y(n)$. If F_s is the sampling frequency (Hz), by searching for the global maximum of $c(n)$, the desired pitch (F_0) is obtained as

$$\hat{F}_0 = \frac{F_s}{T_0}, \quad T_0 = \underset{m}{\arg \max} [c(n)] \quad (9)$$

3. RESULTS AND DISCUSSION

The performance of the proposed method is evaluated using the *Keele* reference database [2]. This database is of studio quality, sampled at 20 kHz with 16-bit resolution. It provides a reference pitch at a frame rate of 100 Hz with 25.6 ms window. The noisy speech with SNR varying from 5 dB to ∞ dB is considered for Simulations, where white noise from the *NOISEX'92* database is used. In order to use the *Keele* database, we have chosen the same analysis parameters (frame rate and basic window size). For windowing operation, we have used a normalized hamming window. In the estimation of

Table 1. Percentage gross pitch-error for white noise-corrupted speech at SNR = 5dB

Methods	Female	Male
Proposed Method	6.98	10.79
CEP Method	24.56	26.80
ACF Method	16.54	19.75
AMDF Method	19.40	28.75

\hat{a}_k parameters by (6), S is chosen as $5p$.

For performance evaluation, we have used the voiced/unvoiced labels included in the database as well as the true pitch value F_0 . As our performance metric, we defined percentage gross pitch-error which is the ratio of the number of frames giving ‘‘incorrect’’ pitch values to the total number of frames multiplied by 100. As reported in [3], estimated \hat{F}_0 is considered as ‘‘incorrect’’ if it falls outside 20% of the true pitch value F_0 .

We have compared the performance of the proposed pitch estimation method with the conventional cepstrum (CEP), autocorrelation function (ACF), and average magnitude difference function (AMDF) methods [1]. For a speaker group, the percentage gross pitch-error is calculated considering two male (or female) speakers. In Table 1, the percentage gross pitch-error for female and male speaker groups are summarized considering the white noise noise-corrupted speech signals at an SNR = 5 dB. It is evident that in comparison to the other methods, percentage gross pitch-errors of the proposed method are significantly reduced for both female and male speakers in the presence of a white noise with a low SNR value. The lower values of percentage gross pitch-errors obtained from the proposed method for all speaker groups in a noisy environment are the testimony of its accuracy against a background noise.

4. CONCLUSION

In this paper, a new method based on residual and cepstral features is presented for pitch estimation from speech corrupted by a white noise. We generated a Hilbert envelope (HE) of the residual signal (RS) and argue that the DFT based power cepstrum (DFTPC) of the HE is more capable of reducing the pitch-errors in a difficult noisy condition. Simulation results using naturally spoken sentences have shown that the proposed method can estimate pitch in a noisy environment with a superior efficacy for both female and male speakers compared to some of the existing methods.

REFERENCES

- [1] D. O’Shaughnessy, *Speech communications: human and machine*, IEEE Press, NY, second edition, 2000.
- [2] G. Meyer, F. Plante and W. A. Ainsworth, ‘‘A pitch extraction reference database,’’ *EUROSPEECH’95*, pp. 827-840, 1995.
- [3] Alain de Chevegne, and Hideki Kawahara, ‘‘YIN, a fundamental frequency estimator for speech and music,’’ *J. Acoust. Soc. Amer.*, vol. 111, pp. 1917-1930, 2002.

BAND-ADAPTIVE FORMANT FREQUENCY ESTIMATION FROM NOISY SPEECH IN CORRELATION DOMAIN

Shaikh Anowarul Fattah, Wei-Ping Zhu, and M. Omair Ahmad

Dept. of Electrical and Computer Engineering

Concordia University, 1455 De Maisonneuve Blvd. W., Montreal, Quebec, Canada H3G 1M8

1. INTRODUCTION

Formant frequency is one of the most important speech parameters, which helps in better understanding human speech production mechanism. Estimation of formant frequencies from observed speech signal has applications in several areas, such as, speech synthesis, recognition, and compression. As an acoustic feature, formant offers phonetic reduction in speech recognition and it plays a vital role in the design of some hearing aids [1]–[3]. Formants are associated with peaks in the smoothed power spectrum of speech. Most of the formant estimation methods, so far reported, deal only with noise-free environments [1], [3]. For example, the formant estimators based on the linear predictive coding (LPC) or autocorrelation function (ACF) exhibit significant performance degradation in the presence of noise. Formant estimation from noisy speech is very difficult but essential for practical applications. In [2] and [4], formant frequency estimation methods have been proposed in order to handle noisy environments. The method in [2] utilizes an adaptive filter bank (AFB) and its performance depends on initial estimates. The method in [4] introduces a residual optimization technique based on a correlation model of voiced speech signals.

The objective of this paper is to estimate formant frequencies accurately under a severe noisy condition. In order to overcome the detrimental effect of noise, instead of using the conventional ACF, we introduce a band-limited repeated ACF (RACF) of the observed noisy speech. It has been shown that the RACF is pole-preserving and capable of drastically reducing the effect of noise. First, a band-adaptive filter-bank is employed on a zero lag compensated ACF to separate each formant frequency region prior to the formant estimation. Autocorrelation operation is then repeated on each of the resulting band-limited ACFs. Finally, a spectral peak picking method is employed to each of the band-limited RACFs to extract formant frequencies. The proposed algorithm is tested on synthetic and natural speech signals in the presence of noise and experimental results demonstrate a very satisfactory performance.

2. PROPOSED METHOD

2.1 Formant Estimation in Correlation Domain

The overall human vocal-tract can be represented by a P -th order all-pole system with a transfer function given by

$$H(z) = \frac{G}{A(z)} = \frac{G}{1 - \sum_{i=1}^P a_i z^{-i}} = \frac{G}{\prod_{i=1}^P (1 - p_i z^{-1})} \quad (1)$$

where G is a gain factor, $\{a_i\}$ the autoregressive (AR) system parameters, and p_k the system pole. Free resonances of the vocal tract system are called formants. In order to model each formant, a pair of complex conjugate poles is required. Formant frequency (F_k) and bandwidth (B_k) can be computed from the pole magnitude r_k , angle ω_k , and sampling frequency F_S as

$$F_k = \omega_k (F_S / 2\pi) ; \quad B_k = -(F_S / \pi) \ln(r_k) \quad (2)$$

Filtering speech signal $x(n)$ by the vocal-tract filter $A(z)$, one can obtain an error or residual signal and minimizing the mean error leads to following relation

$$r_x(\tau) = \sum_{k=1}^P a_k r_x(\tau - k), \quad \tau = 1, 2, \dots, L \quad (3)$$

where an estimate of the ACF $r_x(\tau)$ of N length data $x(n)$ can be computed as

$$r_x(\tau) = \frac{1}{N} \sum_{n=0}^{N-1-|\tau|} x(n)x(n+|\tau|), \quad \tau = 0, 1, \dots, L' - 1; L' \leq N \quad (4)$$

It is to be noted that the ACF $r_x(\tau)$ is a pole preserving function as seen from its z -transform for nonnegative lags

$$R_x(z) = \sum_{k=1}^P \frac{D_{1k}}{1 - p_k z^{-1}} \quad (5)$$

where $\{D_{1k}\}$ are partial fraction coefficients. Estimating parameters $\{a_i\}$ from the least-squares (LS) solution of (3), system poles as well as formants are extracted. In order to reduce the estimation variance, a combination of more than P equations can be used in the LS estimation.

2.2 Repeated Autocorrelation in Noise

Under noisy condition, observed speech is given by

$$y(n) = x(n) + v(n) \quad (6)$$

where, the additive noise $v(n)$ is assumed to be zero mean with variance σ_v^2 . The ACF of $y(n)$ can be expressed as

$$\begin{aligned} r_y(\tau) &= r_x(\tau) + r_v(\tau) \\ r_w(\tau) &= r_v(\tau) + r_{vx}(\tau) + r_{xv}(\tau) \end{aligned} \quad (7)$$

Here the ACF $r_v(\tau)$ of $v(n)$ mainly affects only the zero lag and the effect of crosscorrelation terms are negligible. Thus, the effect of noise in the correlation domain is relatively less pronounced than that in the signal domain. If the autocorrelation operation is repeated on $r_y(\tau)$, in the resulting once-repeated ACF (RACF), the effect of noise term will be drastically reduced. It can be shown that the RACF, like the ACF, preserves the poles of the vocal-tract AR system.

Since the effect of $r_v(\tau)$ is mainly pronounced at the zero lag, in the computation of the RACF, $r_v(0)$ is excluded, which also offers a significant noise reduction. Instead of conventional ACF, the RACF can be used in (3) to estimate the system poles as well as formants. However, in this case, under severe noisy condition, there is a tendency of missing weak poles in the LS solution. In order to avoid such a situation, in what follows we propose a scheme where formant estimation is performed in a band-limited region utilizing the repeated autocorrelation.

2.3 Band-pass Filtering

In order to extract each formant frequency accurately under noisy condition, instead of using directly the RACF, we propose to employ a two stage technique. First, an adaptive filter-bank is employed on the zero lag compensated ACF of the observed speech to separate each formant frequency region in the correlation domain prior to the formant estimation. The autocorrelation operation is then repeated on the resulting band-limited filtered ACF in order to achieve the advantageous features of the RACF. Thus, instead of estimating all formants together from the RACF described in Sec. 2.2, we propose to estimate each formant separately from a band-limited RACF. One important advantage of using filter-bank is that in each of the band-limited RACFs the energy of the neighboring formants is greatly reduced and it contains energy primarily from only one formant. The band-pass filters also attenuate the energy at the pitch frequency in their outputs. Thus, use of band-limited RACFs is expected to offer a better noise immunity at low levels of signal to noise ratio (SNR). In order to estimate three formants, in the filter-bank, three band-pass filters are designed depending on the conventional frequency range of the formant. These filter coefficients are then used to filter the compensated ACF into three band-limited spectral regions. The band-pass filter frequency regions are updated over time based on the previous formant frequency estimates in order to obtain a smooth formant track. The first three formant frequencies of voiced speech segments are estimated from the three band-limited RACFs using spectral peak picking method. In this case, the search zone for estimating a formant frequency is restricted within the band-pass filter frequency region. Finally, from each band-limited RACF, one formant frequency is estimated.

3. SIMULATION RESULTS

The proposed formant frequency estimation algorithm has been tested using various synthetic vowels synthesized using the Klatt synthesizer [1] and some natural vowels extracted from the North-Texas standard databases [1], [5]. For the performance comparison, the 12th order LPC [1] and the AFB methods [2] are considered and the percentage root-mean-square error (RMSE) at different noise levels are computed where each noise level consists of 20 independent trials of noisy environments. In our implementation, we perform the formant estimation every 10 ms with a 20 ms window applied to overlapping voiced speech segments. In literature, the region of formant frequencies has been well-

Table 1. %RMSE (Hz) for Synthetic Vowels

Vowels			0 dB			5 dB		
			Prop.	LPC	AFB	Prop.	LPC	AFB
Male	/a/	F1	10.49	23.63	31.29	7.81	15.87	11.74
		F2	10.36	27.78	34.82	3.92	15.53	9.51
		F3	10.34	19.28	19.34	6.13	13.19	8.23
	/i/	F1	21.42	28.53	28.16	13.76	19.28	16.25
		F2	4.01	9.68	4.27	2.71	7.54	3.33
		F3	5.25	13.29	7.75	3.68	7.82	3.97
Female	/a/	F1	11.43	17.76	16.76	5.33	9.76	15.67
		F2	9.22	19.43	14.39	5.12	8.68	7.49
		F3	3.71	9.26	4.57	2.09	3.18	2.34
	/i/	F1	23.31	39.81	32.27	15.48	28.27	19.14
		F2	10.93	26.21	13.78	5.27	12.43	6.19
		F3	3.12	15.83	3.78	2.32	7.63	2.78

Table 2. %RMSE (Hz) for Natural Vowels

Vowels			0 dB			5 dB		
			Prop.	LPC	AFB	Prop.	LPC	AFB
Male	/a/	F1	10.27	14.33	13.93	7.23	9.57	8.54
		F2	16.76	44.93	28.76	13.12	28.27	16.29
		F3	14.24	38.01	23.34	11.43	23.67	18.31
Female	/i/	F1	10.17	21.19	16.84	5.47	8.61	5.95
		F2	11.28	23.78	19.49	6.75	11.28	10.21
		F3	13.31	31.29	24.82	5.22	21.92	14.58

studied [1]. In the filter-bank, for the first formant fourth order and for other formants sixth order butterworth filters have been used [1]. In Table 1, the estimated %RMSE(Hz) is shown for two synthesized vowels at SNRs 0 dB and 5 dB. It is found that the proposed method provides lower %RMSE (Hz) for both male and female speakers. In Table 2, %RMSE (Hz) for natural vowels /a/ and /i/ (contained in words “hod” and “heed”) is shown, which indicates a better estimation accuracy obtained by the proposed method.

4. CONCLUSION

The proposed scheme offers an attractive feature that it estimates formant frequencies from band-limited formant regions utilizing advantages of the RACF. It has been clearly observed from experimental results on synthetic and natural speech signals under noisy conditions that the proposed method provides a high degree of estimation accuracy at a moderate to low levels of SNR.

REFERENCES

- [1] D. O’Shaughnessy, (2000). *Speech Communications: Human and Machine* (2nd ed.). *IEEE Press, NY*.
- [2] K. Mustafa and I. C. Bruce, (2006). Robust formant tracking for continuous speech with speaker variability. *IEEE Trans. Audio Speech Lang. Processing*, 14, 435–444.
- [3] L. Deng, A. Acero, and I. Bazzi, (2006). Tracking vocal tract resonances using a quantized nonlinear function embedded in a temporal constraint. *IEEE Trans. Audio Speech Lang. Processing*, vol. 14, no. 2, pp. 425–434, Mar. 2006.
- [4] S. A. Fattah, W.-P. Zhu, and M. O. Ahmad, (2007). An approach to formant frequency estimation at low signal-to-noise ratio. *ICASSP’07*, 4, 469–472.
- [5] J. M. Hillenbrand, L. A. Getty, M. J. Clark, and K. Wheeler, (1995). Acoustic characteristics of American English vowels. *J. Acoust. Soc. Am.*, 97, 3099–3111.

DIGITAL EARPLUG FOR BRAIN PLASTICITY RESEARCH

Marc Schönwiesner¹, Jérémie Voix², and Philippe Pango³

¹International Laboratory for Brain, Music and Sound Research (BRAMS), University of Montreal, 90 av. Vincent d'Indy, Montreal, QC, Canada, H2V2S9, marc.schoenwiesner@umontreal.ca

²Sonomax Hearing Healthcare Inc, 8375 Mayrand, Montreal, QC, Canada, H4P 2E2, jvoix@sonomax.com

³Vitasound Audio Inc., 14 Connie Crescent (19-20), Concord, ON, Canada L4K 2W8, ppango@vitasound.com

1. ABSTRACT

This paper will present the feasibility of utilizing a miniaturized, real-time, in-ear, digital signal processing devices to investigate experience-dependent brain plasticity in the humans. An important component of this trial is the use of a recently developed digital hearing protector (from Sonomax, Montreal, QC) made with a custom earpiece that is instantly fitted to the user's ear, tested for attenuation and then equipped with a miniaturized set of microphone, receiver and Digital Signal Processor. The DSP is a versatile audio platform, originally designed for hearing aid applications, but that has also been successfully programmed for several other applications like a non-linear earplug (offering more attenuation when the ambient noise is higher) and as a musician's earplug (offering a constant attenuation over a wide frequency range together with a loudness correction). The central idea of the current study is to use such digital earplugs to change a person's sound perception, in real-time, in- and outside of the laboratory. Various time and frequency manipulations will be performed on the signal pick-up by the microphone and transmitted to the subject's ear by the receiver, while monitoring the brain plasticity with neuroimaging techniques. Preliminary results using a notch filter demonstrate tonotopic reorganization following sensory modification in the human auditory cortex.

2. INTRODUCTION

On the basis of a previous study (Pantev et al., 1999) we hypothesize that the removal of a frequency band from auditory input to both ears for one week may result in a reduction of the cortical surface responding to that band within tonotopically organized auditory cortex, and possibly an enlargement of the representation of neighboring frequency bands. We additionally predict that behavioral frequency discrimination thresholds will increase for the neighboring bands, in correspondence with the increased cortical representation.

3. METHOD

Figure 1 shows a photograph of a pair of digital earplugs. The DSP utilized is a "Voyager", a versatile programmable audio processing platform originally

designed by Gennum Corporation (and now distributed by Sound Design Technologies, Burlington, ON) featuring 2 audio inputs (32 kHz, 20 bit ADC), 2 amplified audio outputs (16 kHz, 20 bit DAC) and 4 specialized audio processing units (Input, Output, Time domain and Frequency domain clusters) with a 2.048 MHz clock rate.

The steps of the experiment were: 1) fitting of the ear plugs, 2) programming the processor of the plugs, 3) validating that the earplugs transform incoming sound as intended, and 4) wearing the earplugs and acquiring several fMRI scans and behavioral data from the participant to measure changes in sound perception and cortical mapping of the acoustic parameter that is studied.

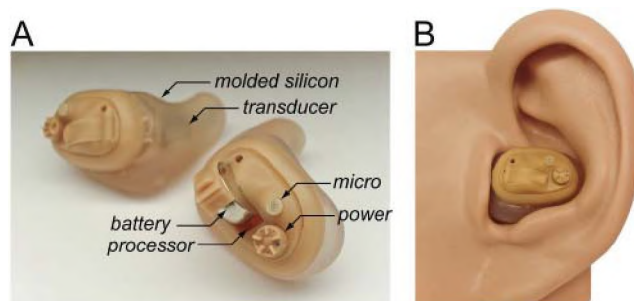


Figure 1: A) Components of a pair of digital earplugs custom-molded to a research participant. The signal processor is located in the center of the plug and programs can be uploaded through the opening of the battery enclosure with a small cable. B) Position of the plug in the pinna. The plug fits comfortably behind the tragus within the concha. The ear is part of an acoustic mannequin (KEMAR, Grass Inc., Denmark) used to measure and validate the frequency response and transfer function of the earplug.

For the present experiment the earplugs were programmed to run a simple notch filter with a center frequency of 1 kHz (Fig 2A). The function of the earplugs was verified by measuring the in-ear frequency response. In addition, standard clinical audiometry was performed with the participant, with and without the earplugs, to evaluate the effect of the earplugs on pure tone detection thresholds. The participant wore the earplugs over the day continuously for a period of 8 days (average 14 hours per day, 112 hours in total). The participant underwent two identical fMRI sessions, one before and one after the period of wearing the earplugs. In each session, high-resolution tonotopic maps of

the participant's auditory cortices were acquired with a protocol one of us developed and verified previously (Schönwiesner et al., 2007). Briefly, responses to 5 frequencies are measured at an in-plane resolution of 1.5 x 1.5 mm. The responses are combined to a tonotopic map by finding the frequency that evoked that largest response (best frequency) at every voxel. This map is superimposed onto extracted temporal lobe surfaces, resulting in an accurate visualization of the frequency representation across the cortical surface of the auditory cortex.

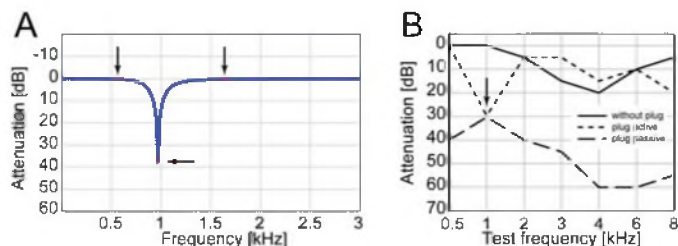


Figure 2: A) Frequency response of the notch filter implemented on the ear plugs. The notch is centered at 970 Hz, with a center rejection of 38 dB and a 3 db-bandwidth of 300 Hz. This design is based on the 5 frequencies tested during fMRI scanning to map the tonotopic organization of the auditory cortex (350, 590, 970, 1670, and 2800 Hz). The middle test frequency is effectively removed by the filter, while the neighboring frequencies are unattenuated (arrows). B) Hearing thresholds were measured from the participant to verify the performance of the earplugs. The plug achieves a passive attenuation of about 30 dB (difference between the solid and the dashed lines). When the plug is switched on, the filter described in (A) becomes active and hearing thresholds return to approximately normal levels, except around 1000 Hz, where the full 30 dB attenuation is maintained.

4. RESULTS AND DISCUSSION

The participant reported that about two days were needed to get completely used to wearing the earplugs, and that after this initial period, use became second nature and did not interfere with daily activities. Audiometric measurement confirmed an attenuation of the target frequency of about 30 dB (Fig. 2B). We compared histograms of the number of voxels responding at each frequency and tonotopic surface maps from the 'before' and 'after' sessions to identify changes in the cortical representation caused by the removal of a frequency band from the sound input. The number of voxels that respond to the frequency that was removed from the auditory input for 8 days is clearly reduced, whereas the neighboring frequencies show an increase (Fig. 3A). This corresponds to the expected pattern of results for a reorganization of a tonotopic map. There was no change at the highest frequency. The lowest frequency also shows a reduction, but we think that this is a result of the non-linear attenuation properties. In this pilot experiment, no attempt was made to correct the earplugs' frequency response electronically, but this will be done in the future studies.

We overlaid color-coded best frequency maps onto the cortical surface of the temporal lobes extracted from the participant's structural scan to find regions of auditory cortex that exhibit a change in the best frequency between the before and after scans. The most prominent difference is visible in the planum temporale where a strip of cortex (Fig. 3B, between arrows) changes its frequency preference from the removed band to neighboring bands.

CONCLUSIONS

These preliminary results are a direct demonstration of short-term tonotopic reorganization in the human auditory cortex. They illustrate the power of using miniature signal processing devices that can be worn in the ear canal to change a person's sound perception in real-time, in- and outside of the laboratory. In light of these initial findings, further research is warranted utilizing a larger group of subjects. Future experiments should also take advantage of the implementation of other filtering schemes onboard the digital earplug's DSP.

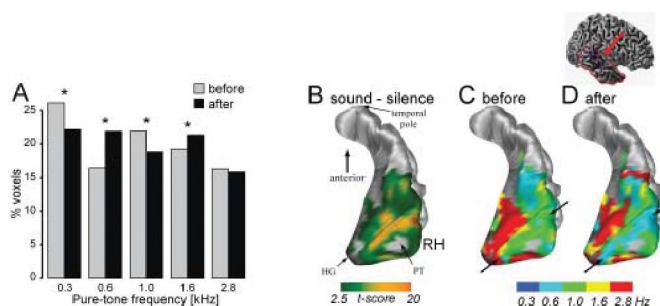


Figure 3: A) Histograms of the percentage of voxels responding maximally to each tested frequency, before and after removal of the frequency band around 1 kHz with digital earplugs. Asterisks indicate a significant difference ($p < 0.05$, permutation test, Bonferroni-corrected for 5 comparisons) between the 'before' and 'after' session for each frequency. B) Surface of the right superior temporal plane (the inset illustrates the point of view), extracted from the participant's structural scan using software developed at the Montreal Neurological Institute, with the color-coded statistical significance of the response to all frequencies vs. a silent baseline superimposed. This illustrates the area of cortex that responds to sounds. Within this area, for each patch of cortex, the frequency condition that elicits the maximal response is color-coded for the session before (C) and after (D) wearing the earplugs.

REFERENCES

Pantev C, Wollbrink A, Roberts LE, Engelien A, Lütkenhöner B (1999) Short-term plasticity of the human auditory cortex. *Brain Res.* 18;842(1):192-9.

Schönwiesner M, Dechent P, Krumbholz K (2007) Basic coding mechanisms in the human auditory cortex: a view through high-resolution fMRI. *Proc. 8th HBM, NeuroImage* 36, suppl. 1.

ACOUSTIC SHOCK DETECTION AND CONTROL

Henry Luo

Unitron hearing, 20 Beasley Dr. Kitchener, Ontario, Canada, N2G 4X1 henry.luo@unitron.com

1. INTRODUCTION

Acoustic transient noise or impulse noise is referred as acoustic shock and defined as a sudden increase of sound pressure level within a very short time [1]. Such sudden signal level increases can result from interfering noises such as dishes breaking, hammering, a door slam, a gun shot, or any other kind of impulse noise. Acoustic shock is a well-known problem in public address systems, handsets and other telecom products as well as hearing aids. The problem of acoustic shock can be much worse for the user of hearing aids since it may cause more discomfort or further hearing damage if not handled properly by the hearing instrument [2].

Many different approaches have been developed to address the detrimental effects of such acoustic shocks. MPO (Maximum Power Output) in the frequency domain can be applied to prevent overshooting, but it is too slow to be effective. Peak-clipping in the time-domain is effective and fast, but it usually causes serious distortion of sound quality.

A sub-band-based acoustic shock algorithm has been presented in a patent application [3]. The shock detection module in the frequency domain detects if shock has occurred in a particular sub-band, and determines the sub-band energy measurement to be used for the gain calculation. The acoustic shock phenomenon is eliminated by applying the appropriate gain reduction to the signal in each sub-band. The additional time-delay is also required for the system to work.

In some shock detections, high pass filters are used since the transient noise has most of the energy at high frequencies. Low-pass filters are also often used to attenuate the transient noise without significantly affecting speech content of the signal [4]. The difficulty for the shock detection and control based on filters are to reliably separate speech from the transient noises without side effect.

What is desired is a method that can detect acoustic shock reliably and quickly with minimum computational cost, and can adaptively attenuate or cancel the shock while, at the same time, maintaining the input signal quality unaffected. This paper presents a new technology referred as AntiShock which consists of smart shock detection and adaptive shock control. AntiShock technology has been successfully applied in new generation of digital hearing aids.

2. METHOD

2.1. Acoustic Shock Definition

If a shock has a peak level of L (dB) at t_1 , duration $T=t_2-t_0$ where t_0 is defined as the starting point of shock and t_2 is

defined as the half-way point between peak level and the signal floor, it looks as shown in Fig. 1.

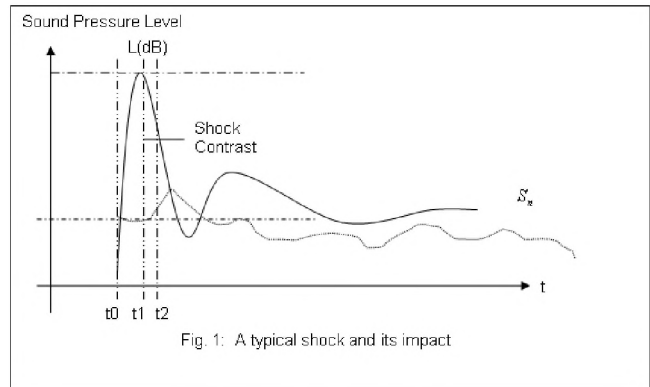


Fig. 1: A typical shock and its impact

The shock contrast level (dB) is $dL = L - S_n$, where

$$S_n = 10 \log \left[\frac{1}{t_1 - t_0} \int_{t_0}^{t_1} |s(t)| dt \right]$$

is the fast averaging $s(t)$ over a short duration of about 2ms so that it can reflect the normal speech signal or the music signal change over time. The higher the shock peak level, the stronger the shock will be perceived. For the same source of shock, the perceived shock strength depends on the actual contrast level dL . With the same shock peak level, the lower the signal floor, the higher the shock contrast that will result in a stronger shock perception. Also, the longer the duration of the shock, the stronger the shock will be perceived. Therefore, it is critical to detect shock contrast in order to determine the level of shock impact and the necessary solution in controlling the shock with an appropriate anti-shock strategy.

2.2. Smart Shock Detection

The absolute shock impact level (SIL) is $SIL = \frac{L}{t_1 - t_0}$

and the shock energy strength (SES) is $SES = \frac{L}{t_1 - t_0} * (t_2 - t_0)^2$, which means the higher L , the

shorter (t_1-t_0) and the longer (t_2-t_0) are, and the stronger the shock energy strength.

For shock detection, the following variables are included:

- Signal floor S_n (dB), which varies per acoustical environment along time;
- Peak level L (dB);
- Shock setup time (t_1-t_0) ;
- Shock duration (t_2-t_0)

The shock contrast level (dL) is defined as $dL = L - S_n$ and the relative shock impact level can be expressed as the Shock Index: $S_{Index} = \alpha \frac{L - S_n}{t_1 - t_0}$, where α is the

coefficient for Shock Index normalization. The shock index normalization constant α can be defined according to the individual's preference. In one of our preferable systems, we define α by referring to a typical dish transient noise with a shock level $L=70\text{dB}$ in quiet ($S_n = 40 \text{ dB}$) with 0.2ms attack time and 20ms duration. Then we get $\alpha = 0.2 \text{ ms} / 30 \text{ dB} = 0.0067 \text{ (ms / dB)}$ to have $S_{Index} = 1$ as normalized Shock Index. For this same dish noise, if the environment signal level increases, the shock Index will drop.

Two thresholds, minimum shock contrast level and minimum shock Index, are used for shock detection. These thresholds can be determined through a self-learning process or pre-determined measurement so that daily life non-transient signals such as speech, music, normal acoustic sound are not detected as shock, and that a transient sound such as a gun shot or a door slam will be detected as shock. A stronger and sharper shock will generate a stronger Shock Index. The duration of shock, $T=t_2-t_0$, will be used together with Shock Index as the measurement of shock strength, which is used for shock control.

The shock detection runs in real-time with the use of the thresholds of minimum shock contrast level and minimum Shock Index. According to the shock contrast level and Shock Index, it can decide whether a shock happens and how strong the shock is. Since the shock detection runs continuously, the shock can be detected anytime as long as it meets with the shock detection criteria.

2.3. Adaptive Shock Control

The objectives of shock control are:

- To reduce or minimize the shock effect;
- To keep the shock sound as natural as possible to allow awareness by the user of the type of shock event;
- To keep the relative loudness of shock so that the user can perceive the shock level;
- To keep the shock within the comfort range of the user.

The shock control will apply Gain Reduction $g(t)$ adaptively to the input signal $s(t)$ to get a new signal $x(t)$. As one typical application, we may use a fast shock attack and a slow anti-shock release for the peak shock duration ($t \in (t_1, t_2)$) so that the shock can be efficiently controlled. In another typical application, we may need to have fast anti-shock release for the peak shock duration ($t \in (t_1, t_2)$), so that the useful signal is less affected. In

other applications, we may want to use the same anti-shock attack and release speed in order to simplify the design.

3. RESULTS

An effective technology for acoustic shock detection and control has been invented for the new generation of digital hearing aids. In one evaluation, 30 hearing aid wearers provide preferences for AntiShock On/Off: Speech, Transients Speech + Transients. As shown in Fig. 2, the clear preference exists for the "AntiShock On" than "AntiShock Off" for acoustic shocks and there is no obvious preference for AntiShock On/Off since AntiShock has not affected any speech context.

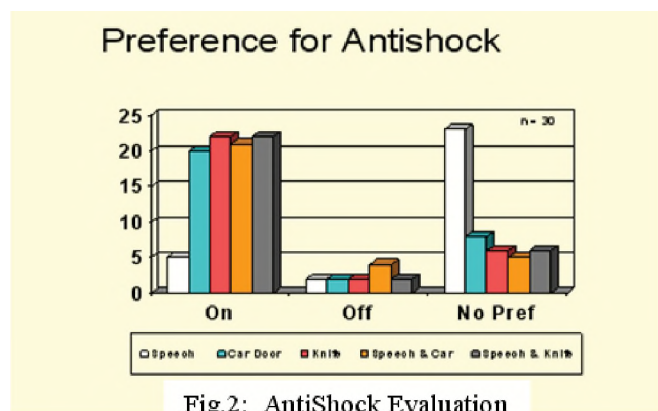


Fig.2: AntiShock Evaluation

While reducing acoustic shock adaptively, AntiShock technology keeps the natural sound quality of shock events for environmental awareness and not hampers the user's safety. AntiShock technology is capable of detecting and controlling acoustic shocks in an optimized way to cancel shock completely or keep shock in comfortable level according to applications without affecting non-shock signals such as speech and other acoustic sounds.

REFERENCES

- [1]. ANSI S12.7 – 1986, "Methods for measurement of impulse noise." American National Standards Institute, 1986
- [2]. D. Henderson and R.P. Hmernik, "Impluse noise: Critical review", J. Acoust Soc Am, vol. 80, 1986
- [3]. Anders Gingsjo, "On Transient Noise and Its Reduction in Hearing Aids", 1997.
- [4]. Schneider, Todd et al.: "Method and system for acoustic shock protection", European Patent Application EP 1 471 767 A2, DSPFactory, 2004.

SOUND CREATED FORM

EXPLORING THE INFLUENCE OF SOUND ON ARCHITECTURAL FORM AND THE AURAL QUALITY OF SPACES IT CAN CREATE

Ben Gaum and Ramani Ramakrishnan

Department of Architectural Science, Ryerson University, Toronto, Ontario

The Montreal Jazz Festival is currently the largest and arguably one of the most important jazz festivals in the world. For 30 years this festival has been drawing massive crowds and continuously attracting the biggest names in the music world despite never really having a proper central location for operations. My thesis design project will explore the influence of sound on architectural form and the aural quality of spaces it can create, by designing a new “Maison du Festival” that will act as the main welcome center, archive/museum and operations center for the Montreal Jazz Festival of the future.

The original schematic design of the building called for two “amphitheatred” sections of the plan that would allow for performances on the festivals two main stages to be projected onto the building itself, essentially acting in the same way as the back of a traditional amphitheatre. This would allow the festival organizers to be able to better contain and control the sound being projected from the stages.

The overall form of the building was mainly derived from

taking various wav file recordings of the streets surrounding the proposed site and then converting those files into a three dimensional representation. This was achieved by using a software called soundplot 1.0™ to capture the wav files that were then generated into 3D form in Rhino. Several recordings of each street were taken at various times of the day both during the festival and after the festival was over to try and get a visual representation of the types of sounds that were directly affecting the proposed site itself. Three dimensional ‘strips’ were then selected from the hundreds of forms generated that would most accurately match the proposed schematic formal design and provide the “amphitheatred” sections required.

After the initial wav strips were selected, they were pieced together and remodeled to complete a uniform building envelope. This building envelope was then analyzed using CATT Acoustics software to test the acoustic properties of the buildings form. Initial analysis showed several interesting problems with the formal design itself. By using the CATT

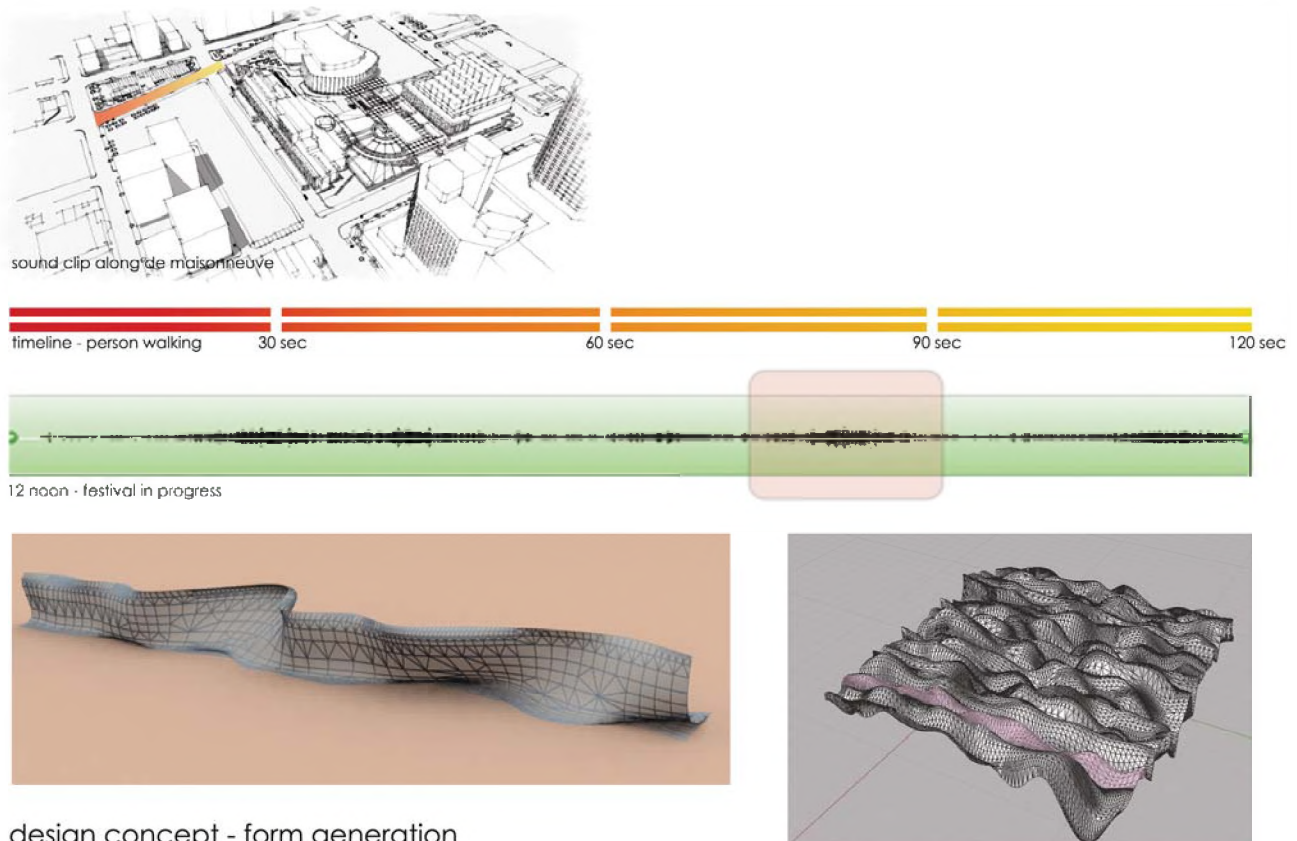


Figure 1 - 3D generated wav files



Figure 2 – Compiled 3D wavy files

Acoustics software as well as information gathered from previous studies of outdoor performance spaces, it was determined that the form of the original building had to be altered in order to satisfy desired acoustical parameters that were required for the musical performances. The results from the CATT Acoustics analysis showed that the north amphitheatred section ended up being too small and was not able to contain or control enough of the projected sound to satisfy the festivals requirements, thus the section had to be enlarged drastically horizontally as well as increased in height. The overall size of the section was increased to allow for a full 15 degrees on either side of the performances center stage. The south amphitheatred section although smaller in overall size, proved to satisfy its requirements due to its slightly

curved roof section that projected the sound back down onto the spectators watching the performance. The CATT Acoustics software proved instrumental in providing the acoustical analysis information that would help accurately transform the buildings form to properly satisfy the acoustical requirements of the festival organizers.

REFERENCES

- 1 Soundplot tm, Michael B. Pliam, Copyright PliaTech Software 2000-2008.
- 2 CATT Acoustics V8.0g, Copyright CATT 1988-2007.

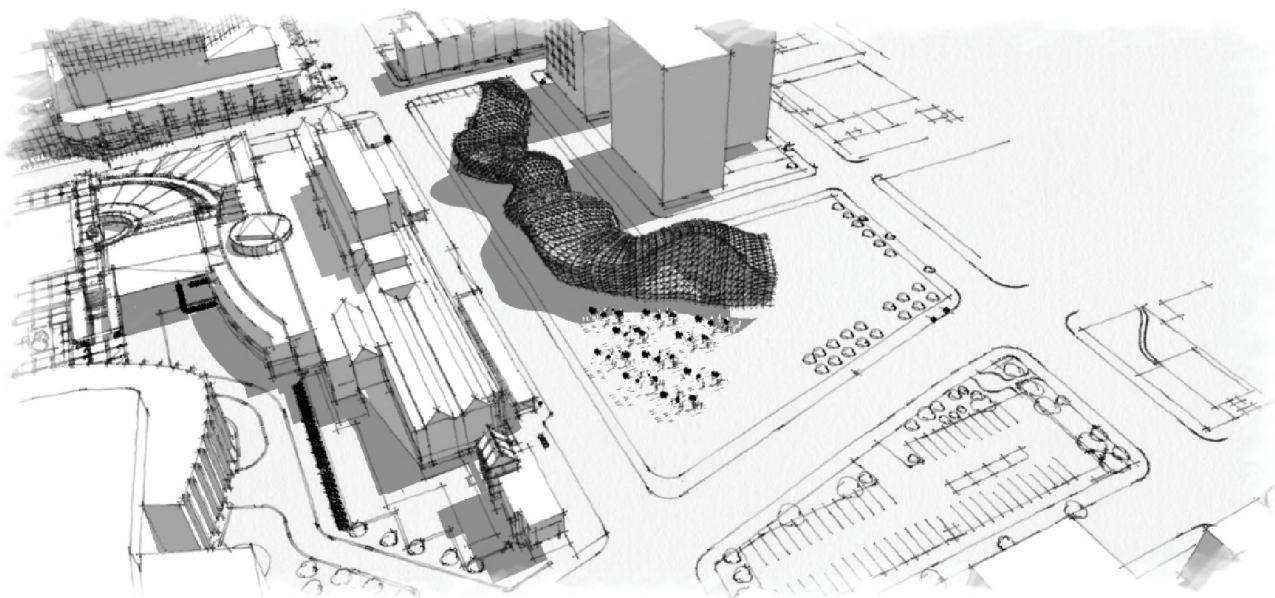


Figure 3 - Final form showing north amphitheatre section

THE SOUNDING MUSEUM

CULTURAL SOUNDSCAPES AS A TOOL IN MUSEUM EDUCATION

Hein Schoer, MA

Research Group Arts in Society, Fontys School of the Arts, Tilburg (NL) / Maastricht University (NL), h.schoer@fontys.nl

1. TUNE IN

Picture yourself standing on a cliff of a frozen fjord in Nunavut, on the south-eastern edge of Baffin Island. Eyes closed, there is no sound to be heard except the cutting wind that constantly blows through the icy desert of the Arctic.

A lonely traveler catches your attention. His snowshoes make crunchy sounds on the frozen ground as he passes by. The ringing of bells from a small Anglican church mixes with howling sled dogs on the other side of the fjord. The scene changes as you approach a small village. Inside one of the small houses we listen to the Inuit women's traditional throat singing, the *katajjait*, rhythmic, with an air of mysticism, and of an excited merriness. It is a playful way to compensate for the harsh environment that the people in these parts have learned to cope with.

Out in the open again, you accompany the hunters on their *ski-doo*s as they ride out onto the frozen sea. They cut wholes into the ice to lure seals in search for air as they dive under the floe. Summer comes along with a flock of wild geese under the mild beams of the midnight sun. A raven croaks. A walrus colony rests on the stony beach. With a small boat you enter a cavity within an iceberg.

All of a sudden you find yourself inside the ice itself. You hear it crackle and gurgle, creak and rustle, you are captured inside the slow evolving and passing of bizarre, ever changing endless forms. A roaring thunder far away bears witness to the great forces at work in nature, as a calving glacier releases its children into the sea to float as mighty icebergs in the Arctic Ocean.

1.1 Listening to Intangible Heritage

For the peoples of the First Nations of North America, whose culture is strongly attached to their oral heritage, acoustic communication always was and still is of high importance.

But another dimension makes the art of listening a vital tool for survival: The ear is the first and the last watch-post in case of danger, it stays on duty after nightfall and even while we are asleep. We have become used to draw our data from texts, charts and tables, also are we no longer in risk of becoming prey of wild animals creeping up in the dark.

In our Sound Chamber we want to re-establish listening as an equal partner among the components of sensual perception.

1.2 The Soundscape

We are surrounded by a world of sound that we call *soundscape*, in reference to the visually perceived landscape, as established by R. Murray Schafer in his milestone publication "The Tuning of the World."

We react to that soundscape, and, to an ever-growing extent, we create it ourselves. The soundscape is a crucial factor to the wellbeing of humanity. It also has strong influence on the cultural development of a society. By assessing it, we can gain access to the strangeness and the familiarity of our own as well as of foreign or even past cultural and natural environments. The museum can be the place to mediate sono-cultural awareness to the public.

2. NONAM

The research project "The Sounding Museum" revolves around the Sound Chamber at the NONAM (Nordamerika Native Museum) in Zurich, CH. The Sound Chamber is a room designed to present the acoustic aspects of the cultures of North America's First Nations by means of sound scenario productions composed of the natural and artificial soundscape of the respective cultural spaces. Deprived of visual stimuli the visitor can fully concentrate on the sonic experience in a low-reverberation quadrasonic setup, and, given an adequate introduction, learn how perception, identity and culture are inseparably intertwined with each other by the link of communication.

The practical implications to be considered when developing such a tool as the Sound Chamber are at least equally important as the conclusions that can be drawn from the evaluation of the data collected during visitor surveys and interviews.

3. SOUND CHAMBER

Early in 2008, the former video cabin on the second floor of the NONAM, next to the permanent exhibition space, saw a great deal of change happening. In accordance with the acoustic concept and technical sketches of Richard Schuckmann the small inconspicuous booth made of pressed wood was turned into a chamber of unique spatial-acoustic experience.

By applying highly absorbent acoustic foam, the chamber's boundaries were acoustically removed, thus allowing the

illusion of distance, as in the endless icy deserts of the Arctic, because the room's own reverberation attributes could not unmask its presence anymore.

Visual stimuli were to be reduced to a minimum. Black sound-permeable curtains, softly illuminated from below, around a circular visitor's platform surrounded by a metal security railing and slightly above the almost invisible ground, create a feeling of being in a safe concealed capsule in the middle of nowhere.

Behind this black horizon beats the technological heart of the installation. Four high-end loudspeakers, one in every corner of the room, and a subwoofer allow sounds to come from any desired direction on the horizontal plane. A DVD-A(audio)-player operated by the museum staff to play back high-quality sound productions is hidden behind a small hatch beside the entrance.

For a special exhibition on arctic cultures, a soundscape production was made that utilises all the merits of the Sound Chamber. The endless polar void, sound events coming from all directions, sounds that continuously change position as they progress, some that are very faint and some that shake the very foundations of the earth (or so it seems), all these merge in the surreal atmosphere of the Sound Chamber into a new, alien reality.

The recordings provided by Philippe Le Goff and other experts of the arctic cultural circle had to be ordered thematically and chronologically and to be technically mastered. Now the selected sounds could be fitted into the timeline for the Sound Chamber presentation. The changing seasons were the chronological basis of the score.

From the studio the final version went to the Sound Chamber at the NONAM, where last adjustments in terms of editing and mixing were made hands on by listening to the production at the place of its intended presentation.

4. THE SOUNDING MUSEUM

"The Sounding Museum" follows (in loose order) a three-stage design.

The theoretical background to perception and education in the cultural field has to be established. This includes recapitulation of the ongoing debates on philosophy of nature, perceptive psychology, cultural theory and museum pedagogy, matched with the demands of acoustic ecology as framed by Schafer and Truax. I chose the aesthetic-atmospheric approach as suggested by Böhme, which has to be aligned to the reality of exhibition design and visitor attitude.

In the next stage the actual Sound Chamber and the soundscapes to-be-presented have to be created in such a

manner that they meet the demands established on the basis of the theoretical and practical prerequisites.

After an extensive testing phase with the Inuit Soundscape I am currently working on the Northwest Coast Soundscape in collaboration with the U'mista Cultural Centre at Alert Bay, BC. In the end, all five cultural spaces presented in the Museum's permanent exhibition will be supported by Soundscape productions, along with special productions focussing on music, speech, etc. These will be presented at the Sound Chamber in workshops that link the culture of the museum's objectives to an educative approach on perception by the example of hearing.

The third stage is the evaluation of the data that will be collected during the workshops, in episodic interviews and visitor surveys.

5. REPRISE

My approach tries to generalise on cultural soundscape production and presentation in museums. The result will be a handbook on how to produce and present soundscape compositions made for anthropologic uses. Composition of course always bears the mark of alteration of reality according to, in our case, anthropologic, aesthetic, and educative requirements. However, I want my soundscapes to be good enough to receive approval from their original authors, the people whose natural and cultural environment produced the sounds I record and use for composition.

The Inuit who live close to the taiga have a saying that the wind, singing in the trees, is a language that just hasn't yet fully come into being.

I will teach you to understand it.

REFERENCES

- Böhme, G. (2000). Acoustical Atmospheres. *Soundscape Journal* Vol. I, No#1, p.14-19
Schafer, R.M. (1977). *The Tuning of the World*. Knopf.
Truax, B. (2001). *Acoustic Communication*, 2nd ed. Ablex Publishing.

ACKNOWLEDGEMENTS

My thanks go to Prof Dr Maaïke Meijer, Paul de Bruyne and Dr Pascal Gielen, Denise Daenzer, Karin Isernhagen, Heidrun Löb and all at the NONAM, as well as to Prof Sabine Breitsameter, but most of all to the late Richard Schuckmann, to whom I wish to dedicate my work.

AUTHOR NOTES

Hein Schoer is a PhD researcher at Fontys College for the Arts in collaboration with Maastricht University. He teaches at Hochschule Darmstadt on sound engineering and acoustic ecology and is a musician and soundscape composer.

TAPPING TO PITCHED AUDITORY FEEDBACK TONES: EFFECTS OF PITCH CONTOUR AND INTERVAL SIZE ON INTERTAP-INTERVAL AND TAP FORCE

Paolo Ammirante¹, William F. Thompson¹, and Frank A. Russo²

¹Dept. of Psychology, Macquarie University, New South Wales, Australia, 2109 paolo.ammirante@psy.mq.edu.au

²Dept. of Psychology, Ryerson University, Toronto, Ontario, Canada, M5B 2K3

1. INTRODUCTION

Some auditory perception studies have reported an interaction between pitch contour (the pattern of pitch direction changes over time) and temporal judgment (Boltz, 1998; MacKenzie, 2007). For example, Boltz (1998) presented participants with pairs of isochronous melodies. Comparison melodies were judged to be slower than standards when they contained more contour changes and faster when they contained fewer pitch contour changes. This effect persisted even for melodic pairs in which tempo was actually held constant. These data imply that *contour-preserving* tones are perceived to be shorter in duration than *contour-violating* tones.

The *imputed velocity hypothesis* proposes that change of position between successive stimuli may be perceived as the trajectory of a single moving object (Jones & Huang, 1982). Thus, Boltz (1998) suggests that the effect of pitch contour on tempo perception might reflect the listener's experience with lawful trajectories of moving objects in the world. She cites the example of a land animal. When running a zigzagging (i.e., contour-violating) course, it must *decelerate* when changing directions in order to maintain balance; when running unidirectionally, it may *accelerate*.

We attempted to replicate this interaction in a finger-tapping task. Participants tapped a key at a steady pace. Auditory feedback was provided in the form of a tone that coincided with the end of each tap. Pitch contour of the feedback tones was continuously and randomly varied. We predicted that participants would dynamically respond to task-irrelevant pitch changes by shortening intertap-intervals (ITI) initiated by contour-preserving feedback tones relative to ITIs initiated by contour-violating tones.

In addition to ITI, we analyzed tap force in order to test a "strong" version of the imputed velocity hypothesis: that velocity implied by pitch contour will influence the velocity of participants' taps. Interference effects between stimulus magnitude along a task-irrelevant dimension and motor response magnitude have been previously reported (e.g., Andres et al., 2004). Thus, we predicted greater tap force for ITIs initiated by contour-preserving tones relative to ITIs initiated by contour-violating tones.

The magnitude of frequency separation between successive tones (Δf) was varied *between* trials. Previous studies have demonstrated that Δf is associated with spatial extent in the listener. For example, in a 2AFC speeded response task, Kadosh et al. (2008) found a response compatibility effect on reaction time between Δf and the spatial separation

between response keys. We predicted that the magnitude of ITI acceleration would covary with Δf because, according to the imputed velocity hypothesis, at a steady pace greater contour-preserving Δf implies greater velocity.

2. METHOD

35 undergraduates (28 females) from Macquarie University participated for course credit. Average age was 22.0 ($SD = 4.9$). Number of years of individual music lessons ranged from 0 to 10 ($M = 1.74$; $SD = 2.66$).

Feedback tones were 250 ms sine tones. There were five conditions: monotone, 25, 50, 150, and 350 cents Δf . For the monotone control condition, the feedback tone was C5 (523.25 Hz). For each pitched condition, tones were a fixed Δf apart; tones 1 and 2 and tones 4 and 5 were above and below C5 (tone 3), respectively. Sequences began on any of the five tones and randomly ascended and descended stepwise (e.g., tone 2, 1, 2, 3, 4, 3, etc.). Pitch range of sequences covaried with Δf and was equal to four times Δf .

Participants heard feedback through headphones and tapped the highest key on a MIDI keyboard with their index finger. They were instructed to maintain contact between their fingertip and the key and to give an equal weight to all taps. It was stressed that the goal was to maintain a steady beat at the rate provided by the pacing signal and that any variation in pitch between feedback tones should be disregarded. Setting the output parameter to a fixed value eliminated variation in feedback tone intensity between taps. We used the continuation tapping task (Stevens, 1886). For each trial, participants synchronized their taps with the pacing signal (15 ms ticks with 500 ms IOI) for 20 taps after which the sound of the pacing signal stopped and was replaced by feedback tones. Participants continued to tap at the tempo set by the pacing signal until the end of the trial (30 additional taps). There were 35 trials in seven blocks, with each block consisting of one trial for each condition. Order of trials within a block was randomized. The task took approximately 30 minutes.

Only the final 25 continuation ITIs of each trial were subjected to analysis. These were defined as the time difference between MIDI "note on" events registered when the key contacted the key bottom. Tap force values were read from MIDI velocity data for each "note on" event in arbitrary units from 1 to 127. Preliminary analysis ruled out the *absolute* pitch height of feedback tones (which was not held constant between conditions) as a factor in mean ITI.

3. RESULTS

As there were no learning effects, ITIs from each trial were pooled by Condition (7 trials x 25 ITIs per trial = 150 ITIs per participant for each of the 4 pitched conditions). ITI_n values were classified according to the pitch direction of the feedback tones initiating both ITI_{n-1} (Previous Tone Direction) and ITI_n (Current Tone Direction). For example, AA means ascending contour-preservation; AD means a descending contour-violation at the onset of ITI_n.

Mean ITI values according to Condition, Previous Tone Direction, and Current Tone Direction were entered into a 4 x 2 x 2 repeated measures ANCOVA with Years of Training as covariate. A main effect of Current Tone Direction, $F(1,33) = 5.89, p < .02$, was qualified by a Current Tone Direction x Previous Tone Direction interaction, $F(1,33) = 8.22, p < .01$. Post hoc analysis revealed that, as expected, contour-preserving ITIs were accelerated relative to contour-violating ITIs (AA < AD, $p < .02$; DD < DA, $p < .01$; AD = DA, $p < .17$, N.S). When analyzed separately by condition, this interaction did not remain significant at 25 cents Δf . In addition, descending contour-preserving ITIs were unexpectedly accelerated relative to ascending contour-violating ITIs (DD < AA, $p < .02$). There was also a main effect of Condition, $F(3,99) = 6.50, p < .001$. Post hoc analysis revealed a significant linear contrast, $F(1,34) = 18.25, p < .001$, indicating mean ITI negatively covaried with Δf . There were no significant interactions with Years of Training. Mean ITI values by contour and condition are displayed in Figure 1.

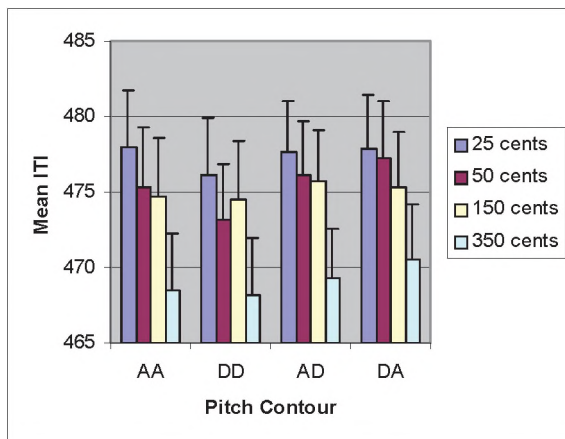


Figure 1. Mean ITI values by pitch contour and condition.

Mean tap force values according to Condition, Previous Tone Direction, and Next-to-Previous Tone Direction were entered into a 4 x 2 x 2 repeated measures ANCOVA with Years of Training as covariate. Only the Previous Tone Direction x Next-to-Previous Tone Direction interaction was significant, $F(1,33) = 12.684, p < .01$. Post hoc analysis revealed that, as expected, taps had greater force following a contour-preserving tone (AA < AD, $p < .04$; DD

< DA, $p < .001$; DD = AA, $p < .28$, N.S; AD = DA, $p < .19$, N.S). At 25 cents Δf , this interaction did not remain significant. We speculate that the inconsistencies between mean ITI and tap force findings for 1) linear effect of Δf and 2) directional asymmetry for contour-preserved tones might reflect smaller tap force effect sizes and the limited sensitivity of the apparatus to detect tap force variation.

Finally, we investigated whether the accumulation of accelerating ITIs would result in an accelerating tempo over the course of a trial. The slope of the linear regression of ITI against tap number was calculated for each participant's 35 trials. Slope values averaged across the seven trials of each condition significantly deviated from zero except in the monotone condition. Means for pitched feedback conditions were all negative, indicating accelerated ITI with increasing tap number. Slope values were entered into a repeated measures ANOVA and revealed a significant main effect of Condition, $F(4,132) = 15.25, p < .001$. Acceleration over the course of a trial increased with Δf , $F(1,33) = 38.3, p < .001$.

4. DISCUSSION

As predicted, when initiated by contour-preserving tones, intertap-intervals (ITI) were accelerated and of greater tap force relative to ITIs initiated by contour-violating tones. Contour-preserving ITI acceleration 1) was initiated rapidly (within ~500 ms); 2) did not vary with years of music training; 3) persisted with Δf of 50 cents or greater; 4) increased in magnitude with greater Δf , and 5) persisted over the course of a trial, the cumulative effect being an increase in tempo. An unexpected finding was that descending contour-preserving ITIs were significantly accelerated relative to ascending contour-preserving ITIs. These data support a strong version of the imputed velocity hypothesis, suggesting that participants were unable to disambiguate increases in velocity implied by pitch contour from the velocity of their finger trajectory.

REFERENCES

- Andres, M., Davare, M., Pesenti, M., Olivier, E. & Seron, X. (2004). Number magnitude and grip aperture interaction. *NeuroReport*, 15(18), 2773–7.
- Boltz, M. G. (1998). Tempo discrimination of musical patterns: Effects due to pitch and rhythmic structure. *Perception & Psychophysics*, 60, 1357–73.
- Jones, B., & Huang, Y. L. (1982). Space-time dependencies in psycho-physical judgment of extent and duration: Algebraic models of the tau and kappa effects. *Psychological Bulletin*, 91, 128–42.
- Kadosh, R.C. Brodsky, W., Levin, M. & Henik, A. (2008). Mental representation: What can pitch tell us about the distance effect? *Cortex* 44, 470–7.
- MacKenzie, N. (2007). The kappa effect in pitch/time context (Doctoral dissertation, Ohio State University, 2007). *Dissertation Abstracts International*, 68, 132.
- Stevens, L.T. (1886). On the time sense. *Mind*, 11, 393–404.

TEMPORAL DYNAMICS OF SELECTIVE ATTENTION DURING DICHOTIC LISTENING

Bernhard Ross

Rotman Research Institute, Baycrest Centre, Toronto, ON, Canada, M6A 2E1, brross@rotman-baycrest.on.ca
Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada

1. INTRODUCTION

In a multi speaker environment we can direct our attention to one conversation and ignore the voices of other simultaneous speakers. Nevertheless, our attention can still be captured if somebody outside of the focus of attention mentions our name. How is the brain able to work through these complex and often conflicting tasks?

A long standing debate is whether selective attention facilitates the sensitivity in the attended sensory channel and suppresses non-attended input [1,2] or alternatively all environmental information is pre-processed and stimulus selection takes place later at a higher order stage [3].

Advanced techniques of source analysis of magnetoencephalographic data provide spatial information about which brain area is modulated by attentional control. Moreover, MEG provides high resolution time courses of activity. We reconstructed the time courses of cortical source activity to study the neural processes underlying stimulus selection during dichotic listening.

2. METHOD

Twelve healthy normal hearing young university students participated in the study and listened to streams of dichotic sounds during whole head magnetoencephalographic (MEG) recordings.

The stimuli were amplitude modulated (AM) tones of 600 ms duration presented in random order to the left and right ear with stimulus onset asynchrony of 900-1100 ms. The modulation frequency of 40 Hz was infrequently changed to 20 Hz in both ears and listeners attended for the duration of a recording block of 7 min either to the left or right ear stimuli and responded to the targets in the attended ear with a right hand button press. A non-modulated 'filler' tone was presented opposite to the AM sound to reduce bottom-up effects of attention switching between ears. Thus, the participants to simultaneous streams of sound in both ears. Different carrier frequencies of 400 and 700 Hz for both ears supported maintaining attention to one ear.

MEG was recorded with a 275 channel whole head system. Auditory evoked magnetic fields were transformed into volumetric maps of source activity using the MEG beamformer approach. This resulted in time courses of % activity change for each volume element separately for standard and target stimuli and attention focused to the left and right ear. Repeated measures ANOVAs were performed

for each volume element to identify the effects of attention and stimulus type on the evoked brain activity as well as interactions between attention and stimulus type. Structural magnetic resonance images were recorded for the overlay of maps of MEG source activity on the individual anatomical image. The same spatial transformation of individual MRI onto an atlas brain was applied to the MEG data for group analysis in standardized Talairach coordinates.

3. RESULTS

Listeners were able to focus their attention on the tones in the required ear and accurately detected the targets. However, false positive responses were made to 3.6% of targets in the unattended ear, 0.2% of attended standards, and 0.06% of unattended standard stimuli. Thus, false positive responses were predominantly a result of interference by the contralateral deviant stimulus. The detectability for left ear sounds of $d' = 3.15$ was in group mean higher than for the right ear ($d' = 2.96$, $t(11) = 3.23$, $p = 0.008$). The median reaction time was 663 ms with respect to the stimulus and slightly faster for left than right ear stimuli ($t(11) = 2.4$, $p = 0.035$) consistently with the observed left ear advantage.

The auditory evoked responses showed predominantly a P1 wave and a sustained response, lasting for the duration of the stimulus, whereas the N1-P2 complex was only small because of the fast stimulation rate. The sustained response showed strong modulation with attention. In general responses were larger for target than standard stimuli and responses to both stimuli were larger when presented in the attended ear. Cortical sources were identified as spatially distinct areas with strongest contrast between responses to attended versus non-attended stimuli: bilateral Heschl's gyri, (HG) the location of primary auditory cortex, bilateral posterior superior temporal gyri (STG) and inferior parietal lobules (IPL), as well as bilateral inferior frontal gyri (IFG) and the central part of the medial frontal gyrus, location of the supplementary motor area (SMA).

Larger responses to stimuli presented in the attended ear regardless of stimulus type were observed as early as 150 ms after stimulus onset in bilateral HG and simultaneously in bilateral IFG. A similar effect of attention became significant at around 300 ms in bilateral STG and IPL however, the response increase under attention was much stronger expressed in the posterior sources than in HG. The over time developing evoked activity was in general larger for target than standard stimuli. This effect of the stimulus type became significant at around 300 ms latency in HG and IFG and shortly later at around 400 ms in STG and IPL. The

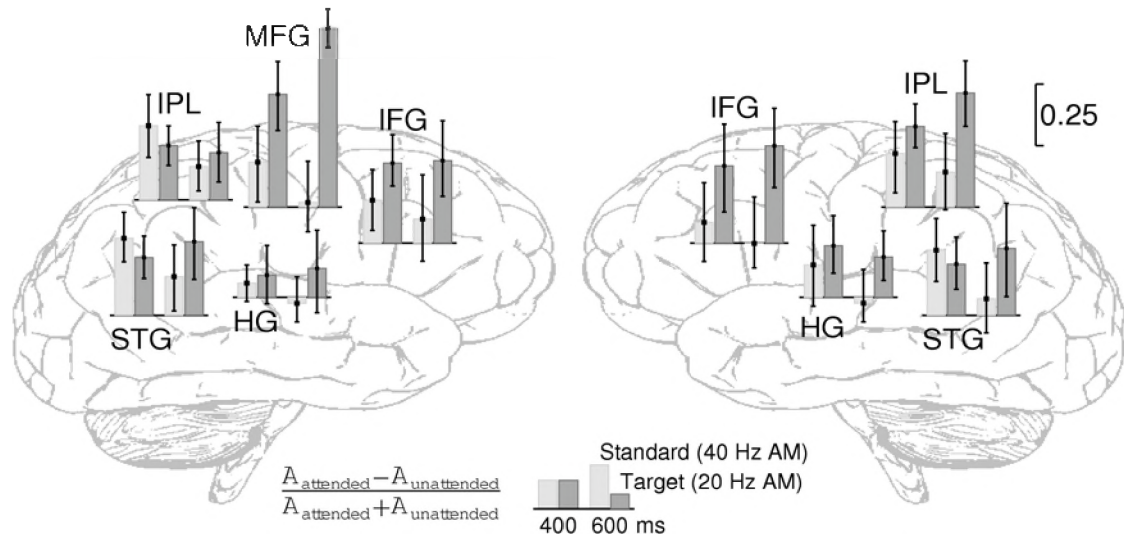


Figure 1. Effect of attention, expressed as relative amplitude increase for stimuli in the attended ear, on responses to standard and target stimuli at various brain regions at 400 and 600 ms latency. The effect of attention increases over time for the targets and decreases for the standard stimuli, indicating the stimulus selection process predominantly in posterior brain regions.

effect of larger responses to targets increased over time and reached a maximum at around 600 ms just before the subjects responded. Interactions between attention and stimulus type, expressed as larger effects of attention on target than standard stimuli were significant in posterior sources at around 400 ms latency. The interaction was strongest at 600 ms in SMA, where attended target stimuli only resulted in strong activity increase.

We expressed the effect of directed attention as the normalized amplitude difference between responses to stimuli in the attended and unattended ear and compared this measure between the stimulus types at latencies of 400 and 600 ms. The hypothesis was that the stimulus selection process would be expressed as initially large effect of attention for both stimuli but the effect would diminish over time for standards and increase for targets. Indeed, the attention effect increased significantly for the targets and decreased for the standards. This interaction is demonstrated in the bar graphs shown in Fig. 1. The effect of ‘time’ was significant for all sources, whereas the interaction between ‘stimulus type’ and ‘time’ was significant for the posterior sources in IPL and STG indicating that the posterior sources were strongly involved in the process of stimulus selection.

In similar way we analyzed the response increase for targets, which may reflect a bottom-up capture of attention by the infrequent targets. No significant interaction between ‘attention’ and ‘time’ was found, indicating that the salience of targets increased over time for stimuli in both attended and unattended ears. This temporal dynamic was most pronounced in the posterior sources. Whereas the contrast between attended target and standard responses increased over time through suppression of responses to standard sounds, the responses to unattended targets even increased.

4. Discussion

Using advanced MEG data analysis approaches we identified spatial maps of cortical source activity underlying attention control during dichotic listening. More importantly, we obtained time courses of activation with high temporal resolution. Directing attention to one ear enhanced the sensitivity in the sensory channel and was associated with increased activation in HG and IFG at latencies 150-500 ms. A later effect of attention was different on target and standard stimuli and identified IPL and STG as the location of stimulus discrimination during the 400-600 ms latency interval. The responses to targets presented to the unattended ear showed an increase of activity during the time interval of stimulus discrimination indicating that despite the higher sensitivity in the selected channel discrimination still proceeded in the ignored channels. Time courses of early activation in IFG as well as early onset of effects of attention and effects of stimulus type suggest a role of IFG in monitoring auditory input, maintaining attention to the selected sensory channel, and controlling stimulus discrimination.

5. REFERENCES

- [1] Fritz JB, Elhilali M, David SV, Shamma SA (2007) Auditory attention--focusing the searchlight on sound. *Curr Opin Neurobiol* 17: 437-455.
- [2] Giard MH, Fort A, Mouchetant-Rostaing Y, Pernier J (2000) Neurophysiological mechanisms of auditory selective attention in humans. *Front Biosci* 5: D84-94.
- [3] Treisman A (1998) Feature binding, attention and object perception. *Philos Trans R Soc Lond B Biol Sci* 353: 1295-1306.

ACKNOWLEDGEMENTS

Funded by the Canadian Institutes of Health Research

HEARING PROTECTORS STANDARDS

Alberto Behar¹ and Willy Wong¹

IBBME, University of Toronto, 164 College Street, Toronto, ON,
Canada, M5S 3G9. alberto.behar@utoronto.ca

1. INTRODUCTION

A Standard is a written document published by a Standard institution that can be national or international. Examples of national institutions are the Canadian Standard Association (CSA) or the American National Standard Institute (ANSI) while the International Organization for Standardization (ISO) is an example of an international institution. In this presentation we will deal exclusively with standards from those three organizations.

All standards writing organizations are non-government, although they may get some funding for governmental sources. They provide mainly supporting functions (secretarial, publishing, etc.) while the writing is done by experts from the academia, manufacturers and users, all of them working ad honorem. Standards are reviewed periodically, usually once every 5 years by the same groups of experts.

Hearing protectors' standards can be divided in three groups:

- a) Measurement Standards, that deal with the characteristics to be measured,
- b) Applications Standards, that describe what to do with the results of the measurements, and
- c) Combination of both.

The principal issue with hearing protectors is to determine the sound level of the protected ear that is the sound level that reaches the tympanic membrane when the individual wears his protectors. For this purpose, what has to be measured is the hearing protectors' attenuation.

2. MEASUREMENT STANDARDS.

They describe how to measure the sound attenuation at the threshold of hearing protectors. The attenuation is defined as the difference in sound levels that elicit the threshold of hearing when the subject is protected and when he is not wearing the protector. Testing is done in a semi-anechoic room, using 1/3 octave bands of white noise at the seven audiometric frequencies between 125 and 8,000 Hz. The measurement is repeated three times on a number of subjects that vary between 10 and 20, depending on the type of protector being tested.

Thus the result of the measurement is a table containing attenuation and Standard Deviation at all 7 frequencies

Two standards belong to this category:

- a) ANSI/ASA S12.6-2008 Methods for Measuring the Real-Ear Attenuation of Hearing Protectors, and

- b) ISO 4869-1:1990: Acoustics - Hearing protectors - Part 1: Subjective method for the measurement of sound attenuation

A variation of the above measuring process is described in:

- a) ISO 4869-3:2007 Acoustics - Hearing protectors - Part 3: Measurement of insertion loss of ear-muff type protectors using an acoustic test fixture.

This standard requires the use of an ATF, also known as an artificial head. The characteristic that is measured in this case is the insertion loss defined as the difference between the sound levels outside of the head and the one measured with the microphone contained in the device, when the protector is worn.

- b) ISO/TS 4869-5:2006 Acoustics - Hearing protectors - Part 5: Method for estimation of noise reduction using fitting by inexperienced test subject.

In this case, the test subjects are persons that have no experience in using hearing protectors. The protocol is also changed, so they do not receive instructions other than the one printed on the protectors' containers and,

- c) ISO/TR 4869-4:1998 Acoustics - Hearing protectors - Part 4: Measurement of effective sound pressure levels for level-dependent sound-restoration ear-muffs

The type of hearing protectors described here (mainly muffs) have a microphone on the outside, that captures the environmental sound level, that is amplified and fed into a loudspeaker inside the protector. The amplification is selective and it is equal to 0 when the inside noise level exceeds 85 dBA. Those protectors are useful in situations where the noise level is moderate, but can eventually exceed the 85 dBA limit. By amplifying the sound it improves the intelligibility of speech with the protector on.

3. APPLICATION STANDARDS

The standards in this group use the results from the measured standards to calculate the noise level of the protected ear in dBA.

- ISO 4869-2/Cor1:2006 Acoustics - Hearing protectors - Part 2: Estimation of effective A-weighted sound pressure levels when hearing protectors are worn - Corrigendum

This standard describes three different ways of this calculation, each one with a different precision. The more precise is also the more complex. It requires the

measurement of the ambient noise in 1/1 octave bands. The difficulty consists of obtaining a measurement that can be representative for the entire work day. The assumption being made is that the person is exposed at the same noise level and by the same spectral content continuously throughout the whole day. A second approach provides a calculation that limits the result to only three numbers that have to be combined to obtain the desired outcome. Finally, there is also the third procedure, much simpler than the previous two, where a one-number-estimate is subtracted from the ambient noise level measured in dBC.

ANSI/ASA S12.68-2007 Methods of Estimating Effective A-Weighted Sound Pressure Levels When Hearing Protectors are Worn.

This Standard also provides three methods, one of which is similar to the 1/1 octave band, mentioned above. The other two provides procedures for the calculation of two estimates. Each one is to be subtracted from the ambient noise level measured in dBA. Each one can predict the attenuation that given % of the protected population will experiment. It is recommended to be used for the upper 20% (the best protected) and the 80% for the general population. In such a way the person in charge of managing the Hearing Protectors' program in the workplace, knowing the degree

of knowledge and conscience of the personnel, can decide which number to use for the estimation of the noise level of the protected ear.

The approach in this Standard is accepted as the state of the art and it is possible that the ISO will be moving in the same direction in the near future.

4. COMBINATION STANDARD

CAN/CSA-Z94.2-02 (R2007): Hearing Protection Devices - Performance, Selection, Care, and Use.

This Standard is intended to serve as guidance to manufacturers and users alike. It specifies that the attenuation should be measured following the **ANSI/ASA S12.6-2008** Standard. It proposes two ways of calculating the noise level of the protected ear using two estimates. It is expected that in a revision that is due shortly, this Standard will follow the procedures in the **ANSI/ASA S12.68-2007**.

The Standard contains an extensive section for the user that includes information and orientation regarding the selection, care and use of the protector. In addition, there is information regarding the non-traditional protectors, their usage and application.

MONITORING SOUND PRESSURE AT THE EARDRUM FOR HEARING CONSERVATION

Anthony J. Brammer^{1,2}, Gongqiang Yu¹, Eric R. Bernstein¹, Donald R. Peterson¹,
Martin G. Cherniack¹, and Jennifer B. Tufts³

¹Ergonomic Technology Center, Univ. of Connecticut Health Center, Farmington CT, U.S.A. 06030-2017

²Inst. for Microstructural Sciences, National Research Council, Montreal Rd., Ottawa ON, Canada K1A 0R6

³Dept. of Communication Sciences, Univ. of Connecticut, Storrs CT, U.S.A. 06269-1085

1. INTRODUCTION

It is well known that the sound pressure at the eardrum differs from that measured near the head, or in the absence of the subject. The differences arise from the interaction of structures forming the external ear, and the head and torso, with the sound field. Amplification of sound reaching the eardrum by the pinna occurs at frequencies above ~1kHz, and by the head and torso above ~200 Hz, depending on source characteristics (Brammer and Piercy, 1977). The relationship between the sound pressure at the eardrum and that of an external sound field is further modified, of course, by the use of hearing protectors.

The need to monitor sound pressures at the eardrum for the purposes of hearing conservation has long been recognized. A method for monitoring sound pressures within the ear was described by Brammer and Piercy, who for "practical considerations" mounted a small microphone on the base of the concha, and developed pressure transformations to estimate the sound pressure at the center-head position in the absence of the subject (Brammer and Piercy, 1977). The use of a probe microphone for noise dosimetry was described by Shotland, who inserted the probe tip a pre-determined distance into the ear canal (Shotland, 1996). No attempt was made to determine the pressure transformation to the eardrum. The probe microphone sensed sound pressures up to 4 kHz, and was held in place by a headband, which limited the use of hearing protection. An increase in exposure was observed compared to that recorded at the shoulder, or chest for all subjects. The increase depended on the individual and differed sufficiently between individuals (by up to 6 dB) to suggest external ear gain as a risk factor for noise-induced hearing loss.

In this paper, a probe microphone assembly is described that permits the uncertainty in probe-tip position to be eliminated, and hence the transfer function to the eardrum to be estimated. The microphone is attached to an ear mold designed to self-locate within the concha and ear canal, and so reproducibly position the probe tip. For the purposes of the present work, an upper frequency limit of 6 kHz was selected for the device, and a target accuracy of ± 2 dB.

2. APPARATUS AND METHOD

The device consists of a miniature probe microphone and a custom-fitted ear mold. The ear mold is fabricated from an ear impression that is obtained for each subject

following established audiological procedures. The ear impression provides the contour of the perimeter of the concha and cymba, which is required to produce the 'C'-shaped structure used to position the ear mold external to the ear canal. A third arm of the ear mold rests along the upper surface of the ear canal, and contains a hole through which the probe tube is inserted, as can be seen from Fig. 1.

The probe microphone is constructed from a miniature electret microphone (Knowles FG 23652), to which is attached a 35 mm soft silicone tube, with outer diameter of 0.94 mm. This assembly has a Helmholtz resonance frequency ~1.5 kHz, as well as probe tube length resonances that commence at ~4.5 kHz. The overall response is digitally equalized to within ± 1 dB at frequencies up to 6 kHz (and with somewhat less precision up to 12 kHz). The output of the probe microphone is amplified by a custom-designed, low-noise preamplifier with a flat frequency response from 20 Hz to 17 kHz, and voltage gain of 34 dB.

The device is assembled after first locating the ear mold in the external ear, and then feeding the probe through the hole in its supporting tube. For this procedure, the subject is seated in an anechoic chamber with a sound source on the inter-aural axis, 1.5 m from the ear. The loudspeaker output was flat (on axis) to within ± 2 dB from 100 Hz to 14 kHz (PSB Image 2B). The tip of the probe tube is adjusted by observing the minimum frequency of the standing wave pattern at different probe locations in the ear canal (Gilman and Dirks, 1986, Stinson, 1990). A frequency of 11 kHz was selected for the target standing wave minimum, and corresponded to the smallest separation between probe tip and eardrum considered safely achievable for all subjects, male and female, irrespective of the length of the ear canal. The microphone assembly is glued to the ear mold when the desired frequency was obtained.

All subjects gave their informed consent to participate in the study, following the provisions of the ethics committee of the University of Connecticut Health Center.

3. RESULTS AND DISCUSSION

3.1 Standing Wave Minima in Ear Canal

The microphone output for three penetration depths of the probe tip in a subject's ear canal are shown in Fig. 2. Inspection of the diagram reveals features that have been observed in measurements conducted on all subjects.

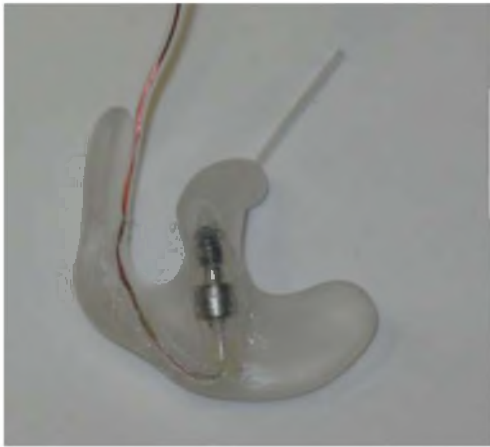


Figure 1. Photograph of probe microphone and ear mold

At frequencies below about 2 kHz, the three spectra are indistinguishable from each other, indicating that the sound pressure within the ear canal is insensitive to the position of the probe tip at these frequencies, as expected (Gilman and Dirks, 1986). All spectra contain common features at some other frequencies, namely a maximum between 2.5 and 3 kHz, and a large minimum between 6.5 and 8.5 kHz. These reflect specific features of the head-related transfer function and are not believed to be associated with the standing wave pattern in the ear canal, which would tend to introduce differences between the spectra. When the common features of the spectra are discounted, three minima at frequencies of approximately 6, 8.8 and 11 kHz can be distinguished, corresponding to the different penetration depths of the probe tip. The resolution of the minimum at 8.8 kHz is poorly defined in Fig. 2, possibly because it overlaps with the dominant feature of the head related transfer function between 6.5 and 8.5 kHz.

3.2 Sources of Error

Variation in position of probe tip on repeated insertion

The variability in the position of the probe tip on repeated insertion in the ear canal is difficult to measure directly, particularly when the probe tip is close to the eardrum. However, a change in penetration depth of the probe tip, and hence position relative to the eardrum, will result in a change in the standing wave pattern. Accordingly, the frequency of the standing wave minimum has been recorded for repeated placement of the ear mold in the ear. For six such repetitions, the lowest to highest minimum frequency ranged from 2% to 4% for different ears. Reference to Fig. 2 reveals that a 20% reduction in standing wave minimum frequency (i.e., from 11 kHz to 8.8 kHz) will introduce a change in microphone output of a maximum of ~4 dB at 6 kHz, and less at lower frequencies. A variation of 2% to 4% will therefore introduce a negligible change in microphone output (<0.5 dB).

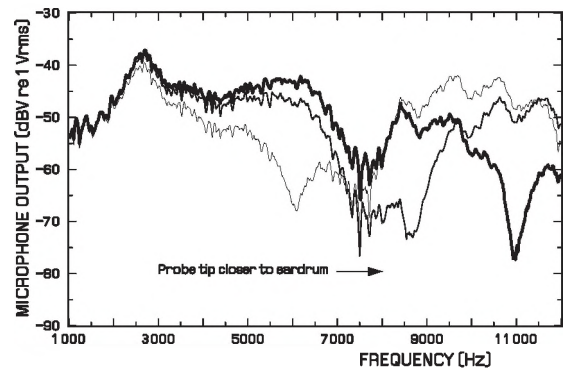


Figure 2. Microphone output for three positions of the probe tip

Influence of probe on sound pressure in ear canal

The perturbation of the sound field by the presence of the ear mold has been evaluated by inserting a second probe microphone with a long, small diameter, probe tube into the ear canal (Etymotic ER-7C). A systematic variation in sound pressure level (SPL) with frequency has been found when the ear mold is first present and then removed, which appears to depend somewhat on the physical dimensions of the external ear. Compensation for this effect appears possible within the error budget of ± 2 dB.

Eardrum impedance

Although the probe tip is positioned close to the eardrum, it is desirable to reconstruct the SPL at the eardrum. Such estimates will be influenced by differences in eardrum impedance. If typical energy reflection coefficients for human ears are employed (Stinson, 1990), the predicted SPLs at the eardrum differ systematically from the measured SPLs by 3.6 dB at 6 kHz, and progressively less at lower frequencies. While the maximum variation in the energy reflection coefficients between ears will introduce an error in the transformed SPLs of <0.5 dB, the error arising from differences in the phase of the eardrum impedance between ears is less certain. Gilman and Dirks estimated a total variation from eardrum impedance of close to 2 dB for a standing wave minimum at 8 kHz. It remains a subject for future work to confirm that the phase error of the transformation remains within the error budget.

REFERENCES

- Brammer, A. J., and Piercy, J. E. (1977). Monitoring sound pressures within the ear: Application to noise exposure. *J. Acoust. Soc. Am.* **61**, 731-738.
- Gilman, S., and Dirks, D. D. (1986). Acoustics of ear canal measurement of eardrum SPL in simulators. *J. Acoust. Soc. Am.* **80**, 783-793.
- Stinson, M. R. (1990). Revision of estimates of acoustic energy reflectance at the human eardrum. *J. Acoust. Soc. Am.* **88**, 1773-1778.
- Shotland, L. I. (1996). Dosimetry measurements using a probe tube microphone in the ear canal. *J. Acoust. Soc. Am.* **99**, 979-984.

ACKNOWLEDGEMENTS

This work was supported by NIOSH (R01 OH008669).

ACOUSTIC LOCALIZATION OF AN AUTONOMOUS UNDERWATER VEHICLE

Nicos Pelavas, Garry J. Heard, Gordon R. Ebbeson, Carmen Lucas, Derek Clark
Defence R&D Canada - Atlantic, 9 Grove St., Dartmouth NS, Canada, B3A 3C5

1. INTRODUCTION

The trend towards melting of arctic ice has resulted in the prospect of the creation of new shipping lanes, and arguably more important potential access to natural resources beneath the Arctic Ocean. Countries with northern regions have begun the task of collecting bathymetry data to be used in determining their sovereign arctic territory subject to the oversight of the United Nations Convention on the Law of the Sea (UNCLOS). Canada has planned missions to map portions of the Arctic Ocean floor using two Autonomous Underwater Vehicles (AUV). The AUVs will collect data along a 400 km path in the neighborhood of Lomonosov and Alpha ridges at depths between 5 and 6 km.

DRDC Atlantic is assisting with the development of the vehicles and providing support during the deployment and recovery phases of the mission. Initially, once the AUV is released through an ice hole it travels downward along a helical trajectory before arriving at operating depth. At this point it is necessary to perform a localization of the AUV in order to aid in calibration of the inertial navigation unit. In the final phase of the mission a similar localization is needed to guide the AUV to the predetermined position of the recovery hole. In this summary we shall briefly review the approach used to implement the localization and present the findings from data collected during an arctic trial based out of CFS Alert.

2. LOCALIZATION METHOD

The term, localization, refers to an estimate of the position of the AUV with respect to a coordinate system set up on the ice surface. Typically, we choose a three dimensional Cartesian coordinate system with origin corresponding to the deployment/recovery hole. Short range localization is conducted using a network of Teledyne Benthos acoustic modems where one of the modems is contained in the AUV and at least four modems are suspended through holes distributed over the ice surface. In this arrangement, one of the surface modems is placed at the deployment/recovery hole and designated the gateway modem. All network communications are controlled through the gateway modem, and using either a broadcast or polling request we can obtain the time of travel for an acoustic pulse between any two modems. By taking into consideration the sound velocity profile (SVP) we can deduce the range between modems. In particular, let us denote the range from the i^{th} surface modem to the AUV as d_i and the coordinates of the i^{th} surface modem to be (x_i, y_i, z_i) . Moreover, setting $\mathbf{p} = (\mathbf{x}, \mathbf{y}, \mathbf{z})$ as the unknown AUV position gives the following nonlinear system of error equations

$$\varepsilon_i = d_i - \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2}, \quad (1)$$

where $i = 1, \dots, N$ indexes the $N \geq 4$ surface modems and $\varepsilon_i = \varepsilon_i(\mathbf{p})$ represents the error associated with the position and range of the i^{th} modem. If the ranges and positions of the surface modems are exact then system (1) will, in general, yield a unique solution when $\varepsilon_i = 0$ for all i . Special cases exist in which degenerate surface modem configurations result in non-unique solutions to (1), these can be avoided by choosing different modem positions or increasing N . However, in practice we seek a solution to (1) that minimizes $\sum_i \varepsilon_i^2$. Let $\tilde{\mathbf{p}}_0 = (\tilde{x}_0, \tilde{y}_0, \tilde{z}_0)$ be an initial guess for the AUV position and denote by $\nabla \varepsilon_i$ the Jacobian matrix of ε_i with respect to $\mathbf{p} = (\mathbf{x}, \mathbf{y}, \mathbf{z})$. Applying Newton's method, we linearize (1) to obtain the following iterative system

$$\varepsilon_i(\tilde{\mathbf{p}}_{k+1}) = \varepsilon_i(\tilde{\mathbf{p}}_k) + \nabla \varepsilon_i(\tilde{\mathbf{p}}_k) \cdot \Delta \tilde{\mathbf{p}}_k, \quad k \geq 0. \quad (2)$$

In equation (2) we have defined $\Delta \tilde{\mathbf{p}}_k = \tilde{\mathbf{p}}_{k+1} - \tilde{\mathbf{p}}_k$, thus for each fixed k we impose constraints, $\varepsilon_i(\tilde{\mathbf{p}}_{k+1}) = 0$, resulting in an overdetermined linear system for the unknown vector $\Delta \tilde{\mathbf{p}}_k$. Since this $N \times 3$ system will generally have no solution we use a singular value decomposition of the matrix $\nabla \varepsilon_i$ to determine a solution with the property that the magnitude of the residual vector, $\varepsilon_i(\tilde{\mathbf{p}}_{k+1})$, is minimized. Solving for $\Delta \tilde{\mathbf{p}}_k$ then gives $\tilde{\mathbf{p}}_{k+1}$. Iterations continue until $\|\varepsilon_i(\tilde{\mathbf{p}}_{k+1}) - \varepsilon_i(\tilde{\mathbf{p}}_k)\| < \delta \|\varepsilon_i(\tilde{\mathbf{p}}_k)\|$ where δ is a sufficiently small positive number. In the applications below we chose to set $\delta = 1 \times 10^{-4}$ and singular values with absolute value less than 1×10^{-2} were set to zero.

Various surface modem configurations were considered in order to simulate deployment and recovery of the AUV. Here we provide two examples from the recovery phase assuming a normal distribution of range noise that is 0.1% of the actual range. Using six surface modems all at the same depth, a coordinate system is chosen such that $z_i = 0$. In addition, one of the modems is always located at the origin. The simulated AUV travels in the $x - z$ plane beginning at $x = 1 \times 10^4$ and $z = -5 \times 10^3$ with subsequent positions having decreasing x and increasing z . In the first configuration, the remaining five modems are positioned on a circle of radius 1000 about the origin starting with $x_1 = 0, y_1 = 1000$ and then placed at $2\pi/5$ angular increments. The second configuration is identical to the first except the circle has radius 100. Graphs in Figure 1 show the results of this simulation, they are arranged so that the first (second) column corresponds to x, y , and z coordinates of the first (second) configuration. The horizontal axis represents samples of AUV positions.

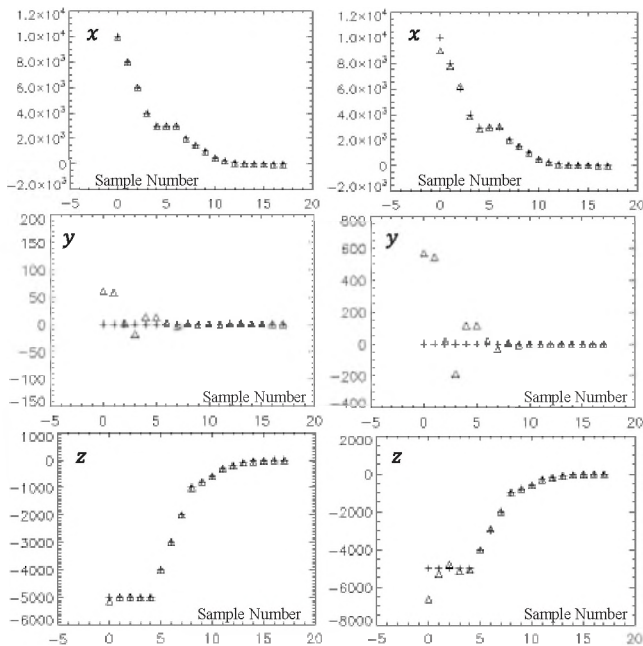


Figure 1. Surface modems at radius 1000 (100) in left (right) column. Actual AUV positions given by + and estimates by Δ As expected, accuracy in position estimation generally increases as the AUV approaches the field of surface modems. Furthermore, it is evident from Figure 1 that larger configurations result in better accuracy than equivalent smaller configurations. Comparing x and z graphs for each configuration we note that deviations in z are typically larger than in x . This may be attributed to the fact that all modems were placed at the same depth whereas the x coordinates were distributed over a circle.

3. RESULTS OF ARCTIC TRIAL

In the spring of 2009 an arctic trial was conducted on the ice near CFS Alert. One of the objectives was to prove the efficacy of the short range localization method and to reveal any potential problems with its implementation.

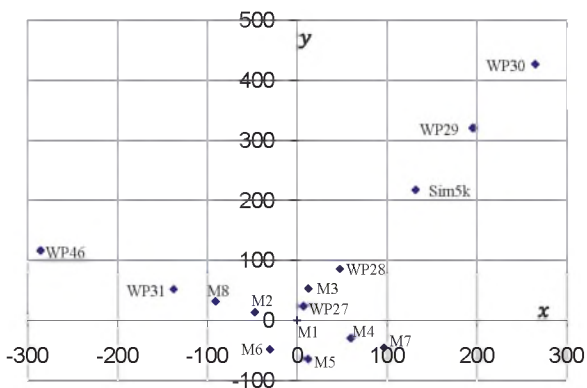


Figure 2. Modem positions labeled M1-M8 along with simulated AUV positions, in meters.

The area over which the experiment was carried out is given in Figure 2; water depth in this region was approximately 100 m. Modems M1 to M8 were used to perform the

localization and the remaining waypoints (WPn, Sim5k) were used as simulated AUV positions. Coordinates of all points were measured using a geodimeter, the vertical displacements were all within 8 cm thus deviations in elevation due to ice roughness are negligible. M1 to M8 were lowered to a depth of 50 m, approximately the center of the water column. At each of the remaining seven waypoints another modem, simulating the AUV, was lowered to depths of 10, 30, 50, 70 and 90 m and the time of travel for an acoustic pulse to M1 through M8 was measured. The corresponding ranges were determined by taking a SVP in this region which gave an average sound speed of 1436 m/s with a standard deviation of 2.5 m/s, indicating that ray paths can be regarded as straight given the short distances considered.

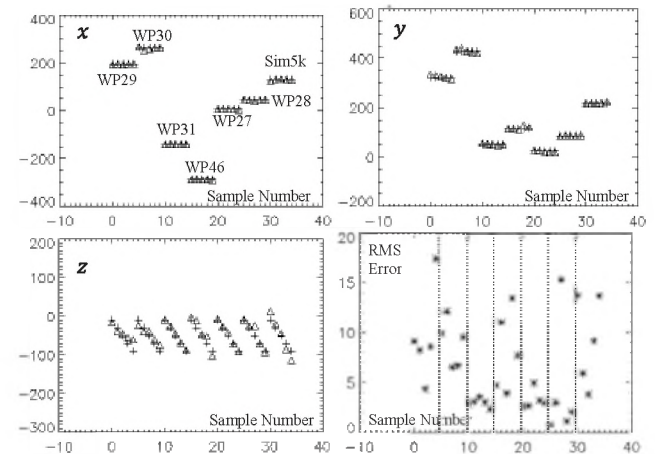


Figure 3. Estimated and actual AUV positions grouped in each graph according to their corresponding waypoints, as well as associated root mean square errors.

It is worth mentioning that at some waypoints and depths a broadcast request resulted in no response from one or two of the modems. In addition, modems that did respond displayed a 10 s delay in returning travel time data. We expect that polling individual modems will reduce this time delay and possibly the number of no responses. The root mean square errors in Figure 3 are all less than 18 m, by considering the ranges of simulated AUV positions and possible multipath effects of shallow water conditions we anticipate that a deeper water environment will give estimation errors less than 5%.

4. CONCLUSION

The singular value decomposition has been used to determine a least squares solution of a system containing errors arising from modem ranges and positions. We have shown that larger spans in x , y , and z of surface modem configurations result in greater accuracy of estimating AUV positions. Arctic trial data supports the feasibility of this method in short range localization.

REFERENCES

Press, W. H. *et al* (1992). Numerical Recipes in C, 2nd ed. Cambridge University Press. Cambridge.

BAYESIAN OCEAN ACOUSTIC SOURCE TRACK WITH ENVIRONMENTAL UNCERTAINTY

Stan E. Dosso and Michael J. Wilmut

School of Earth and Ocean Sciences, University of Victoria, Victoria BC Canada V8W 3P6, sdosso@uvic.ca

1. INTRODUCTION

This paper considers matched-field tracking of a moving acoustic source in the ocean when acoustical properties of the environment (water column and seabed) are poorly known. The goal is not simply to estimate source locations but to determine track uncertainty distributions, thereby quantifying the information content of the tracking process. A Bayesian formulation [1, 2] is applied in which source and environmental parameters are considered unknown random variables constrained by noisy acoustic data and by prior information on parameter values (e.g., physical limits for environmental properties) and on inter-parameter relationships (limits on source velocity). Source information is extracted from the posterior probability density (PPD) by integrating over unknown environmental parameters to obtain a time-ordered series of joint marginal probability surfaces over source range and depth. Given the strong nonlinearity of the localization problem, marginal PPDs are computed numerically using efficient Markov-chain Monte Carlo (MCMC) methods, including Metropolis-Hastings sampling over environmental parameters (rotated into principal components and applying linearized proposal distributions) and heat-bath Gibbs sampling over source locations [1, 2]. The approach is illustrated here using acoustic data collected in the Mediterranean Sea, with tracking information content considered as a function of data quantity (number of time samples).

2. RESULTS

The acoustic experiment considered in this paper was carried out in 130 m of water off the west coast of Italy (near Elba Island) in the Mediterranean Sea (see Fig. 1). Acoustic data were measured at a vertical sensor array consisting of 24 hydrophones at 4-m spacing from 26- to 118-m depth. A ship-towed acoustic source (transducer) at a depth of 12 m transmitted a linear frequency-modulated “chirp” signal every ~0.25 km along a radial track from approximately 2-6 km range. Complex (frequency-domain) acoustic fields at 300 Hz are considered here for source tracking, with a signal-to-noise ratio of approximately 0 dB.

The unknown environment and source parameters considered in the tracking problem are illustrated in Fig. 2. Seabed geoacoustic parameters include the thickness h of an upper sediment layer with sound speed c_s , density ρ_s , and attenuation α_s , overlying a semi-infinite basement with sound speed c_b , density ρ_b , and attenuation α_b . The water

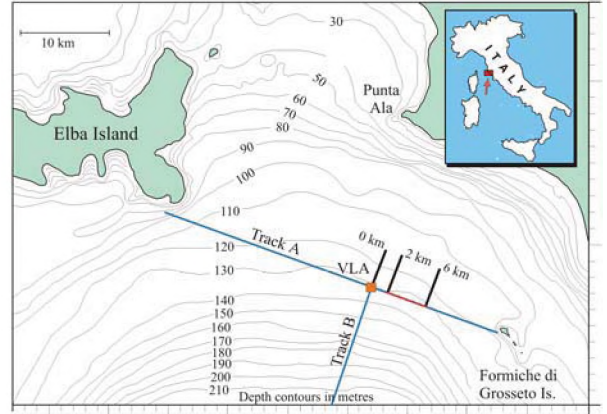


Fig. 1. Location of acoustic experiment.

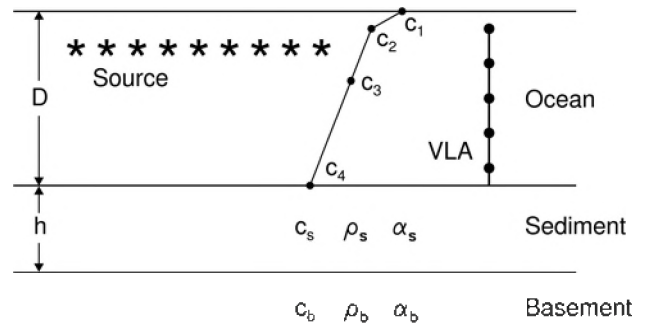


Fig. 2. Experiment geometry and model parameters.

depth is D , and the water-column sound-speed profile is represented by four parameters c_1 – c_4 at depths of 0, 10, 50, and D m. Wide uniform prior distributions (search intervals) are assumed for all parameters.

For efficiency sake, most source tracking algorithms operate in a sequential mode in which the component data sets of a series of acoustic measurements collected at a sequence of times are each inverted independently for the source location at that particular time. However, the total data information content is maximized by simultaneous inversion of all data sets for a sequence of source locations. This is particularly true for source tracking in an unknown environment, since simultaneous inversion brings all data information to bear on constraining environmental parameters, which in turn leads to better estimation of source locations. Further, applying constraints

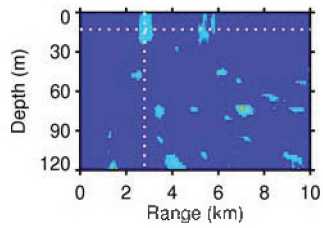


Fig. 3. Joint marginal probability distribution computed for a single source location at the start of the track. Dotted lines indicate the true source ranges and depths.

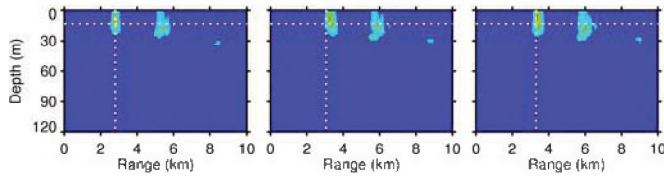


Fig. 4. Joint marginal probability distributions computed for the first 3 source locations along the track. Dotted lines indicate the true source ranges and depths.

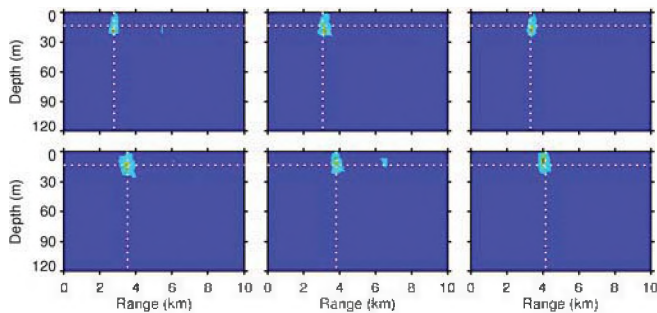


Fig. 5. Joint marginal probability distributions computed for the first 6 source locations along the track. Dotted lines indicate the true source ranges and depths.

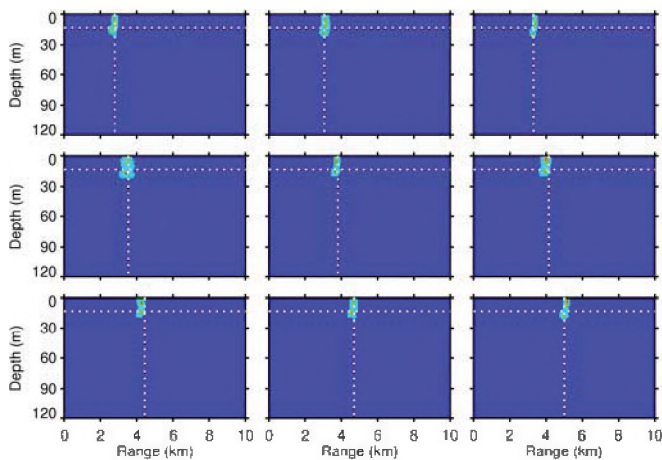


Fig. 6. Joint marginal probability distributions computed for 9 source locations along the track. Dotted lines indicate the true source ranges and depths.

on the maximum horizontal and radial source velocity as prior information provides a link between the various source locations along the track. This has the effect that each source location along the track is constrained by all of the recorded acoustic data, not just the data measured while the source was at that particular location.

The advantages of simultaneous inversion are illustrated in Figs. 3-6. Figure 3 shows the joint marginal probability distribution over source range and depth for the first source location along the track computed by inverting the acoustic data for the first source transmission (i.e., integrating the PPD over unknown environmental parameters via MCMC), with prior limits of 10 and 0.2 m/s for source horizontal and vertical velocity. Although a region of elevated probability (above the background) is associated with the true source location, there are many such local maxima, several with higher probabilities than that at the true location. Hence, the data information content is not sufficient for reliable localization in this case, given data noise and environmental uncertainties. Figure 4 shows marginal probability distributions for three source locations computed by inverting the first three source transmissions. The highest probability peak occurs at the true source location for each transmission, with only two other local maxima in probability, indicating reasonably reliable localization/tracking. Of particular note is the substantial improvement in localization for the first source position over that shown in Fig. 3 for a single transmission. This improvement results from the fact that, in simultaneous inversion with source velocity constraints, the acoustic data for the second and third source transmissions help constrain the source location at the time of the first transmission. This effect is further illustrated in Figs. 5 and 6, which show localization marginal probability distributions computed for six and nine source transmissions, respectively. In each case the ability to localize/track the acoustic source at early times along the track is substantially improved by data collected at later times (which, of course, also leads to improved results at later times). This improvement is only possible through simultaneous inversion.

ACKNOWLEDGEMENTS

The authors thank Dr. Peter Nielsen of the NATO Undersea Research Centre (NURC) for providing the Mediterranean Sea acoustic data.

REFERENCES

- [1] Dosso, S.E. and M. J. Wilmut, 2008. Uncertainty estimation in simultaneous Bayesian tracking and environmental inversion. *J. Acoust. Soc. Am.*, **124**, 82-97.
- [2] Dosso, S.E. and M. J. Wilmut, 2009. Comparison of focalization and marginalization for Bayesian tracking in an uncertain ocean environment. *J. Acoust. Soc. Am.*, **125**, 717-722.

BAYESIAN SOURCE TRACK PREDICTION IN AN UNCERTAIN ENVIRONMENTAL INVERSION

Stan E. Dosso and Michael J. Wilmut

School of Earth and Ocean Sciences, University of Victoria, Victoria BC Canada V8W 3P6, sdosso@uvic.ca

1. INTRODUCTION

This paper considers probabilistic prediction of the future locations of a moving acoustic source in the ocean based on past locations as determined by Bayesian source tracking in an uncertain environment [1, 2]. The Bayesian tracking approach considers both source and environmental parameters as unknown random variables constrained by noisy acoustic data and prior information, and integrates the posterior probability density (PPD) over the environmental parameters to obtain a time-ordered series of joint marginal probability surfaces over source range and depth. The integration is carried out using Markov-chain Monte Carlo (MCMC) methods which provide a large sample of track realizations drawn from the PPD. Applying a probabilistic model for source motion to each of these realizations produces a sequence of source range-depth probability distributions for future times. These predictions account for both the uncertainty of the source-motion model and the uncertainty in the state of knowledge of past source locations, which is itself dependent on environmental uncertainty. The approach is illustrated for range-depth tracking using a vertical sensor array for two source-motion models which differ in the degree of confidence assigned to future predictions based on past localizations.

2. TRACK PREDICTION EXAMPLES

The example illustrates tracking/track prediction for a quiet submerged source in shallow water with little knowledge of environmental parameters. The unknown environment and source parameters are illustrated in Fig. 1. Seabed geoacoustic parameters include the thickness h of an upper sediment layer with sound speed c_s , density ρ_s , and attenuation α_s , overlying a semi-infinite basement with sound speed c_b , density ρ_b , and attenuation α_b . The water depth is D , and the water-column sound-speed profile is represented by four parameters c_1 – c_4 at depths of 0, 10, 50, and D m. Wide uniform prior distributions (search intervals) are assumed for all parameters. Acoustic data are measured at 300 Hz at a vertical array consisting of 24 sensors at 4-m spacing from 26- to 118-m depth (simulated acoustic fields are computed using a normal-mode propagation model). The track consists of an acoustic source at 30-m depth moving toward the array at a constant radial velocity of 5 m/s (~10 kts). Acoustic data are collected at the array once per minute for 9 minutes, corresponding to source-receiver ranges of 6.4, 6.1, ..., 4.0 km. Random complex-Gaussian

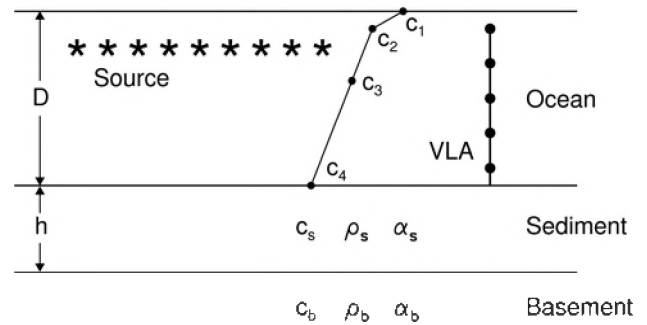


Fig. 1. Experiment geometry and model parameters.

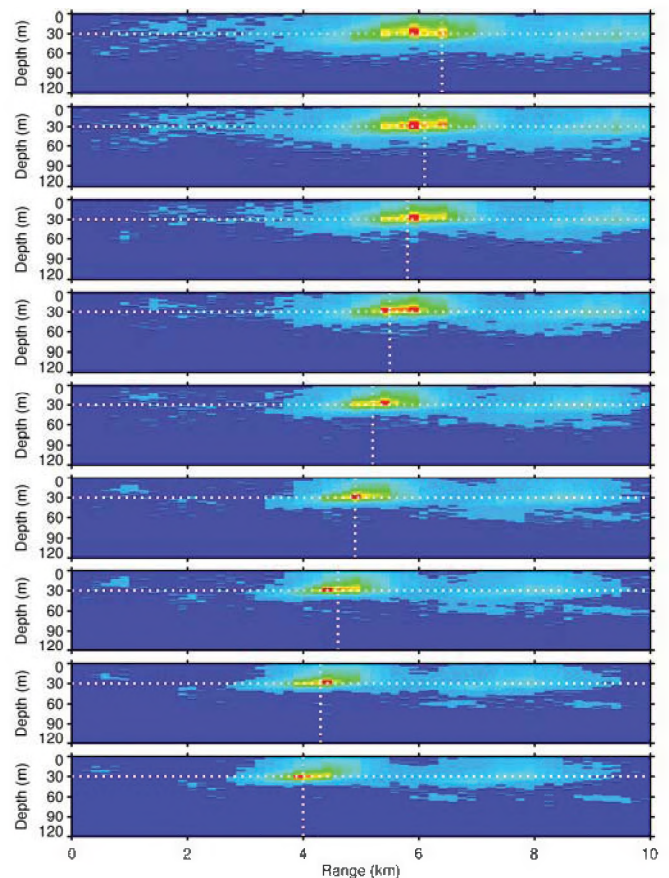


Fig. 2. Joint marginal probability distributions for source location computed via Bayesian tracking (time advances by 1 min/panel from top to bottom). Dotted lines indicate the true source depth and range. (Each panel normalized independently.)

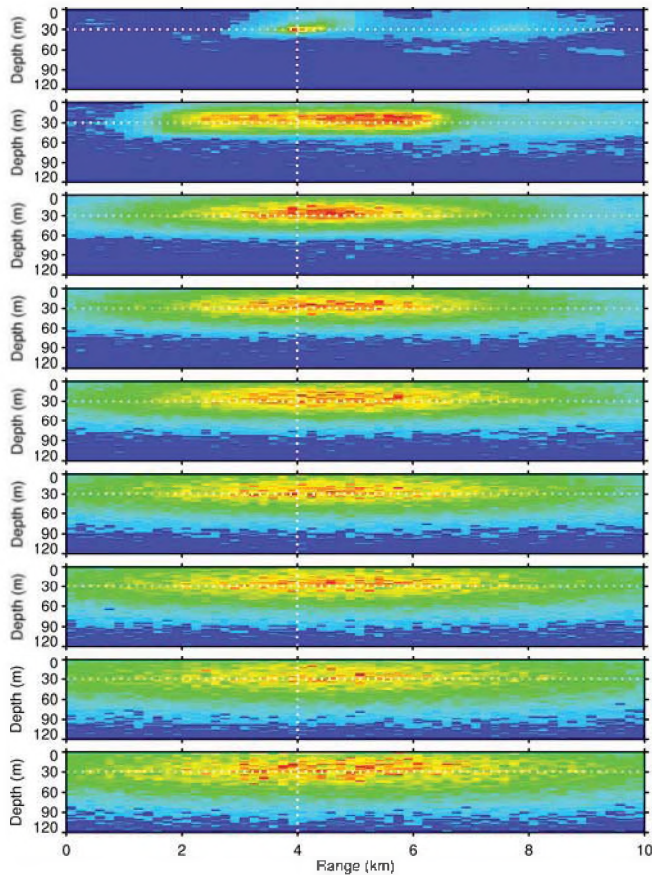


Fig. 3. Probability distributions for predicted source locations using uniform track-prediction model (time advances by 4 min/panel from top to bottom). Dotted lines indicate source depth and range at last tracked location.

errors are added to the data to achieve a signal-to-noise ratio (SNR) that varies from -14 to -8 dB with decreasing range along the track. Figure 2 shows joint marginal probability distributions for source range and depth integrated over all unknown environmental parameters via MCMC, with the additional constraint (prior information) of a maximum source velocity of 10 m/s in the radial and 0.06 m/s in the vertical. Figure 2 shows a strong probability maximum near the true source locations, although weaker secondary maxima are also evident.

The first track-prediction example is based on the pessimistic assumption that estimates of past source motion have no bearing on future motion; source motion is constrained only by the limits on radial and vertical velocity. Under these assumptions, future source positions are modelled as random variables uniformly distributed over the range-depth region allowed by the source velocity constraints. Figure 3 shows the track-prediction results for this source-motion model. The first panel in this figure repeats the last computed source localization probability distribution of Fig.2 (i.e., the take-off point for track prediction); the following panels represent predicted source

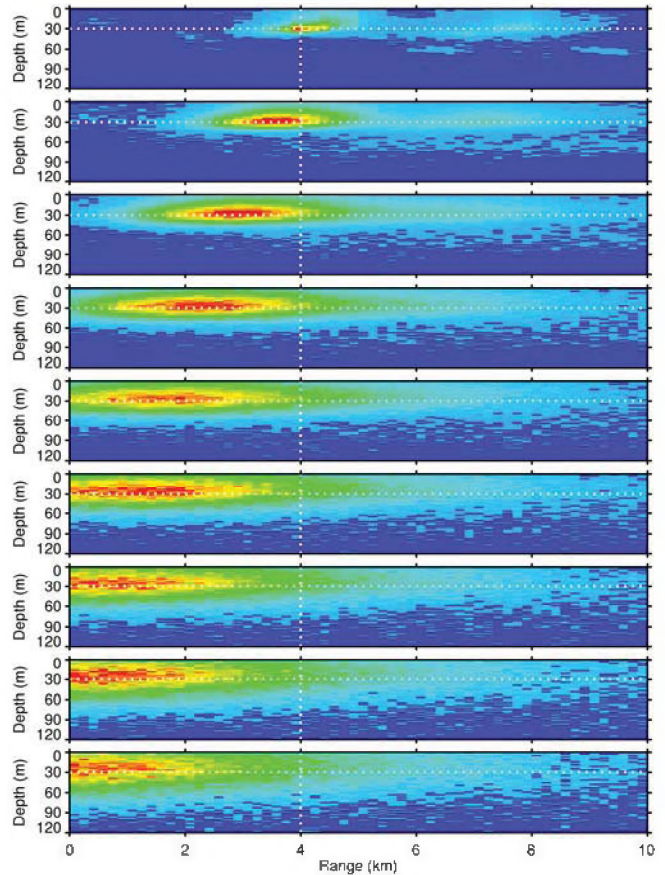


Fig. 4. Probability distributions for predicted source locations using Gaussian track-prediction model (time advances by 4 min/panel from top to bottom). Dotted lines indicate source depth and range at last tracked location.

location probability distributions at time intervals of 4 mins. The second example is based on the more optimistic assumption that estimates of past source motion provide useful information in predicting future motion. In this case the source motion is modelled as a Gaussian-distributed random variable with mean and standard deviation computed from the source locations along the particular track realization of the MCMC sample. The results of this procedure, shown in Fig. 4, generally extend the inward radial source motion detected in tracking results of Fig. 2, although the probability distributions disperse with time due to the uncertainty in the tracking results and the source-motion model

REFERENCES

- [1] Dosso, S.E. and M. J. Wilmut, 2008. Uncertainty estimation in simultaneous Bayesian tracking and environmental inversion. *J. Acoust. Soc. Am.*, **124**, 82-97.
- [2] Dosso, S.E. and M. J. Wilmut, 2009. Comparison of focalization and marginalization for Bayesian tracking in an uncertain ocean environment. *J. Acoust. Soc. Am.*, **125**, 717-722.

MRI MEASUREMENT OF DYNAMICS OF THE LIQUID AND A DISSOLVED GAS DURING ACOUSTIC CAVITATION

I.V.Mastikhin¹, B.Newling¹, N.Hetherington¹, S.Kristoffersen¹

¹Dept. of Physics, University of New Brunswick, 8 Bailey Dr., NB, Canada, E3B 5A3
mast@unb.ca

1. INTRODUCTION

The defining feature of acoustic cavitation is the presence of the gaseous bubbles produced by the oscillating drop of acoustic pressure in the liquid. It is their evolution and collapse that provide the cavitation with the chemical activity, mechanical erosion, degassing, etc. It is, therefore, natural, that the methods most commonly applied to studies of cavitation are optical and acoustical, i.e. the bubbles are looked at and listened to [1].

The complexity of the cavitation cloud as a collection of many thousands of interacting bubbles makes a thorough theoretical analysis almost impossible; besides, in such practical applications of cavitation as sonochemistry and material processing, there is often a need for a conversion of the visual and acoustical information into parameters relevant to those fields such as gas exchange between the bubbles and the liquid and fluid dynamics of the liquid and the gases during cavitation.

In this work, instead of measuring the bubble dynamics, we attempt to obtain information on dynamics of molecules that constitute the bubbles and the surrounding liquid. Such information is directly acquired via the use of the Magnetic Resonance Imaging (MRI) techniques of measuring the displacement and the relaxation parameters of water and an NMR-sensitive highly soluble gas as a function of the acoustic pressure and the water preparation procedures (filtered, non-filtered, with the addition of surfactants). The obtained data are then examined using a two-site exchange model to extract estimates of the average residence time of gas molecules inside the cavitating bubbles (averaged over the volume).

2. METHOD

An MRI scanner with 2.4 T superconducting magnet (Nalorac, CA), 20 cm i.d. bore, was employed for all our measurements. An ultrasonic transducer with the attached water-filled cuvette, satisfying 2 half-wave length standing wave conditions, was placed in an RF probe of the scanner (at resonant frequencies of 1H for measurements of water and 19F for measurements of the dissolved gas (Freon)). A Langevin-type transducer (SensorTech, ON) was employed for the experiments with the acoustic pressure of up to 1.6 Atm (19.7 kHz), and a magnetostrictive transducer with a cylindrical concentrator (Industrial Sonomechanics, Inc) was used in experiments with the acoustic pressure of up to 6 Atm (19.1 kHz).

We worked with two types of water samples: a distilled water, tap water and the water with 1 mM SDS for the bubble stabilization. The water samples were further filtered to reduce the number of impurities that could serve as the nucleation centres. The filtration was done with the syringe 0.2 μ m filter (Sarstedt, Germany).

As an NMR-sensitive gas, Freon-22 was dissolved in the water samples prior to all experiments by bubbling at the ambient temperature and pressure during 20 min, with the resulting dissolution of 0.6 vol/vol (as measured by the NMR signal comparison with the signal from the gas-filled cuvette). Its T2 NMR relaxation times differ by three orders of magnitude for the gas in the free (2 ms) and dissolved (1.5 s) state [2] and hence can be used to differentiate between the two states of the gas molecules.

A PFG SE sequence [3] was employed to obtain propagator (displacement probability) maps both for the water and the dissolved gas. The observation window in the PFG experiments was 88 ms so that only the signal from the dissolved molecules of Freon could be detected. The T2 relaxation was measured with the CPMG [4] sequence for the bulk and with the SE SPI CPMG sequence for the 1-dimensional profile of the cuvette along the acoustic axis.

3. RESULTS

The cavitation onset was detected at 0.6 Atm in the tap water, and at 0.7 Atm in the filtered water by the appearance of the high frequency noise detected by the hydrophone.

MRI Measurements were first performed at the lower acoustic pressures (1.6 Atm) for the samples with the filtered and the unfiltered water. The propagator maps showed a similar dynamics for the motion of the water molecules in the two samples, with the developed acoustic streaming along the acoustic axis.

The distribution of the velocities of the Freon molecules was, however, different in the two samples, with the greater localization of the dissolved gas in the unfiltered water and the greater number of the "stationary" molecules (molecules with the velocities within the resolution limit of the measurement). The displacement of Freon molecules in the filtered water was greater than that of the water molecules, with the almost complete absence of the signal from the "stationary molecules" [5]. The only explanation of the discrepancy in the motions of water and gas molecules dissolved in water was the participation of the gas molecules

in the cavitation process, *i.e.* the diffusion of the gas from water into the cavitating bubbles with which the gas molecules would be able to travel faster than the currents of the surrounding water.

The active participation of the gas molecules in the cavitation meant that a certain fraction of the measurement time they had to spend in the gaseous state inside the bubbles. In that case, the measurement of the T2 of Freon during cavitation should be able to provide us with information on the average duration of the gaseous state, thanks to the profound difference in the T2 between the two states. The changes in T2 during cavitation were reliable yet minor, with a monoexponential decay, indicating that the exchange between the two states was faster than the timescale of the measurement (spin echo time of 0.5 ms). An application of the two-site exchange model resulted in the estimates of very short residence times of molecules of Freon in the gaseous state, on the order of two oscillation periods (~0.1 ms).

At higher acoustic pressures achieved with the other transducer, the dynamics of the water and gas molecules changed again. The motion of the water became more pronounced whereas the motion of the dissolved gas was found to be more and more localized, both at the tip of the concentrator and its sides. At the highest pressures (5.6 Atm), both the gas and water molecules became localized at the tip of the horn, with the active motion of water long the sides.

Measurements of T2 of Freon at the tip of the concentrator showed T2 decrease along the acoustic axis, in the region of the most active cavitation. The application of two-site exchange model to the T2 data yielded the average residence times of Freon molecules of several oscillation periods (100-150 mks).

The addition of SDS to water for the bubble stabilization considerably modified the dynamics of both water and gas molecules. The T2 times of Freon at high acoustic pressures became bi-exponential, indicating two sites with well-defined, quite different molecular environments, which we attribute to the stabilized gaseous bubbles and the surrounding water.

4. DISCUSSION

The observed uncoupling of the water and the dissolved gas dynamics was to be expected, since it is essentially the definition of the cavitation. Still, it was surprising to observe the extent of the involvement of the dissolved gas into the cavitation, with the exchange rate of the total gas content of the vessel on the order of less than a second between the two states.

The observed spatial localization of the gas and water at the elevated pressures can be explained by the increased Bjerknes force action of the acoustic field.

The very short residence times of the gas molecules were unexpected. It is relevant to note that these times are a total time of the gas molecule with all possible translations of the molecule into and out of the bubble over the course of the measurement (~ 1s). With the commonly accepted model of the boundary layer, one would think that the molecules in the proximity of the bubbles will necessarily have a longer exposure to the bubble content. However, the measurements are averaged by volume, and that can mask the boundary layer effect. Also, Freon is a highly soluble gas, so it should be no surprise that as soon as its molecule gets ejected out of the bubble during the collapse phase, it stays in the dissolved state.

The two-site model is quite simplistic, and there is a potential for improving it by adding the boundary layer and the radial diffusive exchange [6]. Another direction of research is the use of other gases with lower solubility: that will be a chance to probe the gas exchange between the bubbles and liquid at higher exchange rate if the experimental problems with the lower signal-to-noise due to the decreased gas concentration will be overcome.

In conclusion, we believe that the presented work demonstrates a potential for the molecule-based approach and MRI measurements in the cavitation research.

REFERENCES

- [1] F. R. Young. Cavitation. Boca Raton: CRC Press, 1990.
- [2] B. W. Dubois, A. S. Evers. 19F-NMR Spin-Spin Relaxation (T2) Method for Characterizing Volatile Anesthetic Binding to Proteins. Analysis of Isoflurane Binding to Serum Albumin. *Biochemistry* 31, 7069-7076 (1992).
- [3] P. T. Callaghan. Principles of Nuclear Magnetic Resonance Microscopy. Oxford: Oxford University Press, 1991.
- [4] H.Y. Carr, E.M. Purcell. Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments. *Phys.Rev.*94, 630-638 (1954)
- [5] I. V. Mastikhin, B. Newling. Dynamics of dissolved gas in a cavitating fluid. *Phys. Rev. E* 78, 066316 (2008).
- [6] M. R. Gherase, J.Wallace, A.R. Cross, G.E. Santyr. Two-compartment radial diffusive exchange analysis of the NMR lineshape of ¹²⁹Xe dissolved in a perfluorooctyl bromide emulsion. *J. Chem. Phys.* 125, 044906 (2006).

ACKNOWLEDGEMENTS

The authors would like to thank the National Science and Engineering Council of Canada for the Discovery Grant (I.M.) and the USRA Award (N.H.).

EIGENFREQUENCY ANALYSIS OF FLUID-FILLED PIPES

N M Alam Chowdhury^{1*}, Zaiyi Liao², Lian Zhao¹ and Ramani Ramakrishnan²

¹Dept. of Electrical & Computer Engineering, Ryerson University

²Dept of Architectural Science, Ryerson University,

*350 Victoria St., Toronto, ON, Canada, M5B 2K3, nmalamc@yahoo.com

1. INTRODUCTION

The eigenfrequency analysis of piping system is used in thermodynamic or chemical engineering process for many years, which are generally subjected to static or time dependent variable loads. In time dependent case, these loads generate acoustic sources which may lead to excessive noise level in the piping system. Many researches [1-4] were performed on this matter for its technological importance and the performance degradation of the piping system. In the prestressed concrete cylinder pipe (PCCP), the acoustic signals are generated through the breakage or sliding of reinforced wires. These signals feature certain frequency spectral characteristics that are close to the eigenfrequency of the piping system. This signal degrades the performance of the condition monitoring of piping system. As far as the authors' knowledge, there is no systematic theoretical analysis on this problem in the open literature. In this work, the eigenfrequency analysis of these events is taken as a principal interest.

The fundamental element of such an analysis is to investigate the eigenfrequency of the wave propagation impacted by the complexity of fluid medium surrounded by a large concrete pipe buried in the ground. The mathematical model is developed using Navier's equation of motion for acoustic wave in frequency domain. The interaction between the fluid and the surrounding layers are modeled based on Newton's law of motion and principle of virtual work. The results of the eigenfrequency analysis are presented and compared with available analytical solutions.

2. MODEL FORMULATION

The wire break or slip generated acoustic signal propagates through the pipe structure and the fluid in the pipe. Let us assume that, the time harmonic acoustic pressure (p) propagates uniformly through a lossless fluid. Therefore, using pressure-velocity relation in Navier's scalar velocity potential equation, the acoustic pressure equation in frequency domain, can be written as [5]

$$\nabla^2 \left(-\frac{1}{\rho_F} \right) p - \frac{\omega^2 p}{\rho_F v_F^2} = 0, \quad (1)$$

where ρ_F is the fluid density, v_F is the speed of the signal in the fluid medium and ω is the angular frequency.

According to Newton's law of motion in equilibrium, this external acoustic pressure forces generate the same internal body forces in the pipe structure. These causes undergo unrelated but consistent displacements and deformations. This displacement in the structure generates the same but opposite amount of traction forces on the pipe

wall. The interactions between different layers are modeled using appropriate boundary conditions [5].

The eigenfrequency of the model can be obtain by solving the homogeneous pressure equation given in Eq.(1). The solution becomes zero except at a discrete set of eigenfrequencies with a well-defined shape of undefined magnitude.

3. NUMERICAL STUDY

Let us consider a uniform and smooth fluid-filled PCCP surrounded by the outer formation (e.g. soil) as shown in Fig.1. For simplicity, assume the pipe structure consist of high-strength concrete only and is filled with static fluid (water). Damping is absent in all medium. The basic properties of the medium are given in Table-1.

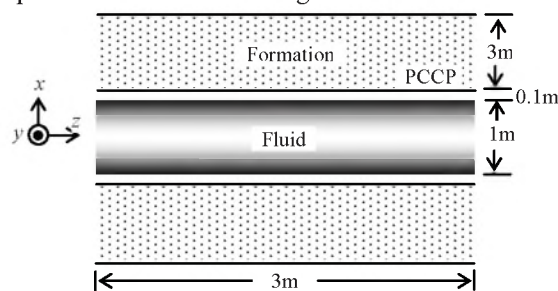


Figure 1. Typical model geometry with dimensions.

Table 1. Basic properties of the medium.

Properties	Water	Concret
Density (ρ), kg/m ³	997	2400
Acoustic wave Speed (v), m/s	1500	--
Elastic modulus (E), Pa	--	40 ⁹
Poisson's ratio (γ)	--	0.33

4. RESULTS AND DISCUSSION

The model is simulated for the rigid and elastic pipe with infinite and finite stiffness, respectively and for the various dimensions of pipe and medium characteristics. For the rigid pipe the simulated results are compared with the analytical solutions [5]. Similar analysis is also performed for elastic pipes, where analytical solutions are unknown. However, the simulated results are analyzed and validated as per theoretical and physical concept.

In this paper, the simulated results of water-filled concrete pipe (elastic pipe with finite stiffness) with different radius (R) are presented (Fig.2), as an example. The acoustic signal propagating in the water has an interaction with the pipe structure, and vice versa. As a result, eigenfrequency, cut-off frequency and phase speed

have variable quantities based on the pipe and medium characteristics.

Figure 2 shows the eigenfrequency analysis of elastic pipe with finite stiffness at different radii. From the figure it is seen that, up to the first cut-off point, the eigenfrequencies are separated at regular intervals. After that, it changes very rapidly. Moreover, increasing the radius reduces the cut-off frequency of pseudo-rayleigh modes. As far as authors' knowledge, there is no analytical solution in the open literature to compare these values. However, it has a similar trend of rigid pipe solution, which can be compared with the analytical solution [5]. This means that, in acoustic emission (AE) monitoring system, it is possible to collect more AE events for the larger diameter pipe by placing the sensors at far distance. On the other hand, this increased radius generates additional vibrating signal which produces complicated dispersion characteristics at the receiving end (sensors). Therefore, care should be taken to indicate the active distress in the pipeline.

At low frequency stoneley mode, increasing the inner radius reduces the pipe stiffness. Consequently, both pipe structure and fluid dissipate its energy due to very low viscosities and material damping, which reduces the phase speed of the propagating acoustic signal. The phase speed of the acoustic wave inside the pipe in the fluid medium can be obtained from the velocity of the first eigen wave. The simulated result is compared with the analytical tube wave [6] velocity (v_T), which is given below in the caption of respective graph, and found a good agreement. In AE monitoring system, this phase speed is important to localize the events.

Similarly, we can observe the effect of pipe thickness, elasticity and outer formation on the acoustic signal propagation using eigenfrequency analysis [5].

5. CONCLUSIONS

The impacts of the pipe radius on the stoneley and rayleigh modes, and the phase speeds of the system are studied. The results show that the radius has significant impacts on the rayleigh modes but not on the stoneley modes of propagation. Extra harmonic eigen waves are observed in the rayleigh modes. This is in controversy with the analytical solutions. These eigen wave solutions are the

fundamentals for a number of more complex wave propagation studies, such as system vibration, pulse propagation and perturbation analysis.

Overall, the speed of the acoustic signals propagating through pipes is not constant and is affected by the pipe profiles and surrounding formations. The acoustic signals recorded by the sensors used in current industry may have been noised by the system resonance and thus the testing data should be analyzed more rigorously. The study presented in this paper produces the fundamentals on which better non-destructive pipe performance testing technologies can be designed.

REFERENCES

- [1] Melo, F.J.Q., Noronha, J.P. and Fernandes, E.A. (1996). The propagation of axisymmetric transverse waves along a thin-walled cylindrical pipe, *International Journal of Pressure Piping*, vol. 65, 109-116.
- [2] Lee, S.I., Chung, J. (2002). New non-linear modeling for vibration analysis of a straight pipe conveying fluid, *Journal of Sound and Vibration*, vol. 254(2), 313-325.
- [3] Pavic, G. (2003). Acoustical analysis of pipes with flow using invariant field functions, *Journal of Sound and Vibration*, vol. 263, 153-174.
- [4] Maess, M., Wagner, N. and Gaul, L. (2006). Dispersion curves of fluid-filled elastic pipes by standard FE models and eigenpath analysis, *Journal of Sound and Vibration*, vol. 296, 264-276.
- [5] Chowdhury, N.M. (2009). Theoretical analysis of acoustic emission signal propagation in fluid-filled pipes, Master thesis (submitted), ELECE Dept., Ryerson University, Toronto, ON.
- [6] Morse, P.M. and Feshbach, H. (1953). *Methods of Theoretical Physics-II*, McGraw-Hill, NY.

ACKNOWLEDGEMENTS

This work has been partially supported by the Ontario Centres of Excellence (OCE) under Grant No. EE50196 and Natural Sciences and Engineering Research Council (NSERC) of Canada under Grant No. DG 293237-04 and 313375-07.

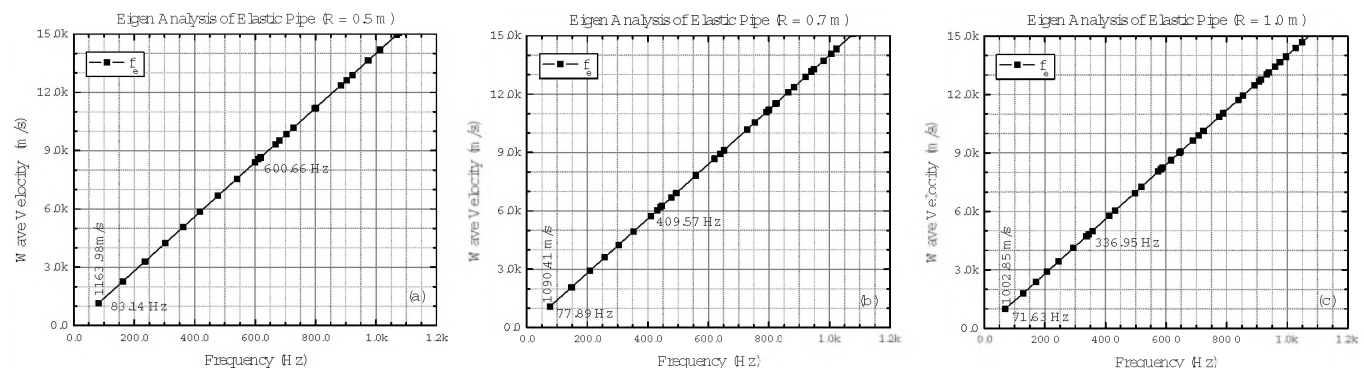


Figure 2. Eigenfrequency of water-filled elastic pipe - (a) $R = 0.5$ m ($v_T = 1164.30$ m/s), (b) $R = 0.7$ m ($v_T = 1094.06$ m/s), and (c) $R = 1.0$ m ($v_T = 1007.35$ m/s).

MODAL SOLUTIONS OF THE ACOUSTIC WAVE PROPAGATION THROUGH FLUID-FILLED PIPES

N M Alam Chowdhury^{1*}, Zaiyi Liao², Lian Zhao¹ and Ramani Ramakrishnan²

¹Dept. of Electrical & Computer Engineering, Ryerson University

²Dept of Architectural Science, Ryerson University,

*350 Victoria St., Toronto, ON, Canada, M5B 2K3, nmalamc@yahoo.com

1. INTRODUCTION

In the pipeline, the guided waves propagate with very complicated wave structures due to a mixture of multi-modes and mode conversion, which prevents guided waves from being widely used. Modal analysis and simulation of guided wave propagation are very useful for solving these problems. In the piping system, the scope of this analysis includes the evaluation of dispersion behavior of guided waves in hollow or liquid-filled cylinders. In long-range inspection of pipes, this guided waves largely reduces inspection time and costs compared to the ordinary point by point testing in large pipeline [1-3].

In the prestressed concrete cylinder pipe (PCCP), the acoustic waves are generated through the breakage or sliding of reinforced wires. These waves have different frequency spectra, which consist of low to high frequency signals, and propagate through the fluid column inside the pipe as a guided wave. At low frequencies, the wave propagates as an evanescent mode with plane wavefront. These waves cannot propagate for long distances, where both pipe structure and fluid dissipates its energy due to very low viscosities and material damping [4]. At high frequencies, it has also propagating mode which exhibits oscillatory amplitude in the fluid and a decaying amplitude in the pipe structure, and propagates a far distance. The number of propagating modes grows proportional to the signal frequency. Consequently, more modes cause complicated dispersion characteristics. In elastic pipe, such waves have strong dispersive phenomena [142]; as a result, there exist several wave modes in the same propagating signal. Therefore, the modal analysis is important to determine the dynamic characteristics of the system.

The mathematical model is developed based on Navier's equation of motion. The model is solved to obtain the modes of propagation and phase wave of acoustic wave in for various radii of the pipe. The finite-element analysis is used to simulate the model and the results are compared with the analytical solutions.

2. MATHEMATICAL MODEL

Let us assume that the wire break or slip generated acoustic emission (AE) wave propagates through a lossless fluid in the pipe. Therefore, using pressure-velocity relation in Navier's scalar velocity potential equation, the inhomogeneous acoustic pressure equation in frequency domain, can be written as [5]

$$\nabla^2 \left(-\frac{1}{\rho_F} \right) p - \frac{\omega^2 p}{\rho_F v_F^2} = g, \quad (1)$$

where ρ_F is the fluid density, v_F is the speed of the signal in the fluid medium, ω is the angular frequency and g is the AE source.

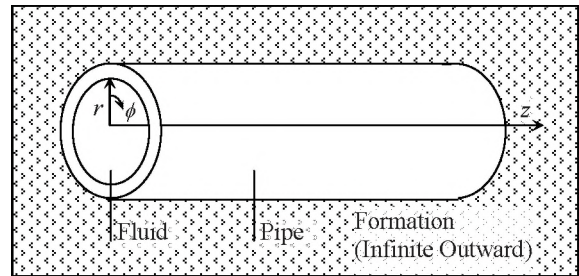


Figure 1. Schematic of fluid-filled PCCP.

2.1 Modal Analysis

Let us assume that the acoustic pressure varies harmonically in longitudinal direction (say z) as shown in Fig.1. Therefore, the modal description can be obtained from Eq.(1) by replacing the pressure with $p(x, y)e^{-ik_z z}$, as

$$\nabla^2 \left(-\frac{1}{\rho_F} \right) p - \left(\frac{\omega^2}{v_F^2} - k_z^2 \right) \frac{p}{\rho_F} = g, \quad (2)$$

where k_z represents the longitudinal wave number in z direction.

By solving Eq.(2) at a given frequency with a nonzero excitation, one can obtain the most axial wave numbers. These values are the propagation constants of evanescent (stoneley) or propagating (rayleigh) waveguide modes. The frequency related to each mode is the modal or cut-off frequency of the guided path. It is the critical frequency between propagation and attenuation, which can be obtained as [5]

$$f_c = \frac{1}{2\pi} \left[\omega^2 - k_z^2 v_F^2 \right]^{1/2}. \quad (3)$$

Theoretically, the modal frequency of fluid-filled cylindrical pipe with infinite stiffness surface can be calculated as [5]

$$f_c = \frac{v_F z_{mn}}{2\pi r} \text{ with } J'_m(z_{mn}) = 0, \quad (4)$$

where r is the radius of the pipe, z_{mn} is the n -th root of the first derivative of Bessel function J'_m .

3. NUMERICAL IMPLEMENTATION

Let us consider a uniform and smooth fluid-filled PCCP surrounded by the outer formation as shown in Fig.1. For simplicity, assume the pipe is filled with static fluid (water) and damping is absent in all medium.

The model is simulated for the rigid and the elastic pipe with different radii. In this paper, the simulated results of rigid pipe are presented and the results are compared with the analytical solutions using Eq.(4). Similar analysis is also performed for elastic pipes, where analytical solutions are unknown. However, the simulated results are analyzed and validate based on the theoretical concept.

The following values of the parameters have been used for the numerical study: density of water, 997 kg m^{-3} ; acoustic wave speed, 1500 m s^{-1} ; inner radius varies as, $R = 0.5\text{m}$, 0.7m and 1.0 m , respectively. To avoid the computational burden, the solutions up to 3.5 kHz were evaluated.

4. RESULTS AND DISCUSSION

The numerical results of cut-off frequency obtained from simulated and analytical solutions are shown in Table-1, for various radii ($R = 0.5\text{m}$, 0.7m and 1.0 m) rigid pipe. From the results it is seen that, increasing the inner radius reduces the cut-off frequency of rayleigh modes, which also satisfies the analytical solution. In this work, the analysis is computed up to 6-th mode of propagation and found only 0.2 to 0.5% variations in the simulated results when compared to analytical solutions. And hence, the numerical solutions provide acceptable results to a reasonable level of accuracy.

The results easily identify the various propagating modes, their excitation frequency as well as the impact of the pipe radius. It is seen that, the number of propagating modes decreases with decreasing pipe radius, and vice versa. This signifies that, in the lower diameter pipe it is difficult to detect the low frequency acoustic waves by placing the sensors at far distance.

5. CONCLUSIONS

The impacts of the rigid pipe radius on the stoneley and the rayleigh modes of the system are described. The modes of propagation and cut-off frequency of acoustic wave in rigid pipes with a range of radii are observed.

The modal analysis is also useful to determine the modes of propagation for elastic pipe with complicated structure and medium, where direct analytical solution might not be known. This information is important for the placement of sensors in AE monitoring system to locate the corroded areas.

REFERENCES

- [1] Mudge, P.J. (2001). Field application of the teletest (R) long-range ultrasonic testing technique, *J. Insight*, vol. 43, pp. 74-77.
- [2] Alleyne, D.N., Pavlakovic, B., Lowe, M.J.S. and Cawley, P. (2001). Rapid long-range inspection of chemical plant pipe work using guided waves, *J. Insight*, vol. 43, pp. 93-96.
- [3] Kwun, H., Kim, S.Y., and Light, G.M. (2003). The magnetostrictive sensor technology for long range guided wave testing and monitoring of structures, *J. Mater. Eval.*, vol. 61, no. 1, pp. 80-84.
- [4] Lamb, H. (1898). On the velocity of sound, as affected by the elasticity of the wall, *Manchester Memoirs*, vol. 17, no. 9, pp. 1-16.
- [5] Chowdhury, N.M. (2009). Theoretical analysis of acoustic emission signal propagation in fluid-filled pipes, Master thesis (submitted), ELECE Dept., Ryerson University, Toronto, ON.

ACKNOWLEDGEMENTS

This work has been partially supported by the Ontario Centres of Excellence (OCE) under Grant No. EE50196 and Natural Sciences and Engineering Research Council (NSERC) of Canada under Grant No. DG 293237-04 and 313375-07.

Table 1. Analytical and simulated results of cut-off frequency of the rigid pipe.

Mode of Propagation (m, n)	Analytical, f_c (Hz)	Simulated, f_c (Hz)	Analytical, f_c (Hz)	Simulated, f_c (Hz)	Analytical, f_c (Hz)	Simulated, f_c (Hz)
	$R = 0.5 \text{ m}$		$R = 0.7 \text{ m}$		$R = 1.0 \text{ m}$	
0,0	0.00	0.00	0.00	0.00	0.00	0.00
1,1	879.01	879.10	627.87	627.92	439.51	439.55
2,1	1458.18	1458.34	1041.56	1041.64	729.09	729.14
0,1	1829.65	1829.64	1306.55	1306.81	914.58	914.75
3,1	2005.83	2006.15	1432.74	1432.84	1002.91	1002.96
4,1	2538.68	2539.71	1813.34	1813.68	1269.34	1269.49
1,2	2545.37	2546.34	1818.12	1818.41	1272.68	1272.81
5,1	3062.94	3065.05	2187.81	2188.40	1531.47	1531.68
2,2	3201.88	3204.09	2287.06	2287.56	1600.94	1601.05
0,2	3349.89	3352.58	2392.44	2393.25	1674.71	1674.94
6,1	3581.46	3585.40	2558.19	2559.10	1790.73	1790.94

CHARACTERISTICS OF WIND TURBINE NOISE

Ramani Ramakrishnan

Department of Architectural Science, Ryerson University, Toronto, ON, Canada- rramakri@ryerson.ca

1. INTRODUCTION

Wind power is one of the most promising renewable energy sources available, due to technical advances in the field and its relatively low cost. Wind turbines can produce a large amount of electricity, but the public surrounding these farms are responding negatively to the installation of these turbines in a rural setting. One of the negative aspects expressed vocally by anti-wind farm groups is the special character of the wind turbine noise. Strong complaints have been lodged concerning the health impacts of the low-frequency components of wind turbine noise as well as the amplitude modulation of the rotating noise. A series of measurements were conducted to address the above issues. Two different types of wind turbines were monitored under different scenarios. A low frequency microphone in the operating range of 1 Hz to 10000 Hz was used for the measurement programme. The measurement results of the case study are presented in this paper.

2. BACKGROUND

One of the main noise concerns raised against wind turbine facilities is the presence of inaudible infrasound and their impact on receptors such as dizziness, heart conditions and other such negative health effects. The first generation of wind turbines from the early 80s had strong low frequency components with an old type gear-box located on the ground. The low frequency components are usually expressed in terms of a dBG scale as contained in ISO Standard 7196¹. The modern wind turbines have very little low frequency components as reported by van den Berg², Howe³, and Levanthal⁴. Levanthal⁴ even concluded that most of the loud and belligerent response from anti-wind farm group literature is deceptive.

3. THE EXPERIMENT

The main aim of the experiment was to evaluate the presence of low-frequency components below 20 Hz as well as the amount of amplitude modulation of the wind turbine noise. Very preliminary results were presented in Inter-Noise 2009 conference held in Ottawa, Ontario⁵.

The following instrumentation was used for the experiments. A Soundbook analyzer, made by SINUS Messtechnik of Germany, was used with GRAS 40AN low-frequency microphone. The whole system is capable of measuring from 0.5 Hz to 10,000 Hz frequency range.

Two areas were chosen for the measurements. The first location, Area A, was serviced by a single wind turbine. The wind turbine characteristics were: 600 kW 3-bladed turbine with 52 m diameter blades. The hub height is 65 m

and the nominal rotational speed is 24.5 rpm. The operating wind speed range is 2.5 m/sec to 25 m/sec. Two locations were chosen, a) right under the wind turbine and b) about 30 m from the wind turbine and in direct line of sight of the turbine. The measurements were conducted between 1 am and 2 am so that the influence of road traffic noise from adjacent roadways would be minimal. The wind gusts at 2.5 m height were minimal.

The second group, Area B, was serviced by more than 50 wind turbines. The wind turbine characteristics were: 1.5 mW 3-bladed turbine with 40 m diameter blades. The hub height is 80 m. Five locations were chosen: i) about 10 m from one wind turbine with minimal influence from other wind turbines; ii) impact from at least 5 wind turbines and the rest of the wind turbines were more than 250 m away; iii) impact from one wind turbine and secondary impact from at least 5 wind turbines and the rest of the wind turbines were more than 250 m away; iv) near a residence with the turbines at least 100 m away; and v) near another residence with one turbine within 50 m and others further away. The measurements were conducted between 4 pm and 5 pm and the wind gusts at 2.5m height were severe at times. The readings were taken with minimal road traffic as much as possible

4. RESULTS AND DISCUSSION

The measurements were recorded for 15 seconds for Area A and 25 seconds for Area B. The results of the measurements are presented in this section.

The results from one 15 second reading at Location (a) and one 15 second reading at Location (b) are presented below.

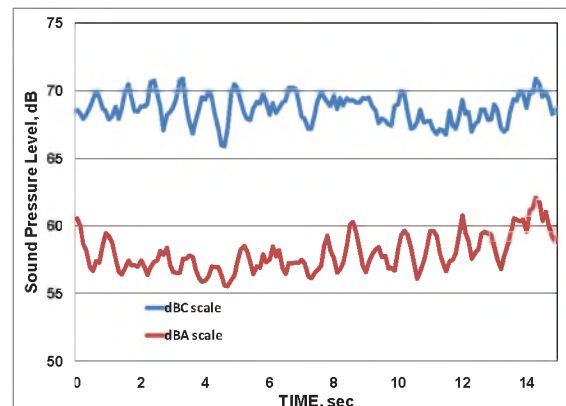


Figure 1. Time History for Location (a), Area A

The results of Figure 1 clearly show that the amplitude modulation can be as high 5 dB right below the 600 kW turbine.

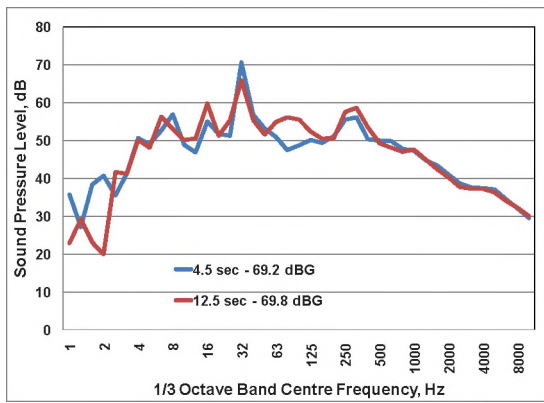


Figure 2. Spectrum for Location (a), Area A

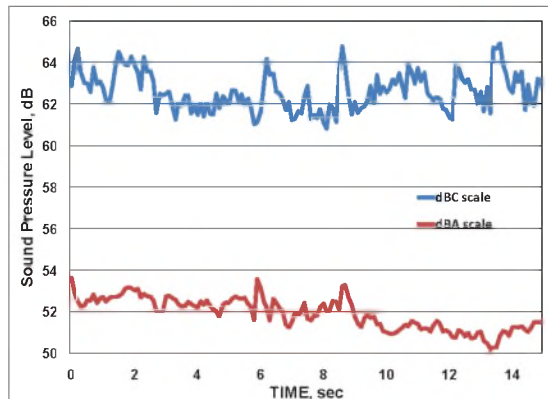


Figure 3. Time History for Location (b), Area A

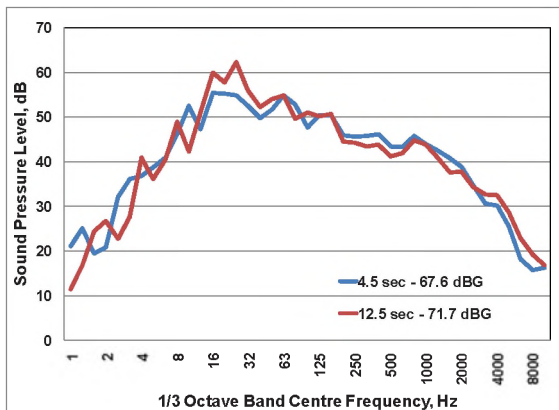


Figure 4. Spectrum for Location (b), Area A

Somewhat similar results are seen for Location b of Area A. The amplitude modulation gets submerged in the background sound and one can discern a 2 dB or less amplitude modulation. The one-third octave band spectra for Location A are presented in Figures 2 and 4. The overall dBG is less than 75 dBG and no frequency components of any major dominance are seen below 5 Hz. The blade-passage frequency is around 1.5 Hz.

Results for Area B are shown in Figures 5 and 6. Similar results to that of Area A are seen even though the turbine produces more power. The ambient wind seems to dominate over the turbine noise. Amplitude modulation of

about 5 dB can be discerned from the dBA plot. The spectral results show that there are no dominant frequency components below 5 Hz and the overall dBG is less than 70 dBG and less than the 75 dBG threshold.

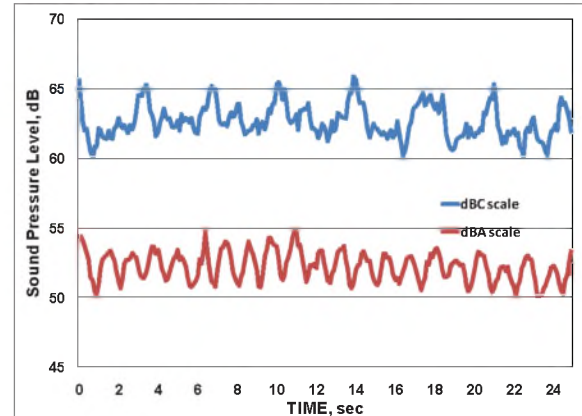


Figure 5. Time History for Location (i), Area B

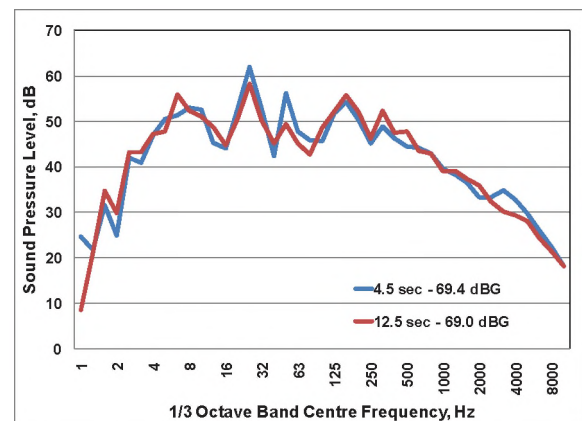


Figure 6. Spectrum for Location (i), Area B

5. CONCLUSIONS

Preliminary results from two different wind turbines show that there is amplitude modulation up to 5 dB, but the infrasound is non-existent.

REFERENCES

1. International Standard, ISO 7196, "Acoustics – Frequency-weighting characteristic for infrasound measurements." International Standard Organization, (1995).
2. G.P. van den Berg. "The Sounds of High Winds: the effect of atmospheric stability on wind turbine sound and microphone noise." Doctoral dissertation, University of Groningen, Netherlands, May 2006.
3. Brian Howe, "Wind Turbine and Infra Sound," Report prepared for CANWEA, Ottawa. 14 Pages (2006).
4. Geoff Leventhall, "Infrasound from Wind Turbines – Fact, Fiction or Deception," *Canadian Acoustics*, **34**(2), 29-36 (2006).
5. Ramani Ramakrishnan, "Wind farms and noise." Proceedings of Inter-Noise 2009, August 2009, Ottawa, Ontario. INCE-USA, Poughkeepsie, NY, USA.

AN APPROACH TO GENERATING AND CALIBRATING A SOUND CONTOUR MAP

Ryan Daneluzzi¹, Steve Bogensberger, Mike Dobrin, Colin Novak², Helen Ule

¹Dept. of Mechanical Engineering, University of Windsor, 401 Sunset, Ontario, Canada, N9B 3P4 daneluz@uwindsor.ca

²Dept. of Mechanical Engineering, University of Windsor, 401 Sunset, Ontario, Canada, N9B 3P4 novak1@uwindsor.ca

1. INTRODUCTION

Urbanized areas in modern cities are typically exposed to a high level of noise pollution from many sources. A prevalent source of noise pollution typically comes from roadways due to high transportation volumes. The City of Windsor has this problem as it has the busiest international border crossing between Canada and the United States. For this study, a noise contour map of a 5 km square area was done as an undergraduate engineering project using Bruel & Kjaer's Lima software.

The process for accurate noise model results can be quite time consuming, however the accurate results are important when used to evaluate problematic areas. A noise contour map can be created based only on traffic count data of vehicle type and volume; however those results may not truly reflect actual noise levels. To calibrate a model, it is beneficial to perform a reverse engineering analysis on the model by using actual noise measurement data acquired along the study area.

2. PHYSICAL MAP GENERATION

The creation of the road emission model in LimA was a tedious task. Geographical data was acquired from several sources; the roads were created using ArcGIS data and the background bitmap image was made by stitching together high-resolution snapshots taken from Google Maps. Generating buildings and structures required the tracing polygons overtop of the background bitmap images and assigning height information. Along with the building heights, elevation contours were implemented into the map in order to properly represent the elevation changes within the study area. These included man made berms.

The road volume data which was inputted into the model was taken from annual average daily road traffic (AADT) data provided by the City of Windsor. In addition to the AADT data, the percentage of heavy vehicles, percentage of daytime/nighttime traffic split, speed limit and road surface type were considered. Data concerning the day night split was not available so an assumption for the two different major types of roads for the area was made. The major roads (EC Row and Huron Church Road) would be assumed to have a heavy traffic concentration of 20% during the day and 80% during the night and the daytime/nighttime traffic split was even by overall volume. This was assumed since

these roads are major transportation routes and are dominated with international heavy truck traffic. Smaller roads in the model were assumed to have very little commercial traffic with only 5% of all traffic expected to be heavy trucks. As most of the traffic on the smaller roads typically occur during the day, it was assumed that 85% of the total traffic volume occurred during the day period.

3. INPUTTING MEASUREMENT DATA

To calibrate the model, sound level measurements were to be taken at chosen locations. Road segments were grouped into emitter groups by observing how the traffic flow occurred with major intersections being the start and end points for the groups. Once the roads were sorted into different emitter groups, it was possible to make general estimations of where different measurement points would need to be taken so that each emitter group would have an independent associated measurement point. In order for the measurement to be useful, it had to be taken within an ideal distance of the emitter group of interest. A good practice is to take one measurement per emitter group in locations where only the emitter group of interest will be the main contributor to the reading. Once the ideal locations were identified, 20 minute L_{eq} readings over a 24 hour periods were made. Using the measured noise data in conjunction with the road traffic data allowed for the program to calibrate the acoustic contour map.

4. CALIBRATION OF MODEL

Once all of the important characteristics of the sound map were entered into LimA the synthesis of calibrated noise contours was possible. The estimated propagation of noise using the British CRTN model produce reasonable results for the study area as it predicted the largest noise levels around the major roads Huron Church Road and EC Row. These two, as predicted, had the highest emissions and the greatest impact in the surrounding areas. Huron Church road had the biggest impact over the area of the study due to it being the center dividing line of the study area as well as the busiest road. Its noise emissions were the only major contributors to many areas of the map as it produced noise levels over 45dB far into residential areas.

The reverse engineering process was carried out using a function of the LimA software. LimA has the option to back calculate the emissions of the measured noise sources.

Ideally, adjusting the road emissions with actual sound measurements will produce a more accurate noise contour since it will correct errors that may have been present in the original model based on simplifying assumptions. The resulting noise contours with and without the reverse engineering is given in Figure 1.



Figure 1. Noise Contour Map of Reversed Engineered Model and Original Model

By inspection of the revised results, it is apparent that after the reverse engineering process, there were minor changes to the predicted road emissions. This also demonstrated is good to see as it showed the original model was a very close approximation of a real life situation.

5. DISCUSSION OF RESULTS

The predicted noise levels in residential areas increased after the calibration process due to the increased levels from the main emitting roads. This change is likely due to the inaccuracy of the provided traffic data. The reverse engineering procedure thus helped to account for discrepancies in the old data as the actual sound pressure level measurements were more current. However, most of the contours had little differences showing that the CRTN model had a fairly accurate calculation of noise propagation.

LimA was able to perform an analysis of the areas which were affected by different noise level ranges. The data showed that after making adjustments from the original road calculation, the noise emissions overall increased. CRTN calculated the L_{10} values which is the maximum sound pressure level which is experienced 90% of the time. This calculated value is different from the typically L_{eq} which is required measurement type by the Ministry of the Environment in Ontario, so direct comparison of the results and Ontario regulation could not be made. The areas affected by higher sound pressure levels over 60 dBA increased slightly. They covered about 40% of the overall area with only 35% of the area actually being in the MOE

recommended range of under 55 dBA and 22% of the area alone being within the range of 55-60 dBA it can be said that there is a noise problem in the study area. Fortunately for the regions with levels below 65 dBA mitigation using noise barriers is possible. However this statistical examination fails to examine how the zoning correlates with the noise data. The industrial areas at the southern part of the study area appears to have a higher noise level present which isn't a major problem; as few people live around this area. The majority of residential areas other than houses directly adjacent to Huron Church road appear to have acceptable noise levels. The area near the Ambassador Bridge also has high emission contours that affect the nearby residential areas as well as the University of Windsor campus which borders it. These emissions would be more difficult to mitigate as the bridge is too high for effect use of noise barriers to block the sound emanating from the bridge.

6. CONCLUSION

A noise contour map can be a very powerful tool if created properly. It provides a graphical examination of problematic areas within a study region and can be used for urban planning and design for noise mitigation in attempt to reduce noise emissions at sensitive residential receptors. The synthesis of the map presented was constructed using the most recent and corrected data available. Upon examination of the reverse engineering results and statistical analysis, it was demonstrated that although there was a change of emission coverage within the study area, the change was small. However, the reverse engineering of the source emissions was an important step in the process as it used actual measured noise data at a locations close to the source to either verify the model based on traffic volume data only or to identify areas where additional analysis was required. Subsequently, through the reverse engineering process, the road noise emission results were successfully back calculated to provide corrected noise emission levels which better correlated to the actual measured noise levels. This was important as it ensured that the noise contour map was correctly adjusted to match reality and is not simply just a theoretical model.

7. ACKNOWLEDGEMENTS

The authors would like to extend their gratitude to Bruel & Kjaer for their in-kind contribution of the Lima software used for this study. Without their support and help, this senior undergraduate project would not have been possible. Thanks is also given to the financial support provided the University of Windsor's Faculty of Engineering as well as the numerous industrial financial sponsors throughout the City of Windsor.

MUSIC AND HEARING AIDS

Marshall Chasin

Musicians' Clinics of Canada

Music as an input to a hearing aid poses some interesting problems both for the hearing aid design engineer and for the hearing health care professional. This is as true for the fitting of hearing aids for musicians, as well as for those non-musicians who like to listen to music. In some cases, the question is “which hearing aid manufacturer would be willing to make subtle changes for individual customers”, rather than “what is the best set of electro-acoustic parameters.” In other cases it comes down to an issue of the number of bits in the front end A/D converter. In order to understand the changes necessary for music as an input to a hearing aid or a cochlear implant, several primary physical differences between speech and music need to be understood. Two factors that have a direct ramification for the setting and/or selection of hearing aids for music are: (i) differing overall intensities, and (ii) crest factors.

DIFFERING OVERALL INTENSITIES

At one meter, speech averages 65 dB SPL (RMS) and has peaks and valleys of about 12 to 15 dB in magnitude. Because speech derives from the human vocal tract, and similar human lungs imparting similar subglottal pressures to drive the vocal chords, the potential intensity range is well defined and also quite restricted – approximately 30 to 35 decibels. In contrast, depending on the music played or listened to, various instruments can generate very soft sounds (20 to 30 dB SPL [e.g. brushes on a jazz drum]) to amplified guitar and even the brass of Wagner's *Ring Cycle* (in excess of 120 dB SPL). The dynamic range of music as an input to a hearing aid is therefore on the order of 100 dB (versus only 30 to 35 dB for speech).

Hearing aids that can handle intense inputs would be better for music than those that cannot. This can be accomplished in several ways- all of which are “hardware” related changes and not “software”. The peak input limiting level is not a hearing aid parameter that is typically reported yet it is the most important element in the transduction of intense inputs, such as music, through hearing aids. Once a hearing aid has been overdriven in the “front end” (e.g. A/D) with a high level input, no amount of software manipulation that occurs later on will restore high fidelity. This can be altered by using less sensitive (and non-broadband) microphones, as well as chips that utilize more than 16 bit architecture.

CREST FACTORS

The crest factor is the difference in decibels between the peak of a waveform and its average or root mean square (RMS). For speech the RMS is about 65 dB with peaks extending about 12 dB higher. The crest factor for speech is therefore on the order of about 12 dB. This is well known in

the hearing aid industry. Compression circuitry and hearing aid test systems use this information. In contrast, a trumpet has no soft walls or lips. The same can be said of most musical instruments and as such the peaks are less damped and “peakier” relative to the average, than is speech. Crest factors of 18 to 20 dB are not uncommon for many musical instruments. Compression systems and detectors that are based on peak sound pressure levels may have different operating characteristics for music as input to a hearing aid as for speech. That is, music may cause some compression systems to enter its non-linear phase at a lower intensity than what would be appropriate for that individual.

Another aspect of the crest factor for instrumental music is that it has ramifications for specifying both OSPL90 and for gain. Loudness discomfort levels for speech are typically used for setting the OSPL90 for a hearing aid yet music and speech have differing crest factors. The peaks of instrumental music are 6 dB higher than for speech (given the same RMS value) since the crest factor for instrumental music is 6 dB greater than for speech (12 dB versus 18 dB). In order to prevent a tolerance problem for instrumental music, the OSPL90 for a “music program” therefore needs to have a 6 dB less intense OSPL90 than for speech. And, given similar compression characteristics between music and speech, this implies that the gain for a “music program” should be 6 dB less than the gain for a broadband speech channel as well.

CONCLUSIONS: THE “MUSIC PROGRAM”

A “music program” or a set of optimal electro-acoustic parameters for enjoying music would include:

1. A sufficiently high peak input limiting level so more intense components of music are not distorted at the front end of the hearing aid.
2. If playing in an intense environment, even for those with significant sensori-neural hearing loss, a non-occluding BTE *and* a high-frequency emphasis microphone should be used.
3. Either a single channel or a multi-channel system in which all channels are set for similar compression ratios and kneepoints.
4. A compression system (similar to the speech-based compression system) with an RMS detector compression scheme with a kneepoint 5 to 8 dB higher if the hearing aid uses a peak compression detector.
5. OSPL90 and gain, for the instrumental music that is 6 dB less intense than for broadband speech.

6. A disabled feedback reduction system, or a feedback reduction system that uses gain reduction or a more sophisticated form of phase feedback cancellation (either one with short and long attack times or one that

only operates on a restricted range of frequencies such as over 2000 Hz).

7. A disabled noise reduction circuit, although because of a long attack time and a short release time, this circuitry may rarely be activated for many forms of music.

EDITORIAL BOARD / COMITÉ EDITORIAL

ARCHITECTURAL ACOUSTICS: ACOUSTIQUE ARCHITECTURALE:	Vacant		
ENGINEERING ACOUSTICS / NOISE CONTROL: GÉNIE ACOUSTIQUE / CONTRÔLE DU BRUIT:	Colin Novak	University of Windsor	(519) 253-3000
PHYSICAL ACOUSTICS / ULTRASOUND: ACOUSTIQUE PHYSIQUE / ULTRASONS:	Werner Richarz	Aeroustics	(416) 249-3361
MUSICAL ACOUSTICS / ELECTROACOUSTICS: ACOUSTIQUE MUSICALE / ELECTROACOUSTIQUE:	Annabel Cohen	University of P. E. I.	(902) 628-4331
PSYCHOLOGICAL ACOUSTICS: PSYCHO-ACOUSTIQUE:	Annabel Cohen	University of P. E. I.	(902) 628-4331
PHYSIOLOGICAL ACOUSTICS: PHYSIO-ACOUSTIQUE:	Robert Harrison	Hospital for Sick Children	(416) 813-6535
SHOCK / VIBRATION: CHOCs / VIBRATIONS:	Li Cheng	Université de Laval	(418) 656-7920
HEARING SCIENCES: AUDITION:	Kathy Pichora-Fuller	University of Toronto	(905) 828-3865
HEARING CONSERVATION: Préservation de L'Ouïe:	Alberto Behar	A. Behar Noise Control	(416) 265-1816
SPEECH SCIENCES: PAROLE:	Linda Polka	McGill University	(514) 398-4137
UNDERWATER ACOUSTICS: ACOUSTIQUE SOUS-MARINE:	Garry Heard	DRDC Atlantic	(902) 426-3100
SIGNAL PROCESSING / NUMERICAL METHODS: TRAITEMENT DES SIGNAUX / METHODES NUMERIQUES:	David I. Havelock	N. R. C.	(613) 993-7661
CONSULTING: CONSULTATION:	Corjan Buma	ACI Acoustical Consultants Inc.	(780) 435-9172
BIO-ACOUSTICS BIO-ACOUSTIQUE	Jahan Tavakkoli	Ryerson University	(416) 979-5000

EVALUATION OF AIDED AND UNAIDED AUDITORY FUNCTIONS FOR ROYAL CANADIAN MOUNTED POLICE (RCMP) MEMBERS

Véronique Vaillancourt¹, Chantal Laroche¹, Christian Giguère¹, Marc-André Beaulieu², and Jean-Pierre Legault²

¹Audiology and SLP Program, University of Ottawa, 451 Smyth Road, Ottawa, Ontario, Canada, K1H 8M5

²Occupational Health and Safety Branch, RCMP, 295 Coventry Road, Ottawa, Ontario, Canada, K1A 0R2

1. INTRODUCTION

Functional hearing involves activities such as sound detection, recognition, localization, and speech perception, which typically occur in noise and depend on binaural hearing [1]. Like many agencies, the RCMP currently uses the audiogram to classify hearing [2]; however, its relationship to functional hearing ability is limited [2-3]. RCMP would therefore like to establish new, scientifically-based, hearing criteria founded on a complete functional hearing assessment – a lengthy process involving the consideration of numerous work-related aspects [3].

This paper reports on the functional hearing assessment of 57 RCMP members facing operational restrictions due to hearing thresholds exceeding current health policy criteria. The primary objective was to assist RCMP in making more informed decisions regarding fitness-to-work in members wearing hearing aids, prior to the establishment of new hearing criteria. A second objective was to verify if hearing aids can improve performance, allowing members to carry out auditory functions required to safely perform their job. It was also hoped that results, together with a description of the hearing aid parameters used, could help identify best practices in hearing aid fittings for optimal functional hearing abilities in the RCMP work environment.

2. METHOD

2.1 Procedures

Members were first required to visit their audiologist to ensure proper hearing aid fit and function, and to gather useful information (settings, amplification strategy, number of programs, microphones, noise reduction algorithms, program used in work environment, etc.) by means of a questionnaire. Following at least one month of regular use, the functional hearing evaluation was performed with the hearing aid program and settings used on a regular basis in the workplace. The testing protocol included a basic audiological evaluation, in addition to unaided and aided measurements of: 1) binaural free field detection thresholds, 2) speech perception in quiet and in noise, using the Hearing in Noise Test – HINT [4-5], and 3) sound localization of a 65-dBA broadband noise (0.25-8 kHz).

Adaptive measurement of speech reception thresholds (SRT) were performed in either English or French, the native or preferred language of members, in quiet and in three conditions of 65-dBA speech spectrum noise: 1) speech in quiet (Quiet), 2) noise from the front (NF), 3)

noise from the right (NR), and 4) noise from the left (NL). A Noise Composite score was also computed $[(2 * NF + NR + NL) / 4]$ to represent overall functional ability for speech perception in noise under binaural listening conditions. Sound localization was assessed in three conditions, with twelve loudspeakers placed behind, to the right and to the left, thereby assessing horizontal localization in the left/right and front/back dimensions. For each condition, the number of left-right or front-back confusions was calculated.

2.2 Data Analysis

Since no scientifically-based hearing standards have yet been established, the following interim criteria were adopted: 1) SRT in quiet no greater than 40 dBA (the level of typical whispered speech at one meter), and 2) noise composite score and number of localization errors no worse than the 5th percentile performance for normal hearing individuals tested with the same protocol, in the same sound field. Individuals meeting the interim criteria were deemed operationally fit; for others, restrictions were maintained until empirically-based hearing standards are established.

In addition to individual data, group data was preliminarily analyzed to meet the second objective. It was anticipated that hearing aids would typically improve speech recognition, but could potentially hinder sound localization by disrupting important localization cues.

3. RESULTS

3.1 Hearing Aid Profile

Since individual RCMP members were evaluated with their own hearing aids, a variety of hearing aid styles, makes and models was seen. Binaural amplification was most prominent (n = 50), with CICs and open-fit BTEs being highly represented (Figure 1).

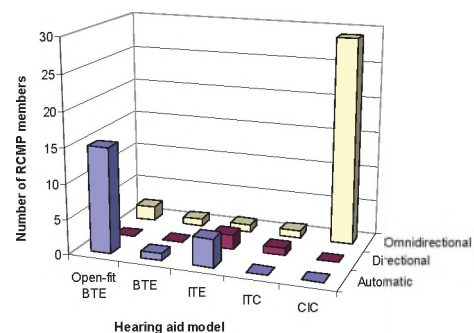


Figure 1. Hearing aid profile for the 57 RCMP members. (BTE: Behind-the-ear; ITE: In-the-ear; ITC: In-the-canal; CIC: Completely-in-the-canal)

3.2 Individual and Group Data

Individual data relative to interim criteria, summarized in Table 1, highlight a key issue for the design and fitting of hearing aids. While they can provide adequate benefits in speech recognition (the SRT in quiet and the noise composite score of 13 and 10 members, respectively, were improved to meet interim criteria), hearing aids significantly hindered front/back localization abilities in 16 cases, altogether changing the outcome from pass to fail. In contrast, left/right localization was not appreciably altered. Such findings are further evident in group data (Figure 2).

Table 1. Summary of unaided vs aided performance.

Task	Performance measure	Aided	Unaided	
			Pass	Fail
HINT	SRT in Quiet	Pass	40	13
		Fail	0	4
	Noise Composite	Pass	27	10
		Fail	0	20
Localization	# of L/R errors (Behind)	Pass	53	2
		Fail	0	2
	# of F/B errors (Side)	Pass	20	1
		Fail	16	20

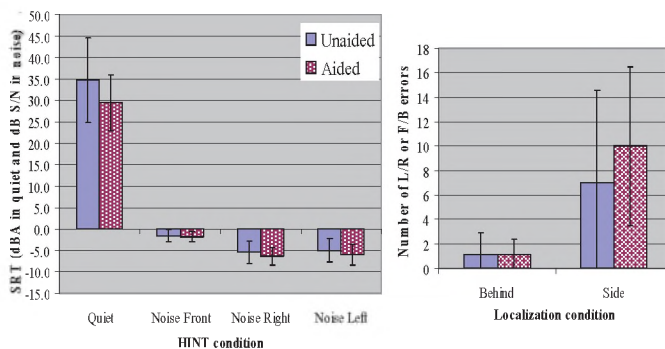


Figure 2. Mean SRTs (left panel) and number of L/R and F/B errors (right panel). Error bars show ± 1 standard deviation.

A 3-way mixed design ANOVA was performed independently for HINT results, and for the number of localization errors. Hearing aid model, the grouping variable, was sorted into five categories: automatic BTEs, omnidirectional BTEs, omnidirectional custom, automatic custom, and directional custom. The repeated measures variables were the use of hearing aids (unaided vs aided) and the testing condition (Quiet, NF, NR, NL for HINT; behind and side for sound localization).

For both hearing abilities, the analyses revealed a significant main effect of condition, main effect of hearing aid use, and interaction between both variables. No significant effect of grouping variable was found, even when analyses were repeated by grouping into two categories (auto + directional vs omni for speech; BTE vs custom for localization).

Apart from NR and NL, all HINT conditions were found to be significantly different from one another. HINT performance was generally better with hearing aids;

however, improvement from unaided to aided was most noticeable in Quiet, and smallest in NF. Localization accuracy was best when speakers were behind and, in contrast to speech recognition, was better without hearing aids, with greater increases in errors from unaided to aided in the side conditions.

4. DISCUSSION

In this sample, hearing aids: 1) improved SRTs; the effects being most prominent in Quiet and least considerable in NF, 2) neither significantly improved nor impeded L/R localization, and 3) in some cases substantially increased F/B errors in localization. Localization was generally better for sources behind than to the side, with fewer L/R than F/B errors, a result consistent with previous research [6]. Combined with previous findings [7], results indicate that hearing aids can considerably affect localization abilities. Additional analyses (not reported here) also point to the limited ability of audiometric data to predict functional abilities, and the need for individual assessments, both unaided and aided, for fitness-to-work purposes.

As members were tested using their own hearing aids, fitted and adjusted independently from this study, makes, models, styles and settings covered a wide range, making it difficult to identify optimal characteristics. Further work is needed to identify best practices in hearing aid fittings for optimal functional hearing abilities, and to develop empirically-based hearing standards for the RCMP.

REFERENCES

1. Soli, S.D. (2003). Hearing and Job Performance. Report for the Division of Behavioral and Social Sciences and Education of the National Research Council for the Committee on Disability Determination for Individuals with Hearing Impairment, 43 pp.
2. Trotter A., Brown J. (1994). Police Health, a Physician's Guide for the Assessment of Police Officers. Supply and Services Canada. Canada Communication Group-Publishing.
3. Giguère, C., Laroche, C., Soli, S., & Vaillancourt, V. (2008). Functionally-based screening criteria for hearing-critical jobs based on the Hearing in Noise Test. *Int J Audiol* 47: 319-328.
4. Nilsson, M., Soli, S.D., & Sullivan, J.A. (1994). Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am* 95: 1085-99.
5. Vaillancourt, V., Laroche, C., Mayer, C., Basque, C., Nali, M., Eriks-Brophy, A., Soli, S.D., & Giguère, C. (2005). Adaptation of the HINT (hearing in noise test) for adult Canadian Francophone populations. *Int J Audiol* 44: 358-369.
6. Keidser, G., Rohrseitz, K., Dillon, H., Hamacher, V., Carter, L., Rass, U., & Convery, E. (2006). The effect of multi-channel wide dynamic range compression, noise reduction, and the directional microphone on horizontal localization performance in hearing aid wearers. *Int J Audiol* 45: 563-579.
7. Van den Bogaert, T., Klasen, T.J., Van Deun, L., Wouters, J., & Moonen, M. (2006). Localization with bilateral hearing aids: without is better than with. *J Acoust Soc Am* 119: 515-526.

AUTOMATIC VOLUME SETTINGS FOR ENVIRONMENT SENSITIVE HEARING AIDS

Rana Rahal, Wail Gueaieb, Christian Giguère and Hisham Othman

School of Information Technology and Engineering, Univ. of Ottawa, 800 King Edward Avenue, Ottawa, Canada K1N 6N5
rraha039@uottawa.ca, wgueaieb@site.uottawa.ca, cgiguere@uottawa.ca, hisham@site.uottawa.ca

1. INTRODUCTION

The development of intelligent devices is becoming a popular trend in the hearing aid industry. Such devices aim at making the user's listening experience more natural and at improving customer satisfaction. One focus of interest in this paper is the automatic adjustment of the hearing aid control settings to minimize the need for manual user interventions. The proposed system is based on computational intelligence tools, namely artificial neural networks and neurofuzzy systems, which have the ability to learn the dynamics of highly nonlinear systems without the need for the explicit knowledge of their mathematical models. Such techniques are adopted here to map the acoustic features (input space) to the desired volume setting (output space) of the hearing aid user. Eventually the proposed system would be integrated into a trainable self-learning hearing aid, such as [1] and [2].

In this paper, two computational intelligence tools, a multi-layer perceptron (MLP) and an adaptive network-based fuzzy inference system (ANFIS) were analyzed on three simulated users with moderate, severe, and profound hearing losses. A hearing aid simulation system provided target volume settings to train and test the learning networks, selected to optimize the speech intelligibility index (SII) in each acoustic situation [3]. The performances of both soft computing models obtained from over 2000 recordings demonstrated a high efficiency of the adopted approach in automatically optimizing volume settings for the three simulated users. The framework of the system is presented in the following section.

2. METHOD

The objective of this paper is to develop an automatic volume control system to optimally match the volume gain preferences of the hearing aid user in different acoustic situations. The automatic volume control system proposed contains the followings stages: a large data set of environmental audio files, feature extraction, selection of the influential features and mapping of features to their optimal volume setting (computational intelligence tool). The system's layout is depicted in Figure 1.

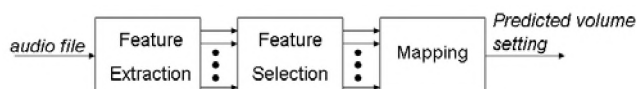


Fig. 1. System's Layout.

2.1 Experimental Data

Audio Files

A virtual environment simulator was used to generate environmental noises. The virtual environment simulator can emulate speech distortion in real environments, thus creating realistic training and testing data for the network. The sound database used for experiments consisted of a total of two thousand noisy speech files, which mimics typical environmental conditions.

Target Volume Settings

Target volume settings are obtained through a simulated hearing aid user assumed to adjust its hearing aid to optimize intelligibility at all times. This set of targets, defined as the volume settings for maximum SII in all environment conditions, is required to train the networks to map features into optimal volume settings. Users with moderate, severe and profound hearing loss (HL) and different uncomfortable listening levels were simulated.

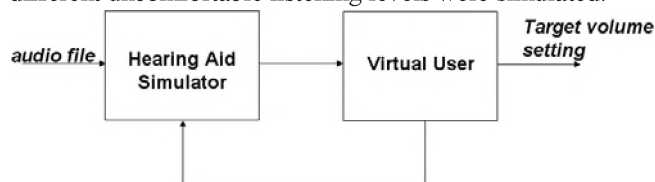


Fig. 2. Generation of target volume settings.

2.2 Feature Extraction

To perform mapping of the feature input space to the output space, it is desirable to extract a set of feature vectors which preserves information that is highly correlated to the output space. Doing so will improve the performance of the network and decrease processing time. Twenty-four real valued features were considered, including a number of frequency-domain and time-domain features. Features were extracted from frames of a sampled audio signal, with no overlap between frames.

2.3 Feature Selection

Feature selection can provide a sub-optimized set of features and reduce the dimensionality of the feature vector. The selection criteria typically involves the minimization of a measured error from models with different inputs. A starting set of 24 features is considered, and then a subset of highly correlated features is selected by using a feature selection method known as the sequential forward search (SFS) [4]. The SFS method was used to select three sets of influential features. Each set corresponds to a certain simulated user (moderate, severe and profound) and includes six influential features.

2.4. Mapping

Two computational intelligence tools were considered to perform the mapping, the MLP and ANFIS models [5], which were analyzed on the three simulated users.

Learning Process

The learning process for both the MLP and ANFIS models is divided into three phases: training, validation and testing. The training process requires a set of network inputs and target outputs for the models to learn (shown in Figure 2). Training allows the models to become familiar with every possible situation. Validation improves the models' generalization during testing and avoids the models from "overfitting" the training data. Finally, the testing stage provides the models with unfamiliar data and will determine the models' robustness and ability to make generalizations on unfamiliar data.

3. RESULTS

For cross-validation, the data set is partitioned into three groups, 1200 files (60%) for training, 400 files (20%) for validation and 400 (20%) for testing. The performance of the models is measured by determining how accurate the predicted outputs of the models are, after training and testing. The predicted output of the model is compared against the target signal and when the mean square error between the two is minimized, the performance of the model is maximized. The performance measure is evaluated by the volume error (VE) and the absolute value of the speech intelligibility index error (SIIE), presented in the following equations:

$$VE(i) = V_{opt}(i) - V_{pred}(i) \quad (1)$$

$$SIIE(i) = |SII_{opt}(i) - SII_{pred}(i)| \quad (2)$$

where i refers to each pattern, the subscript opt and $pred$ refer to the target and the quantity predicted by the network, respectively.

Figure 3 presents the MLP's testing performance. The plot is the volume error (equation 1) for each testing pattern for the moderate HL user. At first glance, there are a number of volume errors greater than 20 dB. However, this may not necessarily result in a high SII error. Indeed, Figure 4 demonstrates that the great majority of the audio files resulted in a low SII error (equation 2).

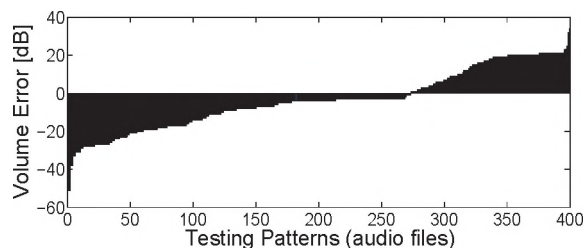


Fig. 3. Volume error plot. MLP's testing performance for the moderate HL user.

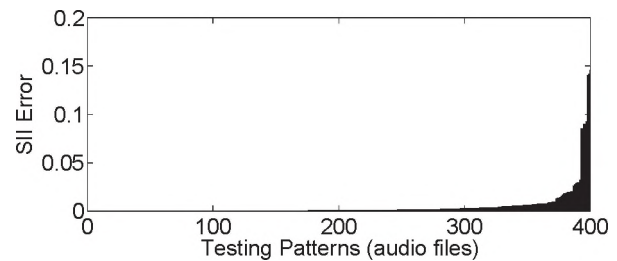


Fig. 4. SII error plot. MLP's testing performance for the moderate HL user.

Figure 5 compares the testing performances of the MLP and ANFIS models for the moderate user. For 95% of the testing patterns, the MLP's performance obtained an SII error of less than 0.02 and ANFIS obtained an SII error of less than 0.005. This demonstrates ANFIS's ability in optimizing the moderate HL user's speech intelligibility more effectively.

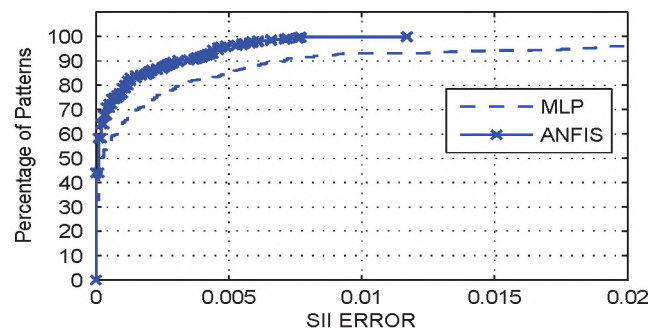


Fig. 5. SII error plot of MLP versus ANFIS testing performances for the moderate HL user.

4. DISCUSSION

From the plots presented in the results, the MLP optimized the volume setting for the majority of the testing patterns for the moderate HL user. As a result the speech intelligibility index for the moderate HL user was optimized as well. In conclusion, both the MLP and the ANFIS showed high accuracy in automatically optimizing the SII for a simulated user. ANFIS performance in terms of SII error is slightly advantageous compared to the MLP's performance. A future direction of this research is to test the proposed system on real human subjects.

REFERENCES

- [1] H. Dillon, J.A. Zakis, H. McDermott, G. Keidser, W. Dreschler, E. Convery, "The trainable hearing aid: What will it do for clients and clinicians?" *Hearing Journal*, vol. 59, pp. 30-36, April 2006.
- [2] S. Chalupper and H. Heuermann, "Learning Optimal Gain in Real World Environments," *International Hearing Aid Conference*, Lake Tahoe, USA, 2006.
- [3] ANSI Standard S3.5-1997, "American National Standards Methods for the Calculation of the Speech Intelligibility Index," American National Standards Institute, New York, 2002.
- [4] A. Jain and D. Zongker, "Feature selection: Evaluation, application, and small sample performance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 2, Feb. 2007.
- [5] F. O. Karray and C. De Silva, "Soft Computing and Intelligent Systems Design: Theory, Tools and Applications," England: Pearson Education Limited, 2004, pp. 249-260.

A NEW FEATURE SELECTION METHOD FOR VOLUME CONTROL IN DIRECT-LEARNING HEARING AID SYSTEMS

Jin Zhou¹, Hisham Othman¹, Hilmi Dajani¹, and Tyseer Aboulnasr²

¹ Dept. of Electrical Engineering, University of Ottawa, 800 King Edward Av., ON., Canada, K1N 6N5, tenniszj@163.com

² Faculty of Applied Science, University of British Columbia, 5000-2332 Main Mall, BC., Canada, V6T 1Z4

1. INTRODUCTION

A desirable feature of modern hearing aids is the ability to automatically adjust its behavior in different acoustic environments. There are two kinds of approaches that can achieve this. One is based on “classification” of environments, another is “direct learning” of preferred hearing aid settings. Our focus is on the second approach, in which an artificial neural network (ANN) learns the preferred volume setting of the hearing aid user. The performance of such a direct-learning system strongly depends on the chosen signal features. While a large number of features have been derived for environment classification (Büchler, 2002; Nordqvist *et al*, 2004), we are now aware of any that have been derived specifically for “direct learning”, which has different requirements. For example, environment classification should not in general be sensitive to the sound volume, whereas direct learning of volume setting not only depends on the volume, but also depends on how this volume affects speech intelligibility and user comfort. Moreover, the preferred volume setting depends on the hearing loss profile, and whether it is profound, severe, or moderate. The goal of this work is to derive suitable features that the ANN will use to set the volume such that it optimizes speech intelligibility. New features are proposed, which are based on measures of speech intelligibility, namely the Speech Intelligibility Index (SII) (ANSI S3.5-1997) and the Coherence SII (CSII) (Kates and Arehart, 2005). The performance of these features is then investigated using a simulator of a hearing aid user (SPOT, 2009).

2. METHOD

2.1 Perceptually-dependent Features

The new features are derived from the calculation of the CSII in third-octave bands (ANSI S3.5-1997; Kates *et al*, 2005), and they reflect the SNR and energy of the speech signal in these bands, weighted by the psychocoustic characteristics of the listener. They are calculated as follows:

- The pure and distorted speech files, $x(n)$ and $y(n)$, respectively, are divided into 50% overlapping segments. Each segment is multiplied by a Hamming window, and then the segments $y_m(n)$ are categorized into three groups for high, medium and low energy segments.

In each energy group, the following steps are repeated to get the CSII in each band.

1. The spectra of $x_m(n)$ and $y_m(n)$, i.e. $X_m(k)$ and $Y_m(k)$ are obtained using FFT, and then the coherence measure for each group is estimated as follows:

$$|\gamma(k)|^2 = \frac{\left| \sum_{m=0}^{M-1} X_m(k) Y_m^*(k) \right|^2}{\sum_{m=0}^{M-1} |X_m(k)|^2 \sum_{m=0}^{M-1} |Y_m(k)|^2} \quad (1)$$

where m is the segment index, k is the FFT bin index and M is the number of segments.

2. Then, the perceptual weighting procedure in ANSI S3.5-1997 is used to calculate CSII in each third-octave band. The hearing loss profile is involved in this procedure. The difference of CSII and SII in the standard is that the SNR is replaced by the signal distortion ratio (SDR) (Kates and Arehart, 2005):

$$SDR(j) = \frac{\sum_{k=0}^K W_j(k) |\gamma(k)|^2 S_{yy}(k)}{\sum_{k=0}^K W_j(k) (1 - |\gamma(k)|^2) S_{yy}(k)} \quad (2)$$

where $S_{yy}(k)$ is the estimated power spectral density of the distorted signal and j is the index of third-octave band. W is the simplified ro-ex filter for the band.

- For each third-octave band, one new feature is calculated by combining the weighted CSII for the three energy groups (high energy level, mid-energy level, low energy level) using the following function:

$$\begin{aligned} c(j) = & -3.47 + 1.84CSII_{Low}(j) \\ & + 9.99CSII_{Mid}(j) + 0.0CSII_{High}(j); \end{aligned} \quad (3)$$

$$New_feature(j) = e^{-c(j)}$$

In total, we have 18 perceptually-dependent features.

2.2 Evaluation of Performance

To test the performance, we used a multi-layer perceptron with 1 hidden layer (14 neurons) and 1 output layer (1

neuron) (Demuth *et al.*, 2008). We used two feature sets: conventional features used in environment classification, which include CGAV, CGFS, Pitchvar, Delpitch, Onstem, Onsetv, Onsetc, Onseth, Beat, Width, Symmetry, Skewness, Kurtosis, Lower Half, MLFS, M1, M2, M3 (Büchler, 2002) and the new features we described in the previous section. The inputs of the neural network were the two feature sets, which were evaluated separately, and the output was the volume gain in dB. Mean square error (MSE) between the volume gain at the output of the ANN and the volume gain that optimizes the speech intelligibility in a simulation of a hearing aid user with profound hearing loss. (This work was done in a simulator to mimic the interactions between the acoustic environment (sound files) and the user behavior (settings) (SPOT, 2009)). One hundred repetitions of the testing and training procedures were used to evaluate the performance.

The dataset for the experiments was generated as follows: We chose 12 30-sec pure speech samples and scaled them to 65 dB SPL. We generated the distorted files by adding white noise to pure speech at different SNR (from -10 to 15dB) and then scaled the files to SPL=65dB. Consequently, we had 312 audio files, from which we randomly chose 200 for training and another 100 for testing. The randomization was done for each of the 100 experiment repetitions.

3. RESULTS

Figure 1 shows the performance of the ANN with the conventional and new feature sets over 100 experiment repetitions. The average MSE obtained from the conventional features is 1.3 (dB based), while that obtained from the features is 1.4 (dB based). With this dataset, we find that the performance of the new features is very similar to the performance of the conventional features.

4. DISCUSSION

Both the new and conventional feature sets performed very well. However, it should be noted that the dataset in this study used only white noise and one SPL level. In the future, we will repeat the experiment using a much more comprehensive dataset that mixes different types of noise, signal, and SNR levels, and with different hearing loss profiles. We expect that the new feature set will perform better in such a challenging situation. One disadvantage of the new feature set is that it requires an estimate of the SNR (or SDR), which may be difficult to obtain at times. We will therefore further investigate features, such as the temporal “modulation level” (Büchler, 2005), that would provide alternative information for the SNR, without the need to estimate it.

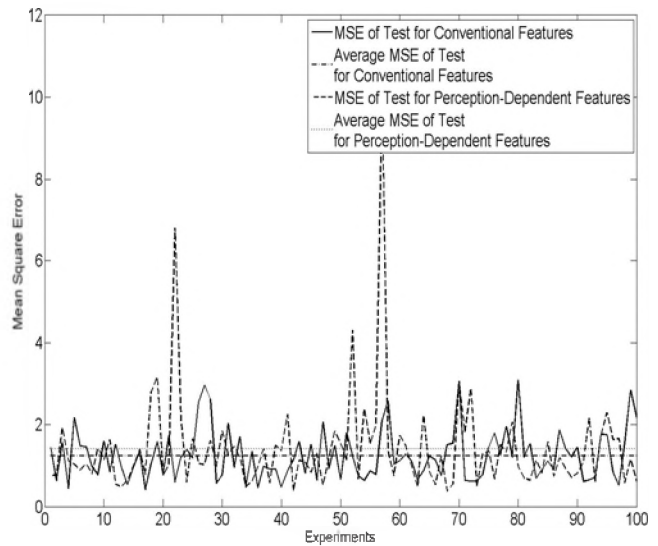


Figure 1. The performance of the conventional and new features.

REFERENCES

- ANSI S3.5-1997, (1997), Methods for calculation of the speech intelligibility index, American National Standards Institute, Inc., New York.
- Büchler, M.C., (2002), Algorithms for Sound Classification in Hearing Instruments, PhD Thesis, Swiss Federal Institute of Technology Zurich, pp. 1-149.
- Demuth, H.; Beale M.; and Hagan M., (2008), Neural Network Toolbox 6 (User's Guide). The Mathworks.
- Kates, J.M.; Arehart, K.H., (2005), Coherence and the speech intelligibility index, Acoustical Society of America, vol. 117, pp. 2224-2236.
- Nordqvist, P.; Leijon, A. (2004), "An efficient robust sound classification algorithm for hearing aids," The Journal of the Acoustical Society of America, vol. 115, pp. 3033-3041.
- SPOT (Signal Processing Oriented Technology) Group, (2009), University of Ottawa, Ottawa, Ontario, Canada, Report on Learning Systems, pp. 1-122.

ACKNOWLEDGEMENTS

The authors thank Professor Christian Giguère and Professor Wail Gueaieb for helpful comments and suggestion.

SUBBAND ADAPTIVE FILTERING FOR ACOUSTIC FEEDBACK COMPENSATION IN HEARING AIDS

Brady Laska¹, Miodrag Bolić², and Rafik Goubran¹

¹Dept. of Systems and Computer Engineering, Carleton University, Ottawa, Canada, {laska,goubran}@sce.carleton.ca

²University of Ottawa School of Information Technology and Engineering, Ottawa, Canada mbolic@site.uottawa.ca

1. INTRODUCTION

Acoustic feedback occurs in hearing aids when the amplified sound signal played out to the user is coupled back to the input microphones. Low level feedback acts as reverberation, masking the desired signal; high level feedback can lead to system instability and high intensity oscillations that manifest as "whistling" or "howling" sounds that are very disturbing to the user.

Methods currently used to prevent feedback include limiting the gain of the hearing aid to prevent a positive feedback loop from forming, and applying adaptive notch filters to attenuate the oscillating frequencies when feedback occurs (Hamacher, 2005). Gain limiting reduces the utility of the device, while notch filtering is reactive and can distort the speech signal.

The most desirable method for feedback control is direct signal compensation. If the feedback path is known, a replica of the feedback signal can be created and subtracted from the microphone signal, enabling arbitrary gain in the hearing aid without signal distortion. The difficulty with this approach is obtaining an accurate model of the feedback path. Since the feedback path changes when the hearing aid moves within the ear, or when telephones, hats and other reflecting surfaces are nearby, dynamic modelling of the feedback path with an adaptive filter is required.

2. SUBBAND ADAPTIVE FILTERING

While the adaptive feedback path modelling is typically performed using a fullband, time-domain signal (Hamacher, 2005), subband adaptive filtering has been used in applications such as acoustic echo cancellation to overcome problems of slow convergence for coloured inputs and high complexity for large adaptive filters. Fig. 1 presents a block diagram of a subband adaptive feedback compensation system within a hearing aid. The input to the hearing aid is the microphone signal $z[n]$ and the output is the loudspeaker signal $x[n]$. The input contains the incoming speech signal $s[n]$ as well as the feedback signal $d[n]$, which is formed when the output signal is coupled to the microphone via the feedback path $H(z)$. During system identification periods, a probe signal $p[n]$ is generated inside the hearing aid.

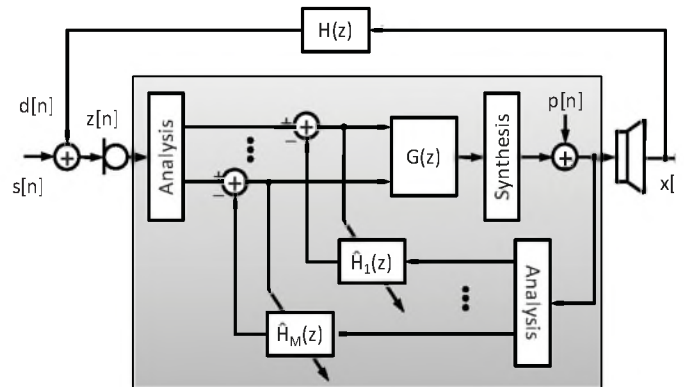


Fig. 1: Subband hearing aid feedback compensation system.

This signal is simultaneously played out through the loudspeaker and passed through a bank of adaptive filters containing estimates of the subband feedback paths. The outputs of the filters are replicas of the subband-decomposed feedback probe signal, which are subtracted from the true feedback probe signal measured at the microphone. The resulting error signal is then used in the normalised least-mean square (NLMS) algorithm to adjust the adaptive filter coefficients to the true feedback path.

Since modern hearing aids currently decompose the input signal to facilitate noise reduction and frequency dependent amplification, adaptive feedback compensation could easily be performed in subbands. However, while acoustic echo paths can require thousands of taps to accurately model, acoustic hearing aid feedback paths are much shorter. The complexity benefits are thus reduced, raising the question of whether subband structures offer any advantages for feedback compensation.

3. SUBBAND/FULLBAND COMPARISON

3.1 Changing feedback paths

Feedback paths can change with movements of the hearing aid caused by ear or jaw movements, or with changes in the arrangement of reflecting surfaces near the outer ear (Hamacher, 2005). Depending on the nature of the path change, characteristics of the frequency response may remain the same despite the disruption. Feedback paths are

typically bandpass, with the 1 - 4 kHz regions experiencing the least attenuation (Puder, 2004). An obstructing object in the feedback path may affect higher frequencies more than lower, because high frequencies are absorbed more easily and because the wavelength of the lower frequencies may be larger than the object. While a fullband filter would have to adapt all tap weights in response to the disruption, a subband adaptive filter would only be required to adjust the weights of the filters in the affected bands. Even if the effects of the feedback path change are uniform across frequency, subband adaptive filters are shorter, so they should be faster to respond to all types of path change.

To compare the performance of feedback compensation structures under controlled conditions, fullband and subband NLMS feedback cancellers were simulated using computer generated white Gaussian noise excitation in a changing acoustic environment at 25 dB measurement SNR. The NLMS adaptation step-size for both algorithms was fixed at $\mu=0.01$. After an initial convergence period an abrupt feedback path change was simulated by changing the feedback path coefficients, and then a more realistic gradual change was simulated by linearly interpolating back to the first response. The mean-squared error (MSE) - the power of the uncancelled feedback signal - averaged over 20 trials, is shown in Fig. 2. The subband structure shows rapid initial convergence and faster re-convergence from the disturbances. During and after the feedback path changes, the subband structure offers consistently better worst-case MSE performance.

3.2 Disturbing speech signal

Since feedback path changes can occur during active speech periods, feedback cancellation systems must be able to adapt in the presence of the incoming speech signal, which acts as a high-level disturbance and can cause the system to diverge. Large step-sizes are desired for faster convergence, but also lead to faster divergence.

Subband adaptive filters can offer more flexibility in dealing with disturbing speech. Since filter divergence occurs when the disturbing signal power is high, and the power of speech signals decays with frequency, a smaller step-size can be used in the low-frequency bands with high incoming speech power, and a larger step-size in the higher frequency bands. Fig. 3 presents MSE curves for fullband NLMS with step-sizes of $\mu=0.005$ and $\mu=0.025$, and for subband NLMS with step-sizes that increase logarithmically with frequency, from 0.005 to 0.025. The large step-size fullband NLMS converges rapidly, but also exhibits MSE fluctuations of up to 15 dB as the interfering speech signal drives the divergence; reducing the step-size greatly reduces the MSE fluctuations, but the rate of convergence is much slower. The subband NLMS with the logarithmically-spaced step-sizes presents a compromise: the initial convergence rate approaches that of the large step-size fullband algorithm, but the divergence during strong speech segments is minimal.

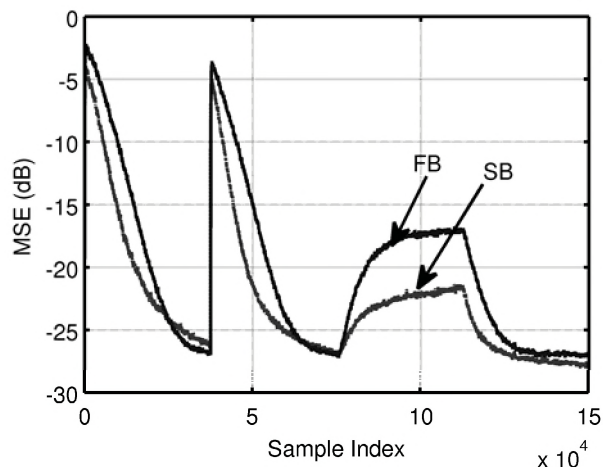


Fig. 2: MSE performance of fullband and subband feedback compensation in changing environments.

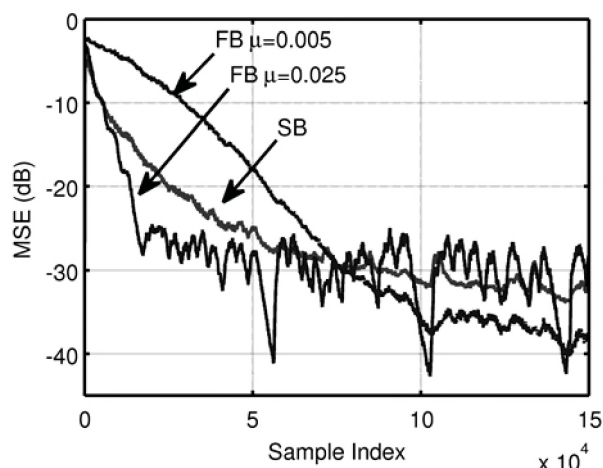


Fig. 3: MSE performance of fullband and subband feedback compensation in the presence of a high-level speech signal.

4. DISCUSSION

Hearing aids use a signal subband decomposition in order to apply frequency-dependent gain and noise reduction strategies. Performing feedback compensation on these subband signals rather than the fullband, allows the system to achieve better tracking of changing feedback paths, and offers additional flexibility in dealing with performance-stability trade-offs.

REFERENCES

- Hamacher, V. et al. (2005), Signal processing in high-end hearing aids: state of the art, challenges, and future trends, EURASIP J. App. Sig. Process., vol. 2005, pp. 2915 – 2929
- Puder, H and Beigel, B (2004) "Controlling the adaptation of feedback cancellation filters - problem analysis and solution approaches," in Proc. EUSIPCO'04.

BINAURAL OBJECTIVE INTELLIGIBILITY MEASUREMENT AND HEARING AIDS

Nicolas Ellaham¹, Christian Giguère², and Wail Gueaieb¹

¹School of Information Technology and Engineering, University of Ottawa, 800 King Edward Ave., Ottawa, ON, K1N6N5

²School of Rehabilitation Sciences, University of Ottawa, 451 Smith Rd., Ottawa, ON, K1H8M5

nellaham@uottawa.ca, cgiguere@uottawa.ca, wgueaieb@site.uottawa.ca

1 INTRODUCTION

Assessment of speech intelligibility is an important area in the design of hearing devices, and consists of two complementary approaches: (1) the subjective measurement of intelligibility scores based on experimental listening tests, and (2) the objective prediction of speech intelligibility under different listening conditions. Objective measures of speech intelligibility date back to the Articulation Index (AI) [1]. Extensions of the AI have led to models of speech intelligibility for normal hearing and hearing-impaired subjects in a variety of listening conditions. Two such models which have gained popularity and become standard in the acoustic literature are the Speech Intelligibility Index (SII) [2] and the Speech Transmission Index (STI) [3], both designed for monaural listening conditions.

In recent years, much attention has focused on binaural hearing. Binaural objective measures incorporate a pre-processing stage to model binaural interactions, followed by intelligibility prediction using a monaural measure like the SII or STI. In [4], Beutelmann and Brand developed one such measure using the equalization cancellation (EC) model [5] and the monaural SII. In a more recent work [6], Wijngaarden and Drullman developed a binaural measure based on the STI. Such measures, however, are not designed to deal with complex nonlinearities. Digital processing inside a hearing aid often involves nonlinear operations such as clipping, compression, and those found in noise reduction algorithms. In this paper, we propose adding an initial stage to deal with such nonlinear processing in order to obtain a new binaural objective measure of speech intelligibility for use with digital hearing aids.

Figure 1 shows a schematic diagram of the proposed 3-stage binaural objective measurement system. Conceptual details of each stage are presented in the subsequent sections of this paper. Inputs to the overall system are the individual speech and noise signals received at the left and right ears (4 signals in total). The left and right hearing aids are enclosed within the first processing stage (HA blocks). The system's output is an objective prediction of speech intelligibility.

2 SIGNAL SEPARATION

The first stage of the binaural measure is based on a simple signal separation scheme used by Hagerman and Olofsson in a monaural study on the effects of noise reduction algorithms in hearing aids [7]. Its incorporation into the proposed model has a dual purpose: (1) it provides separate

access to the speech and noise signals at the output of the left and right hearing aids, a necessary requirement for the subsequent stage in the model; and (2) it deals with nonlinear signal processing and complex signal and noise mixtures inside the hearing aid.

Signal separation is performed by presenting the speech and noise signals simultaneously to the hearing aid twice, the second time with the phase of the noise reversed. For each presentation, the inputs (a_{in} , b_{in}) and outputs (a_{out} , b_{out}) of the hearing aid are summarized in the following equations:

$$a_{in}(t) = u(t) + v(t) \quad a_{out}(t) = u'(t) + v'(t) + e_1(t) \quad (1)$$

$$b_{in}(t) = u(t) - v(t) \quad b_{out}(t) = u'(t) - v'(t) + e_2(t) \quad (2)$$

where $u(t)$ and $v(t)$ denote the input speech and noise signals respectively, $u'(t)$ and $v'(t)$ the output speech and noise signals respectively, and $e_1(t)$ and $e_2(t)$ denote error signals (e.g. internal noise, nonlinear interactions between speech and noise, and distortion due to the noise reduction algorithm). Assuming the error signals are sufficiently small, we can recover separate estimates of the speech $c(t)$ and noise $d(t)$ output signals by adding and subtracting the outputs of the hearing aid from both presentations:

$$c(t) = a_{out}(t) + b_{out}(t) = 2u'(t) + e_1(t) + e_2(t) \quad (3)$$

$$d(t) = a_{out}(t) - b_{out}(t) = 2v'(t) + e_1(t) - e_2(t) \quad (4)$$

This approach was used to study the impact of different noise reduction algorithms, with and without compression under monaural conditions [7]. In the present work, this scheme is applied to retrieve the separate speech and noise signals at the output of the left and right hearing aids (Figure 1). These will be the inputs to the next stage of the proposed binaural measure.

3 EQUALIZATION-CANCELLATION MODEL

The next stage is the central binaural processing element, and is based on the EC theory introduced by Durlach [5]. The EC theory models the way the human auditory system uses binaural cues to improve the perception of a "target" speech signal in the presence of a noise masker. The model transforms the total signal at one ear relative to the total signal at the other ear such that the masking components are equalized (*equalization process*), and then subtracts one from the other (*cancellation process*). If the equalization

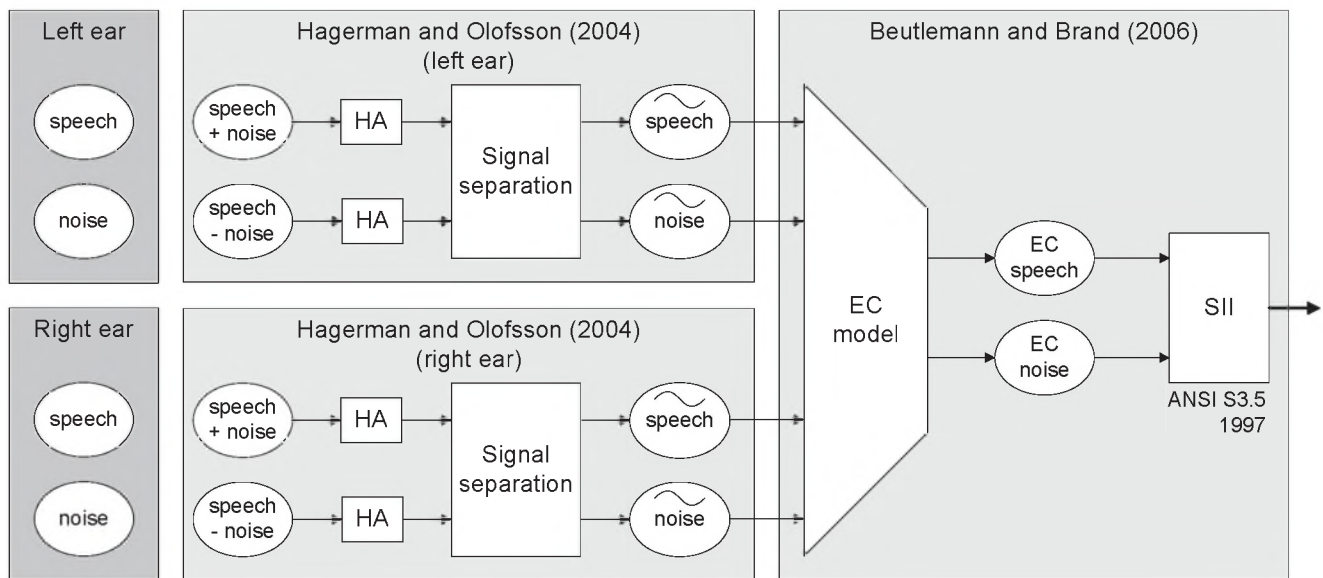


Figure 1: Diagram of the proposed binaural objective measure of speech intelligibility.

transform is perfect, and provided that interaural time and level differences for the target signals are different from those for the noise maskers, the latter will ideally cancel out due to destructive interference. If these interaural differences are identical, the EC processing will cancel out both target and masking components. Otherwise, the residual signal will theoretically have an improved SNR.

Since processing in the EC model is linear, speech and noise signals can be processed separately. This processing is performed independently in each band of the auditory filter, which is modeled using a Gammatone analysis filterbank. The EC model also includes a selection mechanism at the output of each band that selects the signal with the best SNR (left, right or EC-processed). Moreover, since the auditory system is not a perfect processor, the EC model includes various types of errors to account for human inaccuracies. Finally, the processed signals from each frequency band are combined at the output of this stage, to produce speech and noise signals to be used for intelligibility prediction.

4 INTELLIGIBILITY PREDICTION

In this work, prediction of speech intelligibility is done using the standard monaural SII. The SII is “a physical measure that is highly correlated with the intelligibility of speech as evaluated by speech perception tests given a group of talkers and listeners” [2]. It is essentially calculated as a weighted average of the amount of information available to the listener over a chosen number of frequency bands. The ANSI standard provides a detailed description of the SII calculation. Other standard measures could also have been used. The choice of the SII is primarily based on the model presented in [4], as well as the possible future extension of this model to use the coherence function in a manner similar to the work of Kates and Arehart with the Coherence-based SII [8].

5 CONCLUSION

This paper introduced a new objective measurement system for predicting speech intelligibility under binaural listening conditions. The new measure is designed for use with digital hearing aids, which often perform complex nonlinear signal processing. As binaural hearing aids become increasingly popular, the proposed measure will be useful to study the impact of emerging binaural noise reduction algorithms, similar to the monaural work presented in [7]. Our goal is also to use this measurement system to develop and test new algorithms for online learning of user control setting preferences in binaural trainable hearing aids.

REFERENCES

- [1] ANSI-S3.5 (1969), “Methods for the calculation of the Articulation Index,” American National Standards Institute, New York.
- [2] ANSI-S3.5 (1997), “American National Standard: Methods for the calculation of the speech intelligibility index,” American National Standards Institute, New York.
- [3] IEC 60268-16 (1998), “Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index”.
- [4] Beutelmann R. & Brand T. (2006), “Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners,” *J. Acoust. Soc. Am.*, 120 (1), pp. 331–342.
- [5] Durlach NI. (1963), “Equalization and cancellation theory of binaural masking-level differences,” *J. Acoust. Soc. Am.*, 35 (8), pp. 1206–1218.
- [6] Van Wijngaarden, SJ. & Drullman, R. (2008), “Binaural intelligibility prediction based on the speech transmission index”, *J. Acoust. Soc. Am.*, 123 (6), 4514–4523.
- [7] Hagerman B. & Olofsson A. (2004), “A method to measure the effect of noise reduction algorithms using simultaneous speech and noise”, *Acust. Acta Acust.*, 90 (2), 356–361.
- [8] Kates JM. & Arehart KH. (2005), “Coherence and the speech intelligibility Index”, *J. Acoust. Soc. Am.*, 117, 2224–2237

WITHIN-CATEGORY VARIATION IN L2 ENGLISH VOWEL LEARNING

Ron I. Thomson¹ and Talia Isaacs²

¹Dept. of Applied Linguistics, Brock University, St. Catharines, ON, Canada, L2S 3A1, rthomson@brocku.ca

²Dept. of Integrated Studies in Education, McGill University, Montréal, QC, Canada, H3A 1Y2, talia.isaacs@mcgill.ca

1. INTRODUCTION

Second language (L2) pronunciation research typically treats L2 phonological categories as monolithic wholes, without regard for potential within-category variation in learners' productions. Flege [1] argues that L1 transfer most easily occurs when an L2 phoneme has an L1 counterpart in a similar or identical phonetic environment. Some empirical evidence supports this claim (see [2] & [3]). Other factors may also contribute to within-category variation in L2 speech. In this study, we examine whether L2 learners are more apt to accurately produce English vowels 1) when vowels are in more familiar rather than less familiar words and 2) when the learners have access to orthographic representations of those words. We also examine the extent to which the learners' L1 plays a role.

2. METHOD

2.1 Speakers

19 Standard Mandarin (15 female, 4 male; *M* age = 40.1; range = 29-49) and 19 Slavic (12 female, 7 male; *M* age = 38.6; range = 29-49) speakers participated. The Slavic group comprised 13 Russian, three Serbian, two Ukrainian & one Polish speaker. The speakers' mean Length of Residence in Canada was 15.6 months (range = 4 - 40 months). All had been enrolled in intensive ESL classes for an average of 5.2 months (range 2 -16) and were assessed as beginners by the Canadian Language Benchmarks.

2.2 Stimuli

Stimuli comprised a list of 30 English progressive verbs containing 10 target Canadian English vowels. We chose words we thought would vary in terms of their degree of familiarity for beginner ESL learners (e.g., 'keep', 'feed' and 'beat' for the vowel /i/). Words were also chosen to include a range of onsets so as to mitigate any potential contextual biases in performance. The resulting word list was randomized and a male speaker of Canadian English was recorded producing each word, with a five second pause inserted between items. Stimuli were saved to CD.

2.2 Speaking task

L2 productions were elicited in two counter-balanced conditions: 1) after hearing the recorded CD prompts and 2) after hearing the recorded prompts accompanied by the written word list. In a third and final condition, all participants were asked to read the word list without any auditory prompt. Speaker productions were recorded in a quiet room using a high quality Marantz digital recorder.

2.3 Assessment of Word Familiarity

After the recording session, participants were asked to complete a 4-point familiarity judgment for each word where 0 = I don't know it; 1 = I might know it; 2 = I think I know it; and 3 = Yes, I know it.

2.4 Intelligibility Ratings for L2 Productions

Individual words were extracted from all recordings and saved as separate sound files for presentation to two phonetically trained native English speaker judges. Using *Praat* (www.praat.org), all recordings of a given word (e.g., "sitting") were randomly presented and the judges used a mouse click to indicate whether each production was perceived as containing the intended vowel or a different vowel. After one word set was evaluated, recordings of the next target word set were presented for assessment. Multiple sessions were needed to evaluate all 3420 items.

3. RESULTS

Mean familiarity scores for each word were used to assign words containing each vowel category to one of three groups: 1) most familiar, 2) second most familiar and 3) least familiar. For example, with mean familiarity scores shown in parentheses, "cool" (2.8), "fool" (2.4) and "boot" (1.3) were assigned to groups 1, 2 and 3 respectively. In one case, a tie in mean familiarity scores was broken through reference to word frequency in the British National Corpus.

The judges agreed on the identity of 81% of items. Chi-square analyses found no significant differences between judges' intelligibility scores across each of the three speech elicitation conditions, nor for intelligibility scores across each word familiarity group. Responses were pooled across judges to arrive at a mean intelligibility score for each item.

A two-way partially repeated measures ANOVA with Word Familiarity (3 levels) and Speech Elicitation Condition (3 levels) as within-subject factors, and L1 as a between subject factor, revealed a significant effect for Word Familiarity [$F(2,72) = 58.918, p = .000, \eta^2 = .621$] as well as Speech Elicitation Condition [$F(2,72) = 53.689, p = .000, \eta^2 = .599$]. No significant effect of L1 background on vowel intelligibility was found. Nor were there any significant interactions between factors.

Bonferroni adjusted *t*-tests found that vowels in the most familiar lexical context were significantly more intelligible than those in the least familiar lexical context [$t(37) = 8.493, p < .001$], but not significantly more intelligible than vowels in the second most familiar lexical context [$t(37) = 2.403, p = .0215$]. Vowels in the second most familiar lexical context

were significantly more intelligible than vowels in the least familiar lexical context [$t(37) = 8.351, p < .001$].

Post-hoc Bonferroni adjusted t -tests found that vowels in the Auditory + Reading condition were significantly more intelligible than vowels in the Auditory only [$t(37) = 3.22, p = .003$] and the Reading only conditions [$t(37) = 8.59, p < .001$]. In addition, the vowels in the Auditory only condition were significantly more intelligible than those in the Reading only condition [$t(37) = 7.20, p < .001$].

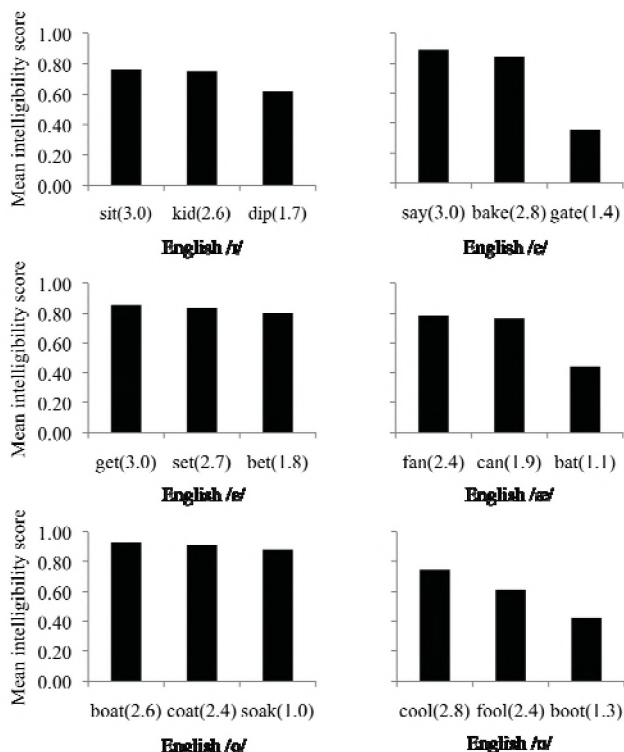


Figure 1. Mean intelligibility scores for English vowels in words that followed the rank order predicted by mean familiarity scores (indicated in parentheses). Results are pooled across Mandarin and Slavic speakers.

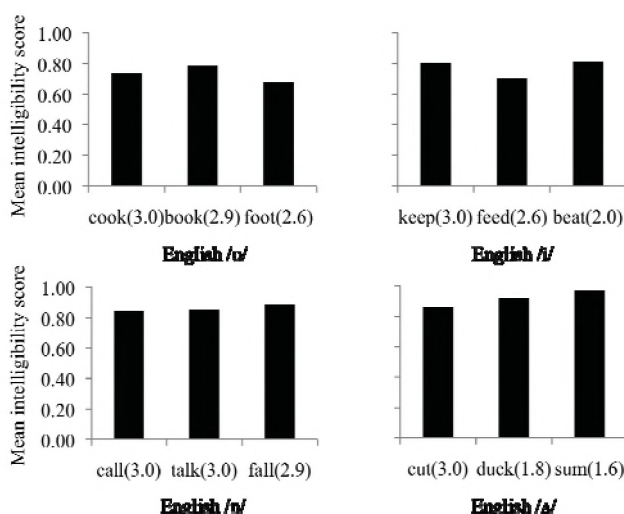


Figure 2. Mean intelligibility scores for English vowels in words that did not follow the rank order predicted by mean familiarity scores (indicated in parentheses). Results are pooled across Mandarin and Slavic speakers.

Mean intelligibility scores for individual words (see Figure 1) provide more specific evidence for the role of lexical familiarity. For six of ten English vowels, the mean intelligibility score decreases as lexical familiarity decreases. In the remaining four cases (see Figure 2), individual items do not follow the predicted pattern. However, if the Mandarin productions of “cook”, “book” and “foot” are examined in isolation, rather than pooled with Slavic productions, intelligibility scores do follow the predicted pattern. Conversely, the Slavic productions of “feed”, “keep” and “beat” examined in isolation also follow the predicted pattern with respect to lexical familiarity.

4. DISCUSSION

The results of this study provide strong evidence that lexical familiarity predicts the intelligibility of L2 phonemes, regardless of the learners’ L1. In the majority of cases, vowels from a single English category were produced less intelligibly when they occurred in less familiar words. Although there was no significant L1 effect, differences between L1 groups in two cases suggest there may be differences with respect to specific English vowel categories, providing some evidence for Flege’s [1] claims regarding L1 transfer being context-specific. The results of this study also indicate that the provision of orthographic information can have a facilitative effect on the intelligibility of L2 speech. Taken together, these findings suggest that L2 speech emerges at the level of lexically conditioned allophones, not as entire categories. Rapid access of semantic and/or lexical information in more familiar words may facilitate greater attention to phonetic form. Reference to orthographic representations, assuming they are relatively transparent as was the case in this study, may also allow for more rapid access of semantic and/or lexical information, and also facilitate attention to form.

REFERENCES

[1] 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.

[2] Munro, M. J. (2008). Variability in Cantonese speakers’ productions of English vowels. *Canadian Acoustics*, 36 (3), 140-141.

[3] Trofimovich, P., Gatbonton, E., & Segalowitz, N. (2007). A dynamic look at L2 phonological learning: Seeking processing explanations for implicational phenomena. *Studies in Second Language Acquisition*, 29(3), 407-448.

ACKNOWLEDGEMENTS

This project was funded by SSHRC and a Brock University, Dept. of Applied Linguistics Research Grant. The authors thank Murray Munro for his input.

VOWEL IMITATION IN SPONTANEOUS PHONETIC IMITATION

Molly Babel

Dept. of Ling, University of British Columbia, Totem Fields Studio, 2613 West Mall, Vancouver, BC V6T 1Z4,
mbabel@berkeley.edu

1. INTRODUCTION

Research on phonetic imitation or phonetic convergence has established that talkers naturally accommodate during interaction (Goldinger 1998, Namy et al. 2002, Pardo 2006). Traditionally, the method of determining imitation has been an AXB judgment task where listeners judge the perceptual similarity of a participant's A and B tokens from a pre-task production and shadowed production, for example, to X – a production from a model talker in the task. At better than chance levels, listeners judge the shadowed production to be more similar to the model talker's production. This method allows listeners to make use of an array of perceptual cues for imitation. It is left unknown, however, what within the phonetic structure is being imitated. What is and what can be imitated have significant implications for the level of detail in linguistic representation. More recently, work by Shockley et al. (2004) and Nielsen (2008) has demonstrated that American English participants imitate VOT. The purpose of the current project is to examine whether spectral characteristics of the vowels are imitated in a lexical shadowing task.

2. METHOD

Fifty monosyllabic low frequency words with the vowels /i æ ɔ o u/ were used as stimuli in a lexical shadowing task. Native speakers of American English ($n = 113$) completed an auditory naming task where they shadowed productions from one of two model talkers. One talker was Black and the other was White.

Participants in the shadowing task were tested individually and were randomly assigned to one of four conditions: Black talker/No Picture, Black talker/Picture, White talker/No Picture, and White talker/Picture. The paradigm for the speech production task is a shadowing paradigm like that of Goldinger (1998). Participants were seated in a sound-attenuated room at a computer workstation. Participants wore a head-mounted AKG C520 microphone positioned about 2 inches to the side of the mouth and AKG K240 headphones. Word productions were digitally recorded to the hard drive of a PC at a 44K sampling rate. Recordings were down sampled to 22K before analysis.

2.1 Procedure

The task proceeded as follows: The first block was a Pre-task Block to establish participants' baseline productions. The words were presented randomly in 36-point font in the middle of the screen. Participants were instructed to read the words as naturally and clearly as possible. In the test blocks, the randomized word list was presented binaurally at 65 dB

(SPL) over the headphones. The test blocks were comprised of three shadowing blocks where words were repeated twice per block. Participants were told that upon hearing the word, they were to repeat it as clearly and naturally as possible. In the Picture Conditions a talker photo was presented on the screen for the duration of the shadowing portion of the task. The post-test block was identical to the pre-test block where participants read the words from the screen. After the task participants in the No Picture conditions were asked to identify the race of the talker. Both talkers were identified as White by participants ($\chi(1) = 0.08, p = n.s.$).

2.2 Data Analysis

First and second formants were extracted from word productions from a series of Gaussian windows spanning the middle 50% of the vowel with a 2.5 ms step size. Formant values were normalized using the Lobanov normalization routine (Lobanov 1971).

The Euclidean distance was calculated from each participant production to that of the model talker. These calculations are a measure of acoustic distance between the model talkers' productions and the participants' productions. To calculate how much a participant modified their production as a result of being exposed to the model talker the original distance for each word was subtracted from the distance for each following instance of that word. The value calculated is the difference in distance (DID). A negative DID value demonstrates that the phonetic distance between the participant and the model talker shrank. A positive value indicates an increase in phonetic distance. A value of 0 demonstrates that there was no change as the result of auditory exposure to the model talker. This DID value is used as the dependent measure in the statistical analysis.

3. RESULTS

DID values were summarized across cells and the means were used in a repeated measures analysis of variance. DID was the dependent variable; Talker Race, Picture/No Picture, and Gender were independent variables; and Vowel and Block were repeated measures. Vowel [$F(4, 396) = 56.2, p < 0.001$] and Block [$F(3, 297) = 60.7, p < 0.001$] returned as main effects. There was also a two-way Vowel x Block interaction [$F(12, 1188) = 20.2, p < 0.001$] and a three-way Vowel x Block x Talker Race interaction [$F(12, 1188) = 3.8, p < 0.001$]. There were also two-way interactions of Vowel x Gender [$F(4, 396) = 2.4, p < 0.05$], Talker Race x Vowel [$F(4, 396) = 8.3, p < 0.001$], and Picture x Gender [$F(1, 99) = 4, p < 0.05$]. The three-way

interaction between Picture x Vowel x Gender was also significant [$F(4, 396) = 5, p < 0.001$].

Figure 1 shows the effect of selective imitation for low vowels. Post-hoc tests show /æ/ and /ɔ:/ are imitated more than /i o u/ ($p < 0.001$). In addition, /æ/ is imitated more than /ɔ:/ ($p < 0.05$). With respect to the Block effect, post-hoc tests find that imitation is cumulative across shadowing blocks. There was more imitation in Block 5 than 4 ($p < 0.05$) and more in Block 6 than 4 ($p < 0.001$). During the shadowing task, there was more accommodation than the post-task block ($p < 0.001$). Normalized formant plots from male and female participants are shown in Figure 2. These figures indicate that the majority of vowel imitation comes from changes within the F1 dimension.

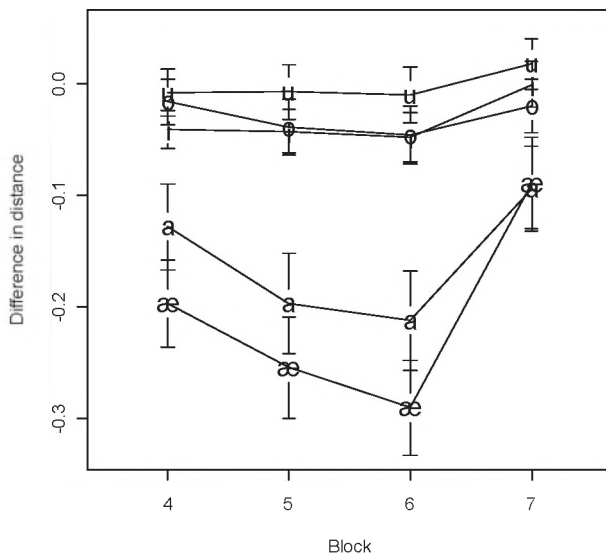


Figure 1. Spontaneous phonetic imitation for all participants by Vowel and Block. The DID measure on the y-axis indicates the amount of phonetic imitation. A negative value demonstrates phonetic imitation and a positive value demonstrates vocalic divergence. Blocks 4, 5, and 6 are Shadowing Blocks while Block 7 is the Post-task Block..

4. DISCUSSION

The results show that talkers accommodate the first and second formants of the model talker in the task, but that not all vowels are imitated to a significant degree. Only the low vowels /æ/ and /æ/ exhibit strong imitation effects, and this effect lies primarily within the F1 dimension. I argue that this is due to the increased repertoire of production variants talkers store for low vowels as a result of the difference in jaw height in accented and unaccented environments (Summers 1987, de Jong 1995).

An important aspect of this result is that it demonstrates the labile nature of linguistic segments with respect to both their

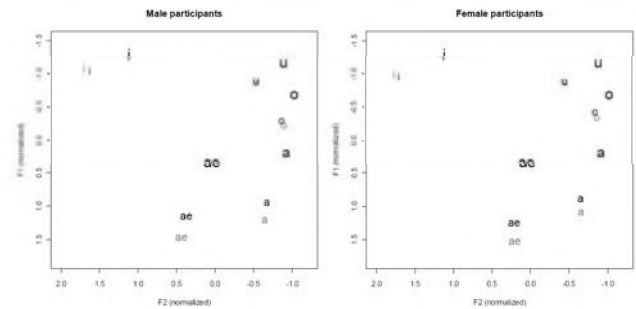


Figure 2. Formant plot displaying the direction of spontaneous phonetic imitation. The mean of the model talkers' vowels are in the slightly larger font and in black. Participants' pre-task vowel means are plotted in a smaller light gray font and their productions from the final shadowing block are in the small black print.

perceptual encoding and their variation in production. First, listeners must perceive the detailed acoustic structure of an utterance in order to have those details influence their production. Second, in speech production, participants alter the characteristics of the output without modifying the categorical identity of the segment they produce. In sum, the exact selection of a speech production variant is determined by auditory exposure.

REFERENCES

- de Jong, K. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *JASA* 97:491-504.
- Goldinger, S. (1998). Echoes of echoes? An episodic theory of lexical access. *Psych. Review*, 105:251-279.
- Lobanov, B. (1971) Classification of Russian vowels spoken by different listeners. *JASA*, 49: 606-08
- Namy, L., L. Nygaard, & D. Sauterberg. (2002). Gender differences in vocal accommodation: the role of perception. *J of Lang & Soc Psych.*, 21:422-432.
- Nielsen, Kuniko. (2008). *The specificity of allophonic variability and its implications for accounts of speech perception*. Doctoral Dissertation, UCLA.
- Pardo, J.S. (2006). On phonetic convergence during conversational interaction. *JASA*, 119:2382-2393.
- Shockley, K., L. Sabadini, & C.A. Fowler. (2004). Imitation in shadowing words. *P&P*, 66:422-429.
- Summers, W.V. (1987). Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses. *JASA* 82:847- 863.

ACKNOWLEDGEMENTS

Thanks to Keith Johnson, Ben Munson, Andrew Garrett, Dasha Bulatov, and Tyler Frawley. This project was supported by the Center for Race and Gender at UC Berkeley and the Abigail Hodgen Publication Fund. This project was conducted as part of the author's Ph.D. dissertation at the University of California, Berkeley.

ACOUSTIC AND AUDITORY COMPARISONS OF POLISH AND TAIWANESE MANDARIN SIBILANTS

Chenhao Chiu

Dept. of Linguistics, University of British Columbia, 2613 West Mall, Vancouver, BC, Canada V6T 1Z4

1. INTRODUCTION

Polish and Mandarin Chinese are two languages well known for the three-way distinction of sibilant fricatives [1]: apical dental [s], laminal flat post-alveolar (retroflex)^a [ʂ], and laminal palatalized post-alveolar (alveolo-palatal) [ʃ]. The spectral measurements of English, Japanese and Mandarin sibilants in Li, Edwards, and Beckman [2] show that sounds with the same IPA transcription can be phonetically different and the differences are dependent on what contrasts are present in the languages inventories. McGuire (2007) [3] points out that Mandarin speakers are better than American English speakers at using frication noise and formant transition as cues to distinguish Polish retroflex and alveolo-palatal sounds. The present study compares the acoustic properties of the three sibilants in Polish (PL) and Taiwanese Mandarin (TM) and discusses the differences in terms of frication noise and the formant transitions. Aside from the auditory and articulatory similarities described in Ladefoged and Maddieson [1], this study will try to answer two main questions: (1) Are the sibilant contrasts in PL and TM the same? If not, how are they acoustically different? (2) What acoustic cues and properties distinguish one sibilant from another in these two languages? This study provides acoustic data, including moments, frequencies, formant transitions and LPC spectra showing that there exist the contrasts of the sibilant types in the two languages.

2. METHOD

2.1 Materials

The dental sibilant and the alveolo-palatal are in complementary distribution in both languages. The alveolo-palatal [ʃ] only occurs before a high front vowel ([i] for PL; [i] and [y] for TM) or a glide ([j] for PL; [j] and [ɰ] for TM) whereas the dental sibilant [s] occurs elsewhere. Both PL and TM retroflex [ʂ] occur in the same environment as the dental sibilant [s]. With these phonological conditions considered, the material environments for [s] and [ʂ] and for [ʃ] are designed to be different. In this study, the materials are real words^b starting with a sequence of sibilant + vowel. [s] and [ʂ] are followed by a low back vowel [a] (PL: [sabʂt] “sabot”, [ʂabla] “saber”; TM: [sa] “to sprinkle”, [ʂa] “to kill”). [ʃ] precedes a glide [j] followed by a low back vowel [a] (PL: [ʃjadaʂ] “to sit”; TM: [ʃja] “blind”). Fifteen tokens of each sibilant type were recorded by one male PL native speaker and two male TM native speakers.

^a The present study follows Ladefoged and Maddieson (1996) in assuming Polish and Mandarin retroflexes to be laminal post-alveolar.

^b Since TM does not have coda consonants (except for the alveolar nasals and velar nasals), only monosyllabic words are recorded.

2.2 Analysis preparation

For each token, the analyses included the centre of gravity (CoG) and F2 transitions of the sibilants along with the dynamic amplitude and frequencies of the maximum amplitude, as applied in Jesus and Shadle [4]. A Hanning window of 100ms from the midpoint of each token was selected to produce LPC spectra. The recording sampling rate was 44.1 kHz. The analysis window length was 25ms. The second formant (F2) of the following vowel was measured in the beginning 25ms and the middle 25ms of the vowel (Hanning windowed). The measurements of the dynamic amplitude and the frequency of the maximum amplitude follow Jesus and Shadle [4]. The dynamic amplitude is the difference between the maximum amplitude between 500 Hz – 20K Hz and the minimum amplitude between 0 and the maximum amplitude. The frequencies of the maximum amplitude were also measured. All the measurements were performed with the phonetic software Praat.

3. RESULTS

3.1 Acoustic analyses^c

Table 1 gives the mean CoG values and the frequencies of the maximum amplitudes of sibilants. For both languages, a main effect of sibilant types is found in CoG (PL: $F(2, 42) = 299.645, p < .001$; TM: $F(2, 42) = 54.056, p < .001$). Pairwise t-tests within places of articulation also reach significance (for [s] pair: $p = .007$; [ʂ]: $p < .001$; [ʃ]: $p < .001$). For frequencies of the maximum amplitudes, substantial differences are found in the retroflex pair and the alveolo-palatal pair (both $p < .001$). For both the CoG and maximum frequencies, within-language analyses show that in PL while dentals are different from retroflexes and alveolo-palatals, retroflexes and alveolo-palatals are not significantly different whereas all TM sibilants are significantly different.

Table 1. Mean CoG and frequency of the maximum amplitude

Context	Center of gravity (Hz)		Max Freq. (Hz)	
	PL	TM	PL	TM
[s]	8442.07	7734.15	8354.03	8079.8
[ʂ]	2287.03	5598.7	2811.74	4597.38
[ʃ]	2788.7	6288.55	3391.62	6098.49

When comparing F2 transitions (as shown in Table 2), a significant difference between PL and TM is found only in the retroflex pair ($p < .001$). Post hoc analyses found

^c All the moment values and frequencies from the two TM speakers are highly correlated, except for the F2 transitions. In present study, the average values of the two speakers are used for analyses.

substantial differences between PL dentals and alveolo-palatals ($p < .001$) as well as among all three sibilants in TM (all $p < .001$). The dynamic amplitudes of the dental pair and alveolo-palatal pair between the two languages are significantly different (both $p < .001$), but not the retroflex pair ($p = 0.121$). In addition, main effects of sibilant types are also found in both languages (PL: $p = .002$; TM: $p < .001$). Post hoc tests show that in terms of dynamic amplitude, PL [s] and TM [ʂ] are significantly different from their counterparts.

Table 2. Mean F2 transition and dynamic amplitude

Context	F2 transition (Hz)		Dynamic amplitude (dB)	
	PL	TM	PL	TM
[s]	58.93	64.18	38.86	29.93
[ʂ]	71.12	170.05	26.65	32.2
[ʐ]	357.44	337	28.47	38.08

The author observed lip protrusion in PL [ʂ] and [ʐ]. For TM, lip protrusion is optional for [ʂ] and is obligatory for [ʐ] only when it precedes [y]. Another fifteen tokens of TM [ʐy] were recorded for comparison (CoG: 5756.67 Hz; Max. Freq.: 5010.21 Hz; dynamic amplitude: 35.8 dB). The acoustic properties of PL [ʂ] and TM rounded [ʐ] are significantly different (CoG: $p < .001$; Max. Freq.: $p = .001$; dynamic amplitude: $p = .002$).

3.2 Spectra comparison

Among all the sibilant types, the retroflex pair in the two languages is the most distinctive. Their CoG, maximum frequencies, F2 transitions, and dynamic amplitude are significantly different. Nevertheless, the LPC spectra show that they are acoustically different (Fig. 1), but auditorily similar (Fig. 2).

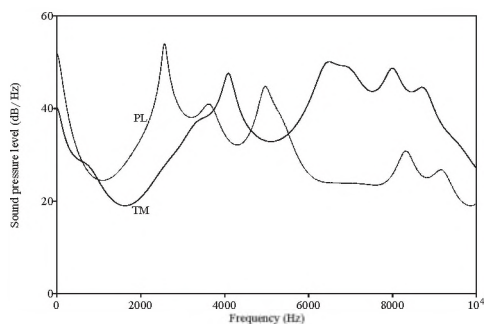


Figure 1. Acoustic LPC Spectrum of PL [ʂ] and TM [ʐ]

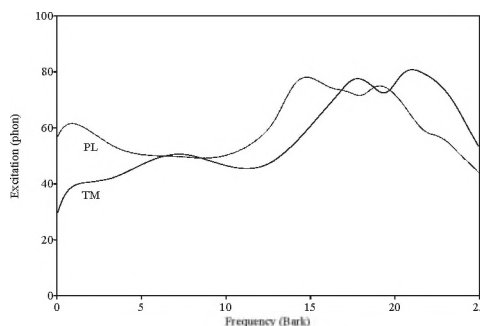


Figure 2. Auditory LPC Spectrum of PL [ʂ] and TM [ʐ]

4. DISCUSSION

Despite the fact that dental sibilants in the two languages are quite similar with respect to CoG and frequency of the maximum amplitude, the acoustic properties are rather different in PL and TM sibilant contrasts, even though the same IPA transcriptions are used for both languages. For both languages, the CoG and frequencies of the maximum amplitude are used to distinguish the anterior sibilant from the other two types and their cross-language differences are yet substantial. The lip protrusion contributes to the F2 lowering and the defining characteristic of stronger spectral energy at lower frequency ranges for PL [ʂ] and [ʐ], distinct from TM rounded [ʐ] and unrounded [ʂ]. These gestural variations seem influential to the acoustics of sibilants. Ladefoged and Maddieson [1] point out the auditory similarities of sibilants in these two languages. However, the acoustic effect of F2 transition stronger in TM than in PL. Although frication noise and the formant transitions are the most salient acoustic contrasts in PL (consistent with the findings in [5]) and TM, whether or not these acoustic cues are determinative to the perception requires more empirical evidence.

REFERENCES

- [1] Ladefoged, P. and Maddieson, I. (1996). *The Sounds of the World's Languages*. Blackwell. Oxford.
- [2] Li, F., Edwards, J., and Beckman, M. (2007). Spectral measures for sibilant fricatives of English, Japanese and Mandarin Chinese. *Proceedings of the XVth International Congress of Phonetic Sciences*, pp. 917 – 920.
- [3] McGuire, G. (2007). Phonetic cue and integrality. *JASA*. 122, 5, Nov., 2971.
- [4] Jesus, L. M. T. and Shadle, C. H. (2002). A parametric study of the spectral characteristics of European Portuguese fricatives. *Journal of Phonetics*, 30, 437 – 464.
- [5] Padgett, J. and Zygis, M. (2006). *A Perceptual Study of Polish Fricatives, and its Relations to Historical Sound Change*. Poster presented at 10th Conference on Laboratory Phonology. Paris, 29.06 – 01.07.2006.

ACKNOWLEDGEMENTS

Special thanks to Amanda Miller and Molly Babel for comments and suggestions, and to Carrie Cutler and Takako Kawasaki for their assistance on statistic analyses.

TALKING WHILE CHEWING: SPEAKER RESPONSE TO NATURAL PERTURBATION OF SPEECH

Connor Mayer¹, Bryan Gick^{1,2}, and Elizabeth Ferch¹

¹Dept. of Linguistics, University of British Columbia, 2613 West Mall, Vancouver, BC, Canada, V6T 1Z4

²Haskins Laboratories, 300 George Street, Suite 900, New Haven, Connecticut, USA, 06511

1. INTRODUCTION

Previous studies of speech motor control have employed various types of mechanical perturbations to investigate the goals of speech production. While the types of perturbations used have varied in terms of being static or dynamic, all have been under external control. Studies of static perturbations have used artificial palates (McFarland, Baum, & Chabot 1996; Aasland, Baum, & McFarland 2006), bite blocks (McFarland & Baum 1995; McFarland, Baum, & Chabot 1996), and dental prostheses (Hamlet, Cullison, & Stone 1979; Jones & Munhall 2002); studies of dynamic perturbations have employed loads on the lower lip or jaw (Abbs & Gracco 1984; Munhall, Löfqvist, & Kelso 1994), horizontal force applied to the jaw (Tremblay, Shiller, & Ostry 2003), and changes in the thickness of an artificial palate (Honda, Fujino & Kaburagi 2002).

These studies have collectively found articulatory, acoustic, and perceptual effects on speech production that underscore both the spatial-motor and acoustic-auditory goals of speech. Tremblay et al. (2003) find that compensation to displacement of the jaw occurs despite the absence of acoustic or perceptual effects, suggesting that “a somatosensory goal is pursued independent of the acoustics” (868). Others, such as Honda et al. (2002) and Jones and Munhall (2002), find evidence that auditory feedback is important, suggesting an acoustic goal; this is consistent with clinical findings and studies of properties such as pitch and vowel quality (Jones & Munhall 2002).

The present study examines the acoustic and articulatory effects of chewing during speech. Chewing, viewed as articulatory perturbation, is a variety of dynamic perturbation that is different from previous approaches in several important ways: It is naturalistic, experienced every day by speakers; it is under speakers’ control; it changes over time, requiring constant readjustment; and it interferes with both the movement of the articulators and the shape of resonating cavities in the mouth.

We propose that speakers in control of this type of highly complex articulatory perturbation during speech will show evidence of optimizing to maintain acoustic-auditory speech goals. Specifically, we expect acoustic distinctions between sibilants to be maintained even when perturbation forces significant articulatory differences in tongue shape.

2. METHODS

Subjects were seated in a modified dentist's chair. An Aloka Prosound SSD-5000 ultrasound machine with a 180 degree probe was used to record midsagittal images of the tongue. Profile video was taken using a Sony Mini-DV Handicam. Subjects wore a pair of sunglasses; two sticks covered in blue construction paper, each with two pink dots affixed, were attached to the sunglasses and to the probe. These dots showed the position of the head relative to the probe, allowing for correction of head movement once tongue shapes had been traced (Mielke, Baker, Archangeli & Racy 2005). The chromakey feature of a Videonics MXPro DV video mixer was used to superimpose the subject's face and the dot positions over the ultrasound video. Audio was recorded using a Sennheiser MK66 short shotgun microphone. The video and audio signals were fed through a Canopus ADVC-110 advanced digital video recorder into a MacPro computer, where the combined video and audio were captured using iMovie.

The stimuli consisted of the carrier phrase "I'm a ___", followed by one of three words containing the phonemes of interest: 'saw', 'shaw', or 'raw'. For the present paper, only the sibilants /s/ and /ʃ/ are considered for comparison. There were two conditions: an experimental condition, in which speakers produced the stimuli while chewing a large bolus consisting of four pieces of Wrigley's™ spearmint gum, and a no-gum control condition. Stimuli were presented in four blocks for each condition, with each block containing four repetitions of each word, resulting in 32 tokens for each word (16 in each condition). These stimuli were produced by 7 native speakers of English. The boundaries and midpoints of /s/ and /ʃ/ were marked using Praat. Acoustic centre of gravity (COG) measurements were made with a 30ms window around the midpoints. Still frames were extracted from the midpoints using ELAN. Palatoglossatron was used to trace tongue shapes and align the tracings.

3. RESULTS

S2, S3, and S4 displayed no significant differences in COG between the two conditions, while S1, S5, and S6 did display significant differences. The remaining subject, S7, displayed significant differences in /s/ but not /ʃ/. These results are summarized in Table 1. Of the three subjects who had significant differences between the two conditions, S6

had COG frequencies that were significantly higher in the with-gum condition, while those of S1 and S5 were significantly lower. Although absolute acoustic targets were compromised by the presence of gum for these three subjects, the relative distances between /s/ and /□/ were maintained. S7 displayed lower COG frequencies for both conditions, but although a distinction was still preserved between /s/ and /□/, relative acoustic distances were not as closely maintained. The relative distances between /s/ and /□/ for subjects with significantly different COG frequencies are shown in table 2. A 2-way factorial ANOVA found no significant interaction effects in any subject.

Table 1. Average COG for bolus (B) and no bolus (NB) conditions with p values. Significant results are bolded.

	/s/ NB	/s/ B	/s/ p	/□/ NB	/□/ B	/□/ p
S1	8863 Hz	8138 Hz	< 0.01	6387 Hz	5717 Hz	0.04
S2	7088 Hz	6763 Hz	0.42	4889 Hz	4920 Hz	0.9
S3	8422 Hz	8160 Hz	0.32	4758 Hz	4628 Hz	0.49
S4	7825 Hz	7546 Hz	0.14	6787 Hz	6643 Hz	0.62
S5	8242 Hz	7643 Hz	0.03	5895 Hz	5454 Hz	0.01
S6	6984 Hz	7731 Hz	< 0.01	5352 Hz	5687 Hz	0.047
S7	7106 Hz	6308 Hz	0.02	4581 Hz	4404 Hz	0.43

SS ANOVA tests were performed on the ultrasound tongue tracings (Davidson 2006). For /s/ and /□/ for all subjects, interaction plots show significant interaction effects with BCI for the no bolus condition. Only S3 /□/ and S4 /□/ had significant interaction effects with BCI for the bolus condition. This is due to increased variability of tongue shape in the bolus condition across subjects. The location of the differences varies somewhat across speakers, indicating different strategies for handling the gum during speech, but tends to be primarily in the blade and root of the tongue.

Table 2. Average COG frequency ratios of /s/ to /sh/.

	NB Condition	B Condition
S1	1.39	1.42
S5	1.40	1.40
S6	1.31	1.36
S7	1.43	1.55

4. DISCUSSION

For all subjects in our study, the presence of a large gum bolus in the mouth interfered with the shapes and movements of the articulators during speech. Despite this perturbation, all subjects maintained the relative acoustic distance between /s/ and /□/ in terms of centre of gravity. In the cases where a speaker's articulatory strategies were not sufficient to maintain normal acoustic targets, relative acoustic targets were still met. S7 provides an interesting case, as her strategy for producing /□/ was sufficient to maintain acoustic targets while her strategy for /s/ was not.

This is not altogether surprising, as the more anterior articulation of /s/ makes it more susceptible to interference from the gum, and adaptation more difficult. As well, /s/ may require more precise tongue control in general (Stone, Faber, Raphael, & Shawker 1992). That all speakers displayed significant differences in tongue shape between conditions strongly suggests that speakers were employing articulatory-acoustic tradeoffs (Guenther, Espy-Wilson, Boyce, Matthies, Zandipour, & Perkell 1999), adjusting their chewing strategies to optimize acoustic-auditory targets. These results show that speakers in control of their own articulatory perturbations – even ones that are highly complex and subject to constant change – adjust their strategies to maintain the acoustic goals of speech.

REFERENCES

- Aasland, W., Baum, S. & McFarland, D. (2006). Electropalatographic, acoustic, and perceptual data on adaptation to a palatal perturbation. *J. Acoust. Soc. Am.*, 119, 2372-2381.
- Abbs, J.H., & Gracco, V.L. (1984). Control of complex motor gestures: Orofacial muscle responses to load perturbations of the lip during speech. *J. Neurophys.*, 51(4), 705-723.
- Davidson, L. (2006). Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *J. Acoust. Soc. Am.*, 120(1), 407-415.
- Guenther, F. H., Espy-Wilson, C.Y., Boyce, S.E., Matthies, M.L., Zandipour, M. & Perkell, J.S.. (1999). Articulatory tradeoffs reduce acoustic variability during American English /t/ production. *J. Acoust. Soc. Am.*, 105, 2854-65.
- Hamlet, S.L., Cullison, B.L. and Stone, M.L. (1979). Physiological control of sibilant duration: Insights afforded by speech compensation to dental prostheses. *J. Acoust. Soc. Am.*, 65, 1276-1285.
- Honda, M., Fujino, A. and Kaburagi, T. (2002). Compensatory responses of articulators to unexpected perturbation of the palate shape. *Journal of Phonetics*, 30, 281-302.
- Jones, J. and Munhall, K. (2002). Learning to produce speech with an altered vocal tract: The role of auditory feedback. *J. Acoust. Soc. Am.*, 113, 532-543.
- McFarland, D. and Baum, S. (1995). Incomplete compensation to articulatory perturbation. *J. Acoust. Soc. Am.*, 97, 1865-1873.
- McFarland, D., Baum, S. and Chabot, C. (1996). Speech compensation to structural modifications of the oral cavity. *J. Acoust. Soc. Am.*, 100, 1093-1104.
- Mielke, J., Baker, A., Archangeli, D., & Racy, S. (2005). Palatron: a technique for aligning ultrasound images of the tongue and palate. *Coyote Papers*, 14, 147-161.
- Munhall, K.G., Löfqvist, A. and Kelso, J.A.S. (1994). Lip-larynx coordination in speech: Effects of mechanical perturbations to the lower lip. *J. Acoust. Soc. Am.*, 95, 3605-3616.
- Stone, M., Faber, A., Raphael, L.J. & Shawker, T.H. (1992). Cross-sectional tongue shape and linguopalatal contact patterns in [s], [□], and [l]. *Journal of Phonetics*, 20, 253-270.
- Tremblay, S., Shiller, D. M. & Ostry, D. (2003). Somatosensory basis of speech production. *Nature*, 423, 866-869.

ACKNOWLEDGEMENTS

Thanks to Donald Derrick, Amanda Miller, Cathy Lin, Denise Tom, and the members of the UBC ISRL.

MAGNETIC RESONANCE IMAGING AND VIDEOFLUOROSCOPY OF PARTIAL GLOSSECTOMEES' SPEECH: PRELIMINARY RESULTS

Janette Quintero¹, Tim Bressmann¹, Katalin Mady² and Ambros Beer³

¹Dept. of Speech-Language Pathology, University of Toronto, 500 University Ave., Ontario, Canada, M5G 1V7

²Institute for Phonetics and Speech Communication, Ludwig-Maximilians University of Munich, 80799 Munich, Germany

³Dept. of Nuclear Medicine, Krankenhaus Rechts der Isar, Munich University of Technology, Germany

Janette.quintero@utoronto.ca

1. INTRODUCTION

A partial glossectomy results in a lingual defect that can have detrimental effects on speech. The speech defects may negatively impact patients' quality of life and well-being. The biomechanical consequences of a partial tongue resection are still not well understood (Bressmann et al., 2004).

In a recent study using ultrasound imaging, Rastadmehr et al. (2008) found that there was an increase in tongue velocity postoperatively, in comparison to tongue velocity measures before the surgery and to tongue velocity measures in normal controls. This finding challenges the assumption that a tongue resection will tether the tongue (Korpjaako-Huuhka et al. 1999).

One of the goals of this study was to replicate Rastadmehr et al.'s (2008) findings, using cine-MRI and videofluoroscopy. The main research objective of this study was to determine the effects of a partial glossectomy on tongue velocity and movement range during speech.

2. METHOD

2.1 Subjects

The six patients in this study were diagnosed with anterior and/ or lateral carcinomas of the tongue. All subjects were male, with a mean age of 55.8 years (SD 12.0 years). They were recruited from the Department of Oral and Maxillofacial Surgery of the Technical University of Munich, Germany. The diameters of the tumours ranged from <1 to 4 cm. The smaller defects were closed using either primary wound healing or local closures. The larger defects were closed using platysma flaps.

2.2 Materials and Procedures

The imaging data sets were originally collected by Mady and Beer. The videofluoroscopies were recorded to VHS video in PAL format with a frame rate of 25 fps. The movies were later digitized for the quantitative analysis. Patients swallowed barium to display the tongue and hypopharynx. The patients repeated a list of ten words three times. Audio recordings were made simultaneously. For each participant, tongue velocity and movement range were calculated for each word production.

The MRI data were recorded using a T1 echo gradient sequence on Phillips ACS NT Gyroscan (TR=4.0ms, TE=1.1 ms, midsagittal slice thickness=10mm), with a frame rate of 8 fps. The movies were converted to .avi format for analyses. The patients repeated each of the ten words for 10 s, resulting in approximately 4-6 repetitions. Audio data was not concurrently recorded. For each patient, tongue velocity and movement range were calculated over the entire 10 s word repetition task, rather than for each word production.

3. DATA ANALYSES

The speakers' tongue movement was analyzed using the Ultrasonographic Contour Analyzer for Tongue Surfaces (Ultra-CATS) software, which permits a frame-by-frame analysis of the recordings. For each frame, the surface of the tongue and the outer contour of the mandible were traced. The program uses a semi-circular grid to make measurements. Figure 1 displays tongue movement over several frames during the production of the word <Theke> 'bar'. Figure 2 shows the corresponding surface plot of the tongue movement observed in Figure 4.

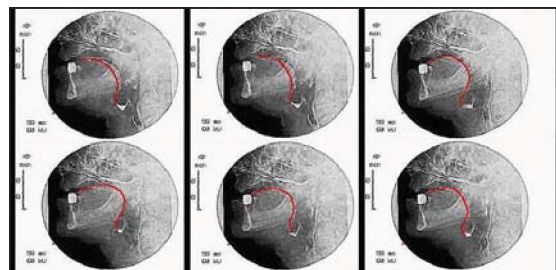


Figure 1: Sequence of videofluoroscopic frames shows tongue movement for the German word <theke> (bar).

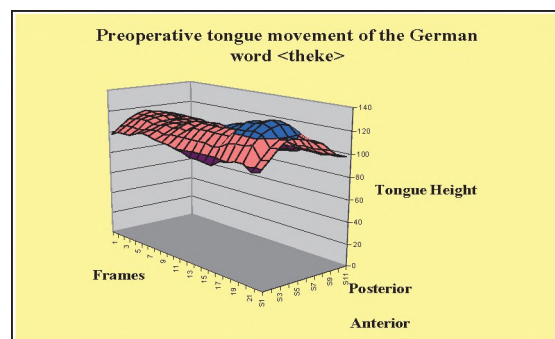


Figure 2: Three-dimensional chart displaying the tongue movement for the videofluoroscopic frame sequence in Figure 1.

For each participant, tongue velocity was calculated by dividing the total tongue distance traveled at various angles along the surface of the tongue by the time lapsed. The difference between the minimum and maximum tongue height was calculated for each angle, giving the tongue movement range for each angle of the measurement grid.

4. RESULTS

Since the project is still in its early stages, only one of the nine words has been analyzed. Principle Component Analyses were used to identify functional segments of the tongue. The analysis revealed three components that corresponded with the anterior, centre, and posterior regions of the tongue. Statistical analyses were limited to the three functional segments.

Multiple paired sample *t*-tests were conducted to compare the preoperative and postoperative tongue velocity measures at the different regions of the tongue. Since there were only a small number of paired comparisons, no Bonferroni-adjustment was made. Due to the small sample size, an alpha level of $p < 0.1$ was set as the criterion for statistical significance. The results suggest that tongue velocity decreased after the surgery in all tongue regions (anterior, centre, and posterior). The decrease in tongue velocity was only significant for the posterior tongue ($t = 2.871$, $df=5$, $p < 0.05$). All patients showed a postoperative deceleration.

There was an increase in the postoperative tongue movement range in the anterior region of the tongue. However, this increase was not significant. The preoperative and postoperative movement range were almost the same in the centre tongue region. The movement range in the posterior region decreased significantly after the surgery ($t=2.158$, $df=5$, $p < 0.1$).

The preliminary results from the MRI data confirm the findings from the videofluoroscopic data. Qualitatively, patients showed a decrease in postoperative tongue velocity in all three functional tongue segments (see Figure 3). Also, there was a reduction in tongue movement range in the anterior and posterior tongue segments, following the partial glossectomy. There was a slight increase in the centre tongue region (see Figure 4).

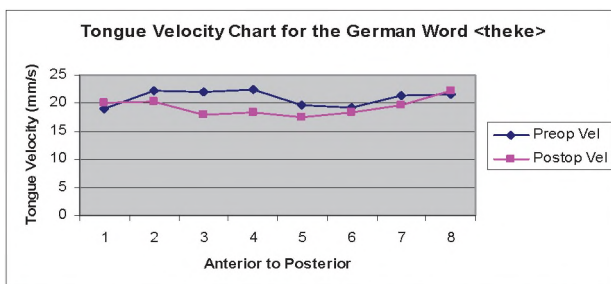


Figure 3: Preoperative and postoperative tongue velocity at various points along the surface of the tongue for patient 6.

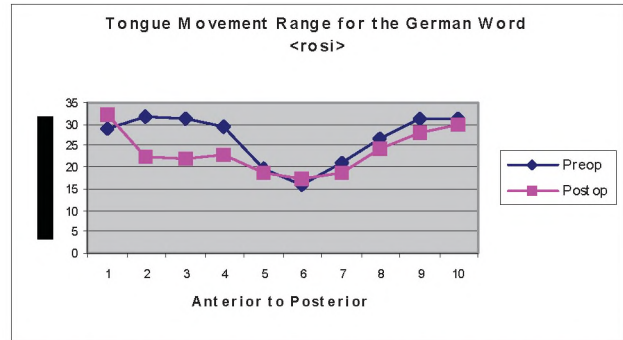


Figure 4: Preoperative and postoperative tongue movement range at various points along the surface of the tongue for patient 3.

5. DISCUSSION

The preliminary results suggest that tongue velocity and tongue movement range decreased after a partial glossectomy surgery. These findings confirm previous assumptions (Korpijaako-Huuhka et al. 1999) that a loss of lingual tissue results in reduced tongue movement and speed. A reduction in tongue speed and movement range could be the cause of the observed speech distortions.

The results from this study contradict the findings by Rastadmehr et al. (2008), who found an increase in post-surgical tongue height and velocity. Further analysis will demonstrate whether the patients in the present study use fundamentally different compensatory strategies. There was also an obvious difference in the speech tasks: Rastadmehr et al. (2008) used a complex reading passage, while the current study investigated the articulation of single words. Finally, the different findings may also be attributable to the differences in patients and glossectomy surgeries.

REFERENCES

- Bressmann, T., Sader, R., Whitehill, T., & Samman, N. (2004). Consonant Intelligibility and Tongue Motility in Patients with Partial Glossectomy. *Journal of Maxillofacial Surgery*, 62, 298-303.
- Korpijaako-Huuhka, A.M., Söderholm, A.L., & Lehtihalmes, M. (1999). Long-lasting speech and oral-motor deficiencies following oral cancer surgery: a retrospective study. *Logopaedics, Phoniatrics, and Vocology*, 24, 3, 97-106.
- Rastadmehr, O., Bressmann, T., Smyth, R., & Irish, J. (2008). Increased midsagittal tongue velocity as indication of articulatory compensation in patients with lateral partial glossectomies. *Otolaryngology-Head and Neck Surgery*, 30, 718-726.

ACKNOWLEDGEMENTS

Janette Quintero would like to acknowledge the valuable support and advice of Dr. Catriona Steele and Dr. Alexei Kochetov.

THE ROBUSTNESS OF INFANTS' EARLY WORD REPRESENTATIONS

Marieke van Heugten and Elizabeth K. Johnson

Dept. of Psychology, University of Toronto, 3359 Mississauga Road N., Mississauga, ON, L5L 1C6, Canada
marieke.vanheugten@utoronto.ca

1. INTRODUCTION

The speech stream contains no one-to-one mapping between acoustic form and lexical entry. Indeed, factors such as idiolect and speaking rate ensure that no word is ever produced in an identical fashion twice. Nevertheless, listeners effortlessly understand spoken language. In this paper, we explore how infants, who have yet to develop a mental lexicon, start recognizing acoustically distinct realizations of words.

Although infants can cope with some natural speech variation (Jusczyk & Aslin, 1995), recent research suggests that word recognition abilities are initially very limited (Houston & Jusczyk, 2000; Singh, Morgan, & White, 2004). At 7.5 months of age, infants familiarized with *bike* and *pear*, for example, listened longer to passages containing these words as opposed to passages containing novel words (e.g., *hat* and *tree*), indicating that they had recognized these words when embedded in context. This effect, however, only holds for words that are acoustically similar during familiarization and test. When the voice changes gender from a male to female speaker or when the affect changes from a happy to a neutral tone, they no longer appear to map the word forms onto the same underlying representations. 10.5-month-olds, in contrast, do generalize across distinct word tokens, indicating that this abstraction ability develops over time. Exemplar-based models have taken this developmental pattern from the early lack of perceptual constancy towards more adult-like word recognition to indicate that words may first be learned in an instance-specific fashion, with generalizations only occurring once a sufficient number of word tokens have been encountered.

In this study, we examine the possibility that past work has underestimated infants' early word recognition abilities. We propose that presenting infants with brief exposure to disembodied unfamiliar voices producing isolated words is not an ecologically valid measure of infants' capabilities. Even adult listeners have difficulty when tested under similar circumstances (Nygaard, Sommers, & Pisoni, 1994). In order to examine this possibility, two groups of infants were tested. The first group was familiarized with passages recorded by their mother. They were then tested on lists of isolated word tokens spoken by their father. The second group of infants heard the same passages and test lists spoken by an unfamiliar speaker (i.e. another infant's parents). Including both groups allows us to test for the possibility that infants can cope with variability in the realization of words more effectively when hearing familiar voices. Following previous studies (e.g., Houston &

Jusczyk, 2000), longer listening times for lists with the familiarized as opposed novel words would indicate cross-gender word recognition. We thus predicted that infants presented with familiar voices would exhibit such pattern and succeed in recognizing words despite the change in gender between familiarization and test. Infants presented with unfamiliar voices, in contrast, should perform in a fashion comparable to past studies and fail to recognize words across different-gender speakers.

2. METHOD

2.1 Participants

Fourty-eight normally developing monolingual English-learning 7.5-month-olds were tested in this study (age range: 220 - 248 days; 25 girls). An additional eleven infants were tested, but excluded from the analysis due to fussiness.

2.2 Materials

Prior to test, twenty-four mothers and fathers whose infants were recruited to participate in the study were audio-taped. Mothers recorded two six-sentence passages in infant-directed speech. Each passage contained a target word (boat, cup, pear, or toque) that occurred once every sentence (see Table 1). For each mother, a familiarization video was created consisting of a picture of the mother zooming in and out combined with three alternating repetitions of each of the two recorded passages. All familiarization videos contained 18 tokens of each of the target words. Fathers recorded test lists with 15 repetitions of the isolated target.

2.3 Procedure

Infants were seated on their parent's lap in front of a TV screen in a sound-attenuated booth. After viewing the familiarization video, the test phase started. An adapted version of the Headturn Preference Procedure (as used by Jusczyk & Aslin, 1995) was used to test infants' recognition of the familiarized words across different genders. First, a red light in front of the infant started flashing. Once the infant oriented towards this light, one of the two lights at the side panels started to flash. As soon as the child turned towards this flashing light, the word list started playing from the loudspeaker mounted underneath the light. Trials were either played until the end of the list or until the infant looked away for two seconds. The computer recorded infants' orientation times to each of the trials.

Table 1. Example passage from familiarization phase

Her boat had white sails. This girl will steer my big boat. That horn on the boat was really loud. He bought himself a new red boat. His boat could go quite fast. We always store your boat in our garage.

2.4 Design

To test the effect of word familiarity, half of the infants were presented with their own parents' voices, while the other half were yoked pairs listening to unfamiliar voices. The target words used during familiarization (*boat* and *tuque*, *boat* and *cup*, *tuque* and *pear*, or *pear* and *cup*) were counterbalanced across conditions. The four test lists were identical across conditions and were presented three times each in quasi-random order.

3. RESULTS

Mean listening times to familiarized and novel words were calculated for each infant separately¹. Across all subjects tested on their parents' voices, the average listening time to lists with familiarized words was 10.18 s (SEM = .57) and to lists with novel words 9.45 s (SEM = .61). For subjects tested on unfamiliar voices, the average listening time to lists with familiarized words was 10.31 s (SEM = .46) and to lists with novel words 9.51 s (SEM = .53; see Figure 1). A mixed 2x2 ANOVA with Word Familiarity (familiarized vs novel words) as within-subject and Voice Familiarity (own parents' voices vs unknown voices) as between-subject factor revealed a main effect of Word Familiarity ($F(1,46) = 4.224$; $p = .046$), indicating that infants listened longer to word lists containing familiarized as opposed to novel words. No other significant main effects or interactions were found (all $F_s < 1$).

4. DISCUSSION

Contrary to our predictions, 7.5-month-olds recognized familiarized words across speakers of different genders, regardless of whether they were presented with familiar or unfamiliar voices. Infants thus seem to generalize across acoustically distinct instances of the same word, suggesting that their early speech perception abilities are more advanced than generally assumed. While this finding is surprising, differences between past and present studies may explain the current findings. Unlike all past studies showing limitations to infants' early generalization capacities, infants in this study were familiarized with passages rather than words, increasing the overall exposure to talker-specific

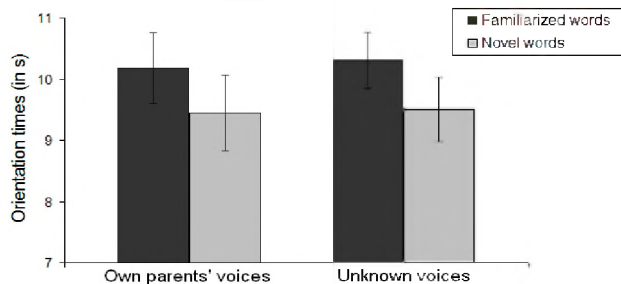


Figure 1. Mean orientation times to familiarized and novel words (and SEMs) broken down by Voice Familiarity

¹ Orientation times more than 2.5 standard deviations away from the infant's mean (3 out of 576 data points) were discarded.

idiosyncrasies. This type of context could be of great help when extracting the intended linguistic form from the speech signal. Contextual information, in other words, may have enhanced speaker adaptation, thereby facilitating word encoding. Support for this view comes from a study with French 8-month-olds, showing that newly presented words are learned better when infants are familiarized with passages than with isolated word tokens (Polka et al., 2008). Another difference between this and earlier studies involves the materials. First, parents' rather than actors' voices were used in this experiment. Second, testing infants on couples' voices may have facilitated word recognition. Vocal convergence might have led to greater acoustic similarity between two partners' than between two randomly selected voices, which in turn may have assisted infants' generalization across the two voices.

Although infants' word recognition in this study was not facilitated by the use of familiar voices, it is very well possible that their robust generalization abilities masked any such effect. In follow-up studies, we are therefore examining the potential advantage of familiar over unfamiliar voices under more challenging listening conditions. Nonetheless, the finding that infants do recognize words across acoustically distinct realizations is intriguing. Models of early speech perception have considered infants' initial lack of generalization as evidence for an episodic memory for speech. When familiarized with more naturalistic stimuli (i.e. passages rather than words), however, infants handle variability in the speech signal more competently than previously thought. Future models of word recognition thus need to take into account infants' abilities under greater acoustical validity.

REFERENCES

- Houston, D.M., & Jusczyk, P.W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception & Performance*, 26, 1570-1582.
- Jusczyk, P.W., & Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 28, 1-23.
- Nygaard, L.C., Sommers, M.S., & Pisoni, D.B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, 5, 42-46.
- Polka, L., Proulx, J., Mersad, K., Iakimova, G., Sundara, M. & Nazzi, T. (2008). *Speech segmentation in French-learning infants: Language-specific and dialect-specific patterns*. Talk presented at the 16th International Conference on Infant Studies, Vancouver, Canada.
- Singh, L., Morgan, J.L., White, K.S. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language*, 51, 173-189.

ACKNOWLEDGEMENTS

This research has been supported by University of Toronto, CFI, and NSERC grants awarded to EKJ and a Prins Bernhard Cultuurfondsbeurs awarded to MvH.

CONTROLLING FOR AGE RELATED HEARING LOSS CAN ELIMINATE AGING DIFFERENCES IN LEXICAL COMPETITION: EVIDENCE FROM EYE-TRACKING AS AN ONLINE MEASUREMENT OF AGE AND NOISE EFFECTS ON LISTENING

Ben-David, B. M.^{1,2,3}, Chambers, C. G.¹, Daneman, M.¹, Pichora-Fuller, M. K.^{1,2}, Reingold, E.¹, And Schneider, B. A.¹

¹ Centre for Research on Biological Communication Systems, Department of Psychology,

University of Toronto Mississauga, 3359 Mississauga Road, N., Mississauga, Ontario, Canada, L5L 1C6

²Toronto Rehab, 550 University Avenue, Toronto, Ontario, Canada, M5G 2A2

³Oral Dynamics Lab, Department of Speech-Language Pathology, University of Toronto, St. George Campus, 500 University Avenue, Toronto, Ontario, Canada, M5G 1V7, boaz.ben.david@utoronto.ca

1. INTRODUCTION

As people age, they experience greater difficulty understanding speech, particularly in environments with background noise. For example, although a 70-year-old grandmother can enjoy her favorite talk radio program at a normal volume in a quiet room, she may have to substantially increase the volume when her grandchildren are playing in an adjacent room, and even then may not accurately hear all of the dialogue. A central question in lifespan approaches to human communication concerns the origin of these speech comprehension difficulties, whether they stem from age-related sensory decline and/or cognitive changes. For example, the difficulties the grandmother is experiencing might be explained by an information degradation account (Schneider & Pichora-Fuller, 2000) whereby the cognitive system was provided with deteriorated sensory (auditory) input. Alternatively, the difficulties could arise from less effective inhibitory mechanisms (Hasher & Zacks, 1988) that work to inhibit the irrelevant noise in the signal, or from a generalized slowing of all cognitive processes (Cerella & Hale, 1994), including those underlying speech comprehension. The goal of the current study is to provide insights into the origins of age-related changes in speech understanding in noise.

Most studies of age-related declines tend to use traditional accuracy-based tests of speech comprehension, e.g., participants are asked to repeat spoken words or sentences presented in different listening conditions. However, research in the field of spoken word recognition has shown that the mapping of speech sounds to lexical candidates begins immediately on the basis of the initial sounds in an unfolding word, often enabling identification before the end of the word has been heard. This phenomenon highlights the fact that speech recognition mechanisms are continuously responding to the unfolding speech signal. Offline tests of accuracy cannot directly reveal these dynamic aspects of the comprehension process. Rather, it is best explored using real-time measures that capture the incremental aspects of speech comprehension.

In recent years, research on spoken word recognition has increasingly used the so-called "visual world" eye-tracking technique, in which listeners hear sentences related to

objects in a visual display (e.g., Tanenhaus, et al., 1995). An eye tracking system records fixations to these objects as speech unfolds. Previous research with young adults has shown that the timing and pattern of eye movements to scene objects can provide a highly sensitive and continuous measure of spoken word processing (see Tanenhaus, et al., 2000 for an overview).

The current study represents the first attempt (to our knowledge) to directly extend this experimental paradigm to a comparison of older and younger listeners' performance in normal and challenging listening situations. We first ensured that overall accuracy was equated. This way, we could directly evaluate whether younger and older listeners differ in terms of the implicit on-line mechanisms that led to correct word identification. We then tested listeners' ability to distinguish the target word from a similar sounding alternative, as the target word unfolded in time.

2. METHOD

2.1 Participants

Twenty-four young adults ($M = 20.2$) and 24 older adults ($M = 70.9$) participated in the study. All participants were native English speakers, had minimum Snellen fractions (visual acuity test) and pure-tone air-conduction thresholds (from 0.25 to 3.00 kHz) appropriate for their age group.

2.2 Materials

Visual Stimuli

For critical trials, we used 32 pairs of clip-art objects, whose corresponding nouns were phonologically related. For 16 of the pairs, their respective nouns overlapped in terms of onset sounds (e.g., *cannon - candy*). For the remaining pairs, the rhyme portion of the nouns overlapped (e.g., *candle - sandal*). On a given critical trial, these phonologically-related objects were displayed with two "unrelated" clip-art objects whose associated nouns matched the other objects in their number of syllables, but did not share onset or rhyme sounds (e.g., *finger* and *zebra*, to accompany *cannon* and *candy*). One of the phonologically-related items was identified as the target object (e.g., *cannon*), and was referred to in a spoken instruction accompanying the visual

display (e.g., "Look at the *cannon*"). The other phonologically-related item constituted the "competitor" object (e.g., *candy*), and was not mentioned in the instruction. On each trial, four images were presented in the four corners of a 3 x 3 grid. An example of a critical trial showing a target-competitor pair with onset overlap is presented in Figure 1. Critical trials were interleaved with 32 filler trials to counteract potential response strategies.

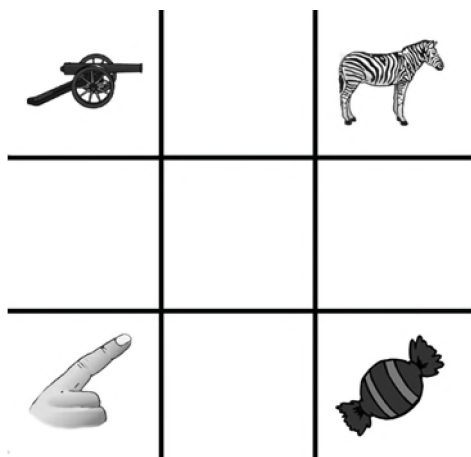


Figure 1. An example of a typical display presented in the experiment. The target (cannon) and the competitor (candy) share onset sounds.

Audio Stimuli

Each display was accompanied by an instruction of the type "Look at the *X*", pre-recorded by a female native speaker of Canadian English. Half the recorded instructions were played in the "quiet" condition and the other half were mixed with speech spectrum noise. The selected signal to noise ratio (SNR) was tailored to the two groups: -4 for younger adults and 0 for older adults. Different SNRs for younger and older adults were used to create comparable overall listening difficulty in noise for both age groups at the sensory level. The audio instructions were played on two speakers located on either side of the participant

2.3 Procedure

Participants were tested individually in a single-walled acoustic chamber. Eye-movements were recorded via a table-mounted eye-tracking system (EyeLink 1000). At the beginning of each trial, a blank grid appeared on the monitor. Participants initiated the presentation of the four clip-art objects by pressing a button. After 2 s, a short tone was played, directing participants to focus on a black fixation cross in the central square. After the system registered cumulative fixations on the central square for 200 ms, the fixation cross disappeared, and the instruction sentence was played. Participants were instructed to look at and maintain gaze on the object denoted in the instruction until a green ("correct") or red ("incorrect") square masked that cell (which occurred after 750 ms of cumulative

fixations within a cell). The objects then disappeared from the grid to signal the end of the trial.

3. RESULTS

3.1 Accuracy

The presence of background noise reduced accuracy, particularly when target-competitor pairs overlapped in terms of onset sounds. However, accuracy rates did not differ by age group ($F < 1$). The absence of an age effect in quiet is characteristic of performance in ideal listening conditions. The lack of an age effect in the noise condition confirms that our selected SNR levels created a similar perceptual (listening) load for both age groups.

3.2 Analysis of Eye Movements

We measured the point in time at which listeners correctly fixated on the target (cannon) and the extent to which listeners momentarily fixated on the acoustic competitor (candy). This measure captures transitory states of indeterminacy that affect the time course of comprehension. Most generally, the eye movement data showed a highly similar pattern of incremental interpretation and target-competitor discrimination for both groups, with greater consideration of the competitor when sentences were presented against a noisy background. These outcomes suggest that after controlling for age-related sensory loss, the cognitive mechanisms underlying on-line spoken word recognition are unchanged over the course of healthy aging.

REFERENCES

- Cerella, J., & Hale, S. (1994). The rise and fall of information processing rates over the life span. *Acta Psychologica*, *86*, 109-197.
- Hasher L., & Zacks, R.T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193-225). San Diego, CA: Academic Press.
- Schneider, B.A., & Pichora-Fuller, M.K. (2000). Implications of perceptual deterioration for cognitive aging research. In F.I.M. Craik (Ed.) & T.A. Salthouse (Ed.), *The handbook of aging and cognition* (2nd ed., pp. 155-219). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Tanenhaus, M.K., Magnuson, J.S., Dahan D., & Chambers C.G. (2000). Eye movements and lexical access in spoken-language comprehension: evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research*, *29* (6), 557-580.
- Tanenhaus, M.K. Spivey-Knowlton, M.J., Eberhard K.M., & Sedivy J.C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632-1634.

ACKNOWLEDGEMENTS

This study was supported by a strategic training grant – Communication and Social Interaction in Healthy Aging, and a group grant on Sensory and Cognitive Aging, both funded by the Canadian Institutes of Health Research.

THE ENVIRONMENT FOR AUDITORY RESEARCH

Bruce E. Amrein and Tomasz R. Letowski

U.S. Army Research Laboratory, Aberdeen Proving Ground, Maryland 21005 USA bruce.amrein@us.army.mil

1. INTRODUCTION

The United States Army Research Laboratory (ARL) recently constructed and made available to their own and other researchers a unique auditory laboratory - the Environment for Auditory Research (EAR). This multi-purpose research facility was designed to conduct signal detection, spatial perception, and communication research in various indoor and outdoor acoustic environments.

The EAR consists of one large integrated Control Room supporting all research functions of the EAR, four indoor listening spaces; Sphere Room, Dome Room, Distance Hall, Listening Laboratory; and outdoor area - the OpenEAR. Indoor listening spaces encompass over 230 m² of configurable laboratory space. Directly adjacent to the indoor facility is the OpenEAR—an outdoor instrumented research space covering more than 4400 m² of natural grassy terrain.

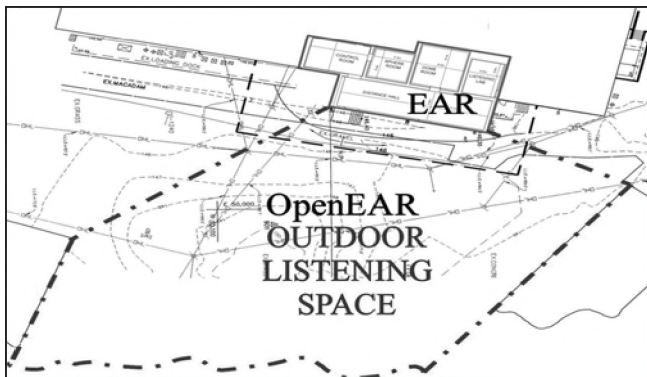


Fig 1. Configuration of the Environment for Auditory Research

2. FACILITY DESCRIPTION

The EAR is an auditory perception and communication research center permitting state-of-the-art spatial perception and communication research in various indoor and outdoor acoustic environments. The EAR, shown schematically in Figure 1, has been designed to address all the needs of auditory spatial research and includes spaces that encompass geometries of most common indoor environments. The four indoor listening spaces (Figure 2) can be used individually or, in some cases, can be combined together to simulate more complex architectural structures. The predominantly anechoic environment of the EAR can be modified, if needed, to represent more realistic environments. The facility is located in the ARL complex at Aberdeen Proving Ground (MD, USA).

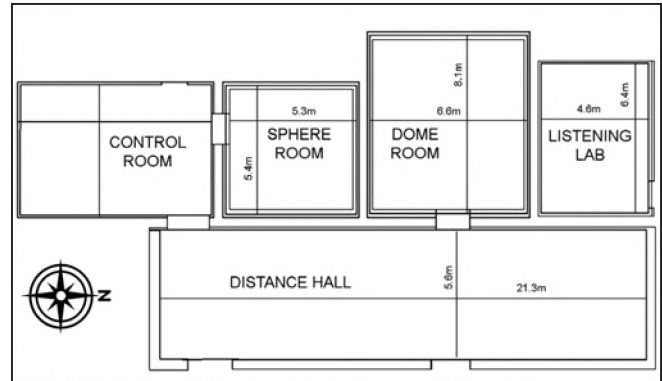


Fig. 2. Schematic configuration of interior spaces.

All indoor listening spaces of the EAR comply with the NC-15 noise criteria resulting in background noise levels close to the threshold of hearing, semi-anechoic listening conditions reducing early acoustic reflections to negligible levels, extensive and flexible means for sound production by 600+ sound sources, and acoustic and electroacoustic means for changing the spatial properties of sound. All listening spaces also have widely adjustable lighting and temperature conditions.

2.1 Control Room

The Control Room (54 m² of floor area) is an integrated control center permitting complete control of instrumentation and research activities in all four indoor spaces and an outdoor listening space. It contains the front-end of all instrumentation and stimuli generation systems. The audio system of the EAR is powered by 4 computers and includes extensive and automatic switching capability. The system is capable of generating up to eight (8) independent audio signals and transmitting them to any or all (approximately 600) loudspeaker and earphone locations throughout the facility.

Functionality of the Control Room enables control and monitoring of as many as four simultaneous experiments conducted in various spaces of the facility from a single location. The networking capabilities of the EAR allow the control functions of the Control Room to be accessed from each of the test spaces allowing a researcher to set up an experiment from within a target space. In addition, audio and video capabilities of the Control Room can be used to provide audio-video demonstrations and instructions for new users, experiment participants, and visitors.



Fig 3. The EAR control center

2.2 Sphere Room

The Sphere Room is a 140 m³ (5.3m × 5.4m × 4.9m) auditory virtual reality space designed to investigate integrity of auditory virtual spaces, realism of complex auditory simulations, effects of Head-Related Transfer Function on auditory perception, and the effect of helmets and other headgear on spatial orientation in a three-dimensional dynamically changing environment (Figure 4 left). The sound reproduction system of the room consists of 57 loudspeakers distributed on a sphere and radially separated by about 25°.

2.3 Dome Room

The Dome Room is a 220 m³ (6.6m × 8.1m × 4.1m) space designed to study the human's ability to localize real or virtual, single or multiple, and stationary or moving sources in a horizontal plane or along two vertical arcs extending from -20° to +40° regarding listener's head position (Figure 4 right). The sound system capabilities support 2° horizontal and 10° vertical spatial resolution.



Fig 4. Sphere Room (left), Dome Room (right).

2.4 Distance Hall

The Distance Hall is a 440 m³ (21.3m × 5.6m × 3.7m) acoustically treated space designed to study auditory distance estimation and the effects of sound source movement toward and away from the listener on sound source detection and identification (Figure 5 left). Acoustic configuration and audio capabilities of the Distance Hall permit extensive investigation of localization and tracking of sound sources moving in a predetermined manner toward and away from the listener, auditory distance and depth

estimation, tracking of sound sources moving above the listener, or detection and recognition of sound sources appearing far away from the listener.

2.5 Listening Laboratory

The Listening Laboratory is a unique multipurpose 140 m³ (4.6m × 6.4m × 3.5m) room for studying the effects of space acoustics (removable acoustic wall treatment) and sound sources configurations on sound perception (Figure 5 right).

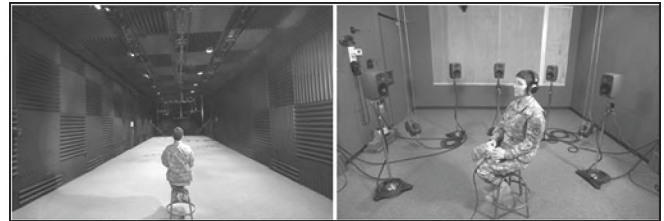


Fig. 5. Distance Hall (left), Listening Laboratory (right)

2.6 OpenEAR

OpenEAR is a 4,459 m² outdoor extension of the EAR complex designed to replicate studies conducted in the laboratory environment in a natural field environment with the same listeners at almost the same time to reduce data uncertainty resulting from laboratory and field studies which are conducted at different times and with different listeners.

3. CONCLUSION

The ARL Environment for Auditory Research is a unique and powerful research tool with indoor and outdoor capabilities that are unmatched at any current military, academic, or industrial facility world-wide. The facility is used by ARL researchers to increase our understanding of Soldier auditory capabilities and challenges on the modern battlefield. However, the EAR is also open to external researchers who are interested in studying spatial orientation, distance and depth estimation, virtual displays design, signature detection and identification, icons and warning signals design, perception of moving sound sources, or similar topics.

REFERENCES

- Henry, P., Amrein, B., and Ericson, M. (in press). The Environment for Auditory Research. *Acoustics Today*.
- Scharine, A. A., and Mermagen, T. (2008). Characterization of the Environment for Auditory Research (EAR) at the U.S. Army Research Laboratory. Proceedings of the 15th International Congress on Sound and Vibration, Daejeon, Korea, July 6-10.

ACKNOWLEDGEMENTS

The authors are grateful to the management of the U.S. Army Research Laboratory for providing necessary funding and support to build the Environment for Auditory Research.

AN INVESTIGATION OF THE WIND INDUCED BACKGROUND SOUND LEVELS IN RURAL AREAS FOR EXISTING AND PROPOSED WIND FARMS

Ian Bonsma, PEng and Megan Munro, BAsC, EIT

HGC Engineering, 2000 Argentia Rd., Plaza 1, Suite 203, Mississauga, Ontario, Canada L5N 1P7

1. BACKGROUND

The increasing prevalence of wind power across the globe has led to a growth in the number of wind turbines in proximity to human settlement. As a result, the sound level impact from wind turbines has become a fairly contentious issue in some areas. The audibility of wind turbine generators is often a factor of low background sound levels, which in rural environments are dominated by natural sounds. Thus, HGC Engineering has considered ambient sound levels as part of a number of acoustic audits and assessments for existing and proposed wind farms.

This paper investigates the variability of wind induced background sound levels in various rural areas by comparing sound level measurements conducted by HGC Engineering in Canada, the United States, and Central America with the reference background sound levels presented by the Ontario Ministry of the Environment (MOE). Specific consideration will be given to the effects of topography and the relationship between hub height wind speeds and ground level wind speeds. This analysis of actual wind induced background sound levels provides a better understanding of the sound level impact from wind turbines in varying rural environments.

2. STUDIES

HGC Engineering's experience has shown that the measurement of background sound levels in conjunction with wind speeds at various heights and locations is an important part of the planning process. A previous study prepared for the Canadian Wind Energy Association (CanWEA) recommended that ambient sound levels be monitored to assist in defining the criteria and to provide a benchmark for any sound measurements conducted after the wind farm is operational [1].

The MOE has recognized wind turbines as unique among other industrial sources because they produce more sound as wind speeds increase and because increasing wind speeds generally result in increasing levels of background sound. The current MOE guidelines, "Interpretation for Applying MOE NPC Technical Publications to Wind Turbine Generators", for the assessment of sound from wind power projects presents sound level limits for wind turbines that are based on reference wind induced background sound levels measured at a particularly quiet site, correlated with wind speeds at a height of 10 m [2]. The reference sound levels are presented in terms of the L_{90} , the sound level which is exceeded 90% of the time during a measurement. The L_{90} sound level eliminates the contribution of short duration local noises.

2.1 An Existing Wind Farm in Ontario

There are a number of challenges associated with auditing the sound level impact from an existing wind farm and comparing the results to appropriate guidelines and the sound level predictions made as part of the initial acoustic assessment. In general, the sound level limits established by the local jurisdiction will typically apply only to the sound

level contribution of the source under assessment (In this case, the wind turbines). Thus, where the sound level measured at a receptor includes significant sound due both to the relevant sound source and to unrelated background sound sources (road traffic, local noises, and wind, for example), some form of evaluation must be made to determine the sound level contribution of the source under assessment in the absence of background sounds.

Several sound level monitors were deployed at representative receptor locations throughout an Ontario wind plant to record the sound levels over a period of approximately one week. A sound level monitor was also deployed at a 'remote location' to record background sounds unaffected by any noise from the wind plant. The monitor was located approximately 3 km north of the closest wind turbine in a relatively exposed area with some tall grass nearby. One of several factors considered in the data analysis was the influence of ambient sound levels. Where trends in the data measured at the remote location mimicked the data recorded at the receptors, it provided an indication that the wind turbine generator(s) were not appreciably affecting the sound levels at the receptors, particularly during periods of high winds and correspondingly high background sounds.

Figure 1 presents the measured L_{90} sound levels and the corresponding 10 m wind speed measured near the monitoring location and also presents the reference wind induced background sound levels (L_{90}) presented in the current MOE guidelines specific to wind turbine generators. The sound level data was considerably higher during the daytime hours due to local activity, therefore only the data recorded during the nighttime hours (19:00 until 7:00) is presented below. Periods of rain were also excluded from the dataset. The grey highlighted area indicates time periods where, at times, the sound level fell below the range of the sound level meter.

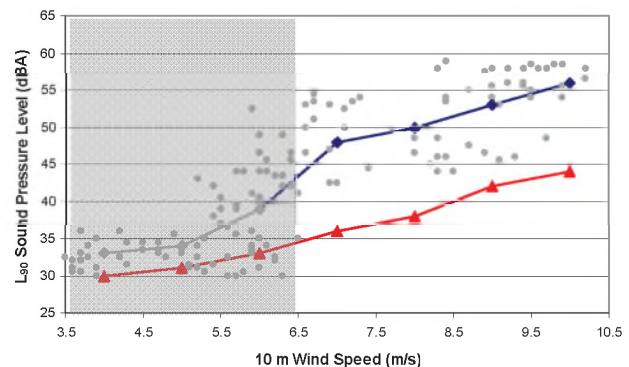


Fig. 1: L_{90} sound levels measured at a wind farm in Ontario during nighttime hours at a remote location (●), average L_{90} sound levels for each wind bin (◆) and the reference wind induced background sound levels defined by the MOE (▲).

Considering 10 meter wind speeds on the order of 7 m/s and higher, the measured background L_{90} sound levels were

typically higher than the reference background sound levels defined by the MOE. This monitoring location was located in a relatively open field that had exposure to wind from all directions. This may not be typical of some residences, which would generally be somewhat shielded from the wind from some directions due to the residence itself and other nearby buildings. Nevertheless, the remote monitor was helpful in determining time periods that were dominated by wind induced background sound at the residences.

2.2 A Proposed Wind Farm in Central America

As part of an acoustic assessment completed for a proposed wind farm in Central America, background sound levels were measured at five representative receptor locations. Locations were selected to capture the sound levels from the major roadways, major topographical changes, and rural areas. Measurements were conducted for a period of approximately 72 hours to obtain a representative data sample. During this time period, wind speeds were also recorded at various locations and heights for correlation with the ambient sound levels.

The site layout was mountainous with elevations ranging from 1,250 to 1,830 meters above sea level. The proposed locations for the wind turbines were heavily influenced by wind resources, which are generally most significant at higher elevations. As a result, the placement of wind turbines was generally along the mountain ridges. This unique site layout and topography played an important role in determining the influence of background sound levels due to wind induced noise.

Analysis of the measurement data showed that the sound levels were not as affected by changes in wind speed, measured near the wind turbines, as the measurements conducted on relatively flat ground for the wind farm in Ontario or as expected based on the reference background sound levels defined by the MOE. The severe topographical variations have a significant impact on the wind speeds at the different receptor locations. As a result, many locations are somewhat protected from high wind speeds.

An evaluation of the background sound and wind speed measurements showed that under certain conditions, it would be possible to have the wind turbines producing full power with background L_{90} sound levels at neighbouring receptors being on the order of 35 dBA.

2.3 A Proposed Wind Farm in Pennsylvania

Background sound levels were measured at six representative receptor locations as part of an acoustic assessment for a proposed wind farm in Pennsylvania. Measurements were conducted for a period of approximately one week to obtain a representative data sample. During the measurement time period, wind speeds were also recorded at two locations and several heights for correlation with the ambient sound levels.

The proposed location for the wind turbines was along the top of a mountain ridge (320 – 400 meters above sea level), with residences located on both sides of the mountain ridge (130 – 260 meters above sea level) at various distances from the proposed turbine locations.

Figure 2 shows L_{90} sound levels measured at one of the monitoring locations and the corresponding 10 m wind speed calculated using the relationship between wind speeds recorded at various heights on the closest metrological tower using the wind shear relationship [3]. This figure also presents the reference wind induced background sound levels (L_{90}) presented in the current MOE guidelines specific to wind turbine generators. Sound levels at this

location did not vary significantly between daytime and nighttime periods. Thus, all measured data, except periods of rain, have been shown in the figure below.

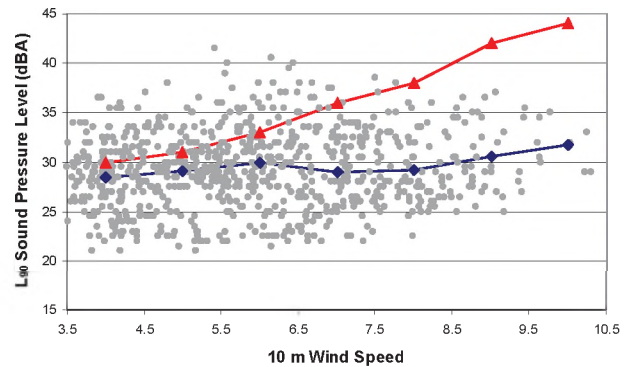


Fig. 2: L_{90} sound levels measured at a wind farm in Pennsylvania (●), average L_{90} sound levels for each wind bin (◆) and the reference wind induced background sound levels defined by the MOE (▲).

As shown in Figure 2, the sound levels recorded at this residence were not significantly influenced by changes in wind speed at the top of the mountain. The ambient sound levels remained fairly consistent throughout the measurement period and were consistently less than the reference sound levels defined by the MOE. An analysis of the data indicates that the wind turbines can be operating at maximum power and the ambient sound levels can be on the order of 30 dBA at very quiet residential receptors.

3. CONCLUSION

HGC Engineering monitored background sound levels as part of a number of acoustic assessments and audits for existing and proposed wind farms. The results showed significant variability in both the magnitude of the measured sound levels and the influence of wind speed, depending on the site layout, orientation, and local topography, shielding effects, and vegetation. Depending on the site, ambient sound levels may be greater than or less than the reference background sound levels defined by the MOE. Elevated ambient sound levels increase the difficulty associated with auditing the sound level impact from the wind turbines. However, they do provide a benefit to the residents in that they tend to mask the sounds from neighbouring wind turbines. The opposite is true for sites with low ambient sound levels.

The results confirm the importance of monitoring ambient sounds at a variety of representative receptors. It is particularly important when the sound level limits are based on ambient sound levels and for sites with significant elevation changes.

REFERENCES

- [1] HGC Engineering, Canadian Wind Energy Association, “Wind Turbines and Sound: Review and Best Practice Guidelines”, 2007.
- [2] Ontario Ministry of the Environment (MOE), “Interpretation for Applying MOE NPC Publications to Wind power Generation Facilities”, 2008.
- [3] IEC 61400-11 International Electrotechnical Commission, “Wind Turbines”, 2009.

INADEQUACY OF WIND TURBINE NOISE REGULATIONS AND THEIR APPLICATION

John P Harrison

Physics Department, Queen's University, Kingston, ON K7L 3N6 harrisjp@physics.queensu.ca

1. INTRODUCTION

As this is being written, the regulations for future wind-energy development in Ontario are being discussed and revised. The reason for new regulation is that in May 2009 Ontario adopted a Green Energy Act which removes municipalities from the approval process and puts the final decision on projects with the Ontario government. The draft regulations for noise at a receptor issued by the Ontario Ministry of the Environment (MOE) are not so different from those established by MOE in October 2008 and, Germany aside, not so different from those in other jurisdictions.

Typically, noise regulations require an $L_{eq} = 40$ dBA limit at a receptor (home, school, institution) as determined by the ISO prediction code 9613, the manufacturer's noise specification for the turbine, a suitable ground effect parameter and an atmospheric absorption coefficient of 0.005 dBA/m. In practice this translates into setbacks from receptors of about 500 m for an isolated modern turbine and about 800 m for a group of three similarly spaced turbines. Two jurisdictions, Ontario and New Zealand, still allow an increase in the limit with increase in wind speed. The new draft regulations in Ontario have dropped this allowance but as yet they are just draft regulations.

There are several recommendations from health and other authorities that these setbacks are far too small (see Harrison 2008). The recommendations are for setbacks in the range of 1.5 to 2 km. An on-going health-impact study in Ontario (McMurtry, 2009) has so far turned up more than 70 victims with health problems, severe enough to require medical attention, caused by the proximity, pre-approved by the MOE, of one or more turbines to their homes. Two field studies, one in Sweden (Pedersen and Perrson Waye 2007) and one in Wisconsin (Bittner-Mackin 2003), have found annoyance with wind turbine noise at the 40 dBA level among 50% of respondents; this compares with 3 to 4% for traffic noise at 40 dBA. It is clear that the turbine noise limits are too lenient to the wind industry and that changes are needed if wind-energy development is to continue to expand. This paper discusses the basis for the inadequacy.

2. INADEQUACIES

The myth of masking from ground level wind noise seems to have been laid to rest thanks to the pioneering work of van den Berg (2004) and the myriad measurements that have demonstrated the large wind speed gradient in the atmosphere at night. This will not be discussed further here.

However, for those people living with the reduced setbacks that resulted from earlier regulations with the masking noise allowance, the problem remains until the offending turbines are shut down. The remaining problems are concerned with the large intrusion of turbine noise above ambient, the characteristic swooshing sound of an operating turbine, the excess low frequency noise due to turbulent inflow and the neglect of uncertainty in noise prediction.

2.1 Intrusion

Rural regions are very quiet, probably below 25 dBA at night. This means that typical guidelines are allowing a 15 dBA intrusion above background and, given the annoying characteristic of turbine noise, this is too much. There is no need to allow this large an intrusion. Germany, which has a population density 20 times larger than that of Ontario and has a well-developed wind energy generation system supplying 6.4% of its electrical energy, has a night-time noise limit of 35 dBA. In another instance, New Zealand, in section 5.3.1 of its draft regulations, is introducing a secondary noise limit of 35 dBA for evening and night-time in low background environments.

2.2 Amplitude Modulation

Wind turbine noise is periodic in the blade passage frequency. It is clear from the work of van den Berg (2005). It is clear from the Salford report (Moorhouse et al 2007) published by the British Wind Energy Authority. It is acknowledged by MOE in its turbine noise regulations published in October 2008. The consensus is that it amounts to about 5 dBA of amplitude modulation. This amplitude modulation is averaged away by regulations based upon an L_{eq} . However, the ear does not average and this swooshing sound adds significantly to the annoyance associated with turbine noise. A 5 dBA penalty is needed to account for the amplitude modulation.

2.3 Turbulence

Many noise complaints draw attention to a component that sounds like a rumble (a dryer or a passing train that never passes!). This is probably excess low frequency noise associated with turbulent inflow of air into the blades. The turbulence has two sources, turbulence in the atmosphere and the turbulent wake from neighbouring turbines. SODAR measurements (Barthelmie 2003) have shown that for $x/D \sim 5$, the turbulent intensity (TI) behind a turbine is comparable to the atmospheric TI (x is the distance behind the blade and D is the blade diameter). They were 5% and

7% respectively. Turbulent intensity is defined as σ/v where σ is the standard deviation of the wind speed v (Wagner et al 1996). The SODAR measurements were made every minute and the averaging time for σ and v was 10 minutes. Low frequency noise requires a faster time scale for the calculation of σ and hence of the appropriate TI. However, the important point is that turbulence about 5 blade diameters behind a turbine is significant. I note that for the Wolfe Island wind farm in Ontario about half of the turbines are within 6 blade diameters of an upwind turbine for the prevailing south-west winds. As an aside, the velocity deficit for the same half of the turbines due to the wake of the upwind neighbours will be 20% (Barthelmie 2003), so lowering the power output efficiency by 50% (0.8^3) from that of the upwind turbines!

Moriarty and Migliore (2003) and Moriarty (2004) working at the National Renewable Energy Laboratory in Golden CO, made a study of inflow turbulence noise from turbines, with both measurements and predictions. Below 1 kHz, the turbulent inflow noise can dominate the total turbine noise. For instance, with a TI of $I = 10.6\%$, at 100 Hz this noise is 30 dBA larger than the combined noise from all other aerodynamic sources. Doubling the frequency decreases the turbulence noise by 5 dBA. The noise power is proportional to I^2 , so that the sound pressure level falls by only 6 dBA as the TI is halved. The noise measurements bear out the predictions apart from the need for an adjustment for the averaging time for the determination of σ .

It is quite clear from measurements of the turbulent wake downwind of a turbine, the close proximity of turbines to each other at wind developments around the shores of the Great Lakes, the predictions of turbulent inflow noise calculations and the agreement with measured noise that it is vital that this noise source be a part of noise regulation. This noise will not go away at night when the day-time atmospheric turbulence gives way to the stable night-time atmosphere. Turbulent inflow noise is predominantly in the low frequency range below 1 kHz, particularly near the lower range of hearing, and where the absorption by the atmosphere is minimal. Enough is known that prediction of turbulence noise can be made both from prior wind speed test tower measurements and from the proposed layout of the turbines. To date, no jurisdiction is requiring turbulence noise in their approval process. This must change.

2.4 Uncertainty

No prediction is going to be 100% correct. The turbine manufacturer quotes an uncertainty of ± 1 or 2 dBA. One of the frequently used prediction codes, ISO-9613, specifically states an uncertainty of $\pm 3\%$. These are independent uncertainties and so will add in quadrature. Therefore the prediction for noise at a receptor will carry an uncertainty of ± 3 to 4 dBA. No self-respecting and responsible engineer would ignore the uncertainty in a design calculation; yet

noise consultants do ignore this uncertainty and, in Ontario, the engineers at MOE allow this neglect.

4. CONCLUSION

Regulations for wind turbine noise presently in force are inadequate to protect rural residents from annoyance and, in many cases, health problems resulting from operating wind turbines. The typical noise limit of 40 dBA needs to be reduced to 35 dBA. There needs to be a 5 dBA penalty for amplitude modulation. There needs to be an analysis of turbulent inflow noise, for both atmospheric and wake turbulence. The uncertainty of noise prediction codes must be included. Together, these essential up-grades to regulation will push setbacks to the 1.5 km range where they should be.

REFERENCES

- Barthelmie, R. J. et al (2003), *J. At. Oceanic Tech.* 20, 466.
- Bittner-Mackin, E. (2003), Excerpts from the Final Report of the Township of Lincoln [Wisconsin] Wind Turbine Moratorium Committee, 12/4/03.
- Harrison, John P. (2008), Proceedings of World Wind Energy Conference (June 2008); see also: Pierpont, Nina <http://www.windturbinesyndrome.com/?p=76>
- Moorhouse, A. (2007), Research into Aerodynamic Modulation of Wind Turbine Noise. [www.bwea.com/pdf/0707%](http://www.bwea.com/pdf/0707%20Moorhouse%20A.pdf)
- McMurtry R. (2009): <http://windconcernsontario.files.wordpress.com/2009/04/deputation-to-standing-committee-mcmurtry.pdf>
<http://windconcernsontario.files.wordpress.com/2009/04/ontario-health-survey-april-22-2009.pdf>
- P. Moriarty and P. Migliore (2003), NREL Report NREL/TP-500-34478
- P Moriarty (2004), Development and Validation of a Semi-Empirical Wind Turbine Aero-acoustic Code. (NREL Report).
- Pedersen, K. and Persson Waye, K. (2007), *Occupational and Environmental Medicine*, 64, 480-486
- van den Berg, G.P. (2004), *J. Sound and Vibration*, 277, 955-970.
- van den Berg, G.P. (2005), *J. Low Frequency Noise, Vibration and Control* 24, 1.
- Wagner, S. et al (1996), *Wind Turbine Noise* (Springer).

SOURCE-RECEPTOR PATH PRIORITIZATION CONSIDERATIONS FOR THE SPECIFICATION OF SILENCERS IN VENTILATION SYSTEMS

Brian Chapnik and Alex Lorimer

HGC Engineering, 2000 Argentia Road, Plaza One, Suite 203, Mississauga, Ontario, Canada, L5N 1P7

1. INTRODUCTION

Large commercial and institutional buildings such as airport terminals, universities, hospitals and office towers have considerable air handling equipment and ductwork so as to circulate heated or cooled air through the occupied spaces. For each large air-handling unit, the related ductwork can be geometrically quite complex, often consisting of one or more main headers and many turns and junctions in both the main ducts as well as secondary and even tertiary ducts branching off the main supply header before arriving at the individual end diffusers. This ductwork "tree" also acts as a channel for directing noise into the same spaces. Noise breakout from the duct walls can also impact spaces along the path of this "tree".

In order to keep noise to an acceptable level, silencers are typically placed near the source of the noise, and/or in each duct path that is predicted to transmit noise levels in excess of the target. The industry standard techniques for determining these noise levels, transmission and radiation effects are specified by ASHRAE [1], which has published methods for calculating the sound along a given duct path to a receiver space. However, ASHRAE does not provide direction as to which paths to consider critical in specifying optimal silencing for a system. Prioritization is important to avoid the need for calculating noise impacts at all spaces along a "tree", which can be an extremely time consuming exercise and to avoid overly conservative simplifications which can result in a more expensive (capital and operating) noise control scheme.

Several important factors are to be considered in prioritizing which source-receiver paths to analyze in specifying silencers, without analyzing all possible paths, such that noise at all locations in the "tree" will be addressed with reasonable confidence. These factors are discussed below.

2. APPLICABLE FACTORS

Some of the most important effects on noise transmitted through a duct system to a receiver room are discussed by Kingsbury [2] and by Hoover and Blazier [3]. These include:

- Attenuation of sound in unlined and lined ducts;
- Division of sound power at duct branch points;
- Attenuation of sound at elbows;
- End reflection losses;
- Break-out from duct walls;

- Radiation of sound into a receiver room ("Room Effect").

While these effects are to some extent true for all source-receptor paths, there is little direction provided by these authors regarding how to prioritize them. Kingsbury notes that "the most critical part of the system is the room closest to the fan serving that system unless other rooms downstream require lower sound levels". This suggests two factors that must be considered in any prioritization:

1. Proximity to the source (fan), and;
2. Noise criteria of the receiver space.

Additional factors which may be important can be derived from the above list of effects, such as:

3. The extent and size of lined versus unlined ductwork between source and receiver;
4. The number and size (area ratio) of branches between source and receiver;
5. The complexity of the path from source to receiver, i.e. elbows or other fittings;
6. The nature of the diffusers in the receiver space (grille, rectangular, or linear), which affects end losses;
7. The geometry of the ductwork (rectangular, round or oval), especially for break-out considerations, and what type of ceiling is below the ducts (exposed, T-bar, drywall).

3. IMPORTANCE OF FACTORS

Clearly, there is no straightforward way to prioritize a given factor globally – the importance of some factors depends on their predominance in a given ductwork system design. However, there are certain trends that can be gleaned from comparing the attenuation of one set of factors to another in typical cases, yielding insight into which paths should be focused on in analysis of these systems in order to specify an appropriate silencer at the fan.

3.1 Rectangular Duct Branches

Typical break-out from a main rectangular duct branch is considerably more likely to result in greater low-frequency noise impact in a receiver space than sound transmission through the supply diffusers. This is because the attenuation of additional duct lengths, branches and elbows downstream of the main rectangular branch, and the end loss at the

distributed diffusers, results in small radiation of sound in comparison with that breaking out of the duct.

Consider a typical case of a 600 mm x 300 mm rectangular duct feeding smaller 300 mm x 150 mm ducts to four rooms, with one 300 mm and one 150 mm elbow in each smaller branch duct. The smaller branches are each fed to two 150 mm round drops to rectangular diffusers in the receiver rooms. Standard attenuation calculations might show the following (some minor effects neglected):

Table 1: Attenuation from Break-Out [dB] (Rect)

Frequency [Hz]	63	125	250	500
Break-out, 600x300x5m, 24ga	3	6	9	12
Ceiling/Plenum loss (T-bar)	3	6	8	10
Room effect (typical)	5	6	7	8
Total	11	18	24	30

Table 2: Attenuation to Diffusers [dB] (Rect)

Frequency [Hz]	63	125	250	500
Duct loss, 300x150x5m	5	3	2	1
Branch loss (25%)	6	6	6	6
Elbow loss (2)	0	0	1	3
End loss (150 mm w/diffuser)	9	7	4	2
Room effect (typical)	5	6	7	8
Total	25	22	20	20

It is clear that for a fan with significant low-frequency energy, which is often the case for centrifugal fans, a silencer designed to limit the break-out from the main duct will be sufficient to address noise transmission from the diffusers, without needing to explicitly consider the latter. Even if the ceiling were drywall instead of T-bar, if the fan energy in the 63 Hz band is dominant, the attenuation due to break-out mechanisms is still likely to be less than those between the main branch and the diffusers. Attenuation factors to the diffusers are lower in the higher frequency bands, but since silencers generally provide more attenuation than needed in these bands, this is usually not of significant concern unless the source itself has high-frequency energy concentration (e.g. high-speed blower or mixed-flow fan) and there is no acoustically lined ductwork in the system.

3.2 Circular Duct Branches

Circular ducts are often used to limit low-frequency break-out, in which case sound transmitted out the diffusers is much more likely to be significant in comparison with break-out, assuming some ceiling below (in the case of exposed ductwork, break-out may still be significant at higher frequencies). Consider the same example as above,

but with approximately area-equivalent circular ductwork instead of rectangular.

Table 3: Attenuation from Break-Out [dB] (Circ)

Frequency [Hz]	250	500	1k	2k
Break-out, 450mmx5m, 24ga	25	17	15	13
Ceiling/Plenum loss (T-bar)	8	10	16	21
Room effect (typical)	7	8	9	10
Total	40	35	40	44

Table 4: Attenuation to Diffusers [dB] (Circ)

Frequency [Hz]	250	500	1k	2k
Duct loss, 250mmx5m	1	1	1	1
Branch loss (25%)	6	6	6	6
Elbow loss (2)	2	4	6	6
End loss (150 mm w/diffuser)	4	2	1	0
Room effect (typical)	7	8	9	10
Total	20	21	23	23

Thus, for a typical system with circular distribution ductwork and a finished ceiling, the silencer may be designed on the basis of noise transmitted through the diffusers of the nearest room, assuming this is the shortest path with similar complexity as other paths in the system. Break-out analysis need not be conducted.

4. CONCLUSIONS

It has been illustrated for two simple cases how reasonable “rules of thumb” can be developed to simplify the task of specifying a silencer for a ventilation system, without needing to consider all paths in the duct “tree”. Some additional guidelines of this nature have been developed which cannot be fully described here. Further work is needed to develop a more comprehensive set of prioritized factors which can be used by the acoustical designer of HVAC systems to reduce the time required to specify silencers for a system, while maintaining a reasonable degree of confidence in the result.

REFERENCES

- [1] HVAC Systems Applications Handbook, American Society for Heating, Refrigeration and Air-Conditioning Engineers (ASHRAE), Atlanta, GA, 2007.
- [2] Kingsbury, Howard (1997). Noise Sources and Propagation in Ducted Air-Distribution Systems. Chapter 85, Encyclopedia of Acoustics, John Wiley & Sons Inc., 1997.
- [3] Hoover, Robert M. and Blazier, Warren E. Jr. (1994) Noise Control in Heating, Ventilating, and Air-Conditioning Systems, Chapter 7, Noise Control in Buildings (ed. Cyril M. Harris), McGraw-Hill Inc., 1994.

COOLING-TOWER NOISE CONTROL FOR THE NIAGARA FALLS CASINO

Thomas Paige

Kinetics Noise Control, Inc., 3570 Nashua Drive., Mississauga, ON, Canada, L4V 1L2 tpaige@kineticsnoise.com

1. INTRODUCTION

The new casino built on the Canadian side of Niagara Falls is located very close to the waterfalls. In the design phase, there was a concern that the noise from the large cooling towers serving the casino would interfere with the natural sound of the falls. Cooling towers usually have multiple large-diameter propeller fans that generate a significant level of low-frequency noise. The intakes for these fans are typically located on the sides of the unit, so fan noise is also radiated horizontally from these openings. Traditional methods for noise control include sound-barrier walls, silencers or noise enclosures. Barrier walls usually have no adverse effect on the performance of the cooling towers. However, silencers and noise enclosures must be carefully designed to minimize airflow resistance and maintain proper operation of the cooling tower. This paper discusses the techniques used to successfully reduce the noise propagation from the casino cooling towers.

1.1 Induced-Draft Cooling Towers

These are the most common type. The propeller fans on top of the unit discharge air upward. Air is drawn into the intakes on the sides of the unit, and passes through a stream of water. Some of the water evaporates into the airstream causing cooling. Figure 1 shows the typical fan noise spectrum. The blade pass tone is typically below 63 Hz. Other noise sources include the spraying and splashing sounds of water, as well as motor, belts and gearbox noises.

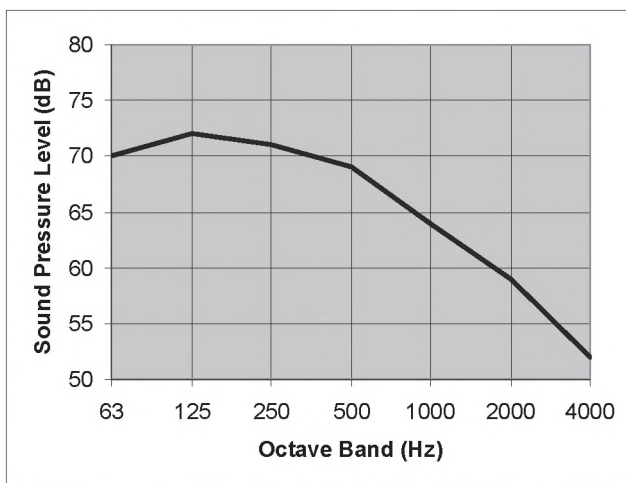


Figure 1. Typical induced-draft cooling tower fan noise spectrum for 2800-kW unit at 15 m (800-ton unit at 50 ft), based on CTI Code ATC-128 (Schaffer 2005).

1.2 Casino Cooling Towers

The casino required 3 double-cell cooling towers with each unit providing 6900 kW (1970 tons) of cooling. In addition to the noise impact of the cooling towers on the natural sound of the waterfalls, there was also a potential impact on the hotel tower that was part of the development. Figure 2 shows the proximity of the cooling towers with the falls.



Figure 2. Proximity of cooling towers with the waterfalls (located near the left side of this picture taken during construction).

1.3 Noise Control Options

Available noise control options for cooling towers include: low-noise fans, reduced fan speed and sound-attenuation components such as noise-barrier walls and silencers. In this case, the decision was made to provide rectangular absorptive silencers on the cooling tower fan discharge and air intakes.

2. SILENCER DESIGN

2.1 Silencer Design Parameters

The 6 discharge silencers each handle an airflow volume of 108,500 L/s (230,000 cfm). Each silencer is 3500 x 3500 x 3000 mm long (11.5 x 11.5 x 10 feet). A rectangular silencer design with a circular bullet was used. The design pressure drop was 15 Pa (0.06 in. wg).

The 12 intake silencers each handle an airflow volume of 54,250 L/s (115,000 cfm). Each silencer is 3500 x 5200 x 900 mm long (11.5 x 17 x 3 feet). These are standard rectangular silencers with a design pressure drop of only 3 Pa (0.01 in. wg).

The installed discharge and intake silencers are shown in Figure 3.



Figure 3. Installed discharge and intake silencers.

2.2 Silencer Insertion Loss

The insertion loss requirements for the silencers were specified by the acoustical consultant and are given below in tables 1 and 2.

Table 1. Discharge silencers required insertion loss.

Octave Band (Hz)	63	125	250	500	1k	2k	4k	8k
Insertion Loss (dB)	4	14	22	25	11	6	2	1

Table 2. Intake silencers required insertion loss.

Octave Band (Hz)	63	125	250	500	1k	2k	4k	8k
Insertion Loss (dB)	6	8	16	29	24	12	11	10

3. MAINTENANCE ISSUES

The silencers were constructed with G 90 galvanized steel to resist corrosion from the discharge air which is 100% saturated with water vapor. Plenums were provided above the fans for access and servicing. Regular inspections and maintenance are required to detect and prevent rusting.

4. STRUCTURAL ISSUES

Separate structural steel frames were provided to support the weight of the silencers. The wind loading on the silencers also had to be considered in the system design.

5. RESULTS

Sound level reading taken after construction confirmed that the cooling tower noise levels met the design criteria. Figure 4 shows an overall view of the cooling tower silencer system.

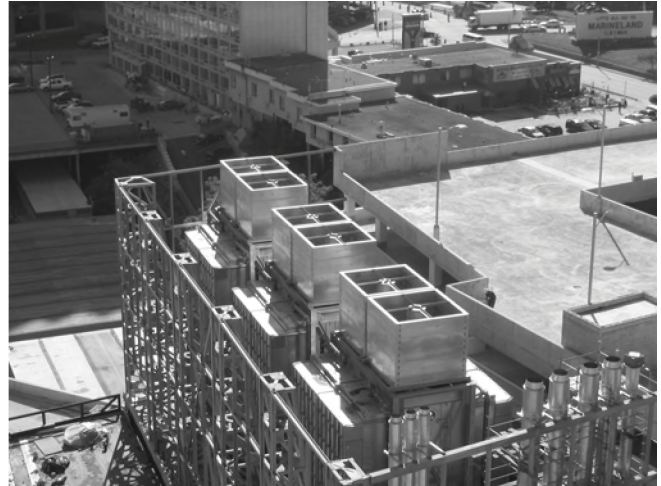


Figure 4. Overall view of cooling tower silencers.

REFERENCES

Schaffer, Mark E (2005). *A Practical Guide to Noise and Vibration Control for HVAC Systems (2nd Edition)*, ASHRAE, Atlanta, GA.

Acoustic Correlates of Neutral versus Angry Affect in Real and Non-Word Sentences

Heather Flowers, M.Ed., M.H.Sc.¹, Tim Bressmann, PhD¹

¹Dept. of Speech Language Pathology, University of Toronto, 500 University Avenue, Toronto, ON, M5G 1V7, Canada

1. INTRODUCTION

To date, only a handful of studies have investigated the acoustic parameters of affect in speech (Hammerschmidt & Jürgens, 2007; Pell, 2001; Banse & Scherer, 1996; Williams & Stevens, 1972). They have demonstrated an increase in fundamental frequency (Williams & Stevens, 1972; Hammerschmidt & Jürgens, 2007; Banse & Scherer, 1996) and amplitude when speakers used a hot angry voice compared to a neutral voice or other affective states (Hammerschmidt & Jürgens, 2007; Banse & Scherer, 1996).

The present study sought to identify changes in three acoustic parameters when participants adopted a hot angry affective state. This study investigated changes in fundamental frequency, amplitude, and nasalance in two sentence types, a real oral sentence (without any nasal phonemes) and a non-word balanced sentence (with oral and nasal phonemes).

We hypothesized that mean fundamental frequency (f_0) and sound pressure level (SPL in dB) would increase in sentences spoken with a hot angry affect. Conversely, we expected nasalance to decrease with a hot angry affect. We posited that increases in vocal cord tension would entail a higher f_0 and dB SPL. Also, synergistic recruitment of velopharyngeal muscles would result in a tighter velopharyngeal port closure and less transmission of acoustic energy through the nasal passage.

2. METHOD

We chose to investigate if the parameters mean f_0 , dB SPL and nasalance conformed to these physiologically driven hypotheses in the speech of normal speakers. We recruited 10 speakers with prior acting experience by posting billboard advertisements at the University of Toronto. The participants were six female and four male speakers with a mean age of 28.9 years (range 18 to 43 years). All spoke standard Canadian English.

The participants read two repetitions of each sentence in the two experimental conditions. The sentences were counterbalanced according to emotional state (neutral versus hot angry) and type (non-nasal versus balanced). Each set contained four sentences, i) the neutral non-nasal sentence, "He had two rock lizards" (Lewis & Watterson, 2003), ii) the neutral balanced non-word sentence, "Hat sundig pron you venzy" (Banse & Scherer 1996), iii) the hot angry non-nasal sentence, "He had two rock lizards" and iv) the hot

angry balanced non-word sentence "Hat sundig pron you venzy".

Participants recorded the sentences using the NasalView headset (22.05 kHz, 16 bit recording, Tiger Electronics Inc., Seattle, WA). This headset has a sound separation plate positioned on the participant's prolabium to permit separate recordings of speech sounds from the nose and the mouth. Preamplification and soundboard settings were constant for both sets and experimental conditions. We used the Praat speech analysis software to analyse mean f_0 and amplitude in dB (Boersma & Weenink, 2004). The NasalView software was used to extract the nasalance values for all 80 sentences.

3. RESULTS

We computed differences in mean f_0 , dB SPL and nasality using repeated measures ANOVAs with two factors, emotion and sentence type (SAS, 9.1; see Table 1).

Table 1. Two factor ANOVA table for three dependent measures (pitch, amplitude and nasalance)

Measure	Emotion	Type	Emotion by Type
Pitch (f_0)	$p < .01$ ($F=278.50$)	NS	NS
Amplitude (dB SPL)	$p < .01$ ($F=109.12$)	$p < .01$ ($F=63.08$)	NS
Nasalance (dB SPL)	$p < .01$ ($F=12.54$)	$p < .01$ ($F=935.01$)	$p < .01$ ($F=268.67$)

There were significant differences in mean f_0 , dB SPL and nasalance for emotion. Similarly, there were significant differences in mean amplitude and nasalance according to sentence type. The ANOVA for nasalance also revealed an interaction between sentence type and emotion (see Figure 1).

A review of the changes in mean values according to sentence type and emotion revealed higher f_0 and dB values with a hot angry affect in both the non-nasal sentence and the non-word phonemically balanced sentence (see Table 2). Nasalance, on the other hand, remained unchanged in the hot angry non-nasal sentence, but decreased in the angry

non-word sentence. An incidental finding was the higher mean dB value in the non-word sentence.

Table 2. Mean f0, dB SLP and nasalance values by sentence type.

Measure	Real Oral Sentence		Non-word Balanced Sentence	
	Neutral	Angry	Neutral	Angry
Pitch (f0)	165.71	218.92	165.64	211.68
Amplitude (dB SPL)	64.42	74.14	67.06	76.21
Nasalance (dB SPL)	23.72	22.04	47.57	38.62

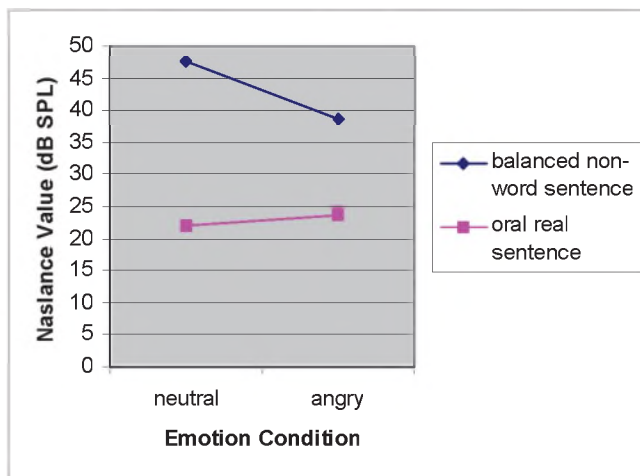


Figure 1. Interaction between sentence type and emotion.

4. DISCUSSION

The findings that mean f0 and dB increase with a hot angry affect support the previous literature (Williams & Stevens, 1972; Hammerschmidt & Jürgens, 2007; Banse & Scherer, 1996) and confirm our hypothesis. The incidental finding of a mean increase in dB in the non-word sentence in both the neutral and hot angry condition may relate to increased effort producing an unfamiliar sentence without semantic content.

The decrease in the mean nasalance in the hot anger condition in the phonemically balanced non-word sentence points to a more tightly sealed velopharyngeal port, also suggestive of increased physiological effort. There was no similar decrease in nasality in the non-nasal sentence for the hot angry condition.

Given that there was no nasal consonant in the real sentence, the velopharyngeal port was already maximally closed in the non-nasal condition. Consequently, an even tighter seal

in the hot angry condition did not lead to decreased nasal emission.

These findings reflect velar-laryngeal interactions. Adopting a hot angry affective state or repeating a non-word phrase resulted in increased amplitude. The hot angry affect led to a concomitant increase in pitch and decrease in nasalance.

Further research needs to be conducted to determine if adopting a more relaxed or calmer affective state alters speech acoustics such that amplitude and pitch decrease with a concomitant increase in nasalance.

In the future, this type of research may prove useful in the remediation of impaired interactions between subsystems, such as the larynx and the velum. Using non-word sentences or adopting a hot angry affective state may provide a means to stimulate increased vocal amplitude or decreased nasalance.

ACKNOWLEDGEMENTS

We thank Jayanthi Sasisekaran for testing participants and for extracting nasalance values for this study.

REFERENCES

- Awan, S.N. (1996). Development of a low-cost nasalance acquisition system. In: Powell T, Ed. Pathologies of Speech and Language: Contributions of Clinical Phonetics and Linguistics. New Orleans, LA: International Clinical Linguistics and Phonetics Association, 211-217.
- Awan, S.N. (1998). Analysis of nasalance: NasalView. In: Zeigler W., Deger K, Eds. Clinical Phonetics and Linguistics. London: Whurr-Publishers, 518-525.
- Banse, R. and Scherer, K.R. (1996). Acoustic Profiles in Vocal Emotion Expression. *Journal of Personality and Social Psychology*, 70, 614-636.
- Boersma, P. and Weenink, D. (2004). Praat. Doing phonetics by computer (version 4.2). [Computer Program]. Retrieved from <http://www.praat.org/>.
- Hammerschmidt, K. and Jürgens, U. (2007). Acoustical Correlates of Affective Prosody. *Journal of Voice*, 21, 531-540.
- Lewis, K.E. and Watterson, T. (2003). Comparison of nasalance scores obtained from the Nasometer and the NasalView. *Cleft Palate Craniofacial Journal*, 40, 40-45.
- Pell, M.D. (2001). Influence of emotion and focus location on prosody in matched statements and questions. *Journal of the Acoustical Society of America*, 109, 1668-1680.
- SAS Software, Version 9.1 of the SAS System for Windows. Copyright (2004). Trademarks of SAS Institute Inc. Cary, NY, USA.
- Williams, C.E. and Stevens, K.N. (1972). Emotions and Speech: Some Acoustical Correlates. *Journal of the Acoustical Society of America*, 52, 1238-1250.

AUTHOR NOTES

This work was conducted at the Voice and Resonance Laboratory, while Heather Flowers was completing M.H.Sc. studies in the Department of Speech Language Pathology at the University of Toronto. The current address is 500 University Avenue, Toronto, ON, M5G 1V7.

UNDERWATER ACOUSTIC LOCALIZATION OF MARINE MAMMALS IN AN UNCERTAIN ENVIRONMENT

Brendan Rideout¹, Stan E. Dosso¹, and David Hannay²

¹School of Earth and Ocean Sciences, University of Victoria, 3800 Finnerty Rd., Victoria, BC, Canada, V8P 5C2, bprideou@uvic.ca

²JASCO Applied Sciences Ltd., 2101-4464 Markham St., Victoria, BC, Canada, V8Z 7X8

1. INTRODUCTION

This paper describes an MSc thesis research project aimed at three-dimensional acoustic localization and tracking of vocalizing marine mammals, with the ultimate goal of providing a more reliable method for assessing temporal variations in walrus population distributions and habitat usage. Large changes in the degree of summer ice coverage in the Chukchi Sea over the last several decades, coupled with recent increased seismic survey activity in the waters surrounding Alaska, have the potential to impact marine mammals, including walrus.

The Pacific Walrus (*O. rosmarus divergens*, see Fig. 1) has been chosen as the focus for this localization study due to their presence in the study area (the Chukchi Sea northwest of Alaska, see Fig.2), and the acoustic characteristics of their calls. Current methods for locating and counting marine mammals are based on visual observations from vessels and aircraft. These visual methods require daylight and good visibility to be effective and only account for animals on the surface. It is hoped that acoustic methods can augment visual methods; for instance, aircraft-based visual methods could be employed over wide areas where acoustic means would be infeasible, whereas acoustic means could be used near feeding areas.

Two approaches to acoustic localization will be considered. The first approach makes use of several independent (but time synchronized) data acquisition systems, each involving a single omni-directional hydrophone and recorder, which are deployed on the seabed at separations of up to a few kilometres. The location of a vocalizing walrus can be estimated using acoustic triangulation if its call is recorded by at least three such systems. This approach is straightforward and fairly standard, but requires the expense and effort of deploying and recovering multiple acoustic systems. The second, more novel, approach involves evaluating the use of a single data acquisition system for three-dimensional localization. In this case, the acquisition system includes a directional sensor to provide bearing, and a vertical array of omni-directional hydrophones which can be used to estimate source range and depth from the multi-path acoustic arrivals. An important issue in this approach is that acoustic localization with a vertical array requires knowledge of the physical properties of the ocean environment (water column sound-speed profile and seabed geoaoustic parameters), which may not be readily available.



Figure 1: Pacific Walrus mother and calf (USGS).

Acoustic field trials for this project will be carried out in the summer and fall of 2009, and hence data are not yet available for analysis. Rather, the remainder of this paper will describe the experiment plan and data analysis methodology proposed for the project.

2. FIELD STUDYS

The first component of the field work will take place in the Chukchi Sea in the summer/early fall of 2009. Three single-hydrophone ocean-bottom data acquisition systems will be

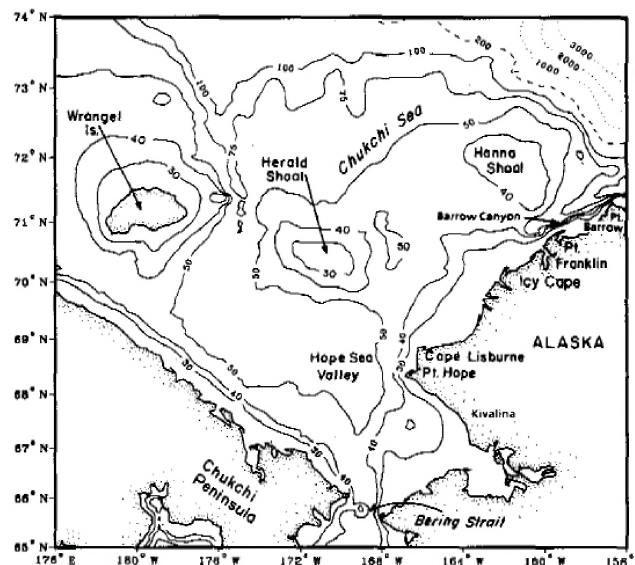


Figure 2: Field study location in the Chukchi Sea [2].



Figure 3 : Vertical array/directional hydrophone data acquisition system.

deployed by ship in the northeast Chukchi Sea in the vicinity of the Hanna Shoal, which is known to be a summer walrus gathering place (Fig. 2). The systems will be time synchronized onboard ship before deployment in early August, and will then autonomously record acoustic data at a sampling rate of 32 kHz for a 3-month period before being recovered in the early fall. The second field-work component will take place in the fall of 2009. A vertical array of omni-directional hydrophones together with a directional sensor (see Fig. 3) will be deployed outside Halifax harbour. Walrus calls will be played over an underwater transducer at a series of known locations. Data recorded in both field studies will be used to develop and test walrus call detection and localization schemes, described in the following section.

3. DATA PROCESSING AND INVERSION

The first aspect of acoustic marine mammal localization requires detecting mammal calls within a noisy time series. A sample walrus click-train waveform and spectrogram are shown in Fig. 4. A Gaussian mixture model, which has been applied to the detection of bat calls [3], will be considered for walrus call detection.

Two approaches to localizing an ocean acoustic source will be considered: matched field processing and ray travel-time inversion. In matched field processing, source range and depth are estimated by matching the measured complex (frequency-domain) acoustic fields measured at an array of hydrophones to replica acoustic fields predicted for a grid of candidate locations using a numerical propagation model. The range-depth grid point which produces the field that most closely correlates to the measured field is selected as the most likely source location. Propagation modeling requires specifying ocean environmental parameters; if these are unknown they can be included in an augmented search procedure usually involving numerical optimization or integration.

In ray travel-time inversion, relative arrival times along ray paths from the source to a series of receivers are inverted for source location. The data can include arrival times for direct-path rays and/or surface- and seabed-reflected rays, if these can be identified in the recorded time series. An iterative linearized inversion scheme, initiated from the results of a rough grid search, will be used for travel-time inversion [1]. If the water-column sound speed and water depth are not well known, these parameters can be inverted for in addition to the source location, provided sufficient acoustic paths exist in the data set.

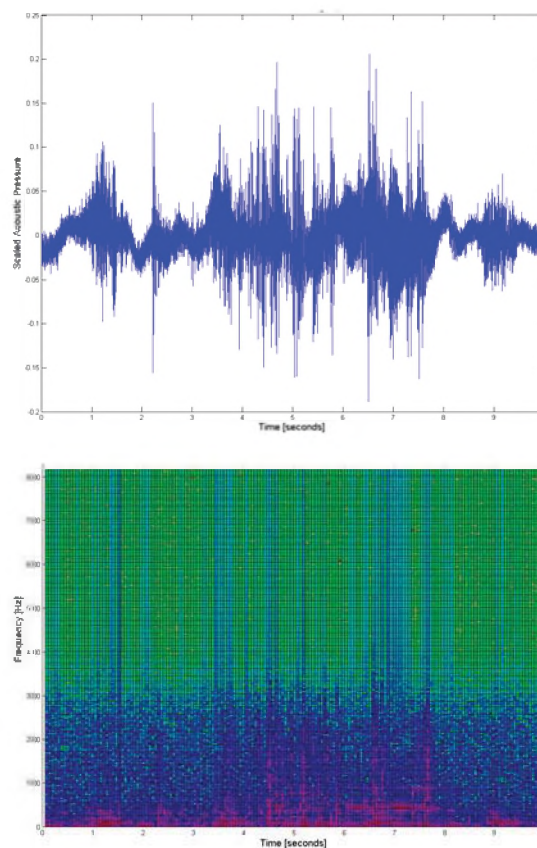


Figure 4 : Walrus click-train time series (upper) and spectrogram (lower).

REFERENCES

- [1] Dosso, S. E. *et al.* (1998). High-precision array element localization for vertical line arrays in the arctic ocean. *IEEE J. Oceanic Eng.*, 23, 365-379.
- [2] Feder, H. M. *et al.* (1994). Mollusks in the Northeastern Chukchi Sea. *Arctic*, 47, 145-163.
- [3] Skowronski, M. D. and J. G. Harris (2006). Acoustic detection and classification of microchiroptera using machine learning: Lessons learned from automatic speech recognition. *J. Acoust. Soc. Am.*, 119, 1817-1833.

ANDEAN ULTRASONICS: BIOACOUSTICS OF TWO TROPICAL MONTANE KATYDIDS

Glenn K. Morris

Dept. of Biology, University of Toronto Mississauga, L5L 1C6 glenn.morris@utoronto.ca

1. INTRODUCTION

Ultrasonics are waves of airborne sound longer than ~1.7 cm (20 kHz) that, by definition, cannot be heard. The term distinguishes sound frequencies above the upper limit of human frequency sensitivity. But of course a dog, a bat or a cricket may hear such waves: animal sensory systems have ‘unhuman’ capacities. Ultrasonic is only a convenient descriptor for ‘wavelengths above the limit of detection’, applied illogically to animal species that hear these waves ‘just fine’.

Some species of the singing insects known as ‘katydids’ (Tettigoniidae) produce and hear ultrasonics. The males make their sounds perching on plants. They rub their forewings together, which excites the oscillation of forewing membranes as sound radiators. Male calls travel through grasses, shrubs or forests to be heard by distant females: these females localize the singer and approach him for mating. Why such carriers should be ultrasonic is puzzling, because ultrasonic wavelengths interact poorly with plant environments; they don’t carry well, and being higher energy at a given sound level, are relatively more costly to produce than audio frequencies. Ultrasonics would not seem a good choice for a beacon directed to potential mates at long range.

Two such katydid species comprise the genus *Myopophyllum*, both with a very high ultrasonic carrier. This is about the physical structure of their sounds and how they have evolved to make them ‘elastically’. At present the ‘why’ of their ultrasonics remains a puzzle.

2. *M. SPECIOSUM* CALL GENERATION

Myopophyllum. speciosum (Fig. 1) has a duty cycle of <2% (Morris et al. 1994): on average, every 8s they produce a call that consists of two brief pulse trains. Together the two trains are about 150 ms in duration. The first train has 9 pulses in 50 ms; after an interval of 50 ms a second train of 12 pulses is given.

The pulses in the trains are quite short, sinusoidal and set relatively far apart in time. They repeat at a rate of about 230/s. The spectrum of the call is a single narrow (high-Q) ultrasonic peak averaging 81,000 Hz. This peak frequency is a stable feature of each individual caller, but varies widely between individuals: from 67 to 95 kHz (Morris et al. 1994).

3. A NEW SPECIES AND ITS SONG

In 2003 on a trip to Ecuador, within the valley system on the eastern side of the Ecuadorean Andes where *M. speciosum* occurs, we discovered a second species of *Myopophyllum* (Fig. 2). As for many tropical katydids, this species remains



Figure 1. Male of *Myopophyllum speciosum* calling from understory near Baeza, Ecuador.

undescribed and unnamed. I refer to it here as *sp. nov.* An adult male was taken at an elevation of about 2000 m, near Cosanga (San Isidro Cabanas). Later nearby, we located several more specimens by the Rio Alisa.



Figure 2. *Myopophyllum sp. nov.* male from montane rainforest above Cosanga, Ecuador.

We recorded and analysed the call of the new species. Like *M. speciosum* it has a low duty cycle, with two short trains of pulses separated by about a tenth of a second (Fig. 3 top trace). Each pulse within these trains is a sinusoid (Fig. 3 middle trace) of variable amplitude. The spectrum of this call has a single major peak of relatively high Q, centred on 73.4 kHz (Fig. 3 bottom trace). Another specimen was recorded and analysed with high-speed video (Montealegre et al. 2006). Together the two males gave an average carrier of 65.5 kHz (Montealegre et al. 2006, see their fig.4).

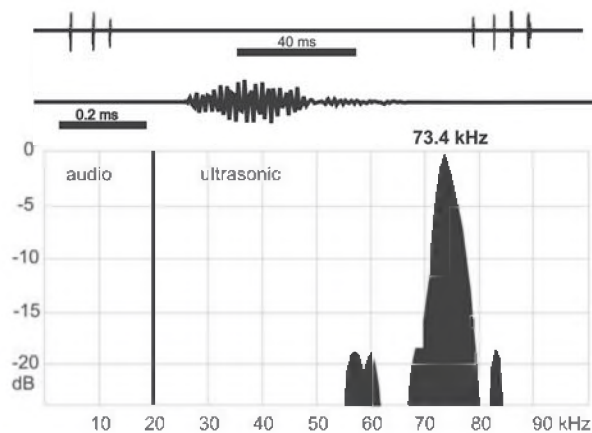


Figure 3. Analysis of one call of *M. sp. nov.* showing time domain and high-Q spectrum.

4. RESONANT VS NONRESONANT

The high-Q spectral peak of both these *Myopophyllum* species indicate generation by a resonant mechanism. Katydid sounds are made by friction of a file and scraper, these structures being on their forewings. Teeth comprising the file on the underside of one wing are traversed by the scraper, an edge of the other wing. This takes place in two ways termed resonant and nonresonant (Elsner & Popov 1978, Montealegre & Morris 1999).

In nonresonant mechanisms each tooth-scraper contact excites forewing membranes that oscillate at inherent frequencies, creating a transient pulse. Many such short complex-wave pulses per wingstroke (one pulse, one tooth) produce a band spectrum. In resonant mechanisms radiator movement gates passage of the scraper along the file. The emitted sound frequency in resonant systems is thus the same as the tooth contact rate: one wave, one tooth.

5. DISTORTION SPEEDS SCRAPER

But how is a resonant mechanism possible at 70 kHz? One cannot contact teeth at rates of 70,000 per second using just the muscle power closing the wings. The rates of tooth contact exhibited by a number of katydid species, notably *Myopophyllum n. sp.* (Montealegre et al. 2006), are beyond the capacity of wing muscles to produce. And high-speed video shows that the bending of the scraper plays an important role. The forewings actually slow during the interpulse intervals and speed up during the pulse (Montealegre et al. 2006). To an extent scraper and wing are decoupled, allowing the scraper to move at different velocities: it lodges briefly behind a file tooth and becomes bent by ongoing progress of the wings. Its distortion stores elastic energy, which when the scraper slips free, contributes to speeding scraper advance. The scraper drives at a rate of 70000 teeth/s at intervals along the file – albeit for only a few teeth at a time.

6. ADAPTATIONS

The songs of the two species are similar and in contrast to most katydid species, it would be difficult to use the physical structure of their calls to distinguish them. So it is unlikely this signal serves to keep females from approaching a wrong-species male. Of course confusion would only arise if the two occupy the same region. They have not yet been found together, and in this mountainous area may well occupy ranges at different elevations.

The best explanation of ultrasonic adaptiveness in *Myopophyllum* remains eavesdropping (Morris et al. 1994). Some bats hunt katydids as food, guided by katydid calls (Belwood & Morris 1987). Ultrasonic frequencies attenuate more severely with distance than audio ones, so using them could reduce singer vulnerability to bats. The paradox of calling to distant listeners with a short-range signal carrier may thus be a trade-off between attracting females and limiting vulnerability to predators. The higher the ultrasonics, the more steeply loudness decreases with increasing distance from the caller. The ears of females of *Myopophyllum* may be tuned to these carriers, helping to compensate for the female's increased localizing difficulties (Mason et al. 1991).

REFERENCES

- Belwood, J.J. & Morris, G.K. (1987) Bat predation and its influence on calling behavior in neotropical katydids. *Science* 238: 64-67.
- Elsner, N. & Popov, A. (1978) Neuroethology of acoustic communication. *Advances in Insect Physiology* 13: 229-355.
- Mason, A.C. et al. (1991) High ultrasonic hearing and tympanal slit function in rainforest katydids. *Naturwissenschaften* 78: 365-367.
- Montealegre, F-Z et al. (2006) Generation of extreme ultrasonics in rainforest katydids. *Journal of Experimental Biology* 209: 4923-4937.
- Montealegre, M-Z & Morris, G.K. (2004) The spiny devil katydids, *Panacanthus* Walker (Orthoptera: Tettigoniidae: an evolutionary study of acoustic behaviour and morphological traits. *Systematic Entomology* 29: 21-57.
- Montealegre, F.Z. & Morris, G.K. (1999) Songs and systematics of some Tettigoniidae from Colombia and Ecuador I. Pseudophyllinae (Orthoptera). *Journal of Orthoptera Research* 8: 163-236.
- Morris, G.K. et al. (1994) High ultrasonic and tremulation signals in neotropical katydids (Orthoptera: Tettigoniidae). *Journal of Zoology London* 233: 129-163.

ACKNOWLEDGEMENTS

Thanks to Fernando Montealegre, Peter Wall and Dita Klimas, working companions on the field trip that revealed the new species. Support for this research was by the Natural Sciences and Engineering Research Council of Canada via a grant to G.K.M.

MEASURING SCATTERING IN APOPTOTIC CANCER CELLS USING ULTRA HIGH FREQUENCY ACOUSTIC MICROSCOPY

Eric Strohm, Michael Kolios¹

¹Dept. of Physics, Ryerson University, 350 Victoria St., Toronto, Ontario, Canada, M5B 2K3, mkolios@ryerson.ca

1. INTRODUCTION

Apoptosis, or programmed cell death, is a method of cellular self-disassembly to minimize damage to nearby cells [1]. The cell breaks into apoptotic bodies to contain the toxic intracellular components, which are then removed by phagocytes. Apoptosis is triggered by toxins, ionizing radiation or physical damage. Proteins begin breaking down the cell, and eventually the cell collapses onto itself. Morphological features such as membrane blebbing and protrusions from the cell are typically observed. Once initiated, the entire apoptotic process can take less than an hour, and is reversible until collapse occurs.

It is important to understand how apoptosis occurs, as deregulation of apoptosis can result in a variety of diseases, such as Alzheimer's, AIDS, autoimmune disorders and tumours. Of particular importance is the study of cancer cells, which have mutated to prevent apoptosis from occurring. Understanding apoptosis is important to treating cancer and other diseases related to apoptotic dysfunction.

High frequency ultrasound (10-60 MHz) has been used to study apoptosis, where an increase in the ultrasound backscatter intensity was observed for cells treated with a chemotherapeutic agent to induce apoptosis, compared to untreated cells [2]. However individual cells cannot be resolved at these frequencies, and the source of increased backscatter could not be determined. Acoustic microscopy uses ultra-high frequencies (100+ MHz) with resolutions approaching 1 μm , which is able to resolve cellular and subcellular features [3]. Ultrasound can be used to image the mechanical properties of the cell, which may give insight into the non-visible changes that are occurring within the cell, an advantage over optical observations.

2. METHOD

MCF-7 breast cancer cells were grown in 175 cm² cell culture flasks (Sarstedt, Newton, NC, USA) using Dulbecco's modified eagle's medium (ATCC, Manassas, VA, USA) with 10% fetal bovine serum and 0.1% insulin. Cells were incubated at 37°C with 5% CO₂ and passed every 3 days to maintain exponential growth. Cells were dissociated using trypsin and transferred to Lab-Tek II chambers (Nunc, Germany). 15 hours prior to experimentation, the medium was replaced with a solution consisting of the DMEM cell culture medium (without serum or insulin), 3 mg/mL of caffeine and 20 ng/mL of paclitaxel to induce apoptosis. During experimentation, cells

were kept at a constant temperature of 36°C with 5% CO₂.

The acoustic microscope (Kibero GmbH, Saarbrücken, Germany) is equipped with an acoustic and optical component, which allows for simultaneous optical and acoustic measurements. The transducer can be visually positioned over a region of interest for a precise alignment of the transducer and the sample. The electronics consist of a monocyte pulse generator at 300 MHz with a 100% bandwidth, 10 V_{pp} amplitude and a 500 kHz pulse repetition rate. The RF signals are sampled at 8 GHz, amplified by a 40 dB amplifier and filtered to remove noise. The transducer used in this study had a center frequency of 375 MHz with a -6dB bandwidth of 42% and a F-number of 1.

Cells were imaged in two different states; before apoptotic collapse (denoted normal) and after collapse (denoted apoptotic). The transducer was positioned over a selected cell, and the ultrasound RF signal was recorded every 10 seconds for a period of 900 seconds. This process was repeated for four cells in each state.

3. RESULTS

An a-scan showing a single RF-signal measurement as a function of ultrasound propagation time is shown in figure 1. The ultrasound backscatter from the cell (A) and the substrate (B) are clearly visible and separated in time. The region from approximately 1500 to 1520 ns, the area including the cell signal but excluding the substrate signal, was integrated to give the backscatter intensity over the cellular region. This calculation was done for each measurement for each cell. The integrated backscatter intensity for a typical normal and apoptotic cell as a function of measurement time is shown in figure 2.

The correlation coefficient of the RF-signal was calculated by comparing the first RF-line, restricted to the cellular backscatter region only, to the other RF-lines at the same location but at different measurement times for that particular cell. The correlation coefficient is a measure of the similarity of two signals; identical signals have a correlation coefficient of 1. The correlation coefficient as a function of measurement time is shown in figure 3 for a typical normal and apoptotic cell. The normal cells had an average correlation coefficient of 0.93 ± 0.05 , compared to 0.68 ± 0.17 for the apoptotic cells.

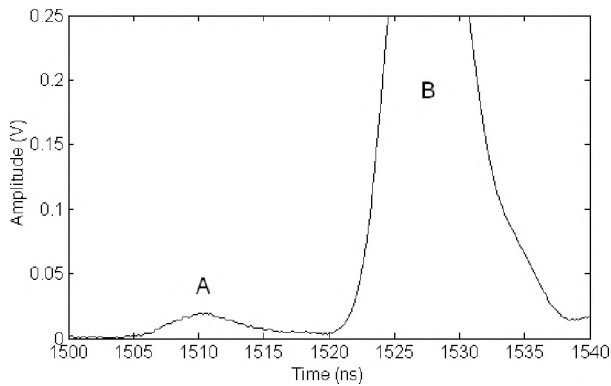


Figure 1. A single RF signal showing the ultrasound backscatter from the cell (A) and substrate (B).

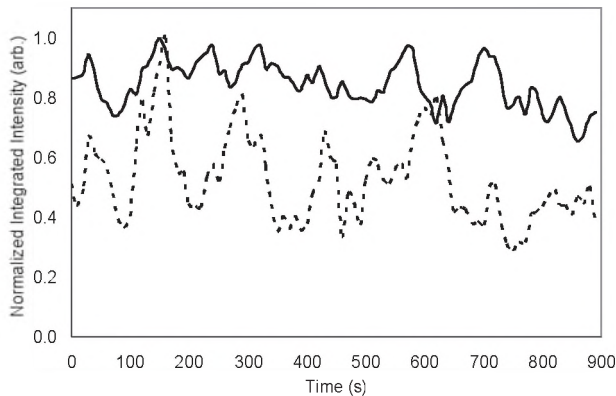


Figure 2. Typical integrated backscatter intensity as a function of time for a normal cell (solid line) and apoptotic cell (dotted).

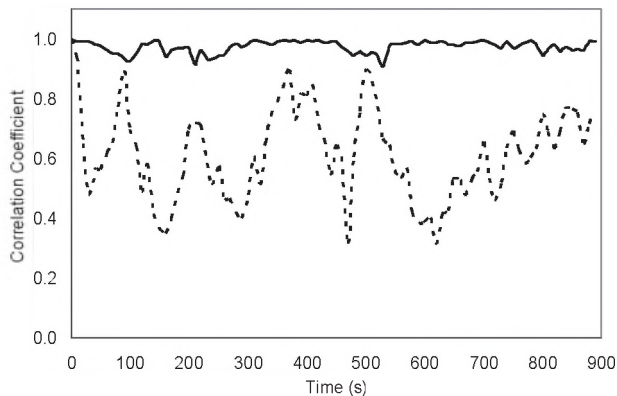


Figure 3. Typical correlation coefficient as a function of time for a normal cell (solid line) and apoptotic cell (dotted).

Table 1. Transient analysis results of MCF-7 cells.

Measurement	Normal Cells	Apoptotic Cells
Correlation Coefficient	0.93 ± 0.05	0.68 ± 0.17
Integrated Intensity	0.75 ± 0.10	0.57 ± 0.14
Maximum Amplitude	0.529 ± 0.003	0.498 ± 0.004

In addition to the cellular backscatter, the maximum amplitude from the substrate was also measured which is an indication of the attenuation in the cell. Table 1 summarizes the average correlation coefficient, integrated intensity and

maximum amplitude for four cells measured in each state.

4. DISCUSSION

The variations of the integrated backscatter intensity were larger for normal cells compared to apoptotic cells, as shown in figure 2. The average normalized backscatter integrated intensity was 0.75 ± 0.10 for normal cells, compared to 0.57 ± 0.14 for apoptotic cells. The larger variations and higher standard deviation for apoptotic cells indicate an increase in activity within the cell during the measurement period.

A more thorough method of measuring the variations in the cell is by looking at the correlation coefficient. An average correlation coefficient of 0.93 ± 0.05 was calculated for normal cells, indicating a small amount of variation of the RF signal over time. For apoptotic cells, an average correlation coefficient of 0.68 ± 0.17 was calculated, which is lower than for normal cells. In addition, the standard deviation is higher, indicating extensive variations are occurring within the cell (Figure 3).

The maximum amplitude from the substrate was 0.529 ± 0.003 for normal cells compared to 0.498 ± 0.004 for apoptotic cells. The low standard deviation suggest the attenuation in the cell is not changing. However variations in the backscatter signal from the apoptotic cells indicate extensive activity occurred within the cell over a time period of only seconds. Morphological changes of apoptotic cells were observed optically in real time, which correlates with the activity observed acoustically. Therefore in addition to surface variations, the interior of the cell also undergoes extensive structural changes during apoptosis. The variation in ultrasound backscatter but not attenuation is expected since backscatter is primarily influenced by changes in structure whereas attenuation by changes in composition. In this work, we have been able to measure these variations in mechanical properties using acoustic microscopy.

ACKNOWLEDGEMENTS

This research was undertaken, in part, thanks to funding from the Canada Research Chairs Program, Canada Foundation for Innovation, the Ontario Ministry of Research and Innovation and Ryerson University. The authors would like to acknowledge Min Rui, Eike Weiss and Arthur Worthington for technical support during this study.

REFERENCES

- [1] Wyllie, A.H., Kerr, J.F.R., Currie, A.R., "Cell death: The significance of apoptosis," *International Review of Cytology*, vol. 68, pp. 251-306, 1980.
- [2] Tunis, A.S., Czarnota, G.J., Giles, A., Sherar, M.D., Hunt, J.W., Kolios, M.C., "Monitoring structural changes in cells with high-frequency ultrasound signal statistics," *Ultrasound in Medicine and Biology*, vol. 31, no. 8, pp. 1041-1049, 2005.
- [3] Briggs, A., *Acoustic Microscopy*, Oxford: Clarendon Press, 1992.

A COMPARISON OF IMAGING MODALITIES TO MONITOR THERMAL AND MECHANICAL ULTRASOUND TISSUE THERAPIES

Arthur Worthington, Sankar Narasimhan, Jahan Tavakkoli, and Michael C. Kolios
Department of Physics, Ryerson University, Toronto, Canada mkolios@ryerson.ca

1. INTRODUCTION

High-intensity ultrasound can be used to produce localized tissue destruction. For long enough pulses, cytotoxic effects are thermal in origin. For higher intensities, shorter pulses, cytotoxic effects result from mechanical tissue destruction (also known as histotripsy). For these modalities to gain clinical acceptance, a noninvasive real-time monitoring technique is required to assess the tissue damage. In this work, conventional ultrasound imaging, transmission ultrasound imaging and photoacoustic imaging are compared as to their ability to discriminate lesions created by high intensity therapeutic ultrasound, resulting in both thermally and mechanically induced tissue destruction.

2. METHOD

2.1 High Intensity Focused Ultrasound

High intensity focused ultrasound (HIFU) was used to create thermal lesions by coagulating the target tissue. A 1MHz, F#0.8, and 125mm aperture diameter transducer (Imasonic, Besançon, France) was used to deliver the therapeutic ultrasound energy. Derived from computer simulations and verified by measurements, the full width at half maximum (FWHM) lateral and axial dimensions of the focal spot in water are estimated as 1.7mm and 6.8mm respectively. A SONIX RP® clinical imaging system (Ultrasonix Medical Corp., Richmond BC, Canada) was used to monitor lesion growth in real time pre-, during and post-exposure via its endocavity 6MHz convex array probe which was coaxially mounted in the centre of the therapy transducer. An AFG3101 arbitrary function generator (Tektronix Inc., Beaverton, OR) provided the bursts of RF to an AG1012 RF power amplifier (T&C Power Conversions Inc., Rochester, NY) which drove the therapeutic transducer. A typical HIFU exposure was 30 s of 200W (electric) delivered in bursts of 1000 cycles every 2 ms. Given the 60% efficiency measured by the manufacturer and the estimated FWHM dimensions, this leads to a spatial peak pulse average intensity (I_{SPPA}) of 5286 W/cm² and peak acoustic pressure (P_{max}) of 12.6 MPa at the focus in water. After the exposures, the tissue sample was also imaged using an ultrasound transmission camera (AcoustoCam®, Imperium Inc. Rockville, MD) and a photoacoustic Small Animal Imager (Imagio®, Seno Medical Instruments, Inc.). The Imagio® uses a Nd:YAG laser to emit 6ns, 50mW, 1064nm pulses to produce thermoelastic expansion in the target tissue. This expansion produces ultrasound pressure signals which are received by a 4 element annular array ultrasound transducer centered at frequency of 4.5MHz. After all the imaging techniques had

been tried, the tissue was cut open so that lesions could be measured and photographed.

2.2 Histotripsy

The same transducer system was used to create histotripsy lesions: short bursts with high peak negative pressure fragmented the target tissue without raising the temperature. A typical exposure was 3 or 4 minutes of 500W (electric) delivered in bursts of 15 cycles at a PRF of 100 Hz. This corresponds to $I_{SPPA}=13215$ W/cm², and $P_{max}=19.7$ MPa at the focus. The exposure was monitored in real time via the B- and Power Doppler modes of the Ultrasonix® scanner. The lesions were monitored post exposure using the imaging systems described in Section 2.1.

3. RESULTS

3.1 Coaxial mounted transducer array

The coaxially mounted transducer array, running in B-mode, clearly showed HIFU lesions forming in real time during and after the exposure. Some of the backscattered signal was due to the thermal coagulation of the tissue and some was due to cavitation. The signal from cavitation faded away in a few seconds after the exposure.

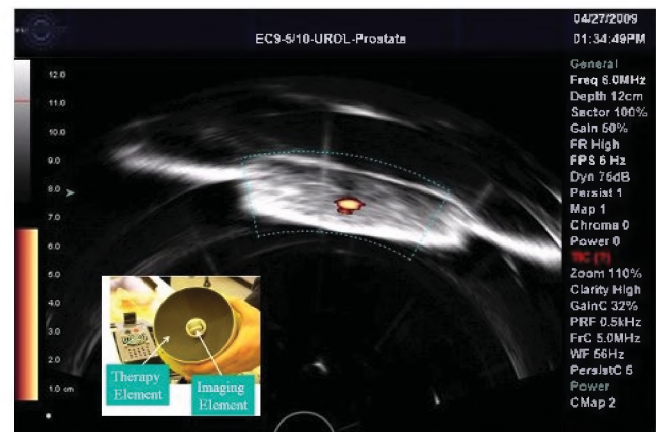


Figure 1. Typical B-mode and Doppler display during histotripsy exposure of pork muscle tissue showing the Doppler signal at the focal point (centre of tissue) as well as a previous HIFU lesion on the left side of the tissue. There are cm markers on the left side. The therapy beam comes from the bottom of the image. The insert shows the coaxial transducers.

When delivering histotripsy, the coaxial transducer running in power Doppler mode showed the location of the focal spot during the exposure, likely due to the motion of tissue

and cavitation bubbles (Fig. 1). It was not possible to see a lesion after a single histotripsy exposure.

3.2 Acoustic Camera

Figure 2 shows the image of a typical HIFU lesion. Although the HIFU lesions were usually detectable using the acoustic camera, it was rarely possible to find a single histotripsy lesion with this technique.

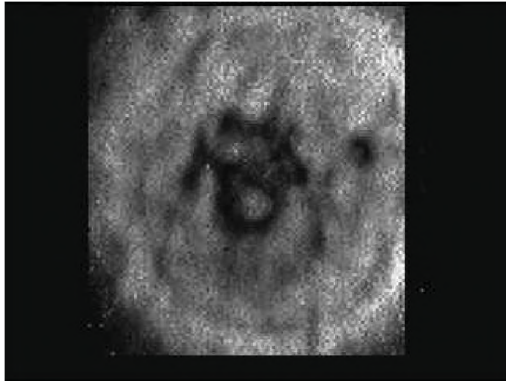


Figure 2. Acoustic Camera image of a HIFU lesion in the pork muscle tissue. The image is 2.4 cm wide × 2.6 cm high. The therapy beam is perpendicular to the image.

3.3 Opto-acoustic Imaging System

The Imagio® opto-acoustic imaging system combines the advantages of optical and ultrasound imaging. The laser-induced ultrasonic waves in tissues are generated by the thermoelastic expansion of optical absorbers, allowing ultrasonic detection of millimeter-sized optical inhomogeneities (Fig. 3).

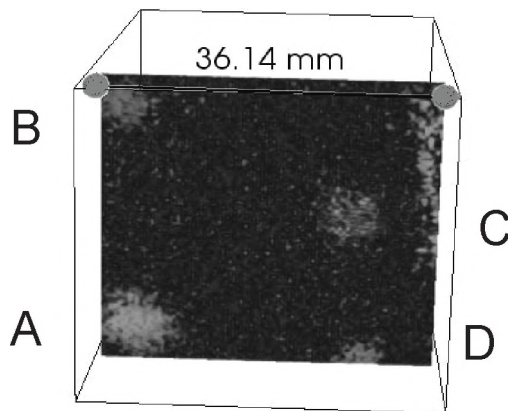


Figure 3. Opto-Acoustic image of four lesions in different depths in pork muscle tissue. A and B are HIFU lesions : C and D are histotripsy lesions.

3.4 Physical examination of tissue.

After employing all imaging techniques, the tissue was cut open and photographed. HIFU lesions were easily observed

(Fig. 4.), but it was never possible to see a single histotripsy lesion.

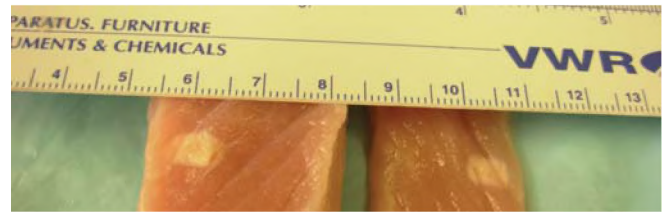


Figure 4. Treated tissue cut and folded open, showing HIFU lesion in its middle, not visible before slicing tissue. The therapy beam axis was horizontal in this photo.

4. DISCUSSION

Two different high intensity ultrasound lesions were formed and three modalities adopted in assessing the lesions. HIFU lesions were clearly shown in both B- and power Doppler modes by the coaxial imaging transducer; both during and post exposure. The same transducer, in power Doppler mode, clearly shows where the histotripsy lesions are being formed during exposure. A single histotripsy lesion is not visible post exposure due to its small size. The AcoustoCam® sometimes showed the HIFU lesions and sometime did not. Occasionally, a histotripsy lesion could be seen with this device. In transmission imaging, the signals through the entire depth of the sample are superimposed at the detector, eliminating the ability to distinguish between small volumes with little difference in signal. An Opto-Acoustic imaging system gave clear images of HIFU lesions and occasionally a poor image of a histotripsy lesion. An array of histotripsy lesions was necessary to create a lesion that was large enough to be seen reliably on any of the imaging systems. After all imaging techniques had been employed, the phantoms and tissues were cut open and the lesions photographed. More work is on progress to increase the sensitivity and specificity of the imaging methods to detect and to image histotripsy lesions.

REFERENCES

- Bailey et al., "Physical mechanisms of therapeutic effect of ultrasound", *Acoustical Physics*, Vol. 49, No. 4, 2003, pp. 437–464
- Winterroth, et al., "Tissue Fragmentation by Pulsed Cavitation Ultrasound-Histotripsy", *IEEE Int.Ultrasonics Symposium*, New York, 2007
- Oraevsky et al., "Optoacoustic imaging of blood for visualization and diagnostics of breast cancer", *Biomedical Optoacoustics III*, *Proceedings of SPIE Vol 4618*, 2002

ACKNOWLEDGEMENTS

This work was partially supported by grants from NSERC, /CIHR, Health Research Project 323745-06, and Ontario Research Fund ORF-RE. We thank Ultrasonix Medical Inc. for technical support.

AN ENHANCED NUMERICAL MODEL TO SIMULATE NONLINEAR CONTINUOUS WAVE ULTRASOUND FIELD

Shahram Mashouf, and Jahan Tavakkoli

Dept. of Physics, Ryerson University, Ontario, Canada, M5B 2K3 jtavakkoli@ryerson.ca

1. INTRODUCTION

Within the realm of therapeutic ultrasound, high intensity focused ultrasound (HIFU) is a rapidly expanding modality with applications in tumor necrosis, hemostasis and immunotherapy [1]. In this method of treatment highly focused ultrasound beams induce a rapid temperature rise around the focal spot due to conversion of acoustic energy to heat. Precise, well defined lesions can be created inside the tissue due to thermal coagulation. One advantage of HIFU over other similar treatment modalities is that it can be performed noninvasively. Selecting the right transducer and excitation parameters ensure that underlying tissue layers remain intact and tissue coagulation happen only around the focal spot.

Due to high acoustic pressure amplitude and intensity produced in focal region, a significant nonlinear distortion can be observed and thus an accurate propagation model needs to include the effect of nonlinearity [2]. The model that we present here is based on a second-order operator splitting method where the acoustic field is propagated over incremental steps taking into account the effects of diffraction, nonlinearity and attenuation. This model is in essence a modified version of the KZK model where the parabolic diffraction term is replaced by a more accurate full diffraction term. This method was first introduced by Christopher et al. [3] for axi-symmetric sources and then improved by Tavakkoli et al. [4] via implementing larger propagation steps. It was then extended by Zemp et al. [5] to general non axi-symmetric problems using angular spectrum method.

In this work, we'll be further refining this method by introducing arbitrary source geometry and excitation definition, full diffraction solution, enhanced pressure calculation, and enhanced power deposition rate and temperature prediction capabilities. The result is a particularly useful tool in carrying out simulations of HIFU beams in tissue including temperature rise predictions. Since a typical HIFU power is usually delivered for the duration of a few seconds at frequencies of a few MHz, a CW simulation will be suitable.

2. METHOD

The KZK equation, which accounts for combined effects of diffraction, attenuation and nonlinearity in propagation of acoustic beam, is given in Eq. (1) below:

$$\frac{\partial p}{\partial z} = \frac{c_o}{2} \int_{-\infty}^{\infty} \nabla_{\perp}^2 p d\tau + \frac{1}{2c_o^3 \rho_o} \left[\left(\mu_B + \frac{4}{3} \mu \right) \frac{\partial^2 p}{\partial \tau^2} + \beta \frac{\partial p^2}{\partial \tau} \right] \quad (1)$$

The first term on the right hand side is the diffraction term in parabolic approximation, the second term reflects the effect of attenuation and the third term is due to nonlinearity. The pressure field can be calculated over propagation planes in incremental steps by bringing $\frac{\partial p}{\partial z}$ to

the left side as in Eq. (1). Also based on the above equation, the effects of diffraction, attenuation and nonlinearity can be applied independently over propagation planes and then added together. This is often referred to as operator splitting method. In the second-order operator splitting method, a certain propagation scheme is maintained which enable larger propagation steps and faster computational time (see Fig.1) [4].

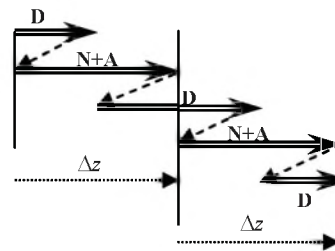


Figure 1. Second-order operator splitting method. D, N, and A represent operators for diffraction, nonlinearity and attenuation respectively.

Note that in this method, nonlinearity and attenuation are combined and propagated in one step. For a CW periodic waveform, the equations of propagations in each step shown in Figure 1 are presented here. For diffraction over the n^{th} harmonic:

$$v_z(x, y, z + \Delta z) = \mathfrak{S}_{2D}^{-1} \left\{ \mathfrak{S}_{2D} \{v_z(x, y, z)\} \times H(k_x, k_y, \Delta z) \right\} \quad (2)$$

where $H(k_x, k_y, \Delta z) = e^{j\Delta z \sqrt{k^2 - (k_x^2 + k_y^2)}}$ and $k = 2\pi(nf_o)c_o$ and k_x, k_y are spatial frequency components. This will be repeated over N harmonics ($n = 1$ to N).

For nonlinearity and attenuation over the n^{th} harmonic:

$$v_n(z + \Delta z) = v_n(z) + j \frac{2\pi\beta f_o}{2c_o^2} \Delta z \left[\sum_{i=1}^{n-1} i v_i v_{n-i} + \sum_{i=n+1}^N m v_i v_{n-i}^* \right] - \alpha_o (nf_o)^b v_n \Delta z \quad (3)$$

which will be repeated over N harmonics as well.

In the enhanced version of the algorithm, user has the capability to define an arbitrary source geometry and input excitation. Since the propagation is done plane by plane, this will require an extra initial step to propagate the field from the surface of transducer to an initial plane. This step is done using the Rayleigh diffraction integral which assumes linear propagation from the source to the first plane. Another method to accomplish this is to propagate the beam from the source onto the initial plane by introducing simple phase shifts. Phase shift methods, however, produce inaccurate results in near field specially when the source surface is highly focused

The normal particle velocity is then calculated on equally spaced discrete points across the initial plane (intersection of solid lines in Fig. 2). The calculated values of v_z is then expanded and assigned to the adjacent squares (dotted lines) to create a 2D array as shown in Fig. 2.

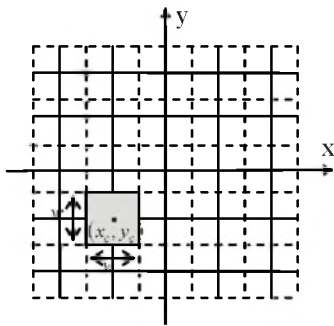


Figure 2. Initial plane as a 2D array.

Since the value of v_z across any given square (e.g. the shaded area shown in Fig. 2) is constant, it can be written in compact form as $v_z = v_o \text{rect}\left(\frac{(x-x_c)}{w}, \frac{(y-y_c)}{w}\right)$ where w is the width of the array element and (x_c, y_c) is the location of the element's centre. The *rect* function has an analytical 2D Fourier transform as below:

$$\mathfrak{F}_{2D}(v_z) = v_o w^2 \text{sinc}\left(w \frac{k_x}{2\pi}, w \frac{k_y}{2\pi}\right) \times e^{-j(k_x x_c + k_y y_c)} \quad (4)$$

Eq. 4 is then added up across all array elements to calculate the Fourier transform of the entire plane. The result is then feed into Eq. 2 to perform the first half step diffraction as illustrated in Fig. 1. After finishing diffraction substep, the result is then converted back to spatial domain using inverse Fourier transform and a nonlinear substep is subsequently performed using Eq. 3. The process is then repeated to propagate the field along the z direction.

3. RESULTS

The results obtained using our method were compared with other methods both in linear and nonlinear regimes. In overall, excellent agreements were observed.

Fig. 3 displays lateral pressure profiles for a concave spherical transducer with effective radius of curvature of 160 mm and aperture diameter of 37.6 mm working at a frequency of 2.25 MHz and with a source pressure of 92.5 KPa. Our results are in excellent agreement with those obtained by Averkiou et al. [6] using the KZK nonlinear model as shown in Fig. 3.

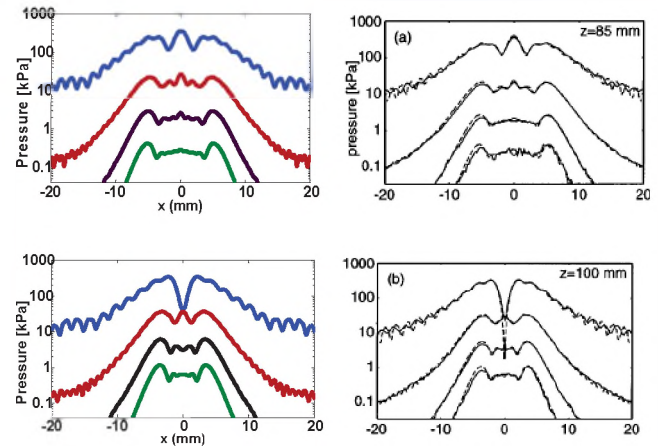


Figure 3. Our simulation results (left column) vs. the KZK nonlinear model (solid and dashed lines represent measurement and simulation results, respectively), for the fundamental and first 3 harmonics.

REFERENCES

1. Bailey MR, Khokhlova VA, Sapozhnikov OA, Kargl SG, Crum LA. Physical mechanisms of the therapeutic effect of ultrasound (a review). *Acoustical Physics*. 2003; 49(4):369-88.
2. Curra KP, Mourad PD, Khokhlova VA, Cleveland RO, Crum LA. Numerical simulations of heating patterns and tissue temperature response due to high-intensity focused ultrasound. *IEEE Trans Ultrason Ferroelectr Freq Control*. 2000; 47(4):1077-89.
3. Christopher PT, Parker KJ. New approaches to nonlinear diffractive field propagation. *J Acoust Soc Am*. 1991; 90(1):488-99.
4. Tavakkoli J, Cathignol D, Souchon R, Sapozhnikov OA. Modeling of pulsed finite-amplitude focused sound beams in time domain. *J Acoust Soc Am*. 1998; 104(4):2061-72.
5. Zemp RJ, Tavakkoli J, Cobbold RSC. Modeling of nonlinear ultrasound propagation in tissue from array transducers. *J Acoust Soc Am*. 2003; 113(1):139-52.
6. Averkiou MA, Hamilton MF. Measurements of harmonic generation in a focused finite-amplitude sound beam. *J Acoust Soc Am*. 1995; 98(6):3439-42.

ACKNOWLEDGEMENTS

This work was partially supported by the Dean's Start-up Fund from Ryerson University.

OLD AND NEW COCHLEAR MAPS

Reinhart Frosch

Sommerhaldenstrasse 5B, CH-5200 Brugg, Switzerland; reinifrosch@bluewin.ch
PSI (Paul Scherrer Institute), Villigen and ETH (Eidgenössische Technische Hochschule), Zurich (retired)

1. INTRODUCTION

In the mammalian-cochlea literature there are two “old” categories of cochlear maps (i.e., curves of frequency f versus distance x from the base), namely the passive-peak (PP) and the low-level active-peak (AP) maps; the AP frequency is often denoted as “characteristic frequency”. At given x , the PP map indicates the frequency yielding the maximal basilar-membrane (BM) velocity amplitude in a cochlea without viable outer hair-cells (OHC) or, approximately, in a healthy cochlea at sound-pressure levels (SPL) of about 100 dB. The AP map indicates, at given x in a healthy cochlea, the frequency yielding the maximal BM velocity amplitude at SPL below about 20 dB. At given $x < 0.3L$ (where L is the total BM length) the AP frequency is found to be higher than the PP frequency by about 0.5 octave (i.e., by a frequency factor of $R = f_{AP} / f_{PP} = 2^{0.5} \approx 1.4$); see Fig. 1 below.

In the two just mentioned SPL regions (i.e., at <20 dB and at ~ 100 dB) the BM velocity v_{BM} at given place x and frequency f is found to be proportional to the sound pressure in the ear canal (EC). It is assumed that also the corresponding sound-pressure difference δp between two points, at the given x , just “above” and just “below” the BM in a two-dimensional cochlear model is proportional to the EC sound pressure. Then a complex specific acoustical BM impedance $Z_{BM} = \delta p / v_{BM}$ can be defined. At given frequency $f > f_{AP}(x = 0.3L)$, impedance functions $Z_{BM}(x)$ such as that shown in Fig. 2 below have been derived from experiments; see de Boer (2006).

In the present study, two “new” categories of cochlear maps related to the just discussed impedance function $Z_{BM}(x, f)$ are presented. In a passive cochlea, that function is related to the angular frequency $\omega = 2\pi f$ and to the BM properties as follows [see e.g. de Boer (1996)]:

$$Z_{BM} = \eta + i \cdot (\omega \cdot M - S / \omega). \quad (1)$$

In Eq. (1), η is the real (i.e., resistive) part of the impedance and is positive if the cochlea is passive; the imaginary part is seen to involve the surface mass density M and the stiffness S of the BM and the cells attached to it; M includes those elements of the cochlear partition which move “up and down” in a travelling wave (TW) but excludes the liquid particles above and below the partition (which in a TW move on elliptical trajectories in x - z planes).

The “BMR map” is defined to yield, at given x , the resonance frequency f_{res} (BMR) of the local BM resonator [spring constant = $S \cdot dx \cdot dy$; mass = $M \cdot dx \cdot dy$] calculated under the hypothesis that the liquid above and below the partition is absent:

$$f_{res}(\text{BMR}) = [1/(2\pi)] \cdot \sqrt{S/M}. \quad (2)$$

For given x , Eqs. (1) and (2) imply that at the BMR-map frequency f_{res} (BMR) the imaginary part of the local passive BM impedance Z_{BM} vanishes.

The “IOCR map” is defined to yield, at given x , the resonance frequency f_{res} (IOCR) of the local internal organ-of-Corti resonator [spring = OHCs and other nearby structures; angle defined by BM and reticular lamina (RL) varies periodically, as observed, e.g., by Fridberger and Boutet de Monvel (2003)]. The IOCR is thought to represent the second degree of freedom enabling the OHCs to “actively” feed energy into the TW, to thus give rise to the active peak, and to make the resistive impedance (i.e., the real part of $Z_{BM} = \delta p / v_{BM}$) negative. For given x , that real part is therefore conjectured to have its most negative value at the local IOCR-map frequency f_{res} (IOCR); see Fig. 2 below.

2. METHODS

The AP (active peak) map can be determined from experimental below-20-dB curves of BM velocity amplitude versus x or f .

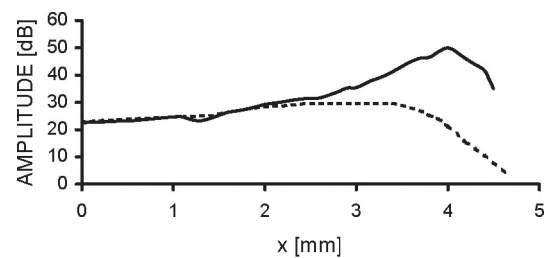


Fig. 1. Experimental BM velocity amplitude (in dB re stapes) versus distance x from stapes; guinea-pig; $f = 17$ kHz; SPL: solid curve 20 dB, dashed curve 100 dB; redrawn from de Boer (2006).

In Fig. 1, e.g. (guinea-pig, 17 kHz), the AP-map place (maximum of solid curve) is seen to be $x(\text{AP}) = 4.0$ mm.

The *PP* (*passive peak*) map can be estimated from experimental healthy-cochlea ~ 100 -dB and post-mortem curves. In Fig. 1, the PP-map place (maximum of dashed curve) is seen to be $x(\text{PP}) \approx 2.8$ mm.

The *IOCR* (*internal organ-of-Corti resonator*) map can be derived from experimental curves of the low-level real part of the BM impedance versus x .

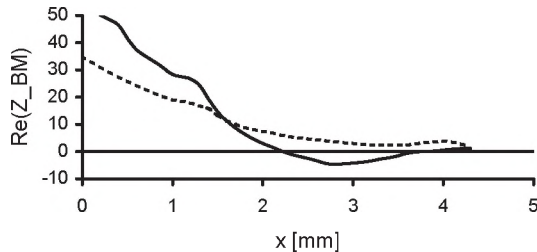


Fig. 2. Real part of the BM impedance (in arbitrary units) versus distance x from stapes; guinea-pig; $f = 17$ kHz; SPL: solid curve 20 dB, dashed curve 100 dB; redrawn from de Boer (2006).

In Fig. 2, e.g. (guinea-pig, 17 kHz), the *IOCR*-map place (i.e., the place of the most negative BM-impedance real part at 20 dB) is seen to be $x(\text{IOCR}) = 2.8$ mm, equal to the *PP*-map place.

In Fig. 2 of de Boer (2006), the *BMR* (*basilar-membrane resonator*) map place (i.e., the zero of the 100-dB imaginary part of the BM impedance for guinea-pig at 17 kHz) is apical of the considered x -range. An extrapolation based on Eqs. (1) and (2) above yields $x(\text{BMR}) = 6$ mm. The guinea-pig map of Greenwood (1990) yields that near the base one octave corresponds to $\Delta x_8 = 2.6$ mm. So $x(\text{BMR}) = 6$ mm is apical of $x(\text{PP}) = x(\text{IOCR}) = 2.8$ mm by 3.2 mm = $1.2 \cdot \Delta x_8$. That *PP*-*BMR* distance of 1.2 octave is confirmed by Figs. 8 and 9 of Kolston (2000) [chinchilla, 10 kHz], and also, in the frequency domain, by Mammano and Ashmore (1993) [guinea-pig, $x = 11$ mm] and by Frosch (2009a,b) [gerbil, $x = 2.4$ mm].

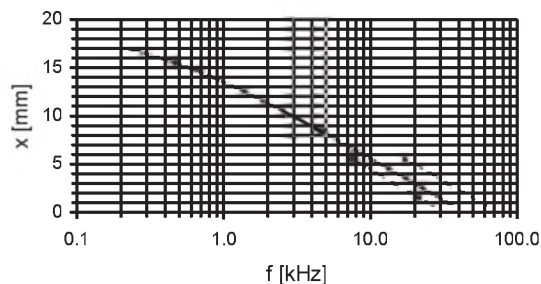


Fig. 3. Maps for guinea pig; $L = 18.5$ mm; solid curve: active-peak (AP) map; lower dashed curve: passive-peak (PP) and internal-organ-of-Corti-resonator (IOCR) maps; upper dashed curve: basilar-membrane-resonator (BMR) map.

3. RESULTS

The guinea-pig maps discussed in Section 2 above are shown in Fig. 3; the active-peak map shown and the mentioned basilar-membrane length, $L = 18.5$ mm, are based on Section IV.A of Greenwood (1990).

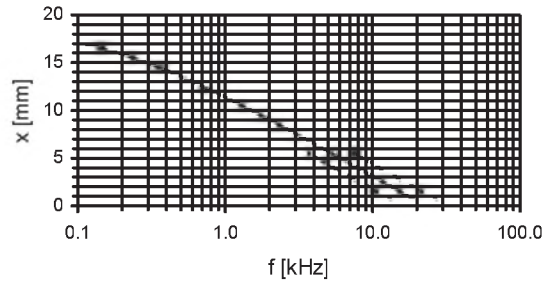


Fig. 4. Same as Fig. 3, for chinchilla; $L = 18.4$ mm.

The chinchilla active-peak map shown in Fig. 4 and the mentioned BM length, $L = 18.4$ mm, are based on Fig. 3 of Greenwood (1990). Further details on the guinea-pig and chinchilla maps, and maps for a third mammalian species (gerbil) are given in Frosch (2009b); the gerbil active-peak map was experimentally determined by Müller (1996).

4. DISCUSSION

A detailed discussion of the “old” and “new” mammalian-cochlea maps introduced above is presented in Frosch (2009b).

REFERENCES

- De Boer, E. (2006). Cochlear Activity in Perspective. In: Nuttall, A. L. et al. (Eds.), *Auditory Mechanisms*. World Scientific, New Jersey, pp. 393-409.
- De Boer, E. (1996). Mechanics of the Cochlea: Modeling Efforts. In: Dallos, P. et al. (Eds.), *The Cochlea*. Springer, New York, pp. 258-317.
- Fridberger, A., Boutet de Monvel, J. (2003). Sound-induced differential motion within the hearing organ. *Nature neurosci.* 6, 446-448.
- Frosch, R. (2009a). DP Phases in Mammalian Cochleae, Predicted from Liquid-Surface-Wave Formulas. In: N.P. Cooper and D.T. Kemp (Eds.), *Concepts and Challenges in the Biophysics of Hearing*, World Scientific, New Jersey, pp. 41-47.
- Frosch, R. (2009b). Old and New Cochlear Maps. Submitted to *Canadian Acoustics*.
- Greenwood, D. D. (1990). A cochlear frequency-position function for several species – 29 years later. *J. Acoust. Soc. Am.* 87, 2592-2605.
- Kolston, P. J. (2000). The importance of phase data and model dimensionality to cochlear mechanics. *Hear. Res.* 145, 25-36.
- Mammano, F., Ashmore, J. F. (1993). Reverse transduction measured in the isolated cochlea by laser Michelson interferometry. *Nature* 365, 838-841.
- Müller, M. (1996). The cochlear place-frequency map of the adult and developing Mongolian gerbil. *Hear Res.* 94, 148-156.

EFFECT OF HEADPHONE TYPE ON LISTENING LEVELS AND LOCALIZATION ABILITIES OF PORTABLE AUDIO DEVICE USERS IN QUIET AND TRAFFIC NOISE

I. Ibrahim, R. Malcolmson, M. B. Jennings, and M. F. Cheesman

National Centre for Audiology, School of Communication Sciences and Disorders,
The University of Western Ontario, London, ON N6G 1H1

1. INTRODUCTION

The sales of different portable audio devices (PAD; such as MP3 players) have witnessed a boom during the last decade (Hodgetts, Rieger, & Szarko, 2007). Their popularity is particularly increasing among young adults. The output of these devices can be high enough to pose a risk of hearing impairment (Fligor & Ives, 2006), depending on several factors, such as the volume setting, listening duration, and music type.

Many PAD owners wear their devices in noisy situations such as in cafeterias, on the street, or while commuting by bus, streetcar, or subway. When worn in noise, wearers typically increase the volume setting to maintain an adequate music-to-noise ratio. Noise-reduction headphones can help reduce the risk of hearing impairment associated with the use of PAD in noisy background by blocking some noise, hence decreasing the need for increased volume levels. However, these headphones may affect the wearer's ability to spatially locate important environmental sounds such as approaching cars. The present two studies examine the effects of headphone type and background noise on sound localization abilities and PAD output levels.

2. GENERAL METHODS

Participants were normal hearing adults, aged 21 to 30, who owned and used PADs for the purpose of listening to music. All completed a hearing history survey and participated in two stages of data collection: real-ear measurements of their preferred listening levels (PLLs) and a sound localization task. All tasks and measurements were conducted in a hemi-anechoic chamber. Participants were seated in the centre of a multi-speaker array at a 1.5 meter distance from all speakers. Stimuli and noises were presented via a 360° subset of 8 speakers (separated by 45°) at the height of the listener's ear.

2.1 Real-Ear Measurements

A probe-tube was inserted in left ear canal of each participant, medial to the output of the headphone and

within 5 mm of tympanic membrane (Audioscan, 2007). Measures of room noise in the ear canal and PLLs in the ear canal (headphones, with music playing, adjusted by user to PLL) were conducted both in quiet and with a background of recorded stereo traffic noise.

2.2 Sound Localization Task

Participants were seated facing forward (0° azimuth) in the centre of the speaker array. Upon hearing a target stimulus, they turned and used an electromagnetic pointing device to indicate which speaker was the source of the target stimulus. Participants returned to the 0° position and the next stimulus was presented.

3. STUDY 1

Four listening conditions were tested: an open ear condition and three lower-priced over-the-counter transducers: an ear bud (Samsung EP370), over-the-ear headphones (Sony - MDR 210-LP), and noise-reduction insert earphones (Skullcandy Smokin' Buds SCBUDP). Twenty participants were tested in quiet (ambient noise < 30 dBA) and in 83 dBA traffic noise.

As expected the participants' PLLs increased when listening to music in the background of traffic noise, as shown in Table 1.

Table 1. Real ear measures of noise (Open ear) and noise combined with MP3 music set to PLL (Music)

	Open ear		Music					
			Ear bud		Over-the-ear		Noise reduction	
	M	SD	M	SD	M	SD	M	SD
Quiet	53.4	0.7	65.3	8.8	66.3	9.6	61	9.1
Traffic noise	83.9	2.3	85.1	5.6	89.5	4.1	77.7	6.9

Localization performance decreased when the participants were listening to music through the MP3 player (Table 2). In particular, there was an increase in the back-front errors (stimuli at the back were perceived to be in front) when over-the-ear phones were worn.

Table 2. Localization errors by headphone type in Quiet and Traffic noise conditions

Headphone type	Quiet			Traffic			Increase M%
	M	SD	%	M	SD	%	
Open Ear (no music)	6.4	3.2	26.7	8.9	3.8	37.1	2.5 (10.4%)
Ear bud	8.5	3.3	35.4	10.4	2.8	43.3	1.9 (7.9%)
Over-the-ear	8.1	2.5	33.8	11.3	2.4	47.1	3.2 (13.3%)
Noise reduction inserts	10.3	2.2	42.9	11.2	2.2	46.7	0.9 (3.8%)

4. STUDY 2

Four listening conditions were used that included an open ear condition and three transducers designed to reduce background noise: ER6i insert phones, Sony MDR-NC6 Noise Cancelling headphones, and Bose QuietComfort 2 Acoustic Noise Cancelling headphones. Participants were tested in quiet (ambient noise < 30 dBA) and in 70 dBA traffic noise.

The PLLs of 19 participants increased in background noise conditions (Table 3). The number of localization errors increased when listening to music through the noise-cancelling headphones (Table 4), and are higher than observed in study 1 in spite of the lower traffic noise levels, suggesting that the noise-reduction headphones degrade localization abilities. Error analyses indicated that back-front errors were the most common error type followed by front-back and lateralization errors.

Table 3. Real ear measures of noise (Open ear) and noise combined with MP3 music set to PLL (Music conditions)

	Open ear		Music					
	M	SD	Bose		ER6i		Sony	
			M	SD	M	SD	M	SD
Quiet	52.5	0.9	63.4	8.7	71.8	11.6	67.9	10.8
Traffic noise	70.1	5	66.9	7.3	75.2	9.5	76.2	7.3

Table 4. Localization errors by headphone type in Quiet and Traffic noise conditions

Headphone type	Quiet			Traffic			Increase M%
	M	SD	%	M	SD	%	
Open ear	7.4	3.2	30.7	9.1	3.6	37.9	1.7 (7.2%)
Bose noise cancelling	12.1	3.9	50.2	14.3	3.8	59.4	2.2 (9.2%)
ER6i inserts	13.9	4.5	58.1	12.5	4.8	52	-1.4 (6.1%)
Sony noise cancelling	13.7	4.4	58.1	14.1	2.8	58.8	0.4 (0.7%)

5. DISCUSSION

The amount of and type noise reduction provided by the earphones varied considerably among the different types as did the wearer's PLL in a background

traffic noise. Some of the earphones provided some sound attenuation by occluding the ear canal (ER6i) while others provided active noise cancellation. Shah et al. (2009) reported that 85% of PAD users are concerned about hearing loss, and willing to protect their hearing. The use of noise reduction headphones may reduce a listener's PLL in noisy environments and thereby reduce the risk of noise-induced hearing loss.

Localization errors were more common when listeners were wearing PADs and slightly higher, on average, when they were wearing them in a background of noise. The type of headphone worn influenced the number of localization errors, with the greatest localization problems occurring while the noise reduction devices were being worn regardless of the background noise condition.

News reports of pedestrians and bicyclists being hit by cars while wearing PADs are not uncommon, nor are they surprising given these results. Wearers should be made aware of the difficulties in identifying the source of warning signals when in an environment posing a physical safety hazard (e.g. walking in traffic).

REFERENCES

- Audioscan (2007). *Verifit Hearing Instrument Fitting Guide (Version 3.0)*. Dorchester, Ontario. [Available: <http://www.audioscan.com/resources/usersguides/Curentverifitguide.pdf>]
- Hodgetts, W.E., Rieger, J.M., Szarko, R.A. (2007). The effects of listening environment and earphone style on preferred listening levels of normal hearing adults using an MP3 player. *Ear & Hearing*, 28, 290-297.
- Fligor, B. J., & Ives, T. E. (2006). *Does earphone type affect risk for recreational noise-induced hearing loss?* The NHCA Conference on Noise-Induced Hearing Loss in Children at Work and Play.
- Shah, S., Gopal, B., Reis, J., Novak, M. (2009). Hear Today, Gone tomorrow: an assessment of portable entertainment player use and hearing acuity in a community sample. *Journal of the American Board of Family Medicine*, 22, 17-23.

ACKNOWLEDGEMENTS AND NOTES

S. Beaulac, L. Coates, J. Cruckley, and D. Grainger provided assistance with this project. The Canadian Foundation for Innovation and the Ontario Research Fund provided financial support. A subset of these results has been presented at the Canadian Academy of Audiology.

The work described in study 1 was conducted while R. Malcolmson was a student at the University of Western Ontario. Her current address is Central Speech and Hearing Clinic, Winnipeg, MB R3T 4J6.

AGE AND NEUROPATHIES IN VIBRATION EXPOSED MANUAL WORKERS

Martin G Cherniack¹, Anthony J Brammer^{1,2}, Ronnie Lundström³, Timothy F. Morse¹, Greg Neely⁴, Tohr Nilsson⁵, Donald R Peterson¹, Esko M Toppila⁶, and Nicholas Warren¹

¹ Ergonomic Technology Center, University of Connecticut Health Center, Farmington CT, U.S.A.

² Institute for Microstructural Sciences, National Research Council, Ottawa, Ontario

³ Department of Biomedical Engineering and Informatics, University Hospital, Umeå, Sweden

⁴ Department of Psychology, Umeå University, Umeå, Sweden

⁵ Department of Occupational and Environmental Medicine, Sundsvall Hospital, Sundsvall, Sweden

⁶ Finnish Institute of Occupational Health, Helsinki, Finland and Department of Otorhinolaryngology, Tampere University Hospital, Tampere, Finland

I. INTRODUCTION

Tingling, pricking or numbness in the hands (paresthesias) are associated with intensive hand use, age, and use of vibratory tools. The Hand-Arm Vibration International Consortium study provides longitudinal data on symptoms, and on some electrophysiological and psychophysical measurements, in four vibration exposed cohorts: shipyard workers, automotive assembly workers, forest workers, and dental hygienists. A raised prevalence of entrapment neuropathy, particularly carpal tunnel syndrome, was associated with age. However, intensive hand use and vibratory exposure predicted hand paresthesias in younger age groups. In industrial settings, where ergonomics and anti-vibration measures were optimized, the prevalence of hand paresthesias and of carpal tunnel syndrome were low. Hand paresthesias are common in certain occupational groups with vibration exposures. There are complex interactions between physiologic age, duration of exposure, and intensity of exposure.

The HAVIC (Hand-Arm Vibration International Consortium) is a multi-national research group, organized to better define exposure-response relationships from segmental vibration through application of longitudinal study design (Cherniack et al., 2006). Study populations include shipyard workers from the United States, dental hygienists using ultrasonics, an inception cohort of Swedish truck cab workers, and Finnish forest workers. The Finnish forest workers and Swedish truck cab workers had been previously studied and the current work incorporates earlier results. The American shipyard workers had been studied previously but with a different set of study metrics. The longitudinal design provided a basis for evaluating disease natural history, providing baseline rates were sufficiently high. The variety of assessment tools – questionnaire and hand diagrams, diagnosis by study physician, past treatment history, and electrophysiological studies – allows for comparative and contingent case definitions.

II. METHODS

Baseline and follow-up questionnaires and structured physical examinations were administered to all study participants. Vibrometry was performed so that mechanoreceptor function could be assessed independently of nerve conduction. Fractionated nerve conduction was

also performed providing comparisons of SNCV in the digits with cross-wrist and longer track measurements. Other measurements included vibrotactile thresholds (VTT), cold challenge plethymography (CCP) and a detailed survey. Hand paresthesias generally, and carpal tunnel syndrome (CTS) specifically were of particular interest. CTS was assessed in three independent ways: 1) by self-reported symptoms and hand diagram; 2) by the study physician using fixed criteria, and by prevalent diagnosis external to the study.

Exposure assessment was performed as part of the HAVIC study. The main feature involved datalogging of force and vibration using a force sensor and accelerometer mounted in tandem on the palm with alignment of the primary axis of activation. Data were sampled and stored over a 15-s at an interval at a frequency of 3 kHz per channel. This was followed by a 45-s interval required for calculation of the vibration exposure and the average, consistent with international standards (ISO 5349:1, 2001; ISO 5349:2, 2001).

III. RESULTS

Table 1 summarizes baseline evaluation for the four incident HAVIC cohorts and presents the rate of CTS diagnoses by the study physician for each cohort. For comparison, control groups of relatively exposure naïve dental hygiene students and non-manufacturing workers are included as control groups. Although hand paresthesias are common in shipyard workers, dental hygienists and forestry workers, diagnosed CTS is relatively common (>5%) in dental hygienists and shipyard workers. Dental hygiene students, who are both young and unexposed to ultrasonic vibration, had low rates of hand paresthesias. Truck cab workers and controls were symptomatically similar. In Table 2, the study groups are examined in terms of total hours of tool or instrument exposure and selected nerve conduction results. Only limited sensory nerve conduction velocities are presented, for wrist-palm and proximal digit-distal digit segments of median nerve of the dominant hand. Abnormalities in the wrist-palm segment, most often associated with CTS were most common in the three oldest and most ergonomically challenged cohorts – dental hygienists, shipyard workers, and forest workers. However, the digital nerve segment, was significantly slowed only for

shipyard workers. In Table 3, data logging results are presented for the three industrial cohorts. There were no comparable measurements for dental hygienists who use ultrasonics only, and for whom instrumentation remains a different and more complicated process. Furthermore, the vibration levels measured for truck cab workers and forestry workers are consistently lower than for shipyard workers

IV. DISCUSSION

Neurological symptoms were relatively common in the three cohorts that had longest tenure of work and the most ergonomic demands. CTS was less common in the two industrial cohorts (forestry workers and truck cab workers) with lower daily vibration exposures. The SNCV-PPDD nerve segment has been described as the part of the peripheral nerve most susceptible to the effects of exposure to hand arm vibration (Sakakibara et al., 1998; Cherniack et al. 2004), so its elevation in the cohort with the apparent highest vibration exposures is notable. At low vibration exposure levels (truck cab workers, forestry workers). CTS was near a background rate, even though hand paresthesias were present. Age and biomechanical risk were relatively high in dental hygienists and forestry workers, although vibratory exposure was reduced. The combined pathology of prevalent symptoms, and high levels of physician diagnosed CTS along with abnormal SNCV velocity in the wrist-palm segment was most apparent in the

shipyardworkers. While vibration control does not address individual or biomechanical risks, it does appear to effectively eliminate HAV as an independent risk factor for clinical CTS.

REFERENCES

Sakakibara H, Hirata M, Hashiguchi T, Toibana N, Koshiyama H, Zhu S-K, Kondo T, Miyao M, Yamada S. Digital sensory nerve conduction velocity and vibration perception threshold in peripheral neurological test for hand-arm vibration syndrome. *AJIM* 1998; 30: 219-224.
Cherniack M, Brammer AJ, Lundstrom R, Meyer J, Morse T F, Nealy G, Nilsson T, Peterson D, Toppila E , Warren N , Fu RW, Bruneau H. Segmental Nerve Conduction Velocity in Vibration Exposed Shipyard Workers. *Int. Arch Occup Environ Health* 2004; 77(3): 159-176.
Cherniack M, Brammer AJ, Lundstrom R, Meyer JD, Morse T, Neeley G, Nilsson T , Peterson D, Toppila E, Warren N. The Hand-Arm Vibration International Consortium (HAVIC): Prospective Studies on the Relationship between Power Tool Exposure and Health Effects. *JOEM* 2007;49(3):289-301.

ACKNOWLEDGMENT

This study-U01 OH 07312—was performed with support from the National Institute for Occupational Safety and Health.

Table I. Characteristics of HAVIC cohorts at Baseline

Cohort	Baseline Test Year	Subjects Tested	Age	+ Hand Paresthesias	CTS Cases Physician Dxed Dominant Hand
Shipyard Workers	2001	217	48.3(6.6)	146 (69%)	74 (35%)
Dental Hygienists (DH)	2002	94	45.5 (8.8)	42 (45%)	14 (15%)
Dental Hygiene Students	2002	66	26.1 (6.4)	6(9%)	3 (5%)
Forestry Workers	2003	61	48.7 (5.3)	17(29%)	3 (6%)
Truck Cab Workers	2002	54	30.1 (3.2)	7 (12%)	2 (4%)
Controls	2002	36	35.1 (5.8)	4 (11%)	0 (0%)

Table 2. Sensory nerve conduction (SNCV) in HAVIC cohorts – median nerve, dominant hand

	Shipyard	DH	Student DH	Forestry	Truck Cab	Control
Exposure Yr.	18.3 (8.8)	17.1 (8.7)	1.0 (1.9)	25.6 (7.3)	9.3 (3.6)	0
SNCV-WP	41.6 (8.1)	43.6 (7.5)	50.3(4.7)	41.9 (7.7)	47.3 (5.8)	49.4 (5.9)
SNCV-PPDD	43.1 (9.4)	46.4 (8.7)	46.9(7.6)	49.9 (7.2)	45.4 (5.8)	46.7 (5.5)

SNCV-WP is the segmental sensory nerve conduction velocity in the dominant hand across the carpal tunnel
 SNCV-PPDD is the segmental nerve conduction velocity in the dominant index or long finger

Table 3. Summary of Data Logging Results Operating Time = minutes RMS= m/s², RMQ = m/s⁴, RMO = m/s⁸

	SHIPYARD (n = 52)		TRUCK CAB (n = 38)		FORESTRY (n = 8)	
	mean (sd)	min - max	mean (sd)	min - max	mean (sd)	min - max
Data Logger Operating Time	243.20 (99.40)	17 - 414	222.9 (64.20)	85 - 312	206.37 (7.54)	193 - 215
Weighted Accel ⁿ – RMS	4.96 (7.27)	0.66 - 35.95	2.92 (2.05)	1.06 - 10.2	2.42 (0.88)	1.64 - 4.29
Weighted Accel ⁿ – RMQ	14.66 (12.29)	2.88 - 63.32	11.85 (5.52)	6.06 - 29.41	7.8 (2.34)	5.97 - 12.72
Weighted Accel ⁿ – RMO	33.06 (16.79)	9.31 - 89.81	29.4 (7.41)	21.35 - 53.85	24.81 (5.66)	21.59 - 38.64

VIBROTACTILE PERCEPTION AS A TEST FOR NUMBNESS AND PAIN

Anthony J. Brammer^{1,2}, Paivi Sutinen³, Sourish Das⁴, Ilmari Pyykkö⁵,
Esko Toppila⁶, and Jukka Starck⁶

¹Ergonomic Technology Center, Univ. of Connecticut Health Center, Farmington CT, U.S.A. 06030-2017

²Inst. for Microstructural Sciences, National Research Council, Montreal Rd., Ottawa ON, Canada K1A 0R6

³Dept. of Physical Medicine and Rehabilitation, North Karelia Central Hospital, FIN-80210 Joensuu, Finland

⁴Statistical and Applied Mathematical Sciences Institute, Duke University, Durham NC, U.S.A. 27709-4006

⁵Dept. of Otorhinolaryngology, Tampere University Hospital, FIN-33521 Tampere, Finland

⁶Finnish Institute of Occupational Health, FIN-00250 Helsinki, Finland

1. INTRODUCTION

Quantitative sensory testing (QST) has long been used in clinical medicine for investigating peripheral neuropathies. The tests were originally primitive – pin prick (for pain), hot/cold objects (for temperature sense), and a cotton wool swab (for fine touch) applied to the skin. While there have been many attempts to devise instruments to quantify symptoms of numbness or pain, none has achieved broad acceptance by the medical community.

In this paper, a vibrotactile test for assessing the presence, or absence, of sensory symptoms in the hands is described based on perception thresholds believed mediated selectively by different types of mechanoreceptors. It is constructed from the summed differences between the thresholds recorded at the fingertip of an individual and the mean values of the threshold for healthy persons at the same stimulation frequencies. The metric is found to be related to reports by subjects of numbness and pain using two statistical tests for evaluating the significance of associations in 2x2 contingency tables (Altman, 1991). The performance of the test is evaluated by adjusting the magnitude of the metric that is to serve as the boundary for the onset of symptoms reported by individual subjects.

A detailed report of this work has been submitted for publication elsewhere (Brammer et al., 2009).

2. APPARATUS AND METHOD

2.1 Apparatus

Vibrotactile thresholds have been determined at the fingertips with the forearm supported horizontally below the shoulder (palm up). The apparatus has been described elsewhere (Brammer et al., 2007), and conforms to method A of ISO 13091-1 (2001). It consists essentially of: 1) a vibration exciter suspended from a beam balance with adjustable fulcrum, to permit the stimulator to be lowered onto a fingertip; 2) an arm rest for the back of the hand and forearm; 3) a 3 mm diameter cylindrical plastic probe, to apply the stimulus to a fingertip with a contact force of 0.05 N; 4) an accelerometer and conditioning electronic circuits, to record the motion of the skin, and; 5) a computer, to

apply the psychophysical algorithm and calculate perception thresholds.

Thresholds were obtained at the fingertips of digits 3 and 5 for both hands at frequencies of 4, 6.3, 20 and 32 Hz. Thresholds at 4 and 6.3 Hz are believed to be mediated by the slowly adapting, type I receptors (Merkel disks), and at 20 and 32 Hz by the fast adapting, type I receptors (Meissner corpuscles) (ISO 13091-1, 2001).

2.2 Method

When all thresholds are expressed in dB (re 10^{-6} m/s²), the summed threshold shift can be expressed as (Brammer et al., 2007):

$$TS_{Sum(SD)} = TS_{SAI(SD)} + TS_{FAI(SD)}$$

In this expression:

$$TS_{SAI(SD)} = [TS_4 + TS_{6.3}] / 2SD_{SAI}$$

and $TS_{FAI(SD)}$ has been similarly constructed. The shift in threshold at a given stimulus frequency, that is, TS_4 , $TS_{6.3}$, TS_{20} , and TS_{32} , is given by the difference between the observed threshold and the mean threshold recorded from the hands of healthy persons at that frequency. As the thresholds of healthy persons appear to approximate a Gaussian distributions with similar variance at each stimulation frequency mediated by the same receptor type, the ranges are expressed by the standard deviations, TS_{SAI} and TS_{FAI} . The last mentioned parameters are used to normalize the threshold shifts. In order to compare values of $TS_{Sum(SD)}$ with the reported symptoms, it is necessary to construct a single metric for each subject, which is taken to be the largest of the threshold shifts recorded from either digit, in either hand.

Tests for the statistical significance of an association between symptoms reported by individuals and values of $TS_{Sum(SD)}$ recorded from their hands have been conducted (Chi-squared test, and Fisher's exact test). For this purpose, 2x2 contingency tables are formed to segregate the reported presence or absence of the symptom of interest, and a fence value, t , is selected for $TS_{Sum(SD)}$ to correspond to the boundary between the presence and absence of the symptom (Altman, 1991). The null hypothesis is that there was no association between the test statistic and the target

symptom, and the alternative hypothesis is that there was a positive association (one-sided test). A probability value of $p=0.05$ was considered significant.

2.3 Subjects

A group of manual workers was selected from forest workers in the Suomussalmi region of Finland (18/109 persons). Their participation was voluntary, and subjects gave their informed consent according to the provisions of the ethics committees of the participating organizations. The work involved operation of lightweight chain saws, except for three subjects who were foremen.

Thirty-nine percent of the subjects reported numbness in the hands, 17% reported nocturnal numbness in the arms, and 33% reported pain in the hands, forearm, or throughout the upper extremity. The symptoms were obtained during a physical examination conducted by a physician, who also performed a neurological examination to screen for polyneuropathies.

3. RESULTS

The association derived from contingency tables between the metric constructed from $TS_{Sum(SD)}$ for each subject and their report of numbness in the hands is given in Table 1. The statistical significance of the association (p-value) is shown for different fence values, t . Each test identifies fence values that produce a statistically significant association, which are underlined. While the tests differ somewhat in the range of fence values calculated to be associated with numbness, values of $3.4 < t < 4.0$ are most significant, with probability values reaching $p < 0.01$ in mid range (i.e., $t = 3.5$).

Results similar to those in Table 1 were obtained for reports of numbness at night and, perhaps surprisingly, for arm pain (but not for neck pain). The fence values in both cases, however, were somewhat greater, namely in the range from 3.5 to 4.5.

4. DISCUSSION

Current clinical practice commonly employs electrically-stimulated nerve conduction for diagnosis of peripheral sensory neuropathies, which: 1) does not include the functionality of the nerve endings (the transducers), and; 2) records the response of the fastest nerves fibers, leaving unanswered questions concerning the biochemical function of most fibers.

The effectiveness of mechanoreceptor-specific vibrotactile perception as a QST for sensory symptoms has been shown by the present work, where the sensitivity ranged from 100% to 80%, and the specificity from 63% to 78%, for fence values possessing a statistically significant association with numbness (Table 1), numbness at night, and pain

(Altman, 1991). The small number of subjects in our study restricts more precise calculation of parameters for assessing the performance of the test. The similarity in fence values for reports of numbness at night, and pain could imply that the two symptoms are related in this group of subjects. As the sensory test is responsive to the tactile end organs, and not those sensing pain, it is possible that the test is reflecting functional changes in nerve fibers innervating the end organs, rather than changes in the end organs themselves.

Table 1. Tests of association for numbness in the hands

$TS_{Sum(SD)}$	p-value		
	Fence, t	Chi-squared test	Fisher's Test
	1	0.945	1.000
	2	0.945	1.000
	2.5	0.244	0.200
	3	0.110	0.077
	3.4	<u>0.044</u>	<u>0.026</u>
	3.5	<u>0.015</u>	<u>0.007</u>
	4	0.067	<u>0.040</u>
	4.5	0.201	0.132
	5	0.200	0.119
	5.5	0.897	0.569
	7	0.388	0.200

Finally, it is interesting to note that the statistically significant fence values obtained when using the magnitude of $TS_{Sum(SD)}$ as a QST encompass the boundary between "normal" and "abnormal" thresholds derived previously for these subjects. The latter boundary corresponded to thresholds for which the probability of occurrence in healthy persons was $p=0.05$.

REFERENCES

- Altman, D. G. (1991). *Practical Statistics for Medical Research* (Chapman and Hall/CRC, Boca Raton).
- Brammer, A. J., Sutinen, P., Das, S., Pyykkö, I., Toppila, E., and Starck, J. Quantitative test for sensory hand symptoms based on mechanoreceptor-specific vibrotactile thresholds. (submitted to J. Acoust. Soc. Am., February 2009).
- Brammer, A. J., Piercy, J. E., Pyykkö, I., Toppila, E., and Starck, J. (2007). Method for detecting small changes in vibrotactile perception threshold related to tactile acuity. J. Acoust. Soc. Am. **121**, 1238-1247.
- ISO 13091-1. (2001). *Mechanical Vibration: Vibrotactile Perception Thresholds for the Assessment of Nerve Dysfunction-Part I: Methods of Measurement at the Fingertips* (International Organization for Standardization, Geneva).

ACKNOWLEDGEMENTS

Work supported, in part, by the Finnish National Board of Forestry, and the Finnish Forestry fund.

PERCEPTION OF MUSICAL SOURCES WITH IMPOVERISHED SPECTRAL ENVELOPES

Michael D. Hall

Dept. of Psychology, James Madison University, MSC 7704, Harrisonburg, Virginia, USA, 22807 hallmd@jmu.edu

1. INTRODUCTION

Musical instrument timbre is multi-dimensional, with spectral, temporal, and spectro-temporal components (e.g., McAdams, Winsberg, Donnadiou, De Soete, & Krimphoff, 1995). Envelope shape has been argued to be a critical spectral dimension (e.g., Krumhansl, 1989). Recent timbre research has shown that listeners rely upon envelope shape instead of brightness (Hall & Beauchamp, 2009).

How much spectral envelope detail is required to permit timbre recognition? One way to evaluate this question is to gradually eliminate spectral peaks and assess when timbre shifts occur. This has been done for timbre discrimination (e.g., see Beauchamp, Horner, & Ayers, 2006), revealing a broad range of performance that depends upon instrument. An alternative method of manipulating spectral detail was pursued in the current investigation. A set of resynthesized musical tones was generated at a common pitch such that all spectral information was eliminated except at harmonics that occurred at average spectral peaks. The impact of these manipulations on timbre identification and discrimination was then evaluated as a function of musical instrument.

2. METHOD

Participants were 9 undergraduates who had a mean of 7.5 years of performance training on a musical instrument (1-11 years). Eighteen tones (44.1 kHz, 16-bit) were derived from A₄ samples from the MUMS database (Opolko & Wapnick, 1987) for piano, vibraphone, electric guitar, tenor trombone, saxophone, and E^b clarinet. Tones were resynthesized using Camel Audio's *Alchemy*, an additive VST plug-in based on spectral modeling synthesis (Serra & Smith, 1990). Phase differences and variation from mean F₀ were eliminated. Tones were equated for loudness and compressed to 1755 ms while retaining amplitude envelopes. Three tones were synthesized for each instrument. One included all harmonics from the original tone. A second version had 7 harmonics—F₀ plus 6 harmonics with higher mean dB than adjacent harmonics. If a tone had fewer spectral peaks, then the remaining harmonics had the highest mean amplitude. A final version had 4 harmonics—F₀ plus three harmonics selected as described for the 7-harmonic series.

Listeners completed two computer-controlled tasks. In the first, timbre identification, listeners indicated from which of the instruments each tone was derived. There were 10 randomized repetitions for each tone. An instrument discrimination task was restricted to all- and 4-harmonic tones. Listeners rated whether tone pairs were from the same instrument. Ratings of 1 to 4 indicated that tones were

from the same instrument, and 5-8, different instruments. Higher ratings indicated greater differences [1 ("identical")-8 ("very different")]. There were 6 repetitions per tone pair.

3. RESULTS

3.1 Instrument Identification

Mean timbre identification accuracy was determined for each stimulus and listener. Grand means and corresponding standard errors are displayed in Figure 1. Reduction to 7 harmonics had minimal impact on identification, and 4-harmonic stimuli were typically identified well above chance. Accuracy decreased for 4-harmonic stimuli (.61 v. .76-.77; Bonferroni $p < .05$), which contributed to a main effect of number of harmonics, $F(2,16) = 11.98, p < .001$.

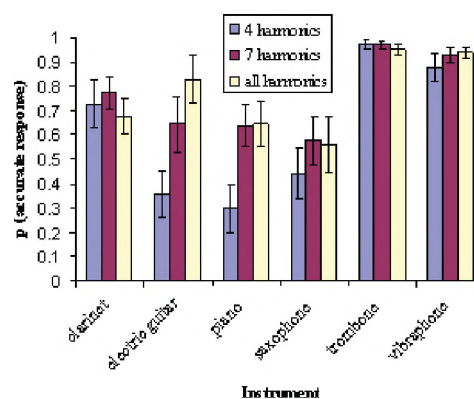


Figure 1. Mean accuracy of timbre identification.

Accuracy differed with instrument [$F(5,40) = 13.08, p < .0001$], and the impact of harmonic reduction depended on instrument, $F(10,80) = 3.89, p < .001$. Accuracy was not significantly reduced with fewer harmonics for trombone, vibraphone, clarinet, or saxophone. Trombone and vibraphone were not confused with other instruments, and clarinet tones were reliably identified. While saxophone tones with few harmonics were increasingly confused with clarinet, similar confusion occurred for all-harmonic tones. Significant reductions in accuracy for 4-harmonic stimuli ($p < .05$) were observed for piano and guitar. For piano there was a corresponding increase in vibraphone responses ($p < .05$). Reducing guitar harmonics produced responses for sources with brief attacks and less spectral irregularity. There were frequent piano responses, and a significant increase in vibraphone responses for the 4-harmonic tone.

3.2 Timbre Discrimination Ratings

Mean ratings of perceptual distance were submitted to multi-dimensional scaling (MDS) using ALSCAL. The

resulting 2-dimensional solution is displayed in Figure 2 ($r^2 > .94$, stress $< .12$). Ratings of intact clarinet, trombone, and vibraphone tones with corresponding 4-harmonic tones were less than for other stimuli ($p < .05$). For the trombone and vibraphone, these ratings were significantly below 4.5 ($X^2 p < .05$ and $.01$), the boundary between timbres. For the vibraphone these ratings did not significantly differ from ratings of identical tones (1.07 v. 1.04). Ratings from same-instrument comparisons also were lowest for the saxophone and guitar. For the all-harmonic piano standard, ratings for the 4-harmonic piano were significantly less than for all tones except guitar tones or the intact vibraphone. The only instrument that may have shifted out of category following harmonic reduction was the saxophone ($M = 4.61$), but its all-harmonic tone also was confused with clarinet (< 4.5).

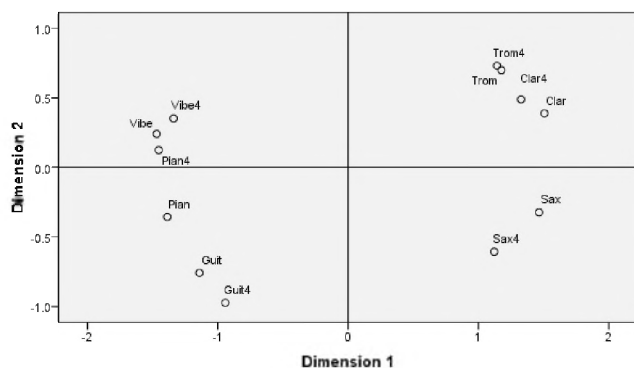


Figure 2. MDS solution based upon timbre discrimination.

MDS coordinates were evaluated for correlations with several acoustic measures, including mean mel-frequency cepstral coefficient (MFCC), mean spectral centroid, and spectral irregularity (see www.vamp-plugins.org), as well as (log) rise time in *ms*. Dimension 1 was best predicted by spectral irregularity ($r = .94$, $p < .0001$), and MFCC ($r = .94$, $p < .0001$; rise time $r = .77$, $p < .01$). Dimension 2 was moderately correlated with rise time ($r = .53$, $p < .05$), but likely reflects responses to spectral differences, including reductions in spectral complexity with the stimulus manipulation. For example, tones with fewer spectral peaks (i.e., with gradual spectral changes) gathered toward the top of the axis. Furthermore, 4-harmonic tones generally are located higher on the axis than their intact counterparts. The exceptions in Figure 2 (guitar and saxophone) reveal stress in the solution; ordering of these 4- and all-harmonic tones was reversed on the corresponding axis in the 3-D solution, which accounted for less than 3 percent more variance.

4. DISCUSSION

Several findings suggest retention of important source information despite a minimal number of harmonics. Stimuli with few harmonics were typically still primarily perceived as the same instrument across tasks, and observed perceptual shifts were limited to instruments (piano and guitar) that had greater spectral complexity than could be approximated by the limited number of frequencies. Furthermore, these timbres were often confused with

instruments that shared aspects of production (e.g., brief attacks followed by gradually decaying spectral information). Thus, it appears that timbre information can be maintained as long as harmonics reflect natural resonances of the instrument that produces them.

These findings are not likely solely due to temporal properties (e.g., rise time). After all, one dimension in the MDS solution from discrimination was best accounted for by spectral or spectro-temporal measures. Although a strong acoustic correlate was not obtained from the included measures for the remaining dimension, it also appears to be best accounted for by spectral information (e.g., diffuseness of spectral peaks). Additionally, reductions in instrument recognition with decreasing harmonics could only reflect responses to changes in spectral envelope shape.

The maintenance of timbre with limited spectral information complements preceding work. For example, it has been demonstrated that timbres become more discriminable from their original counterparts with sufficient smoothing of the spectral envelope, and smoothed envelopes are difficult to distinguish, presumably due to an absence of well-defined spectral peaks (Beauchamp, Horner, & Ayers, 2006). The general findings reported here also can be regarded as a nonspeech analogue to similar speech demonstrations. For example, in sinewave speech phonemes are accurately reported given only three frequencies that correspond to formant center frequencies (e.g., see Remez, Rubin, Pisoni, & Carrell, 1981). In fact, sentence transcription in sinewave speech is more accurate when harmonically related frequencies are used (e.g., Toth & Kocsor, 2003) like those in the current investigation. As a result, the findings reported here could represent further evidence that listeners focus on intense spectral information in vibratory sources.

REFERENCES

- Beauchamp, J. W., Horner, A. B., & Ayers, L. (2006). Effect of critical band data reduction on musical instrument sounds. *Proceedings of the 9th International Conference on Music Perception and Cognition (ICMPC 07)*, Bologna, Italy, 51-52.
- Krumhansl, C. L. (1989). Why is musical timbre so hard to understand? In S. Nielsen and O. Olsson (Eds.), *Structure and Perception of Electroacoustic Sound and Music* (pp. 43-53). Amsterdam: Elsevier.
- McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G. & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities and latent subject classes. *Psychological Research*, 58, 177-192.
- Opolko, F., & Wapnick, J. (1987). McGill University Master Samples [CD-ROM]. Montreal, Quebec, Canada: jwapnick@music.mcgill.ca.
- Remez, R. E., Rubin, P. E., Pisoni, D. B., & Carrell, T. D. (1981). Speech perception without traditional speech cues. *Science*, 212, 947-950.
- Serra, X. & Smith, J. (1990). Spectral modeling synthesis: A sound analysis/synthesis based on a deterministic plus stochastic decomposition. *Computer Music Journal*, 14(4), 12-24.
- Toth, L. & Kocsor, A. (2003). Alternatives to sine-wave speech. *EUROSPEECH-2003*, 2073-2076.

CHANGES IN MELODIC PERCEPTION AS A FUNCTION OF AGE AND LOWEST HARMONIC COMPONENT

Huiwen Goy¹, D. Timothy Ives², Frank A. Russo³, M. Kathleen Pichora-Fuller^{1,4}, and Roy D. Patterson²

¹Dept. of Psychology, University of Toronto at Mississauga, 3359 Mississauga Rd North, Mississauga, Ontario L5L 1C6

²Centre for the Neural Basis of Hearing, Department of Physiology, Development and Neuroscience, University of Cambridge, Downing St, Cambridge, CB2 3EG, United Kingdom

³Department of Psychology, Ryerson University, 350 Victoria St, Toronto, Ontario M5B 2K3

⁴Toronto Rehabilitation Institute, 550 University Ave, Toronto, Ontario M5G 2A2

1. INTRODUCTION

Spectral models of pitch resolution explain pitch perception primarily with reference to the place coding of frequency, in which the maximally stimulated point on the basilar membrane of the cochlea depends on the frequency of a tone which is related to its perceived pitch. Temporal models, however, argue that pitch perception is a result of temporal coding, in which the rate of firing of the auditory nerve corresponds to the periodicity of the stimulus. Both views have been influential in our understanding of pitch perception and recent models have attempted to combine the spectral and temporal aspects of auditory processing [1, 2].

Numerous studies have demonstrated the importance of spectral and temporal coding to speech perception. Deficits in spectral coding associated with cochlear hearing loss compromise speech perception in quiet. Furthermore, relative to younger adults, age-related declines in auditory temporal processing abilities may account for the greater difficulties experienced by older compared to younger adults when they listen to speech in noise. For example, difficulties experienced by normal-hearing older adults in understanding time-compressed or reverberant speech are associated with age-related changes in temporal processing [3]. The effects of age-related changes in auditory temporal processing have also been examined with respect to the perception of specific speech cues. Older adults with clinically normal hearing thresholds were less accurate than their younger counterparts in using differences in voice onset time to distinguish between phonemes [4]. Jittering the temporal fine structure of speech to simulate reduced phase-locking resulted in intelligibility scores in younger adults with normal hearing thresholds that resembled those of older normal-hearing adults [5].

Although temporal processing ability seems to decline with age with consequences to speech perception, the issue of age-related changes in temporal processing and consequences for pitch perception in music have not yet been investigated except in two studies. The first found that older adults were more influenced by irrelevant information (i.e., pitch proximity) than were younger adults [6]. The second found that tonal sensitivity of younger normal-hearing adults listening to jittered tones did not differ from that of older normal-hearing adults listening to unaltered tones [7].

The present study investigated age differences in spectral and temporal mechanisms in pitch perception, using a melodic pitch task adapted from a previous study that tested pitch perception in younger adults [8]. In that study, listeners performed better on melodies with a lower fundamental frequency (F_0) and with fewer harmonics resolved by the cochlea compared to melodies with a higher F_0 and more resolved harmonics, which was consistent with the temporal view of pitch processing.

Following predictions from the temporal view of pitch perception, the performance of both age groups in the present study was expected to decline with an increase in the lowest harmonic component of a complex stimulus. In addition, the performance of older adults was expected to be worse than that of younger adults even when lower harmonics were available. Thus, the reduction in performance with increases in the lowest harmonic component was expected to be smaller for older than for younger adults.

2. METHOD

2.1 Participants

Listeners were six younger (21-26 years, $M = 22.7$) and six older adults (65-75 years, $M = 69.5$). Older adults with good audiograms were selected to match younger adults' hearing thresholds as closely as possible. Both groups had pure-tone audiometric thresholds of ≤ 25 dB HL at the octave frequencies from 0.25 to 8 kHz in both ears, except for two older adults who each had a threshold of 30 dB HL at 8 kHz in one ear. None of the listeners spoke tonal languages and musical training was not a requirement for participation. All listeners gave informed consent and were paid at an hourly rate for their participation.

2.2 Task

In each trial, listeners heard two consecutive four-note melodies, with the tonic note of the melodic scale presented twice before each melody. Melodies were identical within each pair except for one note in the second melody that was shifted up or down by one or two semitones, which listeners had to identify.

2.3 Stimuli

Melodies were generated using a custom MATLAB program. In each trial, an F_0 for the tonic of the melody was randomly selected from a range of 126 to 178 Hz (a half octave range centered logarithmically on 150 Hz). Notes for the melody were then randomly chosen with replacement from the first five notes of the diatonic major scale with the selected F_0 as its tonic. Adjacent notes were never identical. All notes consisted of eight successive harmonic components, with the average, lowest component (ALC) manipulated to be the 4th, 8th or 12th with lower components absent. The lowest component was set to rove by one component throughout the study to prevent the listener from using a single sinusoid to complete the task.

2.4 Procedure

Listeners were seated in an IAC double-walled sound booth. Melodies were played using a Creative SB Audigy 4 sound card and presented binaurally at 60 dB SPL through Sennheiser HD 265 headphones, accompanied by bandpass-filtered white noise at 50 dB SPL ($BW = 50$ -500 Hz) to mask any combination tones and their harmonics.

Prior to experimental trials, listeners were trained using a similar procedure as Ives and Patterson (2008). Immediately after training, a total of 180 experimental trials were completed, with 60 trials in each of three ALC conditions presented in randomized order. Listeners indicated their answers using a computer interface and received visual feedback on whether or not their answers were correct. No response time limit was imposed. All listeners completed the training and experimental trials in a single two-hour session.

3. RESULTS

Figure 1 shows the performance of younger and older adults across three ALC conditions. To test for age differences in performance at each ALC condition, Mann-Whitney U tests with a multistage Bonferroni correction were conducted. The results showed that younger and older adults differed significantly on ALC 4 ($p = 0.004$) and 8 ($p = 0.002$), but not on ALC 12 ($p = 0.180$). To test for differences in performance between ALC conditions, Friedman's test was conducted separately for each age group. The results confirmed that there were differences between ALC conditions in both younger ($p = 0.006$) and older ($p = 0.011$) adults. Wilcoxon's signed ranks test showed that, in younger adults, ALC 4 and 8 were not different from each other ($p = 0.34$), but they both differed from ALC 12 ($p = 0.028$; $p = 0.027$). In older adults, performance on ALC 8 and 12 was not different ($p = 0.673$), but they both differed from ALC 4 ($p = 0.027$; $p = 0.028$).

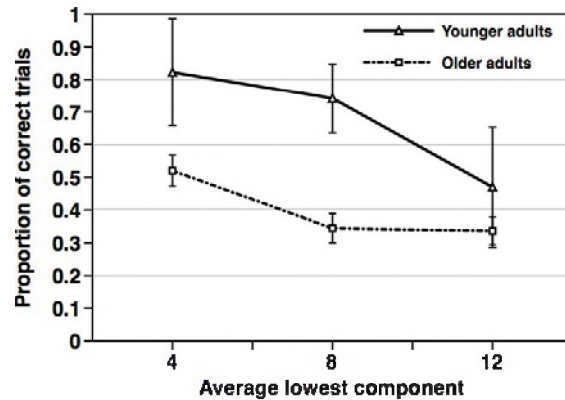


Fig. 1. Average performance of 6 younger and 6 older adults in three ALC conditions. Standard deviations are shown.

4. DISCUSSION

Younger adults performed better than older adults in easier ALC conditions, except in the most difficult condition with the highest frequencies, and the effect of higher frequencies was smaller for older adults than for younger adults. Consistent with the temporal view of pitch perception, task performance declined with higher frequencies. Given that both groups were closely matched on their audiometric thresholds, the results cannot readily be explained in terms of reduced spectral coding due to cochlear damage. This study provides evidence that in addition to speech, age-related declines in auditory temporal processing also reduce the perception of tonality in music.

REFERENCES

- [1] Carney LH. Spatiotemporal encoding of sound level: Models for normal encoding and recruitment of loudness. *Hear Res.* 1994;76:31-44.
- [2] Patterson RD, Allerhand MH, Giguere C. Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform. *J Acoust Soc Am.* 1995;98(4):1890-4.
- [3] Gordon-Salant S, Fitzgibbons PJ. Profile of auditory temporal processing in older listeners. *J Speech Lang Hear Res.* 1999;42(2):300-11.
- [4] Strouse A, Ashmead DH, Ohde RN, Grantham DW. Temporal processing in the aging auditory system. *J Acoust Soc Am.* 1998;104(4):2385-99.
- [5] Pichora-Fuller MK, Schneider BA, MacDonald E, Pass HE, Brown S. Temporal jitter disrupts speech intelligibility: A simulation of auditory aging. *Hear Res.* 2007;223(1-2):114-21.
- [6] Halpern, A. R., Kwak, S. Y., Bartlett, J. D., & Dowling, W. J. Effects of aging and musical experience on the representation of tonal hierarchies. *Psychol Aging.* 1996;11(2): 235-46.
- [7] Minghella D, Russo F, Pichora-Fuller MK. Effect of age on sensitivity to tonality. *Canadian Acoustics - Acoustique Canadienne.* 2007;35(3):58-9.
- [8] Ives DT, Patterson RD. Pitch strength decreases as F_0 and harmonic resolution increase in complex tones composed exclusively of high harmonics. *J Acoust Soc Am.* 2008;123(5):2670-9.

ACKNOWLEDGEMENTS

This study was funded by CIHR MOP-15359 and NSERC RGPIN-138572-05.

DEVELOPMENT AND VALIDATION OF A SENSORY-SUBSTITUTION TECHNOLOGY FOR MUSIC

Carmen Branje¹, Michael Maksimowski², Gabe Nespoli², Maria Karam¹, Deborah Fels¹, and Frank Russo²

¹Center for Learning Technology, Ryerson University, 350 Victoria St., Ontario, Canada, M5B 2K3

²Dept. of Psychology, Ryerson University, 350 Victoria St., Ontario, Canada, M5B 2K3

1. Introduction

The Emoti-Chair is a sensory-substitution technology designed to provide greater access to music for individuals who are deaf or hard of hearing. This is accomplished by converting sound into vibrotactile information that is displayed along the back and legs. The basic conceptual design, referred to as the model human cochlea (MHC), presents discrete bands of frequency as independent points of vibration, allowing for spatiotemporal encoding of the incoming auditory signal. The back is thus treated as a basilar membrane, complete with its own tonotopic map, oriented vertically so as to align with common conceptions of pitch height. Preliminary research with artificially deafened participants has revealed that emotion judgments concerning vibrotactile music presented via the model human cochlea align closely with those of hearing participants, and the judgments are distinct from those made with vibrotactile music that is not frequency-band separated (Karam, Branje, Price, Russo, & Fels, 2007; Karam, Russo, Fels, in press).

The current paper presents progress on a series of psychophysical studies that were designed to ascertain what aspects of music can be perceived through vibration alone.

2. Frequency Difference Limen (FDL)

The first experiment used the method of limits in order to measure ability to discriminate the frequency of vibrotactile stimuli across a wide range of frequencies common to (western) music. In order for vibrotactile music to be a viable undertaking, the skin must possess some degree of frequency discrimination ability. Results from previous research vary, with Pongrac (Pongrac, 2008) reporting the smallest FDL of 18% between 100 and 700 Hz when vibration was applied to the fingertip and forearm. Much of the other work (Goff, 1967; Mahns, Perkins, Sahai, Robinson, & Rowe, 2006) regarding this topic report curves steeper than that found by Pongrac. No work to date has looked at a large contactor applied to the non-glabrous skin on the back.

2.1 Method

A single large contactor (102mm diameter) was placed on the lower back of each of 4 participants. The contactor was secured using a 5.08 cm wide nylon strap

tied around the waist. Participants wore headphones with loud white noise playing in order to artificially deafen them. All anchor and comparison stimuli were equated for subjective intensity in a preliminary experiment. Anchor stimuli spanned a range of fundamental frequencies corresponding to the pitch range from C2 to C6 (with anchors on C's and F#'s). Participants were exposed to 3000 ms of the anchor stimulus, then 500 ms of no vibration and then 3000 ms of the comparison stimulus and were then asked if the two stimuli were the same or different. Initial comparison stimuli started 5 semitones above or below the anchor frequency and moved 1 semitone closer with each successive judgment, until the frequency matched and then proceeded beyond the anchor frequency. Three correct responses resulted in the run ending. Each anchor was presented twice, once in an ascending trial and once in a descending trial.

2.2 Results

Results indicate that frequency difference limens for stimuli presented as a single point of vibration on the back fall between 2 and 3 semitones across the range of vibrotactile sensitivity (5 to 1000 Hz). As seen in figure 1, the FDL curve was shallower than those found by other studies in which vibrotactile information was applied to the fingertip or forearm.

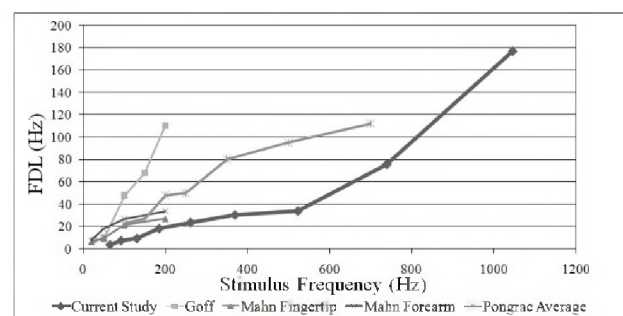


Figure 1 - A comparison of frequency difference limens reported with vibrotactile stimuli.

2.3 Discussion

Since the non-glabrous skin of back has a relatively low density of mechanoreceptors and is not generally considered to be highly sensitive (Mahns et al., 2006), it appears likely that the large contactor size played a

critical role in the relatively shallow FDL curve obtained in this experiment.

3. Timbre

The second experiment examines ability to discriminate between complex vibrotactile waveforms. Because timbre is an important aspect of musical structure (Seashore, 1936), ability to distinguish between different timbres is seen as crucial for perceiving music through vibrotactile channels.

3.1 Method

Complex vibrotactile waveforms generated from an acoustic signal were presented to the back via voice coils embedded in a conforming chair. Participants made same-different judgments based on timbre (piano, cello, trombone). A total of 5 undergraduate students participated in the study, 3 male and 2 females. All participants were deafened using white noise presented through headphones and by wearing *Tactaid* skin transducers taped over their cheek bones to help mask sound transmitted via bone conduction. Each trial consisted of two timbres presented for 1-s each and separated by a 1-s inter-stimulus interval. Stimuli were presented at each of 3 levels of fundamental frequency, 110, 220, and 440Hz, which correspond to pitches A2, A3, and A4, respectively.

3.2. Results

Initial results are promising as a large majority of participant responses across all instrument combinations were correct. All of the instrument combinations showed a significant chi-square result (with alpha set at .05): $\chi^2(4) = 50.438$ for Cello-Cello; $\chi^2(4) = 44.263$ for Cello-Piano; $\chi^2(4) = 14.143$ for Cello-Trombone; $\chi^2(4) = 44.263$ for Piano-Cello; $\chi^2(4) = 40.909$ for Piano-Piano; $\chi^2(4) = 60.842$ for Piano-Trombone; $\chi^2(4) = 17.610$ for Trombone-Cello; $\chi^2(4) = 49.951$ for Trombone-Piano; $\chi^2(4) = 56.977$ for Trombone-Trombone). Participants achieved at least 70% accuracy for all instrument combinations, with most combinations above 80%.

3.3 Discussion

These initial results suggest that humans are able to discriminate complex vibrotactile waveforms perceived through the skin. Moreover, preliminary results from a follow-up experiment that controls for amplitude-envelope variation across timbres is yielding similar results, suggesting that ability to discriminate timbre through vibrotactile channels is based at least in part on spectral attributes.

4. Conclusions

Given the findings of these two preliminary studies, we have a more complete understanding of the specific attributes of vibrotactile music that can be perceived. The ability to discriminate frequencies between 2 and 3 semitones and discriminate complex vibrotactile waveforms may support the experience of structure and emotion in music by the Deaf. Given recent evidence, indicating that judgments of interval size can be made more accurate by providing congruent vibrotactile input (Maksimowski & Russo, 2008), the most promising application of this technology may be in supporting music listening in the hard of hearing.

REFERENCES

- Goff, G. (1967). Differential discrimination of frequency of cutaneous mechanical vibration. *Journal of Experimental Psychology*, 74(2), 294.
- Karam, M., Branje, C., Price, E., Russo, F., & Fels, D. I. (2007). Towards a model human cochlea. *Proceedings of the 34th Graphics Interface Conference*, 267-274.
- Karam, M., Russo, F. A., Fels, D. I. (in press). An ambient crossmodal audio-tactile display. *IEEE Transactions on Haptics*.
- Mahns, D., Perkins, N., Sahai, V., Robinson, L., & Rowe, M. (2006). Vibrotactile frequency discrimination in human hairy skin. *Journal of Neurophysiology*, 95(3), 1442-1450.
- Maksimowski, M., & Russo, F. A., (2008). Audio-tactile integration in music perception. *Proceedings of Acoustics Week in Canada, Canadian Acoustics*, 36, 102-103.
- Pongrac, H. (2008). Vibrotactile perception: Examining the coding of vibrations and the just noticeable difference under various conditions. *Multimedia Systems*, 13(4), 297-307.
- Seashore, C. E. (1936). The psychology of music. III. the quality of tone:(1) timbre. *Music Educators Journal*, , 24-26.

ACKNOWLEDGEMENTS

We acknowledge the Natural Science and Engineering Research Council of Canada for funding support and Christopher Lachine for assistance with data collection.

THE SOUND OF STROOP: ACOUSTIC EFFECTS IN STROOP INTERFERENCE

Boaz M. Ben-David^{1,2,3}, Pascal H.H.M. Van Lieshout^{1,2,3,4}, Alex Nishta^{1,3}

¹Oral Dynamics Lab, Department of Speech-Language Pathology, University of Toronto, St. George Campus, 500 University Avenue, Toronto, Ontario, Canada, M5G 1V7, boaz.ben.david@utoronto.ca

²Toronto Rehab, 550 University Avenue, Toronto, Ontario, Canada, M5G 2A2

³Centre for Research on Biological Communication Systems, Department of Psychology, University of Toronto Mississauga, 3359 Mississauga Road, N., Mississauga, Ontario, Canada, L5L 1C6

⁴Institute of Biomaterials and Biomedical Engineering, University of Toronto, St. George Campus, 164 College street, Toronto, Ontario, Canada, M5S 3G9

1. INTRODUCTION

Selective attention is a fundamental process in everyday life. In almost every task, one has to attend selectively to certain features in the environment while ignoring or actively suppressing others. For example, talking to a friend in a noisy park, one has to focus on the words of the person sitting next to you on the bench, while ignoring or suppressing conversations taking place around you. If you fail to suppress conversations that surround you, will it affect your speech? Generally, are failures of selective attention manifested in speech?

The most common tool for the assessment of selective attention in the lab has been the Stroop colour-word task (Stroop, 1935). Indeed, the Stroop paradigm has been regarded as the golden standard of selective attention for over 75 years (see Melara & Algom, 2003 for a relevant review). Participants are asked to name the colours of printed words, irrespective of their content. The response latency increase for naming the print colour of an incongruent colour-word (e.g., RED printed in blue) over a colour-neutral stimulus (e.g., OOOO printed in blue) is termed Stroop interference (SI), $SI = CN_{\text{incongruent}} - CN_{\text{neutral}}$, where CN stands for response latency in colour-naming. If the participant can completely ignore or inhibit the processing of the lexical content of the colored word when asked to name the color in which it was printed, SI should equal zero. On the other hand, if she or he cannot ignore the lexical content of the word, we should observe $SI > 0$.

The SI is taken to reflect an increase in cognitive load and cognitive stress in colour-naming incongruent words (e.g., Caruso, Chodzko-Zajko, Bidinger, & Sommers, 1994). Failures of selective attention, the semantic conflict presented in an incongruent colour-word, and the process of inhibiting a well-practiced response (reading aloud the printed word) for a less-practiced one (colour-naming) all contribute, in tandem, to the increase in cognitive stress. For example, Renaud and Blondin (1997) found that responding to incongruent stimuli heightened participants' heart rate levels when compared to their heart rate in responding to color-neutral stimuli, indicating higher stress levels.

Stroop effects have been traditionally tested only in terms of response preparation (reaction times and accuracy). However, it stands to reason that these effects go beyond response preparation processes, and have an impact on response production as well, i.e., on the articulation of the oral responses. Specifically, acoustic features of speech have been found to indicate the effect of various emotions and affective conditions, including stress (Scherer, 2003). For example, in the context of our previous example on the conversation in the park, as it takes more effort to distinguish the speech of the person next to you from the other conversations, stress levels are likely to increase as well. This eventually may affect speech, for example, by raising speech intensity.

A few studies (Caruso, & al., 1994; Rothkrantz, Wiggers, Van Wees, & Van Vark, 2004) have investigated features of oral responses in the Stroop paradigm. However, none have actually investigated the acoustic changes in oral responses directly related to the SI. For example, Caruso et al.(1994) compared responses of nine non-pathological participants to incongruent stimuli with responses to congruent ones (RED in red font). However, selective attention fails in both congruent and incongruent trials, facilitating colour responses in the first, while interfering with colour responses in the latter. The results of Caruso et al., therefore, do not indicate the effect of failures of selective attention, but rather an additive effect of interference and facilitation.

In the current study, for the first time, we compared the acoustic characteristics of responses to colour-naming incongruent and neutral stimuli, in an attempt to link the SI in response latency with a parallel Acoustic-SI (ASI), $ASI = AC_{\text{incongruent}} - AC_{\text{neutral}}$, where AC stands for acoustic characteristics.

2. METHOD

2.1 Participants

Twenty-four young adults participated in the study. All participants were native English speakers, had minimum Snellen fraction (visual acuity test) and pure-tone air-conduction thresholds (from 0.25 to 3.00 kHz) appropriate

for their age group. Acoustic data of six participants were not analyzed because of technical problems.

2.2 Procedure and Materials

The study included two blocks of 16 trials, incongruent and neutral. The *incongruent block* included the four colour-words RED, GREEN, BLUE and YELLOW printed in different colours (e.g., RED printed in blue). The *neutral block* included strings of 4 Os printed in green, blue, yellow and red font colours. Participants were tested individually in a single-walled acoustic chamber, seated in front of a computer monitor. In two separate blocks, they were asked to name aloud the font colours of words presented on the monitor speaking into the microphone placed in the table in front of them. In each trial, a single word was presented in the centre of the monitor. After the system registered that a vocal response was initiated, the word disappeared and was replaced by another, one second later.

2.3 Acoustic Analysis.

The digital audio recordings of each participant were split into files that represent single responses. Each response file was compared to the printed stimulus to monitor for the accuracy of the response. Correct responses only were further analyzed using PRAAT software (Boersma, Paul & Weenink, David, 2009). We measured the following acoustic features: fundamental frequency (pitch) and intensity, taken at the midpoint of the vowel; the first, second and third formant frequencies, taken as the mean of values for the entire duration of the vowel; and vowel and word duration. Vowel and word boundaries were manually identified by a trained experimenter while viewing a display showing both a wave form and a broadband spectrogram, with superimposed linear predictive coding based formant tracks (note that in the response “yellow” we analyzed only the first vowel). Acoustic characteristics of single responses were averaged for each experimental block, to allow a within participants comparison design.

3. RESULTS

3.1 Behavioral Data

As commonly found in the literature, accuracy in colour naming was higher for the neutral stimuli than for the incongruent ones, 97.7% and 91%, respectively, $t(23) = 2.72$, $p = 0.01$. A robust SI was found for all participants, with an average SI of 142 ms, $t(23) = 9.9$, $p < 0.001$.

3.2 Acoustic Data

We found significant differences between the acoustic characteristics of responses to incongruent and neutral stimuli (e.g., comparing “blue” responses to RED with “blue” responses to OOOO in blue). Mainly, with

incongruent stimuli pitch values decreased, while the intensity and the second and third formant values increased.

4. DISCUSSION

We found significant acoustic differences in colour responses between incongruent and neutral stimuli in the Stroop paradigm. These differences can be specifically linked to the interference caused by the irrelevant task (reading) on colour-naming, unlike previous studies. Finally, our results provide a parallel acoustic effect to the latency effect of Stroop interference. Future research is needed to elucidate the cognitive and affective meaning of these acoustic differences, comparing them to acoustic effects in other attentional tasks.

REFERENCES

- Caruso, A. J., Chodzko-Zajko, W. J., Bidinger, D. A., & Sommers, R. K. (1994). Adults who stutter: responses to cognitive stress. *J Psychosom Res*, *37*, 746-754.
- Boersma, Paul & Weenink, David (2009). Praat: doing phonetics by computer (Version 5.1.10) [Computer program]. Retrieved July 8, 2009, from <http://www.praat.org/>
- Melara, R. D., & Algom, D. (2003). Driven by information: a tectonic theory of Stroop effects. *J Psychosom Res*, *110*, 422-471.
- Rothkrantz, L., Wiggers, P., Van Wees, J. W., & Van Vark, R. (2004). Voice stress analysis. *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, *3206*, 449-456.
- Renaud, P., & Blondin, J. P. (1997). The stress of Stroop performance: physiological and emotional responses to color-word interference, task pacing, and pacing speed. *International Journal of Psychophysiol*, *27*, 87-97.
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, *40*, 227-256.
- Stroop, J. R. (1935). Studies of interference in serial-verbal reaction. *Journal of Experimental Psychology*, *18*, 643-662.

ACKNOWLEDGEMENTS

This research was undertaken, in part, thanks to funding from the Canada Research Chairs program and a Research Opportunity Program grant from the Faculty of Arts & Science at the University of Toronto in Mississauga. This study was partially supported by a strategic training grant – Communication and Social Interaction in Healthy Aging, funded by the Canadian Institutes of Health Research. The authors wish to thank the following students for their contribution: Tracey D’Chuna, Huiwen Goy, Wu Yan (Lulu) Li, and Skyler Mooney.

Establishing normative voice characteristics of younger and older adults

Jessica Banh¹, Konstantin Naumenko¹, Huiwen Goy¹,
Pascal van Lieshout^{1,2,3,4}, David N. Fernandes^{2,3}, Kathy Pichora-Fuller^{1,3}

¹Dept. of Psychology, University of Toronto, Mississauga, 3359 Mississauga Rd North, Mississauga, ON, Canada L5L 1C6

²Dept. of Speech-Language Pathology, University of Toronto, 500 University Ave, Toronto, ON, Canada M5G 1V7

³Toronto Rehabilitation Institute, 550 University Ave, Toronto, ON, Canada M5G 2A2

⁴Inst. of Biomaterials and Biomedical Engineering, University of Toronto, 164 College St, Toronto, ON, Canada M5S 3G9

1. INTRODUCTION

The use of our voice is an important element of daily social interaction. The mechanism of voicing is complex and there are many disorders and hazards that can disrupt it, such as smoking and degenerative neurological diseases. The loss of voice quality and strength is often seen as a result of physical deterioration, but may in fact be a biological precursor for various pathologies [1]. In order to use voice characteristics for diagnostic purposes clinicians need reliable reference values from normal populations to determine the nature and level of impairment.

An important step towards establishing such reference values is to conduct normative studies on a relatively large sample of healthy individuals across age and gender groups. This project is an important contribution towards this aim. It used a new and unique program for recording voice samples called *SpeechClinic*, which was created by Fernandes and Van Lieshout as a new clinical tool to perform multidimensional voice analysis [2]. When a voice sample is recorded, *SpeechClinic* displays calculated values, such as average pitch, length of phonation and sound pressure level for each task, in comparison with values represented in a unique and flexible database based on a large selection of published references. The algorithms used by *SpeechClinic* in its analyses are well-established, and the parameters measured are therefore directly comparable with most values in the literature. Finally, *SpeechClinic* allows the presentation of this information over time for a given patient, which is useful in monitoring treatment progress.

There are two main reasons for conducting this study. First, normative acoustic data do not exist for all voice characteristics measured by *SpeechClinic*, and the ones that do exist are often based on small sample sizes. Without good normative data, judgments based on these measurements may be inaccurate. Second, computer analysis of voice samples is gaining interest in clinics, but unlike *SpeechClinic*, most of the existing software lacks an intuitive interface, a well-documented reference database that can be easily extended, or facilities to easily document progress and print reports. This study was able to test *SpeechClinic* extensively on these features. The primary purpose of this study was to obtain a large number of acoustic voice measures for healthy younger and older adults to provide critical information on the typical distribution of these voice characteristics. In addition, the study allowed us to analyze age-related and gender differences more specifically. The knowledge of

age-dependent changes in vocal properties is important in establishing norms because it separates the “normal” from “abnormal” population without confounding typical age-related changes with declines due to medical conditions. After a set of norms is obtained for a particular population, the data from an individual belonging to the same group can be compared to the distribution characteristics obtained from that population. Future studies will focus on sampling other populations to broaden the scope of the database. The results presented here are preliminary findings on a part of the targeted population sample.

2. METHOD

2.1 Participants

Data for a total of 169 people are presented here. The younger group contains 111 first year psychology students, including 55 females (median = 18.6 yrs, range 18-23) and 56 males (median = 18.8 yrs, range 18-27). The older group contains 58 participants recruited from a pre-existing volunteer database, including 47 females (median = 70.7 yrs, range 65-81) and 11 males (median = 73.4 yrs, range 66-78). Participants were included if they learned English before the age of 5 years in an English-speaking country, had pure-tone audiometric thresholds ≤ 25 dB HL from 0.25 to 3 kHz in both ears and no voice-altering conditions (e.g., colds, allergies) at the time of the experiment, and were free of self-reported speech disorders and voice pathologies.

2.2 Apparatus

The equipment used to record voice samples included a SHURE Beta 54 wrap-around headset microphone, which was connected to an M-Audio Mobile Pre USB pre-amplifier. The signal was output to a Mac OS X laptop computer running *SpeechClinic*. A Korg CA-30 chromatic tuner and a Source brand sound level meter were used for calibration. Voice samples were recorded in an Industrial Acoustics Company double-walled sound-attenuating booth. A GSI-61 clinical audiometer was used to measure the hearing thresholds of younger participants, while older participants already had data on hearing thresholds from earlier recent screenings.

2.3 Procedure

Each participant was tested on six tasks which were performed three times in a row, with task order counterbalanced. Participants phonated the vowel [a] for

maximum phonation time, at maximum pitch level, and at the lowest sound pressure level (SPL), and they also read a short passage. These four tasks were preceded and followed by phonating the vowel [a] at habitual frequency and intensity for 8 seconds.

3. RESULTS

Table 1 summarizes the means and standard deviations (SD) of acoustic measures for all 169 subjects, separated by age and gender. Table 2 shows the results from a Mann-Whitney U test for gender and age differences. Three measures of jitter (local, RAP, PPQ), four measures of shimmer (local, APQ3, APQ5, and APQ11), harmonics-to-noise ratio (HNR)¹, habitual SPL, and average fundamental frequency (F₀) were taken from the first habitual (baseline) task, based on 4-sec intervals determined automatically by the program.

Table 1. Means of acoustic measures for two age and gender groups (with standard deviations in parentheses).

MEASURE	GROUP			
	Younger Female	Younger Male	Older Female	Older Male
Habitual F ₀ (Hz)	254 (30)	130 (21)	215 (41)	119 (22)
Maximum F ₀ (Hz)	472 (117)	411 (118)	514 (114)	411 (105)
Max / Habitual F ₀	1.87 (0.47)	3.20 (0.96)	2.43 (0.55)	3.50 (0.79)
Jitter local (%)	0.38 (0.29)	0.37 (0.12)	0.45 (0.42)	0.48 (0.19)
Jitter RAP (%)	0.22 (0.19)	0.21 (0.08)	0.26 (0.25)	0.25 (0.12)
Jitter PPQ (%)	0.22 (0.14)	0.21 (0.06)	0.24 (0.23)	0.28 (0.11)
MPT (sec)	14.0 (4.3)	17.8 (5.8)	17.7 (5.8)	17.8 (5.1)
Habitual SPL (dB)	76.2 (3.5)	74.5 (3.2)	74.8 (4.5)	72.1 (4.7)
Minimum SPL (dB)	60.4 (5.8)	55.3 (5.3)	55.3 (9.7)	56.8 (5.8)
(Hab - Min) / Hab SPL	0.21 (0.06)	0.26 (0.06)	0.26 (0.13)	0.21 (0.06)
Shimmer local (%)	2.2 (0.8)	2.6 (1.0)	2.6 (1.5)	4.4 (2.6)
Shimmer APQ3 (%)	1.2 (0.5)	1.4 (0.6)	1.5 (0.8)	2.4 (1.5)
Shimmer APQ5 (%)	1.3 (0.5)	1.6 (0.6)	1.5 (0.7)	2.7 (1.6)
Shimmer APQ11 (%)	1.5 (0.5)	2.1 (0.7)	1.7 (0.7)	3.4 (1.9)
HNR (dB)	24.5 (3.0)	22.9 (2.8)	23.8 (3.7)	21.2 (4.6)

Table 2. P-values from comparisons of means using the Mann-Whitney U test. Bolded values highlight significant age- or gender-related differences ($p < 0.01$).

MEASURE	GROUPS COMPARED			
	Younger F vs. M	Older F vs. M	Female Y vs. O	Male Y vs. O
Habitual F ₀ (Hz)	-	-	< 0.001	0.112
Maximum F ₀ (Hz)	-	-	0.057	0.960
Max F ₀ / Habitual F ₀	< 0.001	< 0.001	< 0.001	0.335
Jitter local (%)	0.058	0.129	0.949	0.081
Jitter RAP (%)	0.380	0.326	0.997	0.264
Jitter PPQ (%)	0.072	0.064	0.949	0.044
MPT (sec)	< 0.001	0.945	< 0.001	0.820
Habitual SPL (dB)	0.010	0.094	0.111	0.048
Minimum SPL (dB)	< 0.001	0.960	0.001	0.672
SPL decrease ratio	< 0.001	0.208	0.011	0.029
Shimmer local (%)	0.014	0.019	0.278	0.012
Shimmer APQ3 (%)	0.121	0.030	0.297	0.014
Shimmer APQ5 (%)	0.012	0.004	0.479	0.008
Shimmer APQ11 (%)	< 0.001	0.003	0.146	0.034
HNR (dB)	0.001	0.036	0.446	0.136

3.1 Gender differences

On average, younger and older adult males showed a greater increase in pitch in the maximum pitch task than their female counterparts. Younger males also showed a longer MPT than younger females, and younger males were able to soften their voice more than younger

¹ The measure of HNR in SpeechClinic is defined as the ratio of the amount of energy in the harmonic component of the signal to the amount of energy in the noise component as determined by autocorrelation [3]. However, this quantity is not comparable to the quantity reported by MDVP (KayPENTAX) as "HNR".

females in the minimum SPL task. Adult males of both age groups showed higher shimmer values than adult females.

3.2 Age differences

The data revealed lower pitch values for older females compared to their younger counterparts, but there was no difference between younger and older males. Older females had a longer MPT than younger females, but there were no differences between younger and older males. Older males exhibited higher mean shimmer values than younger males. In the current study, jitter was generally stable and did not change with age or gender.

3.3 In comparison to the literature

Our MPT values for young adults were lower than most mean values reported from other studies [4], while we had higher values for HNR [5] and female F₀ [6]. Our jitter values were comparable to those in other studies [7].

4. DISCUSSION

Our preliminary collection of voice samples creates a first impression for several voice characteristics that are currently not represented in the literature. Furthermore, our data show age and gender differences that warrant further investigation. Although these are preliminary data, the ultimate sample will be large enough to derive important distributional characteristics for a large number of voice parameters for healthy individuals. These results will be useful for future studies using SpeechClinic (e.g., for evaluating the impact of smoking or vocal training). We also plan to repeat this study with background noise to mimic more realistic situations found in clinical settings and increase the ecological validity of the reference values. Finally, the study has shown the usefulness and robustness of SpeechClinic as a clinical and research tool.

REFERENCES

- [1] Oguz, H, Demirci, M, Safak, MA, Arslan, N, Islam, A, & Kargin, S (2007). Effects of unilateral vocal cord paralysis on paralysis on objective voice measures obtained by Praat. *Eur Arch Otorhinolaryngol*, 264:257-261.
- [2] Fernandes, D & Van Lieshout, P (2008). Assessment software for voice clinicians. American Speech-Language Hearing Association Convention, November 19-22, Chicago, IL. (Poster Presentation)
- [3] Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *IFA Proceedings* 17, 97-110.
- [4] Wuyts, FL, De Bodt, MS, Molenberghs, G, Remacle, M, et al. (2000). The Dysphonia Severity Index: An objective measure of vocal quality based on a multiparameter approach. *J Speech Lang Hear Res*, 43(3), 796-809.
- [5] Wolfe, V, Fitch, J, & Cornell, R (1995). Acoustic prediction of severity in commonly occurring voice problems. *J Speech Hear Res*, 38(2), 273-279.
- [6] Ferrand, CT (2006). Relationship between masking levels and phonatory stability in normal-speaking women. *J Voice* 20(2), 223-228.
- [7] Gelfer, MP and Fendel, DM (1995). Comparisons of jitter, shimmer, and signal-to-noise ratio from directly digitized versus taped voice samples. *J Voice*, 9(4), 378-382.

ASSESSING THE INTRINSIC RELATIONSHIP BETWEEN FACIAL MOTION AND ACOUSTICS IN PATIENTS WITH PARKINSON'S DISEASE

Luyao Ma^{1,2}, Huawei Colin Li^{1,3,4}, Akiko Amano-Kusumoto^{1,5}, Willy Wong^{3,4}, and Pascal van Lieshout^{1,2,3,6}

¹Oral Dynamics lab, Department of Speech-Language Pathology, University of Toronto, Toronto, luyao.ma@alumni.utoronto.ca; ²Department of Psychology, University of Toronto; ³Institute of Biomaterials and Biomedical Engineering, University of Toronto; ⁴Department of Electrical and Computer Engineering, University of Toronto; ⁵OGI School of Science & Engineering, Oregon Health & Science University, Portland, USA; ⁶Toronto Rehab, Toronto

1. INTRODUCTION

Gestural patterns shape the vocal tract dynamics in both visual and auditory ways, so the perception of speech is not bound to the auditory modality. From the perceivers' perspective, the coherence of observed visual information and acoustic signals is very important for comprehension¹. Given the multi-model nature of speech perception, the congruency between facial motility and acoustic signals is an important factor in how clearly a person produces speech and how others perceive the intended message. The study described here focuses on one particular population where both facial motility and voice quality are impaired, namely individuals with Parkinson's disease.

Parkinson's disease (PD) is a common neurological condition, characterized by muscle rigidity, tremor and slowness in physical movement and is prevalent in people about 50 years of age². One of the most severe consequences of PD is the lost of expressiveness in the face. The relatively weak voice in PD speakers in combination with reduced oral and facial movements makes their speech less intelligible and can have serious social consequences³⁻⁵. Furthermore, these adverse acoustic changes in PD which affect prosodic contrast in speech are evident in earlier stages of disease progression³.

One way to study the congruency between speech acoustics and visual information is by mapping the relationship between acoustic data and facial motility. If the signals are highly congruent with each other, this model would provide an accurate prediction of facial motion from acoustic input. This was confirmed in a recent study from our lab using a linear multi-regression (MLR) model with data from healthy young speakers⁶. To date, it remains an open question to what extent this relation is different in people with PD.

The current study uses 3D motion data with time-aligned acoustics acquired from participants with PD, age-matched healthy speakers, and young healthy speakers. In line with previous work, we used a MLR model which has been shown to provide a good predictor model⁶. It can be hypothesized that the congruency between acoustic signals and facial motion in PD may be lower in comparison with age-matched healthy speakers and with young healthy speakers. Apart from providing important theoretical knowledge about the audio-visual relationship in speech of these populations, the results may have implications for the future development of facial motion based speech recognition software.

2. METHOD

2.1 Participants

The experimental group consisted of individuals (N=8; mean age 62.8 years) diagnosed to be in early stages of PD, recruited from the Morton & Gloria Shulman Movement Disorders Center at the Toronto Western Hospital. Two groups of healthy participants were included: an age-matched control group (OC; N=10, mean age 69.1 years) and a young control group (YC; N=10, mean age 26.3 years). Participants were excluded from the study if they shown any history of neurological disorder or disease (other than PD), orofacial musculoskeletal abnormalities, speech disorders, any history of drug and/or alcohol abuse, and hypersensitivity to sunlight, as Blacklight illumination was used for motion tracking. All groups consisted of native Canadian English speakers only.

2.2 Stimuli

The speech stimuli chosen in this experiment were the same as in a previous study done on young adults and consisted of 90 sentences selected from the TIMIT and HARVARD sentence database⁶. These speech stimuli were considered to provide a representative sample of linguistic materials used in daily speech.

2.3 Procedures

Eleven glow-in-blacklight dots of face paints, about 2 mm in diameter, were applied to various locations on the face of participants. Locations include midsagittal positions on forehead, the dorsum of the nose, upper and lower lip, chin, cheeks and lip corners. The forehead position was used as a reference point for head movement correction (Figure 1). With respect to the selected gestures, F_UL represents upper lip motion relative to forehead marker; BC (bilabial closure or lip aperture) represents upper lip versus lower lip motion; F_JAW represents chin movement relative to the forehead. CHEEKS represent the motion of the left relative to the right cheeks and BP (bilabial protrusion) represent left lip corner versus right lip corner movement.

During the experiment, participants were seated in a darkened, UV illuminated room. In order to generate 3D representation, two digital camcorders were used for motion tracking located just to the left-of-center and right-of-center of the face, approximately 1 meter away and at a 45 degree angle.

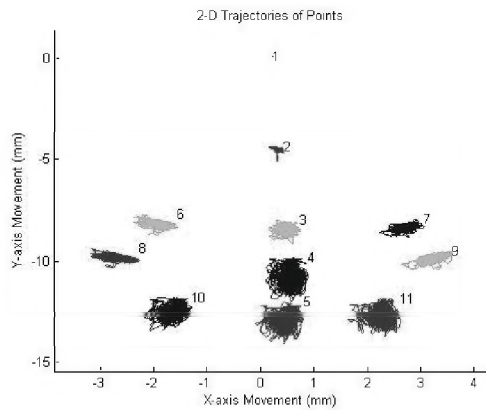


Figure 1, Example of 2-D Trajectories of the 11 markers' position for individual articulators.

The ninety sentences were randomly presented to the participants with no repetition. Participants were instructed to read sentences in a normal speaker manner. Speech acoustics was recorded with a digital voice recorder (Marantz PMD670/U1B) at 22 kHz sampling rate.

Visual and acoustic data were represented in matrix form. The acoustic information was represented by an array of 65 parameters: 16th order Linear Predictor Coefficient and 16th order Line Spectral Pairs and their first derivatives, and the root mean squared energy. In order to predict motion data from acoustic data in each time frame, an MLR analysis was performed⁶. Acoustic data corresponding to one sentence was used for testing, while remaining acoustic and visual data were used for training the MLR model. A correlation coefficient (CC) was calculated between the predicted and acquired movement⁶. This provides an index of the congruency between acoustics and facial motion.

3. RESULTS

We tested for differences in CC for GROUP (PD versus OC versus YC) and GENDER (males vs. females) for each gesture separately (F_UL, BC, F_JAW, CHEEKS, and BP) using repeated measures ANOVA with z-score transformed correlations. The original CC values for GROUP and Gesture are shown in Table 1. For BC, there was a significant GROUP effect, $[F(2,19) = 4.04, p = 0.03]$, showing lower CC values for the PD group when compared to YC but not OC. BP also showed a GROUP effect, $[F(2,19) = 4.51, p = 0.02]$, with YC having significantly lower correlations than OC. No other effects were found significant.

4. DISCUSSION

The findings of the current study show that individuals at an early stage of Parkinson's disease show relatively spared speech related functions, at least with the stimuli set presented in this study. However, PD subjects do show a significantly lower CC value than YC, yet not different from OC. Thus, this effect may be more of an age-related phenomenon. Lip aperture (our BC gesture) is considered the most important component in bimodal speech perception⁷ and a reduced bimodal congruency may impact on speech intelligibility. This would fit with other

changes in the elderly voice as reported in the literature⁸.

Even though the gender effect was not significant, female PD subjects show larger differences than male PD subjects with respect to bilabial closure. Perhaps this reflects the larger fluctuations of loudness observed in female PD patients⁹. We have no clear idea what caused YC subjects to show less congruency in lip protrusion compared to older subjects, but perhaps older subjects due to the decrease in bimodal congruency in lip aperture, compensate with stronger lip movements in the horizontal dimension.

CC(STD)	PD	OC	YC
Gesture			
F_UL	0.45 (0.16)	0.44 (0.14)	0.40 (0.09)
BC	0.54 (0.13)	0.55 (0.10)	0.66 (0.08)
F_JAW	0.52 (0.14)	0.58 (0.12)	0.64 (0.08)
CHEEKS	0.56 (0.09)	0.58 (0.07)	0.46 (0.14)
BP	0.51 (0.10)	0.57 (0.08)	0.48 (0.06)
Mean	0.51 (0.13)	0.55 (0.11)	0.53 (0.14)

Table 1, CCs of individual gestures for 3 groups

ACKNOWLEDGEMENTS

This research was undertaken, in part, thanks to funding from the Canada Research Chairs program.

REFERENCES

- Rosenblum LD, Miller RM, Sanchez K (2007) Lip-read me now, hear me better later: cross-modal transfer of talker-familiarity effects. *Psychological Science* 18:392-396
- Pinto S, Ozsancak C, Tripoliti E, Thobois S, Limousin-Dowsey P, Auzou P (2004) Treatments for dysarthria in Parkinson's disease. *Lancet neurology* 3:547-556
- Cheang HS, Pell MD (2007) An acoustic investigation of Parkinsonian speech in linguistic and emotional contexts. *Journal of Neurolinguistics* 20:221-241
- Miller N, Allcock L, Jones D, Noble E, Hildreth AJ, Burn DJ (2007) Prevalence and pattern of perceived intelligibility changes in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, 78: 1188-1190
- Tickle-Degnen L, Lyons KD (2004) Practitioners' impressions of patients with Parkinson's disease: the social ecology of the expressive mask. *Social Science & Medicine* 58:603-614
- Craig MS, van Lieshout P, Wong W (2008) A linear model of acoustic-to-facial mapping: model parameters, data set size, and generalization across speakers. *Journal of the Acoustical Society of America* 124:3183-3190
- Kaynak M, Zhi Q, Cheok A, Sengupta K, Jian Z, Chung K (2004) Lip geometric features for human-computer interaction using bimodal speech recognition: Comparison and analysis. *Speech Communication* 43:1-2
- Benjamin B (1997) Speech production of normally aging adults. *Seminars in Speech and Language* 18:135-141
- Hertrich I, Ackermann H, Braun S, Spieker S (1996) Gender-specific vocal dysfunctions in central motor disorders. *Sprache Stimme Gehor* 20:169-174

MEN, WOMEN AND LENITION: GENDER DIFFERENCES IN THE PRODUCTION OF INTERVOCALIC VOICED STOPS IN MEXICAN SPANISH

Anna Limanni

Dept. of Spanish and Portuguese, University of Toronto, Victoria College #208, 91 Charles Street West, Toronto, Ontario, Canada M5S 1K7, anna.limanni@utoronto.ca

1. INTRODUCTION

Recent experimental studies suggest that the lenition of Spanish voiced and voiceless stops is a variable process influenced by phonetic factors including prosodic and segmental context [3, 6, 7]. Phonological factors may also play a role. Specifically, it has been suggested that inventory constraints may inhibit the degree of lenition in order to maintain a segmental or class contrast [7]. Other studies have suggested that dialectal differences may also influence lenition and that Spanish varieties do not lenite stops to the same degree. [2, 6]. Thus, a complete account of the lenition patterns for Spanish stops requires the investigation of extralinguistic factors (like dialect) along with phonological and phonetic factors. In studies on other Romance varieties, an additional extralinguistic factor found to influence lenition of intervocalic consonants is gender. In a study of lenition of intervocalic voiced and voiceless stops in Florentine Italian, Villafaña Dalcher [9] found “a much higher incidence of lenition in the female subjects” (p. 157). However, the influence of gender has not been experimentally investigated for Spanish varieties. With the objective of addressing this experimental gap, the present study explores the effects of gender in the realization of intervocalic /b d g/ in Mexican Spanish. We test the hypothesis that significant gender differences exist in the degrees of lenition of these consonants. In a departure from the above studies, the data examined here come from interviews designed to elicit conversational speech.

2. METHOD

2.1 Data Collection

To test the above hypothesis data were collected from a total of six native speakers (three males and three females) of a similar variety of Mexican Spanish (Mexico City and nearby areas). The speakers ranged in age from early-20s to late-40s. Data come from hour-long interviews during which speakers performed the following: (i) a picture identification task and (ii) a story-telling task where speakers were given three sets of sequential pictures and asked to relate the events in each set. These tasks were designed to elicit a casual, conversational speech style. To this end, in the picture identification task, the picture itself did not represent the target word but rather served to introduce questions meant to elicit the target word embedded in a longer stretch of speech, rather than as a single, carefully articulated word. The participants were recorded with a Marantz CDR300 CD recorder and an

AudioTechnica unidirectional condenser microphone in a quiet room (Multimedia Lab, Carr Hall, University of Toronto). The recorded files were then downsampled (22050 Hz, 16 bits) and low-pass filtered to remove frequencies over 11,000Hz.

2.2 Measurement and Analysis

In total, 426 tokens were analyzed (235 for the female speakers and 191 for the male speakers) using Praat 4.4.26. The acoustic parameters used as measures of lenition were duration and speed of consonant release.

Duration has been identified as a very robust indicator of lenition [5, 6, 9] with shorter segments considered more lenited. First, the absolute duration in milliseconds of the intervocalic was taken for each token. The beginning of the intervocalic C was taken to be the point at which there was a rapid drop in intensity and a decreased formant structure from V-to-C and the end of the C was taken to be the point where a rapid increase in intensity as well as a clearer formant structure were apparent, thus signaling the start of the second V. Then, to compensate for possible inter- and intra-speaker variation in speech rate, the absolute duration of the intervocalic consonant was normalized as a proportion of the total duration of the VCV sequence [9]. It is these normalized durations which are reflected in the analysis.

Speed of consonant release is a measure of the difference in intensity between a C and a following V over time. It is measured in decibels per second (dB/s) [7]. A smaller number reflects a lower speed of consonant release and a more lenited, more vowel-like production.

Tokens with no discernible difference in intensity and/or formant structure in the VCV transition or which had a consonant release of less than 1dB/s were deemed cases of deletion and were not included in t analyses. Also not included were cases of intervocalic /d/ which were part of a past participle, *-ado* or *-ido*. The purpose of this omission was to avoid inflating the rate of intervocalic /d/ lenition since /d/ appearing in the context of past participles (especially *-ado*) is subject to morphological as well as phonetic conditioning [1].

Results were evaluated using repeated-measures ANOVAs with the statistical software SPSS 16, with *p* level set at .05.

3. RESULTS

3.1 Duration

Figure 1 shows the results for the relative duration of intervocalic /b d g/. The values reflect a RM-ANOVA with *consonant* as the within-subject factor and *gender* as the between-subject factor. The male speakers exhibit a shorter duration for all three places of articulation and thus appear to lenite more. However, this difference is not large enough to be statistically significant. On average, the relative duration of the female speakers' consonants was approximately 23.5% compared to 21.8% for the male speakers.

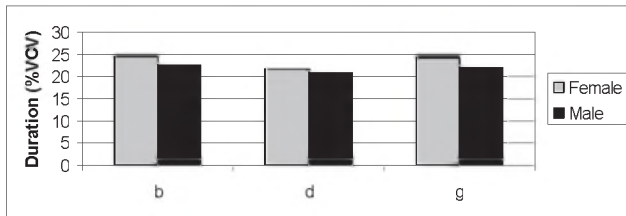


Figure 1. Relative duration of intervocalic /b d g/ in Female and Male Mexican Spanish speakers

3.2 Speed of Consonant Release

The results for speed of consonant release are given in Figure 2. The values again reflect a RM-ANOVA with *consonant* as the within-subject factor and *gender* as the between-subject factor. On this measure we find that males have a lower speed of consonant release across all places of articulation. Here the difference appears greater than with duration, with the female speakers averaging a speed of consonant release of 13.4 dB/s compared to 8.1 dB/s for the male speakers. However, as above, this observed difference does not reach statistical significance.

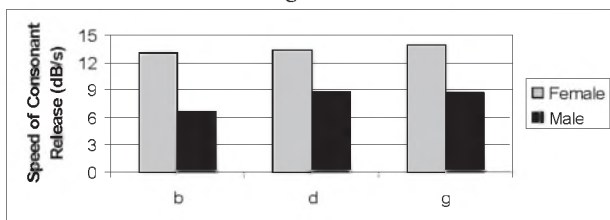


Figure 2. Speed of consonant release of intervocalic /b d g/ in Female and Male Mexican Spanish speakers

4. DISCUSSION

Results show that on both measures the Mexican Spanish males lenite more. This is reflected in a shorter duration and in a slower speed of consonant release in the consonants produced by the male speakers, across all places of articulation. Despite this observed trend, our hypothesis is not supported since the effects of gender on lenition of intervocalic /b d g/ did not reach statistical significance. However, preliminary results from four speakers of Argentine Spanish (two males, two females) show the same tendency. Figure 3 shows that their results for speed of consonant release mirror those obtained for the Mexican

speakers. Thus, a future direction for this study is to analyze more speakers across different dialects to determine whether gender is indeed a significant influence on lenition.

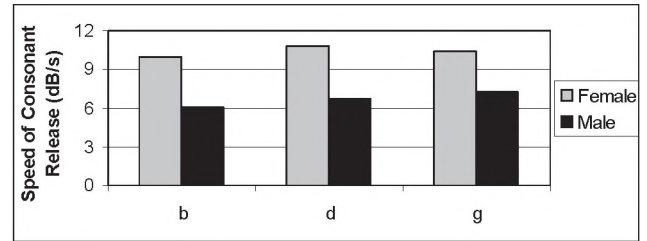


Figure 3. Speed of consonant release of intervocalic /b d g/ in Female and Male Argentine Spanish speakers

Our results also appear to contradict Villafaña Dalcher [9] who found that in Florentine Italian female speakers lenite more. However, we argue that in both cases the results reflect the tendency for females to prefer formal and/or prestige speech variants of stable linguistic variables [4]. Thus, the Florentine Italian female speakers may adopt higher levels of lenition because this lenition is a positive social marker, associated with the Florentine identity [9]. In the case of Spanish varieties, the opposite trend is expected since lenition, while not necessarily a negative social marker, may be associated with a less formal way of speaking [8].

REFERENCES

- [1] Bybee, Joan. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.
- [2] Carrasco, P. (2007). Variability of Spanish Voiced Stop Lenition in Post-consonantal Position. Paper presented at the Hispanic Linguistic Symposium 2007. University of Texas at San Antonio. November 1-4.
- [3] Cole, J., Hualde, J. I. and Iskarous, K. (1999). Effects of prosodic and segmental context on /g/-lenition in Spanish. In O. Fujimura, B.D. Joseph and B. Palek (Eds). *Proceedings of the Fourth International Linguistics and Phonetics Conference*. Prague: The Karolinum Press.
- [4] Labov, William. (2001). *Principles of Linguistic Change: Social Factors*. Oxford: Blackwell.
- [5] Lavoie, L. (2001). Consonant strength: phonological patterns and phonetic manifestations. New York: Garland.
- [6] Lewis, Anthony. (2000). Acoustic variability of intervocalic voiceless stop consonants in three Spanish dialects. In H. Campos et al. (Eds.), *Hispanic linguistics at the turn of the millennium*. Somerville, MA: Cascadilla Press.
- [7] Ortega-Llebaria, M. (2004). Interplay between phonetic and inventory constraints in the degree of spirantization of voiced stops: Comparing intervocalic /b/ and intervocalic /g/ in Spanish and English. In T. Face (Ed.), *Laboratory Approaches to Spanish phonetics and phonology*. The Hague: Mouton de Gruyter.
- [8] Pérez, H.E. (2007). Estudio de la variación estilística de la serie /b-d-g/ en posición intervocálica en el habla de los noticieros de la televisión chilena. *Estudios de fonética experimental* XVI: 227-259.
- [9] Villafaña Dalcher, Cristina. (2006). *Consonant weakening in Florentine Italian: An acoustic study of gradient and variable sound change*. Ph.D. dissertation. Georgetown University, Washington DC.

SEPARATING NORMAL VARIATION IN MOVEMENT AMPLITUDES FROM GRADIENT SPEECH ERRORS

Anneke Slis¹, and Pascal van Lieshout^{1,2,3,4}

¹Oral Dynamics Lab, Department of Speech-Language Pathology, University of Toronto, anneke.slis@utoronto.ca;

²Toronto Rehab, Toronto, Ontario, Canada

³Department of Psychology, University of Toronto Mississauga, Mississauga, Ontario, Canada

⁴Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada

1. INTRODUCTION

The results presented in this paper are part of a larger study that looks at articulatory constraints and its influence on the occurrence of gestural intrusion and reduction errors, as described in a study by Goldstein, Pouplier, Chen, Saltzman, and Byrd¹.

Until recently, speech errors were mostly explained in abstract models of speech production², in which the feature or phoneme is the smallest unit of speech. The idea of the phoneme as the smallest unit of speech, however, is challenged by evidence from recent kinematic studies¹³. In these studies, speech errors, when studied at the level of articulatory gestures, frequently resulted into sub-phonemic speech errors. Most errors consisted of the activation of a gesture, not supposed to be involved in that particular sound production, accompanying the production of the actual target gesture. These unintended activations were often gradual in nature. In order to define a gradual error, Goldstein et al. compared the activation level of a gesture, for example tongue dorsum (TD) in alternating trials (e.g. *cop top*) with the mean activation level of the TD in non-alternating trials (see figure 1).

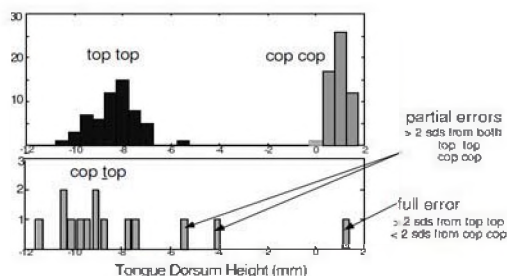


Figure 1. Comparison between amplitudes of TD gestures in alternating and non-alternating trials. From: Pouplier³, p. 37.

They did indicate that this method introduced the risk of comparing speech sequences with different coarticulatory properties, which could affect the analysis. This issue was reinforced in a recent study where it was found that higher production rates could be achieved in alternating trials, *tapa*, compared to non alternating trials, like *tata*, because of less energy consumption for the jaw⁴. This indeed suggests a different production mechanism for alternating trials compared to non alternating trials.

Another aspect that can make it difficult to separate normal amplitude variation from intrusion errors is that the amount of amplitude variability of the articulators differs across types of constrictions and contexts⁵⁶, so different production

constraints could play a role⁶. Finally, the articulators can be activated as a consequence of active control in forming a constriction gesture or they can move as a passive consequence because of the activation of another gesture⁵, in which the amount of passive movement is dependent on the context. Taken all this into account, it is important to determine what the co-articulatory properties of a C1VC C2VC sequence are and how these differ from non-alternating trials C1VC C1VC. This information will help to establish what an intrusion error is, because it allows us to determine the potential contributions of biomechanical constraints, coarticulatory influences, and normal variation⁷. The current study is a first attempt to investigate to what extent changes in movement amplitudes in alternating trials can be related to above mentioned factors. As a first approach, we will use correlations to determine the (in) dependence of gestures and their contributing articulators⁵⁸⁹.

2. METHOD

Speech errors were invoked by a repetitive speech task, comparable to the study described in Goldstein et al¹. Two speech rates were employed, normal and fast, which were individually determined for each participant and controlled by metronome presentation. Movement data were recorded with the 3D EMA system¹⁰. The raw movement amplitudes were normalized such that the maximum amplitude of a constriction per trial was set to 100 % and the minimum constriction per trial was set to 0 %. This way it is possible to compare across trials and speakers. Maxima for TT and TD during each segment were automatically determined using a peak-picking algorithm. In this paper, data are presented for the bisyllable *topcop* in the normal speaking rate condition for 8 participants. For each separate trial, correlations were calculated between the normalized amplitude of the target gesture and that of the co-occurring non-target gesture in the bisyllables *top cop*, *top top* and *cop cop*. Furthermore, correlations between normalized amplitudes of the same gestures at the target and non-target location were used to determine co-articulatory influences across syllable positions. Finally, correlations are reported for the tongue tip and tongue dorsum after subtracting the contribution of the jaw to assess the latter influence on these measures. It was hypothesized that if the correlation between target and non-target is high, TT and TD are not moving independently from each other and biomechanical constraints play a role⁵⁸. Moreover, it was hypothesized that if the gestures were highly dependent due to co-articulatory constraints, high

correlations for the same gesture in target and non-target positions should be found across words.

3. RESULTS AND DISCUSSION

The correlation values are shown in table 1a and b. No consistent significant positive correlations were found between target and non target gestures in the same syllable, except for one participant, C9.

1a

Top Cop												
	Beats per minute	With jaw					Without jaw					n
		TT TD top	TD TT cop	TT TD top	TD TT cop	TD TT top	TT TD cop	TD TT top	TT TD cop	TD TT top	TT TD cop	
C6	80	0.01	0.10	0.54	-0.20	0.09	0.3	-0.2	-0.1	0.16	16	
C7	92	-0.28	-0.06	-0.07	-0.26	-0.23	-0.16	-0.04	0.14	0.17	17	
C8	90	0.34	0.47	0.02	0.10	-0.2	0.08	-0.05	0.4	0.17	16	
C9	115	-0.15	0.89	-0.18	0.40	-0.04	0.21	-0.44	0.17	0.16	16	
C10	90	-0.5	0.03	-0.04	-0.56	-0.59	-0.1	-0.55	-0.48	0.18	17	
C11	100	-0.41	-0.53	-0.27	-0.44	0.18	-0.20	-0.33	0.63	0.17	17	
C14	90	0.20	0.40	0.13	0.28	0.25	0.2	-0.03	-0.35	0.16	17	
C19	90	0.3	0.45	-0.46	-0.42	-0.74	0.25	-0.12	0.05	0.17	17	

1b

top top												
	Beats per minute	With jaw					Without jaw					n
		TT TD top1	TD TT top2	TT TD top1	TD TT top2	TD TT top1	TT TD top2	TD TT top1	TT TD top2	TD TT top1	TT TD top2	
C6	80	-0.04	-0.24	0.02	0.28	0.3	-0.3	-0.14	-0.22	0.22	0.17	17
C7	92	0.75	0.37	-0.3	0.2	0.17	-0.01	0.47	-	-	0.17	17
C8	90	0.47	0.71	-0.13	0.01	0.16	0.75	-0.04	0.33	0.29	0.16	16
C9	115	0.25	0.18	0.55	0.56	0.17	-0.23	-0.12	-0.02	0.08	0.17	16
C10	90	0.33	-0.30	-0.58	-0.37	0.17	0.59	0.49	-	-	0.15	15
C11	100	0.08	0.34	0.12	-0.08	0.17	0.23	0.41	0.32	0.73	0.16	16
C14	90	-0.3	0.65	-0.35	-0.05	0.17	-	-	-	-	0.17	17
C19	90	0.78	0.33	-0.08	0.44	0.18	0.01	0.38	0.18	0.65	0.17	17

Table 1a & b. Correlations for TT and TD in both alternating (top) and non-alternating (bottom) trials. The last column shows the number of repetitions in a trial. The marked boxes are significant at $p < 0.01$.

Figure 2 shows movement data for subjects C9 and C6 (who had very low correlations). The significant correlation between TT and TD in the word cop for C9 may indicate a passive movement of the TT during TD, due to the shared organ, the tongue root¹. However, after subtracting the jaw component, correlations in most cases decreased, suggesting a common influence of the jaw underlying the tongue movements during /k/ especially for C9. In general, correlations between TD and TT were low and non-significant, suggesting that biomechanical constraints of the tongue tip and tongue dorsum are not a major factor in our study. This is also in accordance with Jackson & Singampalli⁸ who showed similar low correlations between TD and TT. They did find a high correlation between jaw and TT, which would fit our findings for jaw-corrected tongue movements. One exception is C6, who showed a very high correlation in the no-jaw condition for TD and TT in the word cop.

With respect to correlations across syllables, no consistent correlations were found between the target gesture and the same gesture in the non target position. This means that the activation of a gesture in the target position doesn't influence the activation of this gesture in a non target position and consequently doesn't contribute to co-articulatory influences. Finally, there are large individual differences, as can be observed in table 1 and figure 2, and it is safe to assume that speakers use individual control strategies that may or may not result in a higher correlation between TT and TD.

In the non-alternating trials more significant correlations between TT and TD were found. The values for the non-alternating trials in the word top were slightly higher. However, the values didn't show a consistent pattern, ruling out a biomechanical or co-articulatory constraint. Again, the jaw plays an important role in the higher correlation values. Once the jaw is left out, most of the significant values become insignificant. The fact that non alternating trials show a different pattern than the alternating trials means that these two conditions cannot be compared. Therefore, the use of non-alternating trials for the estimation of intrusion or reduction errors in alternating trials can affect the analysis and may cause an inaccurate determination of normal variation.

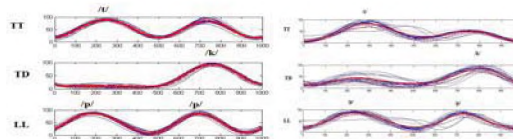


Figure 2. C9 (left) and C6 (right). The y-axis shows normalized amplitude and the x-axis shows normalized time.

REFERENCES

- Goldstein, L., Pouplier, M., Chen, L., Saltzman, E., & Byrd, D. (2007). Dynamic action units slip in speech production errors. *Cognition*, 103, 386-412.
- Levelt W. (1989). *Speaking. From intention to articulation*. MIT Press, Cambridge.
- Pouplier, M. (2003). *Units of phonological encoding: Empirical evidence*. Unpublished dissertation. Yale University.
- Rochet-Capellan, A. and Schwartz, J.L. (2007). An articulator basis for the Labial-to-Coronal effect: /pata/ seems a more stable articulatory pattern than /tapal/. *Journal of the Acoustical Society of America*, 121(6), 3740-3754.
- Browman, C. P. (1994). Lip aperture and consonant releases. *Phonological Structure and Phonetic Form. Papers in Laboratory Phonology III*, Cambridge: Cambridge University Press, 331-353.
- Koenig, L.L., Lucero, J.C. & Lofqvist, A. (2003). Studying articulatory variability using functional data analysis. *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona, Universitat Autònoma de Barcelona.
- Frisch, S.A. (2007). Walking the tightrope between cognition and articulation: The state of the art in the phonetics of speech errors. In Schutze, C.T. & Ferreira, V.S. (Eds.), *Proceedings of the LSA Institute Workshop. MIT Working Papers in Linguistics*, 53, The State of the Art in Speech Error Research (pp. 155-171), MITWPL.
- Jackson, P.J.B., & Singampalli, V. (2009). Statistical identification of critical articulators in the production of speech. *Speech Communication*, 51(8), 695-710.
- Stone, M., Epstein, M. A., & Iskarous, K. (2004). Functional segments in tongue movement. *Clinical Linguistics & Phonetics*, 18, 507-521.
- Yunusova, Y., Green, J., & Mefferd, A. (2009). Accuracy Assessment for AG500, Electromagnetic Articulograph. *Journal of Speech, Language, and Hearing Research*, 52(2), 547-555.

ACKNOWLEDGEMENTS

We would like to thank Dr. Aravind Namasivayam for his assistance in running the EMA sessions.

TRANSFER PATH ANALYSIS USING ENGINE RADIATED SOUND AND MOUNT VIBRATION

Nikolina Samardzic, and Colin Novak

Dept. of Mechanical, Automotive and Materials Engineering, University of Windsor, 401 Sunset Ave.; Windsor, Ontario

INTRODUCTION

One of the main areas of consideration in defining vehicle customer satisfaction is vehicle interior sound and vibration. Engine radiated sound and mount vibration are significant contributors to vehicle interior sound and vibration. Characterizing their contribution requires a detailed and methodical analysis that may be used to establish baseline levels, identify manufacturing deficiencies, and influence design changes for baseline improvements of vehicle interior sound and vibration. Component-level testing such as engine sound and vibration testing is performed frequently because it is more practical and less costly compared to similar vehicle-level testing. Thus, estimating vehicle performance based on engine performance can potentially provide cost reduction benefits at all stages of vehicle design and manufacturing. The objectives of this study are to:

1. Develop a vehicle interior sound and vibration numerical estimation method based on engine radiated sound and mount vibration and verify the estimations by actual vehicle interior sound and vibration measurements at all operating conditions under investigation.
2. Evaluate the contribution of individual engine air-borne and structure-borne sources of sound and vibration to the total vehicle interior sound and vibration.

DATA ANALYSIS METHODS

The vehicle interior sound and vibration is a combination of air-borne and structure-borne transfer paths. The behaviour of principal sources of sound and vibration is dependent on the vehicle design and its operating conditions. Examples of vibroacoustic sources include wind and road, powertrain, intake and exhaust systems and others. The total sound and vibration level in the vehicle interior is a sum of partial pressures and accelerations caused by all air-borne and structure-borne delivery paths from the sources of sound and vibration. The Transfer Path Analysis (TPA) is an experimental technique used to establish the effect of individual air-borne and structure-borne paths on the interior sound and vibration. The TPA method requires the vibroacoustic source and the transfer function data. A target response due to a single path, $r(\omega)$, is a combination of frequency response function, $H(\omega)$, and the magnitude of the source or excitation, $s(\omega)$, as shown in equation:

$$r(\omega) = H(\omega) \cdot s(\omega)$$

The total receiver sound pressure level, $R(\omega)$, is the sum of n partial pressures of the individual transfer paths, i :

$$R(\omega) = \sum_{i=1}^n \frac{r(\omega)}{s_i(\omega)} \cdot s_i(\omega)$$

Figure 1 illustrates an overview of vehicle interior sound and vibration numerical estimation method used in this study.

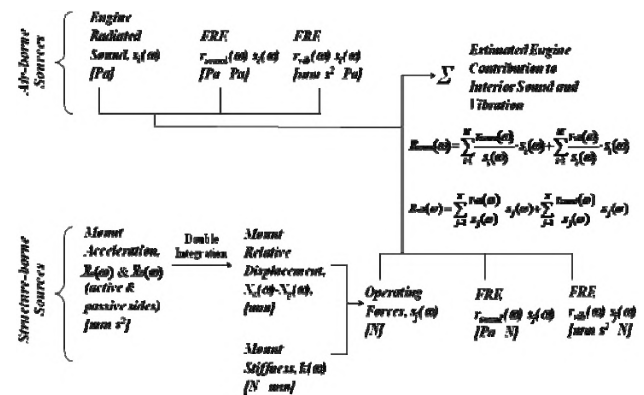


Fig. 1: Vehicle Interior Sound and Vibration Numerical Estimation Method Overview.

The modal analysis was performed to estimate the sound and vibration transmission characteristics from the engine compartment to vehicle interior. It included acoustic and mechanical excitation of the frame and body assembly of a light truck. The calculated frequency response functions are then combined with frequency based source data (engine radiated sound and forces transmitted through the mounts) to estimate the response at the receivers. The response was characterized as the binaural sound pressure and driver's seat track triaxial acceleration. The results were verified by actual vehicle interior sound and vibration measurements.

EXPERIMENTAL DETAILS

The engine radiated sound is a source of air-borne contributions to interior sound and vibration. In order to calculate these contributions it is necessary to quantify acoustical radiation from exterior surfaces of the engine by recording time domain sound data at each operating condition (rpm).

The forces transmitted through engine mounts resulting from engine operation are structure-borne sources of interior sound and vibration. The relative displacement across the engine mounts was required to calculate structure-borne engine operating forces transmitted through the mounts. The relative displacements are proportional to the forces transmitted through the mounts and into the vehicle body at each

operating condition and frequency. The proportionality constant is the frequency-dependent, manufacturer-specified mount stiffness. Initially, the acceleration time data was obtained from the mount vibration test. This acceleration spectra was then calculated for both the active and body sides of the mounts and integrated twice to obtain the displacements. The engine acoustic inputs used to measure FRF's were simulated using white noise input through the loudspeakers. The engine force input through the frame was simulated by impact hammer excitation of the frame, with the engine removed. The interior sound and vibration were measured using binaural head and triaxial accelerometer, respectively at the same five receiver locations used in the FRF measurements.

RESULTS AND DISCUSSION

The results indicate a close correlation between the estimated results and actual measurements (see examples in Figures 2 and 3). This verified the effectiveness of the acoustic and impact excitation methods used in the study to simulate the air-borne and structure-borne engine inputs and the transfer paths for the operating conditions considered.

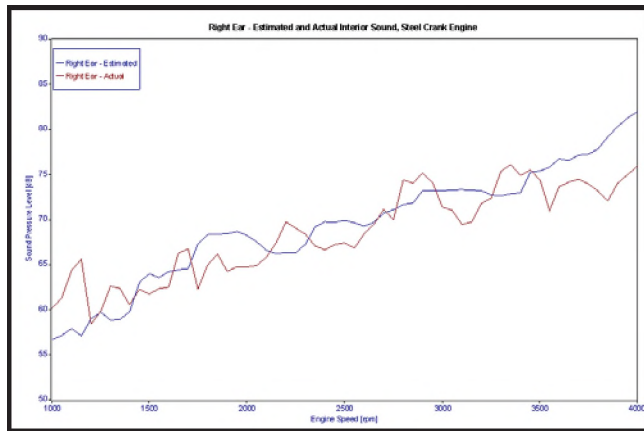


Fig. 2: Estimated and Actual Interior Sound, Right Ear.

The interior sound and vibration are functions of individual contributions of source sound and vibration as well as operating conditions. The transfer path contributions to receivers were calculated for each operating condition and example results are shown in Figure 4.

The contributions of individual engine air-borne and structure-borne paths of sound and vibration to total vehicle interior sound and vibration were evaluated and a ranking method was used for rapid identification of significant paths. The paths were ranked in terms of their average overall contribution to each receiver (Table 1). Common paths of significance included the X-direction of the right mount (consistently the least significant path) and front radiated sound (one of the most significant paths). Table 1 also shows a comparison between transfer path contribution associated with a steel and a cast crank

engine. Any differences in transfer path contributions between the different engines may be used as part of a future interior vehicle sound quality assessment.

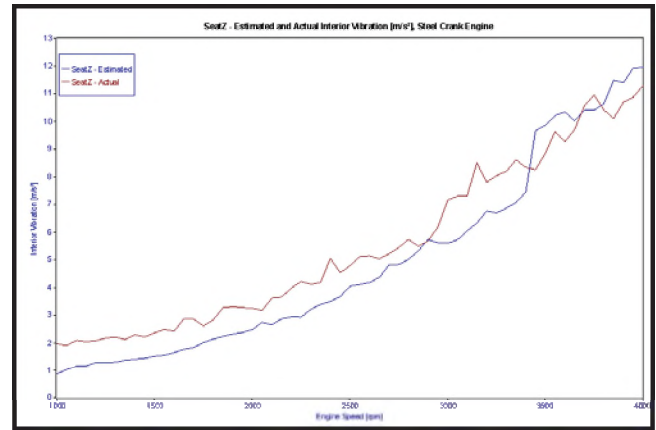


Fig. 3: Estimated and Actual Interior Vibration, Z-direction.

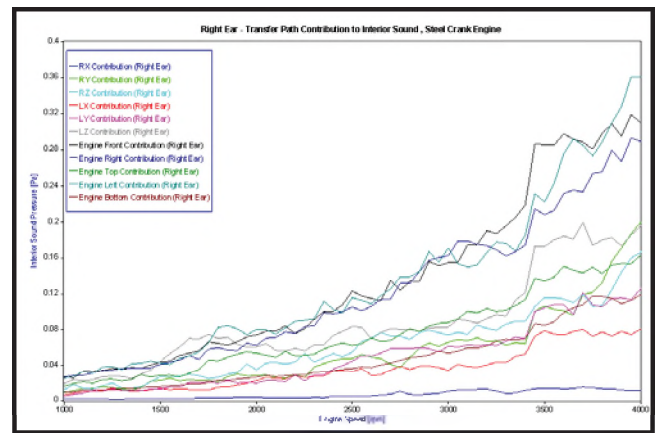


Fig. 4: Transfer Path Contribution to Interior Sound, Right Ear.

Tbl. 1: Transfer Path Contribution Ranking.

		Interior Sound				Interior Vibration					
		Right Ear		Left Ear		X-dir		Y-dir		Z-dir	
		Steel Crank	Cast Crank	Steel Crank	Cast Crank	Steel Crank	Cast Crank	Steel Crank	Cast Crank	Steel Crank	Cast Crank
Air-borne Sources	Front, X-dir	2	1	3	3	1	1	1	1	1	1
	Right, Y-dir	3	3	4	4	4	5	4	4	4	5
	Top, Z-dir	5	7	5	8	5	7	5	5	5	6
	Left, Y-dir	1	2	1	2	2	3	2	2	2	2
Structure-borne Sources	Bottom, Z-dir	9	9	2	1	3	2	3	3	3	4
	Right Mount, X-dir	11	11	11	11	11	11	11	11	11	11
	Right Mount, Y-dir	7	6	8	7	6	9	6	8	6	8
	Right Mount, Z-dir	6	5	7	5	7	4	8	7	7	5
	Left Mount, X-dir	10	10	9	9	10	10	10	10	10	10
	Left Mount, Y-dir	8	8	6	6	6	9	6	7	6	7
Left Mount, Z-dir	4	4	10	10	8	8	9	9	9	9	

Significant:	1 - 4
Moderately Significant:	5 - 7
Insignificant:	8 - 11

REFERENCES

[1] N. Samardzic, "Vehicle Interior Sound and Vibration Numerical Estimation Method Based on Engine Radiated Sound and Mount Vibration", Master's Thesis, University of Windsor, 2005.

AN ACCURATE COUPLED STRUCTURAL-ACOUSTIC ANALYTICAL FRAMEWORK FOR THE PREDICTION OF RANDOM AND FLOW-INDUCED NOISE IN TRANSPORT VEHICLES: ITS VALIDATION

Joana da Rocha¹, Afzal Suleman¹, and Fernando Lau²

¹Dept. of Mechanical Engineering, University of Victoria, PO Box 3055, Stn. CSC, Victoria, BC, Canada, V8W 3P6
jdarocha@uvic.ca, suleman@uvic.ca

²Centro de Ciências e Tecnologias Aeronáuticas, Dept. de Engenharia Mecânica, Instituto Superior Técnico, Av. Rovisco Pais, Lisboa, Portugal, 1049-001 lau@ist.utl.pt

1. INTRODUCTION

Aircraft cabin interior noise and vibration is mostly generated by the external flow excitation and engine noise. In opposition of what happens during takeoff, where the engine noise is the dominant cabin noise source, during cruise flight the airflow sources are the major contribution for the interior noise [1]. Additionally, turbulent boundary layer pressure (TBL) levels on the exterior of the fuselage increases with the flight speed. Similarly, automotive industry is progressively more concerned with passengers comfort. Since major advances have been made to the reduction of sound transmitted to the interior from the engine, transmission, and tires, the reduction of flow-induced noise is becoming more important.

Reduced cabin noise levels are desirable for both comfort and health-related reasons, and they are balanced with cost, complexity, and physical constraints of noise control systems. However, the successful implementation of noise control techniques is a challenging problem, and is far from being a straightforward task. To efficiently design a noise control system, a clear understanding of the mechanisms of sound radiation and transmission of the structural-acoustic system is crucial. Furthermore, when considering the TBL excitation, the noise reduction problem turns into even more complicated, since the TBL-induced pressure has a random and broadband nature. Mathematically, this problem can be simulated by the interaction of three models: (1) an aerodynamic model, which represents the TBL pressure on the cabin structure; (2) a structural model, defining the vibration of the cabin structure; and (3) an acoustic model representing the cabin interior sound pressure level.

The main goal of the current research is to develop an analytical framework for the prediction TBL-induced noise into transport vehicles cabins, and its validation. The knowledge of the characteristics of the TBL excitation, its induced vibration on the structure, and the noise radiated into the cabin is essential for the accurate prediction of the interior noise. The effect of the receiving room space is an important factor for the accurate interior noise prediction.

1.1 Turbulent Boundary Layer Wall Pressure Model

The TBL wall pressure field over a flat plate is usually statistically described in terms of the pressure power spectral density (PSD). Corcos [2], proposed a model which considers the cross PSD of the stationary and homogeneous TBL wall pressure field, over a flat panel, for zero pressure gradient, in a separate form in the streamwise, x -, and spanwise, y -direction, as

$$S(\xi_x, \xi_y, \omega) = S_{\text{ref}}(\omega) e^{-\frac{\alpha_x \omega |\xi_x|}{U_c}} e^{-\frac{\alpha_y \omega |\xi_y|}{U_c}} e^{-\frac{i \omega \xi_x}{U_c}}, \quad (1)$$

where $S_{\text{ref}}(\omega)$ is the reference PSD; $\xi_x = x - x'$ and $\xi_y = y - y'$ are the spatial separations in x and y directions, respectively; α_x and α_y are empirical parameters, chosen to yield the best agreement with the reality. Recommended values for aircraft boundary layers are $\alpha_x = 0.1$ and $\alpha_y = 0.77$.

1.2 Structural Model

Generally, an aircraft fuselage is a conventional skin-stringer-frame structure, with several panels connected between adjacent stringers and frames. The panel was considered to be flat and simply supported in all four boundaries. The plate governing equation, for a given applied external pressure, $p_{\text{ext}}(x, y, t)$, is defined by

$$D_p \nabla^4 w + \rho_p h_p \ddot{w} + \zeta_p \dot{w} = p_{\text{ext}}(x, y, t), \quad (2)$$

in which w is the plate displacement, ρ_p is the density of the panel, h_p is its thickness, D_p is the panel stiffness constant, and the term ζ_p was added to account for the panel damping.

1.3 Acoustic Model

The acoustical physical system consists of a three-dimensional rectangular enclosure, with five fixed walls, and one totally or partially flexible wall. The governing equation of this subsystem is the wave equation, defined by

$$\nabla^2 p - \frac{1}{c_0^2} \ddot{p} - \zeta_{\text{ac}} \dot{p} = 0, \quad (3)$$

where p is the interior pressure, c_0 is the speed of sound inside the enclosure, and the term ζ_{ac} was added to account for the acoustic damping in the enclosure.

2. METHOD FOR SOLUTION

The governing equations for the coupled structural acoustic system are obtained from the combination of the previously described equations for the individual uncoupled systems. To perform this combination, some mathematical manipulations are needed (please refer to [3] for details). The final coupled system governing equation becomes

$$\begin{bmatrix} \mathbf{M}_{pp} & \mathbf{0} \\ \mathbf{M}_{cp} & \mathbf{M}_{cc} \end{bmatrix} \begin{Bmatrix} \dot{\mathbf{q}}(t) \\ \dot{\mathbf{r}}(t) \end{Bmatrix} + \begin{bmatrix} \mathbf{D}_{pp} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{cc} \end{bmatrix} \begin{Bmatrix} \dot{\mathbf{q}}(t) \\ \dot{\mathbf{r}}(t) \end{Bmatrix} + \begin{bmatrix} \mathbf{K}_{pp} & \mathbf{K}_{pc} \\ \mathbf{0} & \mathbf{K}_{cc} \end{bmatrix} \begin{Bmatrix} \mathbf{q}(t) \\ \mathbf{r}(t) \end{Bmatrix} = \begin{Bmatrix} \mathbf{p}_{tbl}(t) \\ \mathbf{0} \end{Bmatrix}, \quad (4)$$

in which \mathbf{M}_{pp} , \mathbf{D}_{pp} , and $\mathbf{K}_{pp} \in \mathfrak{R}^{M \times M}$, \mathbf{M}_{cc} , \mathbf{D}_{cc} and $\mathbf{K}_{cc} \in \mathfrak{R}^{N \times N}$, $\mathbf{M}_{cp} \in \mathfrak{R}^{N \times M}$, $\mathbf{K}_{pc} \in \mathfrak{R}^{M \times N}$; $\mathbf{q}(t)$ and $\mathbf{p}_{tbl}(t) \in \mathfrak{R}^{M \times 1}$, $\mathbf{r}(t) \in \mathfrak{R}^{N \times 1}$; $M = M_x \times M_y$ is the number of plate modes; $N = N_x \times N_y \times N_z$ is the number of acoustic modes; and $w(x, y, t)$ and $p(x, y, z, t)$ were defined, respectively, as:

$$w(x, y, t) = \sum_{m_x=1}^{M_x} \sum_{m_y=1}^{M_y} \alpha_{m_x}(x) \beta_{m_y}(y) q_{m_x m_y}(t), \quad (5a)$$

$$p(x, y, z, t) = \sum_{n_x=1}^{N_x} \sum_{n_y=1}^{N_y} \sum_{n_z=1}^{N_z} \psi_{n_x}(x) \phi_{n_y}(y) \Gamma_{n_z}(z) r_{n_x n_y n_z}(t). \quad (5b)$$

Since the TBL wall pressure field model is expressed in the frequency domain, Eq.(4) is rewritten in the frequency domain, as: $\mathbf{Y}(\omega) = \mathbf{H}(\omega)\mathbf{X}(\omega)$, where $\mathbf{Y}(\omega)$ is the response of the system to the excitation $\mathbf{X}(\omega)$, and $\mathbf{H}(\omega)$ is the frequency response matrix of the system, defined as following:

$$\mathbf{Y}(\omega) = \begin{Bmatrix} \mathbf{W}(\omega) \\ \mathbf{P}(\omega) \end{Bmatrix} \quad \text{and} \quad \mathbf{X}(\omega) = \begin{Bmatrix} \mathbf{P}_{tbl}(\omega) \\ \mathbf{0} \end{Bmatrix}, \quad (6a)$$

$$\mathbf{H}(\omega) = \begin{bmatrix} -\omega^2 \mathbf{M}_{pp} + i\omega \mathbf{D}_{pp} + \mathbf{K}_{pp} & \mathbf{K}_{pc} \\ -\omega^2 \mathbf{M}_{cp} & -\omega^2 \mathbf{M}_{cc} + i\omega \mathbf{D}_{cc} + \mathbf{K}_{cc} \end{bmatrix}^{-1}. \quad (6b)$$

where $\mathbf{W}(\omega)$, $\mathbf{P}(\omega)$ and $\mathbf{P}_{tbl}(\omega)$ correspond to the frequency domain vectors of $\mathbf{q}(t)$, $\mathbf{r}(t)$ and $\mathbf{p}_{tbl}(t)$. One last step, needed to obtain a solution, is to transform the system equations to the PSD domain, as: $\mathbf{S}_{YY}(\omega) = \mathbf{H}^*(\omega)\mathbf{S}_{XX}(\omega)\mathbf{H}^T(\omega)$, where $\mathbf{S}_{XX}(\omega)$ is the PSD matrix of the random excitation, $\mathbf{S}_{YY}(\omega)$ is the PSD matrix of the random response, and superscripts * and T denote Hermitian conjugate and matrix transpose. Using this methodology, it is possible to analytically determine the PSD of the plate displacement, $S_{ww}(\omega)$, and the PSD of the acoustic enclosure pressure, $S_{pp}(\omega)$. Please refer to [3] to access the derived analytical expressions for the PSD functions.

3. RESULTS AND DISCUSSION

Several practical cases available in the literature were used for the validation of our analytical framework. In the present summary, results are shown for the study in [4]. It consists of a rectangular simply supported aluminum panel, excited by a TBL, and coupled with an acoustical enclosure. The value of S_{ref} was provided, and the sound

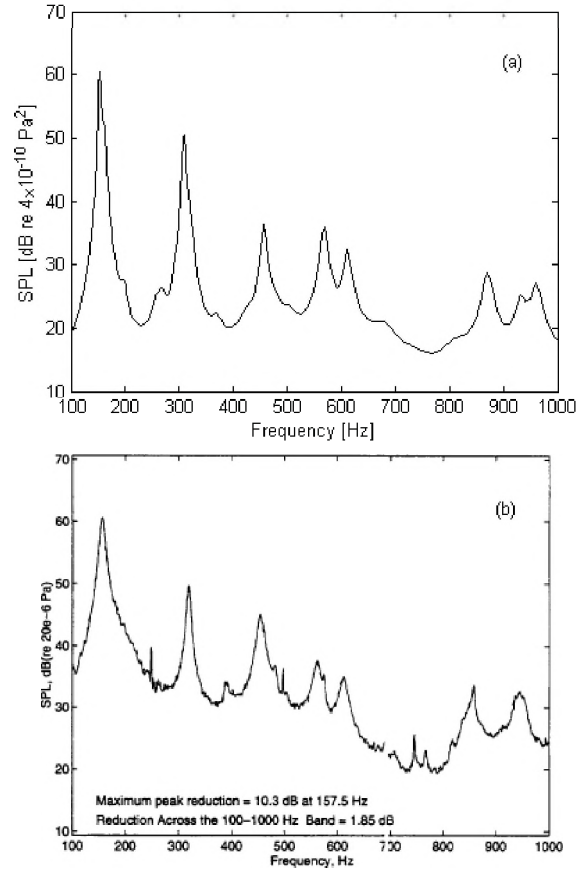


Fig. 1. SPL: (a) from our analytical model; (b) measured, from [4].

pressure level (SPL) was measured inside the acoustic enclosure. To accomplish convergence of the spectral quantities, a total number of $M_x=5$ and $M_y=4$ plate modes, and $N_x=8$, $N_y=6$ and $N_z=5$ acoustic modes were used in our analysis. For the frequency range of interest, [0; 1000] Hz, was necessary to include some non-resonant modes to accurately predict the SPL, showing that the number of plate and acoustic modes plays an important role to obtain a truthful prediction. By comparison of parts (a) and (b) of Fig. 1, it can be concluded that our analytical model provides a good approximation to the experimental data. It is shown that, for an accurate interior cabin noise prediction, is important to consider not only the structural natural modes, but also the cabin acoustic modes, as well.

REFERENCES

- [1] D. E. Bishop, Cruise flight noise levels in a turbojet transport airplane. *Noise Control* (1961), 7, pp. 37-42.
- [2] G. M. Corcos, Resolution of pressure in turbulence. *Journal of the Acoustical Society of America* (1963), 35(2), pp. 192-199.
- [3] C. M. Heatwole, R. J. Bernhard and M. A. Francheck, Robust feedback control of flow induced structural radiation of sound. *NASA Report 0278-1* (1997).
- [4] J. da Rocha, A. Suleman and F. Lau. An accurate coupled structural-acoustic analytical framework for the prediction of random and flow-induced noise in transport vehicles: Its Validation, *Canadian Acoustics* (submitted, July 2009).

EXPERIMENTAL STUDIES ON THE IN-PLANE VIBRATIONS AND SOUND RADIATION IN AN ANNULAR THICK DISK

Salem Bashmal¹, Rama Bhat, and Subash Rakheja

¹Dept. of Mechanical and Industrial Engineering, Concordia University, 1455 De Maisonneuve Blvd. W., Montreal, Quebec, Canada, H3G 1M8, s_bashm@encs.concordia.ca

1. INTRODUCTION

Studies on the in-plane vibrations and sound radiation in circular annular disks are rarely found in literature. Most of the studies are on the out-of-plane vibrations since the contribution from the in-plane vibrations are small compared to out of plane vibrations. However, in thick plates, modes of vibration within the plane of the disks are coupled with the out-of-plane modes, contributing to total vibration and noise radiation (Tzou et al (1998)). Due to this coupling, out-of-plane modes can be excited by forces along the plane of the disk. Moreover, sound radiation from in-plane modes is significant in many engineering applications such as disk brakes and railway wheels (Thompson (2000)).

Non-uniformity in the boundary conditions affects both the in-plane and out-of-plane natural frequencies of the disk. Studies on the out-of-plane vibrations show that some of the natural frequencies have different values for the symmetric and anti-symmetric modes. This observation has been discussed based on analytical and experimental investigations (see for example Bauer and Eidel (2006) and Eastep and Hemmig (1982)). Experimental studies on the in-plane vibrations in circular annular disks subject to non-uniform boundary conditions are not found in literature. Moreover, previous studies on the in-plane vibrations were obtained for disks with uniform boundary conditions only. Review of previous analytical studies on the subject can be found in Bashmal et al. (2008). Few experimental investigations were conducted on in-plane vibrations of circular disks with uniform boundary conditions. Ambati et al. (1976) measured the in-plane vibrations of free annular disks with different radius ratios to study the characteristics of vibrating modes as the annular disk approaches thin ring. The natural frequencies and sound pressure from in-plane vibrations were measured for annular disks with free boundary conditions by Lee (2003). Leung and Pinnington (1987) conducted experiments to measure the in-plane response of stationary disk under rotating edge loads.

The present study presents experimental investigations on the in-plane vibrations of thick plates with non-uniform boundary conditions. The natural frequencies of an annular disk supported at one point are measured and compared with values measured for free support conditions. Moreover, using the Rayleigh-Ritz method, the boundary characteristic orthogonal polynomials are used as admissible functions in a three dimensional analysis to obtain the natural frequencies and the associated mode shapes analytically.

2. THEORY

In this section, a three-dimensional thick disk model is used to investigate the modal characteristics of a stationary circular disk using the Rayleigh-Ritz method. The material of the disk is assumed to be isotropic with mass density ρ , Young's modulus E and Poisson ratio ν . Let the outer radius of the disk be R_o , the inner radius be R_i and the thickness of the disk be h . The radial, circumferential and transverse displacement components of a point in the annular disk are denoted by u_r , u_θ and u_z , respectively. Introducing the non-dimensional parameter, $\xi=r/R_o$, $\zeta=z/h$ the expression for the maximum kinetic (T_{max}) and strain (Π_{max}) energies of the disk in cylindrical coordinates (ξ, θ, ζ) are (So and Leissa (1998)):

$$\begin{aligned} \Pi_{max} = & \frac{1}{4} \frac{Eh}{(1+\nu)} \int_0^{2\pi} \int_{-0.5}^{0.5} \int_\beta^1 \left\{ \frac{2\nu}{(1-2\nu)} \left(\frac{\partial u_r}{\partial \xi} + \frac{1}{\xi} \frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{\xi} \right. \right. \\ & + \frac{R_o}{h} \frac{\partial u_z}{\partial \zeta} \left. \right)^2 + 2 \left[\left(\frac{\partial u_r}{\partial \xi} \right)^2 + \left(\frac{u_r}{\xi} + \frac{1}{\xi} \frac{\partial u_\theta}{\partial \theta} \right)^2 \right] \\ & + \left(\frac{R_o}{h} \frac{\partial u_z}{\partial \zeta} \right)^2 + \left(\frac{1}{\xi} \frac{\partial u_r}{\partial \theta} + \frac{\partial u_\theta}{\partial \xi} - \frac{u_\theta}{\xi} \right)^2 \\ & + \left(\frac{R_o}{h} \frac{\partial u_\theta}{\partial \zeta} + \frac{1}{\xi} \frac{\partial u_z}{\partial \theta} \right)^2 + \left(\frac{R_o}{h} \frac{\partial u_r}{\partial \zeta} + \frac{\partial u_z}{\partial \xi} \right)^2 \left. \right\} \xi d\xi d\zeta d\theta \end{aligned} \quad (1)$$

$$T_{max} = \frac{1}{2} \omega^2 h \rho R_o^2 \int_0^{2\pi} \int_{-0.5}^{0.5} \int_\beta^1 (u_r^2 + u_\theta^2 + u_z^2) \xi d\xi d\zeta d\theta \quad (2)$$

The free in-plane vibration of the disk is assumed to have sinusoidal variations along the circumferential direction of the disk, and may be expressed in the form:

$$\begin{aligned} u_r(\xi, \zeta, \theta) = & \sum_{l=0}^{\infty} \sum_{m=1}^{\infty} \sum_{n=0}^{\infty} [\bar{U}_{c,lmn}(t) \psi_l(\zeta) \phi_m(\xi) \cos(n\theta)] \\ & + \bar{U}_{s,mn}(t) \psi_l(\zeta) \phi_m(\xi) \sin(n\theta) \end{aligned} \quad (3)$$

$$\begin{aligned} u_\theta(\xi, \zeta, \theta) = & \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} \sum_{k=0}^{\infty} [\bar{V}_{c,ijk}(t) \psi_i(\zeta) \phi_j(\xi) \cos(k\theta)] \\ & + \bar{V}_{s,ijk}(t) \psi_i(\zeta) \phi_j(\xi) \sin(k\theta) \end{aligned} \quad (4)$$

$$\begin{aligned} u_z(\xi, \zeta, \theta) = & \sum_{p=0}^{\infty} \sum_{q=1}^{\infty} \sum_{g=0}^{\infty} [\bar{W}_{c,pqg}(t) \psi_p(\zeta) \phi_q(\xi) \cos(g\theta)] \\ & + \bar{W}_{s,pqg}(t) \psi_p(\zeta) \phi_q(\xi) \sin(g\theta) \end{aligned} \quad (5)$$

where \bar{U} , \bar{V} and \bar{W} are radial, circumferential and transverse deflection coefficients, respectively. The subscripts, c and s refer to cosine and sine components of the deflections, respectively. The function $\phi(\xi)$ and $\psi(\zeta)$ are the assumed deflection shape satisfying the geometric boundary conditions in the form of boundary characteristic orthogonal polynomials (Bhat (1985)). The assumed solutions - equations (3) to (5)- are substituted into the energy equations (1) and (2) to predict the natural frequencies using the Rayleigh-Ritz method.

3. EXPERIMENT

An experiment for measuring the natural frequencies of an annular disk subject to different boundary conditions is performed. The disk used in the experiment is an annular steel disk of 300 mm outer diameter, 100 mm inner diameter and 10 mm thickness. An accelerometer is mounted on the outer edge of the disk to capture the acceleration in the radial direction. An accelerometer is mounted on the face of the disk to measure the out-of-plane accelerations. The disk is excited by an impulse hammer along the radial direction. The natural frequencies are measured for the disk supported horizontally on soft cushion to resemble free boundary conditions. Then, the disk is mounted on a vise along portion of its outer edge to represent a point support.

Table 1: In-plane natural frequencies of an annular disk with free edges ($\beta = 0.3, \nu = 0.3$).

Experimental (Hz)	4848	8952	9720	10450
Analytical (Hz)	4821	8865	9602	10332

4. DISCUSSION

The natural frequencies obtained analytically are compared with those measured in the experiment. Table 1 shows the in-plane modes of vibration of a free disk while Table 2 shows the out-of-plane modes of vibration. Out-of-plane modes were measured to distinguish the in-plane modes from out-of-plane modes detected due to coupling between modes. Comparison shows good agreement between analytical and experimental results. The frequency spectrum of the in-plane modes of vibration is also shown in Figure 1. For the disk mounted on a vise, the frequency spectrum, shown in Figure 2, shows that some of the modes split into two different values. This can be attributed to the nature of coupling between different cosine and sine modes due to the non-uniformity of the boundary conditions.

Table 2: Out-of-plane natural frequencies of an annular disk with free edges ($\beta = 0.3, \nu = 0.3$).

Experimental (Hz)	522	936	1326	1908	2354	3486	3598
Analytical (Hz)	528	922	1337	1965	2388	3572	3678

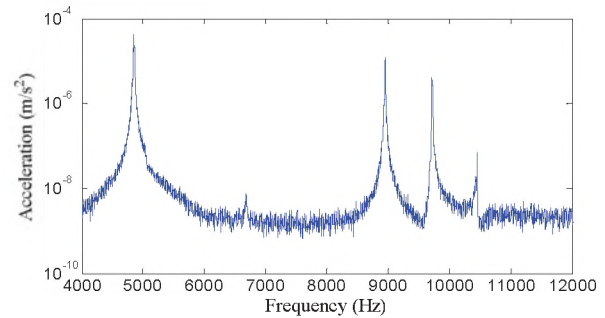


Figure 1: Frequency spectrum for an annular disk with free edges.

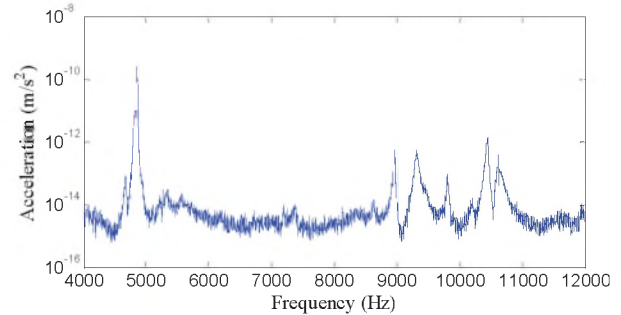


Figure 2: Frequency spectrum for an annular disk with point support at outer edge.

REFERENCES

- Tzou K., Wickert J., Akay A. (1998). In-plane vibration modes of arbitrarily thick disks, *Transactions of the ASME. Journal of Vibration and Acoustics*, vol 120, 384-391.
- Thompson D., Jones J., (2000). A review of the modelling of wheel/rail noise generation, *Journal of Sound and Vibration*. vol 231, 519-536.
- Bauer H., Eidel W., (2006). Determination of the lower natural frequencies of circular plates with mixed boundary conditions, *Journal of Sound and Vibration*. vol 292, 742-764.
- Eastep F., Hemmig F., (1982). Natural frequencies of circular plates with partially free, partially clamped edges, *Journal of Sound and Vibration*. vol 84, 359-370.
- Bashmal S., Bhat R., Rakheja S., (2009). In-plane free vibration of circular annular disks, *Journal of Sound and Vibration*. vol 322, 216-226.
- Ambati, G. (1976). In-plane vibrations of annular rings. *Journal of Sound and Vibrations*, vol 47, 415-432.
- Lee H., (2003). Modal acoustic radiation characteristics of a thick annular disk. PhD Thesis, Ohio State University.
- Leung R., Pinnington R., (1987) Vibration of a rotating disc subjected to an in-plane force at its rim, or at its centre. *Journal of Sound and Vibrations*, vol 114, 281-295.
- Bhat R. (1985). Natural frequencies of rectangular plates using characteristic orthogonal polynomials in Rayleigh-Ritz Method. *Journal of Sound and Vibrations*, vol 102, no 4, 493-499.

MODELING ELECTROSTATIC FIELD IN MEMS DEVICES USING ARTIFICIAL SPRINGS IN RAYLEIGH-RITZ METHOD

Avinash K. Bhaskar¹, Muthukumaran Packirisamy², Rama B Bhat²

¹Industry Analyst, Frost & Sullivan, 2001 Sheppard Avenue E, Toronto, CANADA M2J 1L6

²Optical BioMEMS Laboratory, CONCAVE Research Center, Department of Mech. & Ind. Engineering, Concordia University, Montreal, Quebec, Canada H3G2W1,

pmuthu@alcor.cocordia.ca

1. INTRODUCTION

Micro Electro Mechanical Systems (MEMS) are mechanical devices having moving components of size in the order of tens of micrometers that are integrated with electronic circuitry in the same chip. An electrostatically actuated MEMS device is shown Fig. 1. This device was manufactured through surface micromachining process.

Structural modification of an elastic system such as introducing stiffeners or creating notches would affect the dynamic behavior of the system. The operational environment such as the electrostatic field acts in the form of an elastic foundation and influences the stiffness property of the structure in addition to the structural geometry, support conditions, fabrication process [1-4]. The boundary support conditions obtained for microfabricated systems are not classical and can be somewhere in between pinned and clamped conditions. A method to quantify the above effects will be very useful in designing and modeling the dynamic performance of the system.

Bhat [5] presented a method to obtain the natural frequencies of rectangular plates under various boundary conditions using boundary characteristic orthogonal polynomials in the Rayleigh Ritz method. These orthogonal polynomials were generated using Gram-Schmidt process. The first polynomial is chosen so that it satisfies at least the geometrical boundary conditions of the structure.

In the present work, the plate type microdevice, boundary conditioned with external electrostatic field, has been analyzed for its dynamic behavior using boundary characteristic orthogonal polynomials in the Rayleigh-Ritz method for different end support conditions, namely, clamped (CCCC), simply supported (SSSS) and microfabricated condition (MMMM) that lies between the CCCC and SSSS conditions. A formulation to obtain the static deflection of the microdevice is also presented in this paper. Further, the vibration problem is formulated at the static equilibrium condition by including an elastic foundation stiffness to represent the electrostatic field effects in order to obtain the modeshapes and the Eigenvalues of the plate type microdevice at the static equilibrium position.

2. MODELING USING ARTIFICIAL SPRINGS

The microdevice is supported by translational and rotational springs with stiffness values ' K_T ' and ' K_R ' per unit length at the end supports. The microdevice is the main electrode which is separated by a gap of ' d ' from the substrate that acts as the bottom electrode. When a bias voltage ' V ' is applied between the device and the substrate, electrostatic field is created [6] and it results in attractive electrostatic force that pulls the plate type main electrode towards the substrate.

This electrostatic field has a softening effect on the elastic property of the system, since it acts in opposition to the structural stiffness. In order to apply the concept of boundary conditioning to represent the electrostatic effect, quantification of the softening effect in terms of distributed springs or an elastic foundation with stiffness per unit area of ' K_e ' is introduced [7]. As the deflection increases due to increase in voltage, the value of the stiffness ' K_e ' will indicate more weakening.

It is known that a constant value of the stiffness $K_e(x, y)$ would result in different natural frequencies but same modeshapes when compared to the microdevice without elastic foundation. As the deflection of the microdevice due to electrostatic force results in non-uniform spacing between the electrodes, it results in softening elastic foundation with non-uniform variation of $K_e(x, y)$ as shown in Fig.2. Even though, the electrostatic field results in non-linear dynamic behavior, it is linearly represented with artificial springs as most of the electrostatically operated MEMS devices are in the linear operating range. The static behavior shown in Fig. 3 has been predicted with energy method while the dynamic behavior has been predicted with Rayleigh-Ritz method.

3. RESULTS AND DISCUSSION

The dimensions of the microdevice are $600\mu\text{m}$ in length and breadth for a constant electrode gap of $15\mu\text{m}$. Fig. 4 shows the variation between ' V ' and eigenvalue of the first mode for increasing voltage. Eigenvalue for each support condition reaches zero for a certain voltage called snap voltage as shown in Fig. 4. The graph indicates that the snap voltage of the square microplate for MMMM condition

lies between the CCCC and SSSS condition. The pull-in roughly occurs when the microdevice reaches 1/3 of the gap between the two electrodes. Using snap voltage as a basis, the microdevice has been analysed using a normalized voltage parameter defined as,

$$\sigma = \left(\frac{V}{V_{\text{snap}}} \right) \times 100\% \quad (1)$$

where, 'V_{snap}' is the pull-in voltage for the corresponding end condition and 'V' is the applied voltage. This normalized parameter is used to study the effect of electrostatic field on the modeshapes and natural frequencies of the microdevice. As the voltage is increased the eigenvalue decreases.

REFERENCES

- [1] P. Muthukumar, I. Stiharu, R.B. Bhat, Vibration of micromachined condenser microphone plate diaphragms in electrostatic fields, CANCAM 97, 16th Canadian Congress of Applied Mechanics, Quebec, Canada, June 1997, pp. 87-88.
- [2] R.B. Bhat, G. Mundkur, Plate characteristic functions to study sound transmission loss through panels, Second International Congress on Recent Developments in Air and Structure-Borne Sound and Vibration, 1992, pp. 461-468.
- [3] P.G. Young, S.M. Dickinson, On the free flexural vibration of rectangular plates with the straight or curved internal line supports, Journal of Sound and Vibration, 162 1 (1993) pp. 1643-1647.
- [4] A. Achong, The steel pan as a system of non-linear mode-localized oscillators Part1: theory, simulations, experiments and bifurcation, Journal of Sound and Vibration 197 4 (1996) pp. 471-487.
- [5] R.B. Bhat, Natural Frequencies of Rectangular Plates Using Characteristic Orthogonal Polynomials in Rayleigh-Ritz Method, Journal of Sound and Vibration, 102 4 (1985) pp. 493-499.
- [6] K.B. Avinash, P. Muthukumar, R.B. Bhat, Effects of electrostatic field on the dynamics of plate type microsystems using characteristic orthogonal polynomials in Rayleigh-Ritz method, VETOMAC-3 and ACSIM 04 Conference, New Delhi, India, 2004, pp.6-9.
- [7] P. Muthukumar, R.B. Bhat, I. Stiharu, Boundary conditioning technique for structural tuning, Journal of Sound and Vibration 220 5 (1999) pp. 547-557.

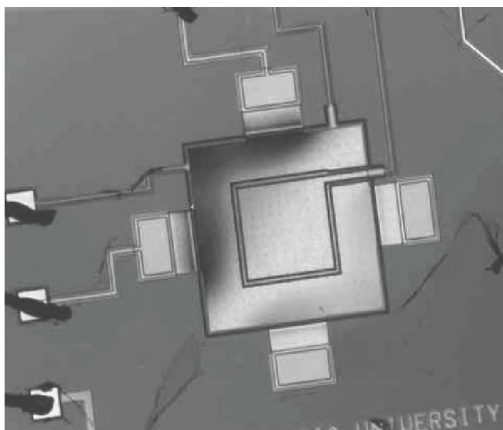


Figure 1 Surface Micromachined Capacitive Microdevice

Size of the microplate: 600µm x600µm

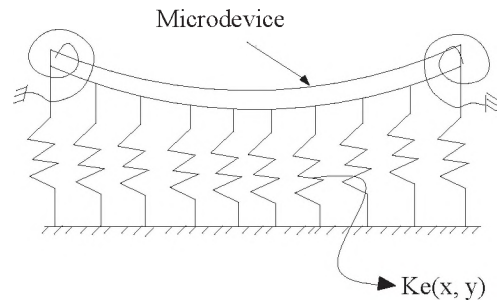


Figure 2 Electrostatic field modeled as non-uniformly distributed spring

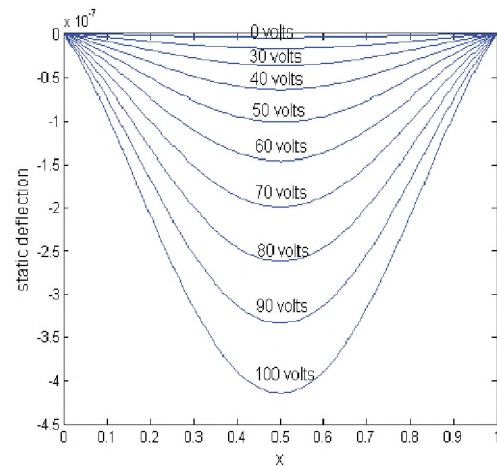


Figure 3 Static deflection in µm at y=0.5 for M MMM condition at different voltages.

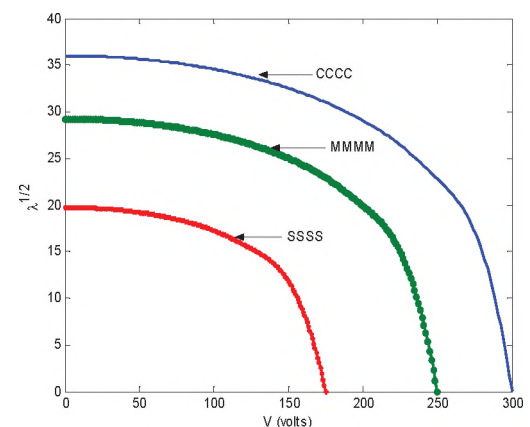


Figure 4 Variation of fundamental eigenvalues against bias voltage indicated end support conditions

DYNAMIC BEHAVIOR OF MICRO STRUCTURE IN MICRO FLUIDIC CHANNEL

MD Shakhawat Hossain¹, Muthukumaran Packirisamy², Subhash Rakheja³

^{1,2,3}Optical BioMEMS Laboratory, EV.13.235, CONCAVE Research Center, Department of Mech. & Ind. Engineering, Concordia University, 1515 St. Catherine (W), Montreal, Quebec, Canada H3G2W1, sojib1612@yahoo.com

1. INTRODUCTION

The potential of Micro-fluidics devices such as micro pumps, micro valves, micro viscometer, biomedical related micro fluidic chip, etc, draws the vital attention towards the investigation about fluid-solid interaction issue in micro level. However, there are several energy dissipations that dominant the performance of those Micro-fluidic devices. Among those energy dissipation viscous fluid damping leads above all and it creates the most significant change in the response of the micro devices operating in the fluid [1]. The result of relevant literature for the micro structure operating in fluid showed that there is dramatic shifting and broadening in the fundamental resonance peak to structural vibration in gases compared to the inviscid model [2]. Cantilever-fluid model is the simplest model for the investigation of microfluid-microstructural interaction. The experimental [3] and analytical [4] both types of studies indicate that micro or nano cantilevers operating in liquid is a heavily damped system, with a large shift in the resonant frequency away from the natural frequency of the system. If the cantilever is excited by other source rather than fluid, then the quality factor in air is typically higher than the estimated quality factor in water for the same resonator. Therefore, lower quality factor is found in air rather than other liquid when the cantilever is excited by any kind of fluid excitation such as Brownian force [5], which is the opposite result to the earlier one.

The current study presents the finite element modeling of micro cantilever excited by the fluid force. The quality factor of the cantilever is greatly influence with the pressure force loading and viscosity of the fluid. However, the length of cantilever also controls the quality factor for particular fluid. This paper explores all this issues comprehensively with the microfluid-microcantilever Finite element model.

2. THEORY

In the current model the cantilever is excited by a fluid pressure force created by the fluid flow, which is kind of step input action on a cantilever beam. For this type of model the equation of motion will be

$$M\ddot{y} + EI \frac{d^4 y}{dx^4} + c\dot{y} = F$$

Where, M is beam mass per unit length, y is beam deflection at its free end, E is Young's modulus of elasticity, c is damping coefficient per unit length and F fluid force per unit length. This force is caused by pressure difference as well as flow difference between the top and bottom side of the cantilever. However for un-damped system the natural frequency is,

$$\hat{\omega}_n = (\beta_n L)^2 \sqrt{\frac{EI}{ML^4}}$$

Where, β_n is a coefficient and product $\beta_n L$ values are 1.8751, 4.6940, 7.854 radians correspond to first 3 modes, respectively.

3. FINITE ELEMENT MODEL

3.1 Finite Element Model

COMSOL Multiphysics is used to develop the Finite Element model where, 3 different Modules are used. They are micro fluidics module (Incompressible Navier Stokes transient analysis), structural mechanics module, and deform mesh (moving mesh with transient analysis) module. In fluid module the inlet flow used as input boundary condition. The load developed due to this flow is used as input for structural module. The moving mesh application mode makes sure the flow domain is deformed along with the cantilever.

3.1 Boundary Conditions and Geometrical Parameters

Two different types of finite element modeling is done with the same geometry, which is showed in figure 1. In the first one for a particular length of cantilever and particular flow of fluid different quality factor and response is characterized with respect to the viscosity of different fluid. In second model for the particular fluid (air) and flow, quality factor is characterized with respect to different Aspect ratio of cantilever.

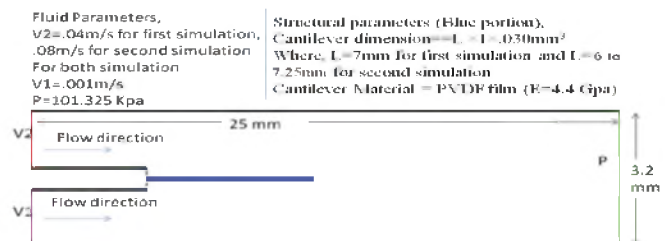


Figure 1: Geometry and boundary conditions

4. RESULT

4.1 Viscosity and Quality Factor

The quality factor with respect to different fluid is shown in figure 2.

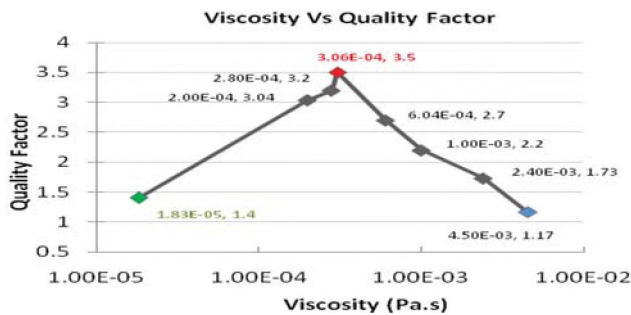


Figure 2: Viscosity Vs Quality factor. Highest quality factor is seen for Acetone indicated with red color and lowest is seen for Fluid XP indicated in blue color

Table1: Fluid parameters and frequency of the cantilever

Fluid	Viscosity (Pa.s)	Density (kg/m ³)	Frequency (Hz)	Pressure Difference Across Cantilever Tip (Pa)
Air	1.83E-05	1.2	156	0.01
Freon-12	2.00E-04	5.11	136	0.14
Water at 100°C	2.80E-04	958	16.1	1.67
Acetone	3.06E-04	790	17.12	0.42
Benzene	6.04E-04	878.6	15.5	2.83
Water at 20°C	1.00E-03	1000	13.89	3.86
Isopropyl Alcohol	2.40E-03	786	15.7	2.89
Fluid XP	4.50E-03	1029	12.6	4.54

The density and viscosity of fluid both control the quality factor for the fluid excitation to the cantilever. The fluid density is responsible for pressure load which creates the dynamic motion of cantilever. However, viscosity tries to resist this dynamic motion. Though air has the lower viscosity which is a good condition of high quality factor but air has very low density comparative to the other fluid. This creates lower amplitude where, higher amplitude is another significant condition for higher dynamic behavior or good quality factor. In figure 2 it is clearly shown that highest quality factor is found for the response in Acetone which has higher density than air but lower than water. However, when the cantilever starts moving (due to the stiffness) opposite direction to the flow after initial deflection by pressure loading (due to the flow) then the pressure and viscosity both resist the motion of the cantilever. Therefore, the amplitude decrement is faster in hot water (100°C) rather than Acetone though hot water has lower viscosity but has higher density.

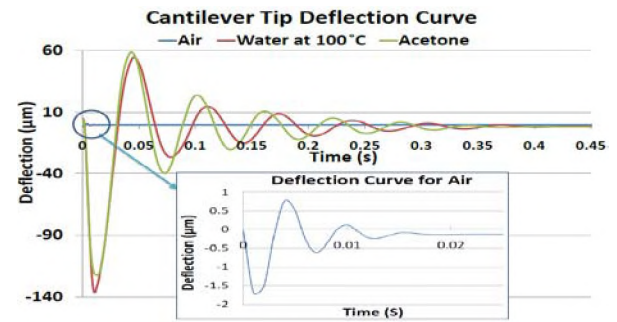


Figure 3: Deflection curve in Air, Hot water, and Acetone. Response in Air is shown as zoomed view in the bottom box

4.2 Aspect Ratio and Quality Factor

The quality factor is also characterized with respect to aspect ratio (length/width) which is showed in figure 4. However the peak value was found in 7, that was 7mm long cantilever as the width was consider as 1mm.

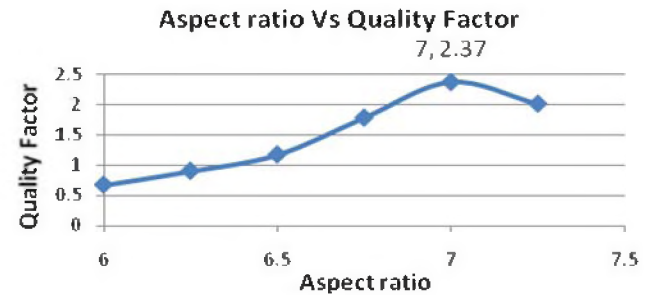


Figure 4: Aspect ratio Vs Quality factor.

5. DISCUSSION

It is clear that acetone is the good fluid as an actuation element for micro structure. In figure 4 after peak quality factor the value decreases as because of the size of the cantilever increases which causes increment of damping.

REFERENCE

- [1] Lin R. M. and Wang W. J., "Structural dynamics of microsystems - Current state of research and future directions," *Mechanical Systems and Signal Processing*, vol. 20, pp. 1015-1043, 2006.
- [2] Sader J. E., "Frequency response of cantilever beams immersed in viscous fluids with applications to the atomic force microscope," *J. Appl. Phys.*, vol. 84, pp. 64-76, 07/01. 1998.
- [3] Hirai Y., Mori R., Kikuta H., Kato N., Inoue K. and Tanaka Y., "Resonance Characteristics Of Micro Cantilever In Liquid," *Microprocesses and Nanotechnology Conference, 1998 International*, pp. 89-90, 1998.
- [4] Martin M. J. and Houston B. H., "Frequency response of nanoelectromechanical cantilevers operating in fluid," in *2008 8th IEEE Conference on Nanotechnology (NANO)*, 2008, pp. 303-6.
- [5] Sukuabul S. and Sood D. K., "Analytical models of resonant rectangular cantilever type chemical sensors for applications in fluids," in *ICST 2008*, 2008, pp. 604-9.

ABSTRACTS WITHOUT SUMMARY PAPERS

AEROACOUSTICS

Partially Spinning Behaviour of Flow-Excited Diametral Modes of Internal Shallow Cavities
Kareem Aly & Samir Ziada

The change of the azimuthal orientation of aerodynamically excited ducted cavity diametral modes has been studied experimentally. The diametral modes of ducted cavity are antisymmetric acoustic modes, which are trapped to the cavity. The diametral modes are excited by the oscillation of the grazing flow over the cavity axisymmetric mouth. This excitation process does not possess a preferred azimuthal orientation. To determine the azimuthal behaviour, a ducted axisymmetric shallow cavities have been tested over the range of cavity length to depth ratios from 1 to 6 and at Mach numbers up to 0.4. A set of pressure transducers flush mounted to the cavity floor is used to determine the acoustic mode amplitude and orientation. The observed azimuthal behaviours are classified into spinning, partially spinning and stationary diametral modes. A stationary mode is characterized by fixed orientation while the pressure amplitude is fluctuating by the resonance frequency. A spinning mode is characterized with constant amplitude with the orientation is turning over the circumference with the resonance frequency. A partially spinning mode is a combination of the spinning and the stationary scenarios. It was found that the partially spinning behaviour is most common. The characteristics of the partially spinning behaviour are detailed in this paper. An analytical model based on the superposition of two orthogonal modes is developed to reproduce the experimental data, and thereby provides a physical explanation of the observed behaviour of the diametral modes.

Robust Multi-Channel Mixer & Spatial Post Filter
Phil Hetherington

High quality voice communication is often hampered by the presence of noise. In hands free solutions, especially in the automobile, this can not be easily avoided as the choice, placement, and number of microphones is often constrained by cost and pragmatics. In some cases the solution required is some form of beam-forming (e.g., two microphones on a mirror facing only the driver). In other cases it is more like mixing (e.g., one microphone over the driver and another over the co-driver). And in still others it is a combination of the two (e.g., two microphones spaced 20 cm apart on the overhead console splayed to point slightly at the driver and co-driver). This heterogeneity of use cases and requirements coupled with a preference for a “one-size fits all” solution was the motivation for a robust, generic, yet powerful, solution that automatically behaves like a beam-former when the microphones are closely spaced, like a mixer when the microphones are well separated, and a combination of the two when the microphones are placed somewhat close but with a splayed orientation. This “complex mixer” solution has proven very successful with several car manufacturers and can be generalized across any number of microphones and can also be used in non-automotive environments.

Extracting Dual Spectra from Measured jet Noise Data
Werner Richarz

Subsonic jet noise appears to be generated by two distinct source mechanisms. There are several formalisms that describe these phenomena. An early version due to Ribner, based on Lighthill's acoustic analogy postulates that the two spectra a similar, but shifted in amplitude and frequency. More recently, Tam proposed an alternate two spectrum that combines physical arguments with observed properties of jet noise. A comprehensive data base of carefully measured scale model jet noise data has been analyzed. The two spectra have been extracted using the formalisms due to Ribner and Tam. The results suggest that either approach appears to be valid.

ARCHITECTURAL ACOUSTICS

Acoustical Performance Criteria and Treatment Protocols for Learning Spaces at a Large Institutional Teaching Facility
Mandy Chan & Bill Gastmeier

This paper investigates the application of the recently released ANSI Standard, Acoustical Performance Criteria, Design Requirements and Guidelines for Schools (ANSI S12.60-2002), to a wide variety of generally rectangular learning spaces during the conceptual design and design developments of a large institutional teaching facility. Criteria for Reverberation and Background Sound are developed and the acoustical analysis is summarized. Acoustical ceiling and wall treatments, volume sizes and furnishings are considered in the context of the proposed occupancies and activities in order to meet the acceptable reverberation levels. The results of this research indicates that the requirements of the ANSI standard can generally be met, with some compromise, although significant acoustical treatment is required and should be allowed for in the project budget at the outset. The supporting case study, analysis and discussions are presented.

Passive noise attenuation for natural ventilation system
Jean-Philippe Migneron

In noisy environments, acoustical impact is a limiting factor in using of natural ventilation. When there is a significant demand for passive or hybrid systems of ventilation, the development of some devices with good noise attenuation becomes a great challenge. That is the main goal of the research project. As part of this project, it is needed to analyze various possibilities of passive noise attenuation, by experimental measurements made in laboratory. A correlation between sound insulation and airflow capacities (and pressure drop) can be established. Different types of silencer were studied, as example, simple absorptive materials, baffles, or multiple resonators. Those were tested in the same aperture device which is planned to be installed in the thick of the facade of a building. That system also allowed adjustments of opening. Many conditions were then compared in term of third octave bands noise reduction and sound transmission class index. This work of validation may help the design and the good integration of those ventilation systems in buildings.

AUDITORY SCENE ANALYSIS

Neural Generators Underlying Concurrent Sound Segregation
Stephen Arnott

Hearing is a dynamic activity that uses a variety of acoustic cues to segregate sounds of interest from the temporally-overlapping background sounds. Despite increasing research on auditory scene analysis however, relatively little is known about the neural substrate underlying this ability. In this talk I will present results from an event-related synthetic aperture magnetometry (ER-SAM) analysis of concurrent sound segregation. Unlike traditional source analysis techniques, ER-SAM does not require a priori assumptions about the number or location of brain sources. Using this unconstrained, data-driven approach, concurrent sound segregation was found to be associated with sources proximal to N1m generators in the auditory cortex, as well as in the medial portion of the anterior temporal lobe. Similar analyses in the aging brain will also be discussed.

Neural Correlates of Rapid Speech Identification Learning
Sandra Campeanu, Boaz Ben-David, Kelly Tremblay & Claude Alain

Auditory perceptual learning is accompanied by a rapid improvement in performance that may reflect changes in perception, changes in response criterion or both. In two experiments, we used scalp recording of auditory evoked potentials (AEPs) to assess whether rapid improvement in speech identification coincides with changes in the sensory and/or response pathways. Participants practiced a speech identification task during which they also heard a broadband noise control stimulus (Experiment 1) or also performed a pure tone identification task (Experiment 2) to determine if changes in AEPs were specific to the trained speech cue. In both experiments, behavioural improvement in speech identification coincided with a decrease in P2 amplitude, which was followed by an increased negativity at frontal sites that was localized in the left pre-frontal cortex using distributed source modeling in the second study, and a late positive parietal complex that increased in amplitude in the first study and decreased in the second. While the first study indicated both learning-specific and general attention-related modulations, the second clarified AEP learning-specific changes. These experiments reveal rapid physiological changes in a fronto-temporal network, suggesting that learning to identify speech tokens may involve neural interactions between auditory centers as well as areas involved in speech production, articulation, and comprehension.

Something or nothing? Electrophysiological variations in encoding and retrieval during the processing of complex auditory scenes
Ben Dyson

Task load is a critical influence on cognitive processing but the degree to which increases in task load may be offset by advance preparation is less well understood. The current data provide insights into the neural correlates of auditory working memory by examining variations in scalp recorded auditory event-related potentials (ERPs) as a function of both encoding selectivity (cued, uncued) and retrieval load (0, 1, 2) during complex auditory scene presentation. At the time of encoding, ERP patterns were best characterised by the degree of perceptual selectivity rather than magnitude of task load per se. At the time of retrieval, ERP patterns provided evidence for cognitive closure as a principle applicable across the entire processing enterprise: the greater the selectivity during encoding, the smaller the workload required during retrieval. The present study confirms the utility of ERP in collectively addressing central questions regarding the encoding and retrieval of auditory memory, and also demonstrates how the study of something and nothing, is something rather than nothing.

The perceived continuity of a tone through noise and silence.
Nicholas Haywood, Valter Ciocca & Julie Chang

If an inter-tone interval is filled by a louder noise, listeners typically hear a single, continuous tone through the noise. This

'perceived continuity' is attributed generally to perceptual compensation for the masking effect of the noise (Warren et al., 1972), but there is some evidence that illusory continuity can occur through brief silences (Remijn et al., 2007). Experiment 1 - listeners used a nine-point scale to rate the continuity of six discrete 'flankers' through 5 interleaved bursts of white noise (flankers = 1 kHz pure-tone, 55 dBA, 500 ms; noise = low-pass filtered 10kHz, 80 dBA, 100/140 ms). For these arrangements, perceived continuity was strong, but listeners were unsure of continuity when 20 ms silent intervals were inserted at the mid-point of each noise-burst. Listeners were also unsure of whether a short target tone was physically present during the 20-ms intra-noise silent interval; performance improved for 40 and 60 ms silences. Experiment 2 - participants heard two flanker-noise-X-noise-flanker sequences (see Experiment 1, except noise = band-pass filtered 891-1122 Hz, combined duration of 300-ms duration; X was either a pure-tone target or a silence). Listeners indicated which sequence contained the target (2AFC). Preliminary data indicate that for 20 or 40 ms targets, most listeners do not perform the task accurately (< 70% correct). Performance improves (> 95%) when flankers are removed. Therefore, impaired performance is likely due to perceived continuity through both the noise-bursts and the intra-noise silent interval(s), and cannot be attributed to masking effects.

An object-based account of auditory scene analysis

Ada Leung

Object-based account of auditory attention suggested that listener's attention is allocated to auditory objects rather than its features. We tested this by examining the detection of a brief signal (i.e. gap) embedded in a complex sound. We used a mistuned harmonic paradigm with 16% mistuning and participants were required to make a judgment regarding the presence/absence of a gap in the signal regardless of harmonicity. The idea is that the mistuned harmonic "popped out" as a distinct auditory object that would easily capture attention and the gap could be viewed as an added feature of this new auditory object. According to the object-based theory, one might expect that performance in detecting the gap would be enhanced when the gap is inserted in the mistuned rather than the tuned harmonic. Surprisingly, our results showed the opposite with reduced gap detection for the mistuned stimuli compared to the tuned stimuli. This effect remained regardless of the duration of the gap and the tonal position of the mistuning. We conducted another experiment in which a low and a high harmonic were mistuned by 6%, such that the former promoted sound segregation while the latter did not. We found that gap detection was poor in the mistuned harmonic only in the former condition. Intriguingly, these results suggest that concurrent sound segregation impairs the processing of the attended auditory object which is supposed to be within the realm of one's attention. This research presents new insights on theories explaining attentional deployment during auditory scene analysis.

BIOMEDICAL ACOUSTICS

MODELLING HIGH FREQUENCY ACOUSTIC BACKSCATTER FROM BIOLOGICAL CELLS

Omar Falou, Min Rui, Ahmed El-Kaffas, Carl J. Kumaradas & Michael C. Kolios

The measurement of the ultrasound backscatter from individual biological cells is required for applications ranging from tissue characterization to molecular imaging. However, performing such a measurement remains a challenge. For instance, the presence of air bubbles in a suspension of cells during the measurements may lead to the incorrect interpretation of the acoustic signals. This work attempts to confirm the fluid nature of acute myeloid leukemia (AML) cells by measuring the ultrasonic backscatter response from individual AML cells and comparing it to theoretical predictions. This was done by combining a Xenoworks microinjection system (Sutter, Inc., Los Angeles, CA) with a co-registered Olympus IX71 inverted microscope (Olympus America, Inc., Center Valley, PA) and a VEVO770 Ultrasound imaging device (VisualSonics, Inc., Toronto, ON). This allowed the measurement of the ultrasound backscatter response from a cell under optical microscope guidance. The AML cells were imaged at 25 and 55 MHz to get information for a wide band of frequencies (12-58 MHz). The foci of the optical lens and the transducer were aligned to obtain optical and ultrasonic images of the same region. The cells initially attached to the micropipette and then released from the micropipette (using positive pressure) while simultaneously imaging both optically and ultrasonically. The ultrasonic backscatter responses vs. frequency were plotted and compared with theoretical predictions of fluid sphere scattering model. A very good agreement was found between the experimental and theoretical results suggesting that AML cells are of fluid nature. Finally, the implications of these results on the prediction of ultrasound backscatter from cells are discussed.

DIGITAL SIGNAL PROCESSING

Objective Assessment of Speech Enhancement Algorithms

Nazanin Pourmand & Vijay Parsa

Measures of speech quality are important in benchmarking the performance of speech enhancement algorithms. While subjective opinions are the most reliable criterion for speech quality, they are also expensive, time-consuming and not

repeatable. Objective computational metrics of sound quality that correlate highly with average subjective ratings are therefore attractive. Objective metrics of sound quality rely on two functional blocks: a feature extraction block that extracts perceptually relevant features and a “cognitive modelling” block that combines these features to generate a speech quality estimate. In this paper, we investigated the performance of several objective quality measures for predicting the quality of speech enhancement algorithms. In particular, the performance of two cognitive models, viz. multivariate adaptive regression spline (MARS) and multivariate Bayesian model, was evaluated. The feature set included Loudness Pattern distortion (LPD), ITU standardized Perceptual Evaluation of Speech Quality (PESQ), Itakura-Saito (IS), Weighted Spectral Slope (WSS) and Log-Likelihood Ratio (LLR). A noisy speech corpus, NOIZEUS, has been used for evaluating the objective measures. This dataset includes two SNR levels (5 and 10dB), four different types of background noise and speech/noise distortions introduced by 13 different speech enhancement algorithms. The relative performance of the two models in prediction accuracy, computational complexity, flexibility, and overfitting will be discussed.

Robust Multi-Channel Mixer & Spatial Post Filter Phil Hetherington

High quality voice communication is often hampered by the presence of noise. In hands free solutions, especially in the automobile, this can not be easily avoided as the choice, placement, and number of microphones is often constrained by cost and pragmatics. In some cases the solution required is some form of beam-forming (e.g., two microphones on a mirror facing only the driver). In other cases it is more like mixing (e.g., one microphone over the driver and another over the co-driver). And in still others it is a combination of the two (e.g., two microphones spaced 20 cm apart on the overhead console splayed to point slightly at the driver and co-driver). This heterogeneity of use cases and requirements coupled with a preference for a “one-size fits all” solution was the motivation for a robust, generic, yet powerful, solution that automatically behaves like a beam-former when the microphones are closely spaced, like a mixer when the microphones are well separated, and a combination of the two when the microphones are placed somewhat close but with a splayed orientation. This “complex mixer” solution has proven very successful with several car manufacturers and can be generalized across any number of microphones and can also be used in non-automotive environments.

ENVIRONMENTAL NOISE

Windmachine Noise Diagnostics Michael Medal

Windmachines prevent frost damage to fruit trees and grape vines. A large, tower mounted propeller fan blows the relatively warmer air towards the ground. The fan revolves about the tower axis so that a large area receives the benefit of warm air. Although used infrequently, wind machines are a source of complaints as they impact listeners accustomed to quiet rural sound scapes. Weather conditions for wind machine use, are not at all ideal for acoustic measurement equipment; so there is limited measured data. Scale model experiments, where Helmholtz and Reynolds numbers are sufficiently close to the full-scale values assure similitude, overcome these logistical difficulties. The relative contribution of blade passage tones, boundary layer noise and tower-wake interaction noise has been quantified by measurement. The influence on source strength from the number of blades, rotational speed, and rotor-tower separation can then be deduced, and used to form the basis of a preliminary design for a quiet wind machine.

SIMULTANEOUS MEASUREMENTS OF SOUND LEVEL AND WIND PROFILE Al Lightstone, Guangsheng (Sam) Du, Ian Matthew & Irshad Rizvi

Simultaneous, continuous measurements have been made of sound levels at 3 m above ground and wind speeds at a number of heights from 3 m to 50 m, at a major wind farm site in an agricultural area. The measurement dataset includes two large subsets - over one year of monitoring the ambient condition (wind turbines not operating) and several months of monitoring the operating condition (wind turbines functional). The purpose of the study is to analyze the “real world” relationship between wind speed and sound level both for the operating case and for the ambient case. The study also includes an analysis of the nature of the actual wind profile from close to the ground to higher elevations. This analysis of empirical data is relevant to the controversy as to whether regulatory wind farm noise criteria that vary with wind speed, such as are used in Ontario, are appropriate.

Airport Noise Modelling For The New Quito International Airport Scott Penton, Kevin Carr, Craig Vatcher & Russ Lewis

The New Quito International Airport (NQIA), is a much-needed replacement of the existing Mariscal Sucre airport, and is to be located in mountainous terrain in one of the highest cities of the world. This paper outlines the approaches and challenges

overcome in conducting noise modelling for this airport.

HEARING CONSERVATION

The Effects of Hearing Loss and Linear vs. Nonlinear Hearing Aid Fitting Algorithms on Auditory Spatial Attention
Gurjit Singh, Kathy Pichora-Fuller, Thomas Behrens & Tobias Neher

To understand better how hearing loss and different hearing aid fitting algorithms affect spatial listening abilities in complex multi-talker situations, we investigated auditory spatial attention in 8 older adults with normal audiometric thresholds below 4 kHz and 8 older adults with bilateral sloping sensorineural hearing losses, ranging from moderate to severe in the higher frequencies. In a condition with real spatial separation, a target sentence from the CRM corpus was presented from one spatial location and competing sentences from two different locations (00, 450, 900 azimuth), with pre-trial cues specifying the target's identity and location. In a condition with simulated spatial separation, corresponding perceived spatial locations of the target and competitors were achieved through exploitation of the precedence effect. Seven different probability specifications indicated the likelihood of the target being presented at the three locations (100-0-0, 80-0-20, 60-0-40, 0-0-100, 20-0-80, 40-0-60, and 50-0-50). All listeners completed 10 hours of baseline testing and then listeners with hearing losses were fitted with linear and nonlinear hearing aids using a within-subject, within-device crossover design that included 8 weeks of acclimatization to each set of hearing aids. At the conclusion of each acclimatization period, listeners completed an additional 10 hours of testing. As expected, hearing-impaired listeners performed more poorly than normal-hearing listeners across all listening conditions, but especially under test conditions with less target location certainty. Implications for clinical practice, hearing aid development, and future research will be discussed.

Electroacoustic Measures of Frequency Lowering Hearing Aids
Julie Ladiges, Vijay Parsa, Susan Scollie & Danille Glista

There has been a renewed interest in frequency lowering technologies in digital hearing aids as an alternative amplification strategy for severe to profound high frequency hearing loss profiles. Frequency lowering technologies transfer high frequency information to lower frequency spectral regions with better residual hearing. This information transfer can take place either through frequency transposition where a band of high frequency spectral content is moved down and overlaid on top of a low frequency band, or through nonlinear frequency compression where spectral content beyond a certain cutoff frequency is compressed into a narrower bandwidth. While both these techniques have the potential to enhance the audibility of high frequencies, they can also negatively impact the sound quality and can compromise phonemically important spectral cues (e.g. /s/ and /sh/ spectra may become indistinguishable after frequency lowering). In this paper, a number of speech-based electroacoustic indices of the performance of frequency lowering devices are discussed including (a) the aided audibility index (AAI) which quantifies the increase in audibility due to frequency lowering, (b) spectral overlap metrics which quantify the amount of overlap between /s/ and /sh/ phoneme spectra after frequency lowering, and (c) formant shift metrics which compute the formant structure distortion due to the frequency lowering process. Results obtained from both frequency transposition and nonlinear frequency compression hearing aids fitted to a variety of audiograms are shown, which demonstrate the utility of the electroacoustic measures in characterizing the performance of frequency lowering devices.

Direct measurement of actual listening levels of digital music players in the field
Cheng Qian & Maximilian Touzel

There is widespread concern about risks of hearing loss as a product of today's digital music players. As a result, there is also a growing need for proper measurement of listening levels in the field. Compared to lab experiments, field measurements are particularly valuable as they are unaffected by behavioural changes induced by a laboratory setting. While field measurements have been conducted for many years, they have generally faced serious problems with earphone coupling, or have had to measure electrical output and correlate this to sound output. We propose a more direct method, in which an artificial ear fitted with a pinna is taken into the field, in this case the streets of Toronto. Operating as a hearing damage awareness experience, our setup is visited by pedestrians who stop out of their own interest. While informing the public, we also measure their preferred listening levels for 30 seconds and gather data on typical listening durations. In the presentation, we will discuss benefits and drawbacks of the approach, and provide some preliminary results.

MUSICAL ACOUSTICS AND SOUNDSCAPES

The Vibroacoustic Behavior of Harp Soundboxes
Shira Daltrop, Andrzej Kotlicki & Chris Waltham

Modern harp soundboxes typically consist of a thin soundboard made of an acoustic wood with a definite set of resonances,

and a relatively heavy conical or polygonal back in which there are several holes. These holes exist to allow access to the string attachments but also have important acoustical properties. Harps more than 200 years old frequently have small holes in the soundboard itself, and none in the back. The quality of the sound radiation from a harp is determined in large measure by the vibrational behavior of the soundboard and the motion of air in the soundholes. Using an impact hammer, accelerometer and acoustic-velocity probe, we have measured the response of a harp soundboard, and the air in the soundholes, to excitation at the string attachment points. The relationship between these motions is complex but also crucial in determining the quality of the instrument.

MUSICAL PERCEPTION

The role of acoustic scale in the perception of musical notes and instruments

Roy Patterson, Etienne Gaudrain & Tom Walters

This talk is about the sounds made by the sustained-tone instruments of the orchestra and how we perceive those sounds. It is intended to explain the basics of musical note perception; that is, why instruments come in families; what determines 'register' within families; and why we hear distinctive differences between members of a given instrument family - even when they are playing the same note. On the surface, the answers to these questions may seem obvious; one could say that brass instruments all make the same kind of sound because they are all made of brass, and the different members of the family sound different because they are different sizes. But there is a deeper explanation involving three acoustic properties of musical sounds, as they occur in air. The talk describes these properties (with audio demos) and explains why they are particularly useful in (a) summarizing the physics of note production by instruments, on the one hand, and (b) explaining the dimensions of musical note perception, on the other hand.

Does altered auditory feedback sound like feedback?

Peter Pfordresher & Justin Couchman

Auditory feedback refers to the sounds one creates during sequence production. Alterations of auditory feedback (e.g., delayed auditory feedback) can disrupt production. However, it is not clear whether altered feedback actually functions as feedback. In two experiments, participants performed short novel melodies from memory on a keyboard. On certain trials auditory feedback was altered with respect to its synchrony with actions and/or its pitch contents. Alterations varied with respect to how closely related the feedback pattern was to the produced sequence (e.g., participants could experience a fixed or random amount of delay). After each trial participants rated their experience of self-agency: whether they experienced sounds as feedback. In experiment 1, participants performed alone. Self agency decreased with increasingly unrelated feedback sequences. Disruption (error rates, slowing) was nonlinearly related to agency, with greatest disruption occurring when feedback yielded an intermediate level of self-agency. In a second experiment, the experimenter performed on a keyboard in view of participants, who were told that sounds could result from either their performance or the experimenter's. This manipulation reduced the self agency for trials with moderate alterations of feedback but not for more extreme alterations, and increased participants' susceptibility to disruption from altered feedback. These results suggest that altered auditory feedback disrupts production because its relatedness to actions is ambiguous. Disruptive effects may reflect a tendency for actions to be attracted to the pitch/time patterns formed by a related yet distinct external sound pattern.

PSYCHOLOGICAL ACOUSTICS

Cognitive skills and reading ability in children with CI

Björn Lyxell, Malin Wass, Birgitta Sahlén, Tina Ibertsson & Lena Asker-Arnåson

Children with CI have a different course of development of basic academic skills such as language- and reading skills, compared both with children with normal hearing and children with severe hearing impairment or deafness. Increased knowledge about the cognitive development in children with CI is necessary in order to adjust their situation in various settings. In the present study, working memory capacity, lexical access, phonological ability, prosodic skills and reading ability was examined in 35 children with CI (age: 6 – 10 years) and compared to 120 children with normal hearing. The results demonstrate that children with CI have a visuospatial and general working memory capacity equivalent to the hearing children. They had lower performance levels on most of the other cognitive and prosodic tests. The differences were particularly prominent in tasks that required phonological working memory and/or extensive phonological processing. Seventy-five percent of the children with CI had a performance level comparable to hearing children in the tests of reading comprehension. The results will be discussed with respect to development of phonological skills and the role of phonology in reading for children with CI.

Segregation of simultaneous auditory objects by harmonic mistuning in infants Nicole Folland, Laurel Trainor & Nicholas Smith

One aspect of auditory scene analysis involves parsing a complex sound wave into representations of multiple simultaneously-sounding objects. Despite claims that auditory scene analysis involves basic bottom-up processes, little research addresses how infants process simultaneous sounds. For adults, a mistuned harmonic segregates from a complex tone and is perceived as a separate higher-pitched auditory object. In Experiment 1, adults and 6-month-olds were tested on discrimination of a mistuning of the third harmonic in a 240 Hz complex tone (6 harmonics, random phase, \hat{A} - \hat{a} “6dB/oct). Adult thresholds were between 1 and 2%. Six-month-old infants heard the complex tone repeating from a speaker on their left and were conditioned (rewarded with dancing toys) to turn their head when the third harmonic was mistuned. After completing 24 trials with the 8% mistuning infants were tested with one of a 6%, 4% or 2% mistuning (8 infants/group). Results show that 6-month-olds are able to discriminate mistunings of 8%, 6% and 4%, with performance decreasing between 6% and 4%. We are currently testing 2% mistunings. In Experiment 2 we will confirm whether infants are actually hearing two objects in the case of mistuned harmonics with a preferential looking paradigm in which we observe whether they prefer to look at a single large ball bouncing or at a large and a small ball bouncing together in synchrony with the repeating complex tone. We expect that they will prefer the single ball in the in-tune case, but two balls when the third harmonic is mistuned.

Development of voice discrimination: Perceptual narrowing through experience in infancy Rayna Friendly & Laurel Trainor

The ability to discriminate and identify people by their voice is important for social interaction in humans. In the present study, we are investigating the role that experience plays in the development of voice discrimination. Learning to discriminate a number of socially-relevant stimuli has been shown to follow a common pattern of experientially-driven perceptual narrowing. For example, young infants initially discriminate phonemes from any category used in human language, but before the end of their first year after birth they become specialized at processing the phonemic categories of their native language. So, young infants discriminate foreign phonemes more easily than older infants and adults. Similarly, infants are initially better than adults at discriminating faces from foreign races and species, but become specialized for own-race faces before one year after birth. We are investigating whether a similar narrowing pattern occurs for the processing of voices. We are testing English-speaking adults' and 6-month-old infants' ability to discriminate either native-language (English), foreign-language (Mandarin) or foreign-species (primate) vocalizations. Results to date show that 6-month-olds can readily discriminate both English and primate voices whereas adults perform much better with human than primate voices. Data is currently being collected with Mandarin voices. It appears that young infants are able to discriminate individuals on the basis of their voices across a range of human and non-human species, but that through exposure to primarily human voices, this ability becomes specialized for the dimensions of voice variation that apply to human voices.

Temporal representation of pitch in auditory cortex Blake Butler & Laurel Trainor

Pitch is derived in the auditory system through complex spectrotemporal processing, but it is not known whether both spatial and temporal mechanisms are operating in young infants. However, in adults the temporal code is particularly important for pitch perception. We first examined whether there is a temporal representation of pitch at the cortical level in adults using EEG and iterated rippled noise (IRN). The IRN samples consisted of a segment of frozen white noise, repeated in 5 and 6 ms intervals to create 450 ms sounds with 200 and 167 Hz pitch sensations, respectively. The stimuli were high pass filtered to remove the resolvable harmonics. On 85% of trials (standards) the 167 Hz sound was presented, while the remaining 15% (deviants), consisted of the 200 Hz sound. EEG waveforms were averaged separately for standard and deviant trials in each individual. We found a mismatch negativity (MMN) component in response to the occasional changes in the perceived pitch of the IRN at 250 ms after deviant sound onsets. MMN is thought to reflect the updating of auditory sensory memory traces and indicates that adults were able to discriminate the pitches of the two IRN stimuli. Since both stimuli contained information in the same spectral range, and there were no resolvable harmonics present, the brain must have based this pitch sensation largely on the temporal information coded in the IRN stimuli. We are now conducting this experiment with infants in order to investigate the development of the temporal mechanism for pitch.

Neural mechanism of phonemic restoration Antoine Shahin

Phonemic restoration occurs when speech is perceived continuous through noisy interruptions. This illusory process enhances comprehension in adverse acoustical environments and illustrates how contextual knowledge can improve perception. In a functional MRI study, we show that this illusory filling-in process is based on two dissociable neural mechanisms: the matching of the speech signal to templates in memory, termed illusory continuity, and the sensory repair mechanism that

restores the neural representations of the missing speech. Regions mediating illusory continuity include the superior temporal sulcus (STS) and left posterior angular gyrus (AG). Regions of the repair network include Broca's area, bilateral insula, and pre-supplementary motor area (pre-SMA). The left AG/STS and the repair regions show evidence for word-level template matching and communicate more as degradation of the speech signal is increased. The Brain applies prior knowledge to restore degraded sensory representations through two-path neural process.

Musicians experience less age-related hearing loss
Benjamin R. Zendel & Claude Alain

It is well known that hearing ability declines with age. Age-related hearing loss, or presbycusis, is multi-faceted and highly variable between individuals. It is also known that being a musician can improve certain aspects of auditory processing. The current study investigated whether the observed auditory processing enhancements in musicians could offset some components of normal age-related hearing loss. To test this hypothesis, four hearing tests (gap detection threshold, mistuned harmonic threshold, speech in noise & pure-tone thresholds) were used to assess age-related hearing loss across participants ranging from 18-86 years of age. In addition, participants were split into two groups based on musical training, where one group had no musical training, while the other group had extensive musical training and were active musicians at the time of testing. Performance on all four tests declined with age. For the gap-detection task, the rate of decline was shallower in musicians compared to non-musicians. For the mistuned harmonic task the rate of decline was similar between groups, but overall, musicians performed better than non-musicians across the lifespan. For the speech in noise task, musicians performed better overall, and showed a shallower rate of age-related decline compared to non-musicians. There were no group differences on pure-tone thresholds. These findings suggest that musical training can offset some aspects of normal age-related hearing loss at the central level. More importantly, these findings may have implications for treatment and rehabilitation of presbycusis.

Measuring auditory attention networks
Kathleen Corrigan & Laurel Trainor

Attention research has largely focused on the visual modality. For example, in the visual Attention Network Test (ANT; Fan et al., 2002) each trial consists of a fixation point, a visual-spatial cue, and then a target. The ANT measures reaction time and accuracy to the target and can assess three attention networks: 1) Alerting, or maintaining an alert state, occurs when the cue prepares participants for the target, but does not predict where or when that target will occur, 2) Orienting, or shifting attention, occurs when the cue does predict where or when the target will occur, and 3) Executive control involves selectively attending to one aspect of a target (central arrow) while ignoring conflicting information (flanking arrows). Our goal was to create an auditory ANT that could also measure these three attention networks. In our auditory ANT, we measured alerting by presenting an auditory warning signal at a variable duration before the target, orienting by presenting auditory cues that predicted when the target would occur, and executive control by comparing congruent and incongruent trials on an auditory stroop task. Preliminary results with 13 participants on the auditory ANT revealed significant effects of alerting [$t(12) = 7.34, p < .01$], orienting [$t(12) = 2.20, p = .05$], and executive control [$t(12) = 5.87, p < .01$]. Our results suggest that these three attention networks can be measured in the auditory modality. We are currently testing whether auditory alerting, orienting, and executive control can also be measured using frequency cues.

Tone envelope: A cue for audio-visual unity
Michael Schutz & Michael Kubovy

Contrary to the predictions of established theory, Schutz and Lipscomb (2007) have documented a musical illusion in which visual information influences the perceived duration of concurrent sounds. This illusion is at odds with previous work on sensory integration demonstrating that vision does not influence auditory judgments of event duration (Walker & Scott, 1981). Recent work has shown that causality plays a key role, and suggests that tone envelope is crucial in triggering the perception of cross-modal causal links (Schutz & Kubovy, in press). This series of experiments demonstrates that exponentially decaying (also known as *apercussive* or *damped*) envelopes trigger a form of audio-visual binding not observed in previous research – which generally focuses heavily on sounds using flat envelopes. The first two experiments demonstrate that impact gestures integrate with tones shaped with (event-congruent) percussive envelopes, but not with spectrally matched tones shaped with flat (i.e. *off-on-off*) envelopes. The third experiment demonstrates that this privileged form of cross-modal binding is specific to decaying sounds by showing increasing tones (i.e. percussive tones played backwards) do not integrate with impact gestures. The fourth experiment suggests envelope shape does not appear to affect audition's influence on vision – percussive and flat tones influence visual perception of event duration equally. Therefore, sounds with percussive envelopes allow vision to influence auditory perception of event duration (contrary to previous research), but have no effect on the degree to which audition influences the visual perception of event duration.

Rural Noise and Noise Control

Absorptive Properties of Magnesium Panels for Automotive Dash Panel Applications
Michael Bowie, Colin Novak & Helen Ule

The level of a customer's attraction to a vehicle is dependant on their perception of quality. The psychoacoustics of powertrain noise emitted during normal idle and moving scenarios plays a critical role in this. For structural dash panels, the material in which they are constructed from can have an affect noise which is transmitted into the vehicle cabin, and thus, the psychoacoustics of this transmitted noise. Currently, magnesium alloy AZ31B (3% aluminium, 1 % zinc, balance magnesium) has been proposed as a potential material for the dash panel's main layer of many modern vehicles. The airborne excitation of the powertrain noise causes the dash panel to vibrate microscopically. Due to the excitation of powertrain noise travelling through the dash panel, it has been found that the resulting vehicle cabin noise is low in sharpness, moderate in annoyance, high in roughness, moderate in harshness, and high in loudness. These are the results of tests conducted inside a semi-anechoic chamber with a reverberation room underneath, containing the source. In addition to reflection, magnesium alloy AZ31B has a certain amount of absorption for each powertrain noise frequency within 200-4000 Hz to cause these psychoacoustic perceptions. Within this range, absorption coefficients jump from low to relatively high values frequently.

SINGING (AIRS)

Imprecise singing is widespread
Steven Brown & Peter Pfordresher

There has been a recent surge of research on errors in singing. To date, this research has failed to address an important distinction in motor control: that between accuracy and precision. Accuracy (typically the focus of recent research) refers to an overall tendency to miss one's goal in a motor task, for instance a tendency to sing "sharp" in general. By contrast, precision refers to the consistency with which the production occurs. The current study was a first attempt to relate the occurrences of imprecision and inaccuracy in singing. A group of 45 musically untrained participants was asked to vocally imitate novel 5-note melodies as well as to sing familiar songs (e.g., Happy Birthday) from memory. To measure accuracy for the novel imitations, we computed the difference in cents between the produced pitches and the target pitches, and did the same for intervals. We measured precision by examining the variability with which each pitch- and interval-category was produced on repeated occurrences within an individual. A criterion of 100 cents was used to define both inaccuracy and imprecision. The results showed that the frequency of imprecision within the group was more than 50% of participants, whereas fewer than 20% of participants were inaccurate. Imprecision is thus widespread and is much more prevalent than inaccuracy among non-musician singers. Finally, performance on the imitations of novel melodies (using intervals only) was highly correlated with performance on the familiar melodies, both for accuracy and precision.

The role of mimicry in perception of sung emotion
Lisa Chan & Frank A. Russo

Recent research by Livingstone, Thompson, and Russo (2009) suggests that facial mimicry may be involved in perception of sung emotion. In their study, participants were asked to observe and then reproduce songs that were emotionally expressive. The songs were presented audio-visually and the onset of reproductions was cued so as to occur one full bar after the song presentation. Using facial electromyography (f-emg), muscle activity was observed over four epochs: perception (presentation of target), planning (pre-imitation), production (imitation), and post-production (after imitation). Evidence of mimicry was found across all epochs, including perception. The mimicry observed in the perception epoch is somewhat surprising and may be the manifestation of a facial-feedback network that serves emotional communication. However, it is not clear as to whether mimicry would have been observed in the absence of production demands. To investigate the role of mimicry in perception of sung emotion more directly, a perception-only condition is needed. In this paper, we present findings from an adaptation of the Livingstone et al. (2009) study, in which the production demands have been removed.

Development of singing: The current state of our knowledge and an outline of critical questions
Rayna Friendly, Laurel Trainor & Steven Brown

Singing, like speech, is a universal human behaviour. Yet, we know little about its development. In this paper, we review literature about singing development and begin to define a direction for future research. The first question we consider is the developmental trajectory of singing in terms of how accuracy and precision improve with age. The second question concerns the factors that influence the development of singing ability, such as sex, musical experience in the home environment, and formal instruction at school. The third question concerns the relationship between perception and production, exploring, for

example, whether singing accuracy in children is constrained by perceptual skills, vocal factors, or sensorimotor interactions. The fourth question concerns the relative trajectories of speech and song development, in terms of both the progression of their developmental stages and the ages at which milestones are achieved. The fifth question deals with how best to measure singing ability and generate a classification of singer types among developing singers. Finally, the sixth question concerns the extent to which inaccurate singing can be ameliorated through training, and whether there exists a sensitive period for singing development.

Holding a note: Sensorimotor control during singing

Jeffery A. Jones, Colin S. Hawco & Dwayne Keough

Singing requires accurate control of vocal pitch (F0). Previous work has demonstrated that F0 control involves an interplay between closed and open loop control. In a series of three studies, we examined the acoustic-motor representation of the mapping between F0 feedback and the vocal production system. In Experiment 1, we compared responses to brief pitch perturbations before and after participants adapted to hearing their feedback shifted upward one semitone. The results we observed suggest a change in control strategy at utterance onset versus mid-utterance when singers hold a note. At utterance onset, auditory feedback is compared to predicted feedback to ensure the pitch goal is achieved. After utterance onset, the control strategy changes and stabilization is maintained by comparing feedback to previous F0 production. These findings are supported by event-related potential data obtained in Experiment 2. A mismatch-like negativity was observed when speakers heard their F0 suddenly shifted mid-utterance. This finding is different than the N100 differences typically observed when feedback is shifted from the beginning of an utterance. Finally, in Experiment 3, we show that trained singers can rapidly adapt to different perturbations to ongoing auditory feedback. Participants heard three musical notes and reproduced them in succession. Two of the notes were pitch shifted in opposite directions. We found that singers easily adapted to alterations and that after effects were note specific. Together, our findings indicate that the strategies for correction of errors detected via auditory feedback are task specific.

A test battery for singing suited for lifespan and longitudinal studies

Annabel Cohen, Jenna Coady, Marsha Lannan, Emily Gallant & Vickie Armstrong

We describe the development and use of a test battery of singing skills. The purpose of the battery is to efficiently provide information about abilities associated with singing as well as comparison and contextual information about speaking. The singing skills examined include determination of range; singing certain musical intervals, triads, and scale segments; singing a familiar and learn an unfamiliar song; creating a song, creating the ending of a song, and repeating a song heard earlier in the battery. The test was administered at 5 monthly intervals to 4 children of ages 3, 5, and 7 years, and to young adults 4 of whom had musical training and 4 of whom did not. It was also administered on two occasions (generally) to 4 healthy elderly persons and 6 persons with probable Alzheimer's disease. Preliminary quantitative analysis to date has focused on the singing of pitches of the familiar song. The sensitivity to the repeated (aabbccdd) structure of the song was evident even in the youngest children, and pitch accuracy was comparable for young and older adults. We expect to report on additional test components at the meeting. The data obtained from the test are very rich, and the positive outcome of repeated testing suggests that a similar test could well serve longitudinal investigations over durations of several years in connection with the AIRS (Advancing Interdisciplinary Research in Singing) SSHRC MCRI project.

SOUND QUALITY

Acoustic Imaging for BSR in Non-Quiet Environments using Spherical Beamforming

Gary Newton Jr., Colin Novak, David Bogema & Helen Ule

The intrinsic sound quality of an automobile is directly related to the consumer's overall perceived quality of the vehicle. As such, manufacturers have made great efforts to improve interior noise through the control of unwanted and annoying noises characterized as Buzz, Squeak and Rattle (BSR). Detecting the sources of BSR is extremely challenging as the automotive interior cabin is a very complex space comprised of many components. The common approach to identify sources of BSR is through either subjective driving tests or testing of the vehicle within a quiet laboratory environment using traditional acoustical and vibration acquisition methods. These include noise and vibration analysis using FFT, sound intensity mapping and statistical energy analysis. While these techniques can provide very good results, they can also be time consuming and require a high degree of expertise and experience to interpret the results. For this investigation, a spherical beamforming technique was investigated to create an accurate acoustic mapping of the vehicle interior in a non-quiet environment to allow for the visualization of sources of BSR. By doing so, this work demonstrated a fast, accurate and easy to interpret method for the quantification of BSR sources which does not require an anechoic environment thus reducing acquisition time and cost. It also has the benefit of giving a visual result of BSR sources which are easily interpreted.

SPEECH

Compensatory response to perturbed voice onset time feedback

Takashi Mitsuya, Ewan MacDonald & Kevin Munhall

Talkers listen to their own voice while they speak and use that as feedback to monitor and control fine details of speech production. Consequently, when auditory feedback is perturbed in real time, talkers spontaneously alter their speech production in order to compensate for the perturbation. This compensatory response has been observed for both suprasegmental and segmental classes of speech signals. Most of the research in real-time altered auditory feedback has focused on spectral manipulations of vowels and little attention has been devoted to temporal or consonants. In the present study, we examine the role of acoustic feedback in control of voice onset time (VOT). We first recorded native English speakers' productions of the words "TIP" and "DIP" and selected several of their most representative productions. In the altered feedback trials, the talkers were again asked to repeatedly produce either "TIP" or "DIP". During these productions a real-time processing system was used such that when talkers said one word, they heard their own voice saying the other word simultaneously. Results showed that the speakers compensated for the VOT perturbation such that they shortened their VOT for /d/ when the VOT of the feedback was longer (/t/). Similarly, they lengthened their VOT for /t/ when the VOT of the feedback was shorter (/d/). The results will be discussed in terms of models of the temporal control of articulation.

Auditory Feedback and Maintenance of English Vowel Production

David Purcell & Kevin Munhall

Hearing one's own voice is important for the acquisition of speech. Adult talkers however also use the sound of their own voice in the ongoing maintenance of accurate speech production. Auditory feedback has been demonstrated to play a role in the utterance to utterance control of voice loudness, pitch, and the acoustic characteristics of vowels. Tactile and proprioceptive sensory feedback also contribute to the control of speech. Relatively little is known about how the information from these different sensory modalities is combined or weighted. In this study, auditory feedback provided to talkers through headphones was manipulated in real-time while they spoke words in an /hVd/ context (examples include the words "head", "heed", and "hid"). A real-time computer was used to filter the vowels and change the first formant in the acoustic feedback either upwards or downwards 200 Hz. This change is approximately a shift of vowel category for most English talkers. The second formant and above were not altered significantly by the manipulation. Vowels from across the English vowel space were changed in both directions and the effect on speech production was observed. Results show that auditory feedback has less influence on the production of vowels near the extremes of the vowel space and normal articulation. Near these limits, the speech motor control system may have access to more robust information from non-auditory sensory modalities. Since we did not manipulate other senses, the system may better detect cross modality incongruence in these cases and weigh the auditory information less heavily.

VIBRATIONS

NOISE AND HAND-ARM VIBRATIONS: A DANGEROUS MATCH

Alice Turcot, Richard Larocque, Serge André Girard, Marilène Courteau & Samuel Tétrault

Workers exposed to hand-arm vibrations generated by vibrating tools are also exposed to high noise levels. Studies have shown that workers exposed to hand-arm vibration suffer from both deafness and vibration-induced white fingers. It has also been shown that workers exposed to vibrating tools presented a more severe loss of hearing when they also presented white fingers. This loss is more pronounced at high frequencies. Quebec has an audiometric exam data-base, conducted for auditory health screening in mobile laboratories, for auditory evaluation of workers under standardised conditions. Our objective is to verify if the decrease in noise-exposed workers' hearing of those showing white fingers is more pronounced than in workers exposed to the same noise levels alone. This is done by a critical exam of all audiometric frequencies with recourse to the data of ISO 7029 standards, something that studies conducted in the 1980's did not allow. We are presenting the results of this study. Vibrations definitely appear as a potential aggressor for the degradation of hearing. The results reinforce the necessity to reduce exposure to these two aggressors.

Sound and Vibration Exposures from Polishers and Scalars in a Dental Hygienist Population

Donald Peterson, Takafumi Asaki, Anthony Brammer & Martin D. Cherniack

Dental hygienists are commonly exposed to noise and vibration when operating various dental instruments and the impact of these exposures on human health remains unclear. Since direct measurements of vibrations with high-frequency components on small and light-weight dental instruments using traditional accelerometry can be problematic, a new method is presented involving a non-contact, two-dimensional laser Doppler vibrometer. The most commonly used rubber-cup polishers and sonic

and ultrasonic scalers were identified from a population of experienced and new dental hygienists (n=163) and the weighted and un-weighted one-third octave band frequency spectra of noise and vibration between 251 Hz and 63.095 kHz were measured using a microphone and the laser vibrometer. The dental instruments were measured using two bench-top mounting conditions: mounted at the opposite end of its tip and mounted using a simulated pinch grip at its typical pinch point. Resulting sound and vibration levels varied and the spectra indicated small variations in frequency components between mounting conditions. In addition, the vibration perception thresholds (VPT) were measured on the dental hygienists and extensive self-reported information was also collected, including dental instrument use and estimated duration of vibration exposures. Experienced hygienists were observed to have elevated VPTs at the FAII mechanoreceptors and higher reported average weekly use of vibratory instruments.

Whole-body vibration exposure: understanding, evaluating and predicting negative health risks
 Tammy Eger, Jim Dickey & Michele Oliver

A person can be exposed to whole-body vibration (WBV) when standing, sitting, or lying on a vibrating surface. WBV can have a negative impact on the health, comfort, or performance of the individual exposed. The development of health problems is dependent on a number of factors including vibration exposure magnitude, direction, frequency, and duration. The body harmlessly attenuates most vibration; however, frequencies between 1-20 Hz can cause the body to resonate which can result in a wide range of health problems including muscular fatigue, low back pain, gastrointestinal tract problems, dysfunction of the autonomic nervous system, degraded circulatory functioning, loss of hearing, and nausea. When determining health risks associated with WBV exposure, standards set out in ISO 2631-1 are generally consulted. ISO 2631-1 provides guidance on the quantification of WBV in relation to human health and comfort, the probability of vibration perception, and the incidence of motion sickness. Another approach used to assess health risks considers the transmission of vibration through the body to specific regions of interest. Results from our research will be presented to illustrate the potential negative impacts of WBV on health and comfort. Predicted health risks associated with equipment operation in mining, forestry and construction will be summarized. Changes in reported comfort associated with exposure to complex vibration profiles will illustrate the influence of vibration on comfort and seat-head WBV transmissibility findings will illustrate the complex relationship between vibration exposure frequency, magnitude, and participant posture on the amplification of vibration as it travels through the body.

WHAT'S NEW in Canada ??

*Promotions
 Deaths
 New jobs
 Moves*

*Retirements
 Degrees awarded
 Distinctions
 Other news*

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

QUOI DE NEUF en Canada??

*Promotions
 Décès
 Offre d'emploi
 Déménagements*

*Retraites
 Obtention de diplômes
 Distinctions
 Autres nouvelles*

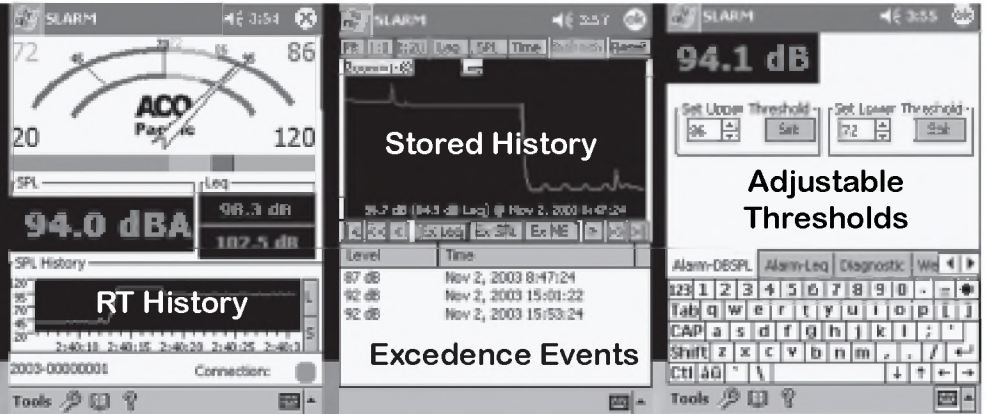
Avez-vous des nouvelles que vous aimeriez partager avec les lecteurs de l'Acoustique Canadienne? Si oui, écrivez-les et envoyer à:

Jeremie Voix - Email: voix@caa-aca.ca

Noise Pollution

The SLARM™ Solution

PDA & Laptop
Displays
Wired
Wireless



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

Typical Applications Include:

Community

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

Industrial

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

FEATURES

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

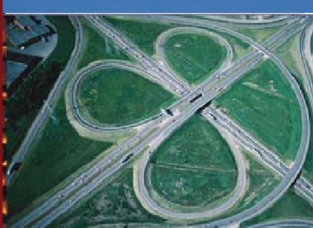


2604 Read Ave., Belmont, CA 94002 Tel: 650-595-8588 FAX: 650-591-2891
www.acopacific.com acopac@acopacific.com

ACOustics Begins With ACO™

831 sound level meter/real time analyzer

- Consulting engineers
- Environmental noise monitoring
- Highway & plant perimeter noise
- Aircraft noise
- General Surveys
- Community noise



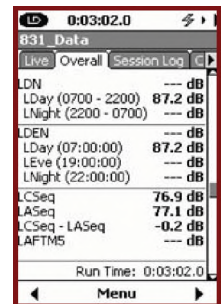
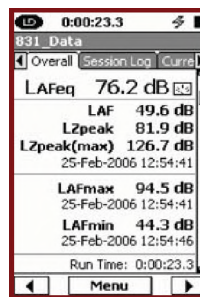
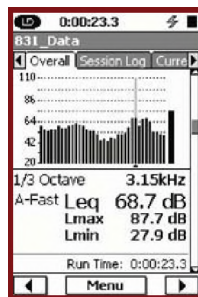
FEATURES

- Class 1/Type 1 sound level meter
- Small size with large display. Ergonomic
- User friendly operator interface
- 120MB standard memory expandable up to 2GB
- Single measurement range from 20 to 140 dB SPL
- Up to 16 hours of battery life
- Provided with utility software for instrument set-up and data download
- Field upgradeable
- AUX port for connection to USB mass storage & cellular modems



MEASUREMENT CAPABILITIES

- Real time 1/1 & 1/3 octave frequency analysis
- Simultaneous display of several noise measurements—ANY DATA (Leq, Lmax, Spectra, etc)
- Automatic logging of user selectable noise measurements (Leq, Lmax, Spectra, etc...)
- Exceedance logging with user selectable trigger levels
- Audio and voice recording with replay



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 technical 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 manuscrit 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 photos 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.



Application for Membership

CAA membership is open to all individuals who have an interest in acoustics. Annual dues total \$70.00 for individual members and \$30.00 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 August 31 will be applied to the current year and include that year's back issues of *Canadian Acoustics*, if available. New membership applications received after August 31 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 \$70.00. Many companies and institutions prefer to be a Sustaining Subscriber, paying \$300.00 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 August 31 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 | 5. Psychological / Physiological Acoustic | 9. Underwater Acoustics |
| 2. Engineering Acoustics / Noise Control | 6. Shock and Vibration | 10. Signal Processing / Numerical Methods |
| 3. Physical Acoustics / Ultrasound | 7. Hearing Sciences | 11. Other |
| 4. Musical Acoustics / Electro-acoustics | 8. Speech Sciences | |

For student membership, please also provide:

 (University) (Faculty Member) (Signature of Faculty Member) (Date)

I have enclosed the indicated payment for:

- CAA Membership \$ 70.00
- CAA Student Membership \$ 30.00

- Institutional Subscription \$ 70.00 plus mailing surcharge outside Canada:
 \$8 to USA, \$15 other International

- Sustaining Subscriber \$ 300.00 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 application and attached payment to:

Secretary, Canadian Acoustical Association, PO Box 74068, Ottawa, Ontario, K1M 2H9, Canada



Formulaire d'adhésion

L'adhésion à l'ACA est ouverte à tous ceux qui s'intéressent à l'acoustique. La cotisation annuelle est de 70.00\$ pour les membres individuels, et de 30.00\$ pour les étudiants. Tous les membres reçoivent *l'Acoustique Canadienne*, la revue de l'association. Les nouveaux abonnements reçus avant le 31 août 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 31 août s'appliquent à l'année suivante.

Abonnement pour la revue *Acoustique Canadienne* et abonnement de soutien

Les abonnements pour la revue *Acoustique Canadienne* sont disponibles pour les compagnies et autres établissements au coût annuel de 70.00\$. Des compagnies et établissements préfèrent souvent la cotisation de membre bienfaiteur, de 300.00\$ 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 31 août 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 31 août s'appliquent à l'année suivante.

Pour obtenir des formulaires électroniques, visitez le site Web: caa-aca.ca

Pour correspondance administrative et financière:

Nom / Organisation _____
Adresse _____
Ville/Province _____ Code postal _____ Pays _____
Téléphone _____ Téléc. _____ Courriel _____

Adresse postale pour la revue *Acoustique Canadienne*

Nom / Organisation _____
Adresse _____
Ville/Province _____ Code postal _____ Pays _____

Cocher vos champs d'intérêt: (maximum 3)

- | | | |
|---|-------------------------------|--|
| 1. Acoustique architecturale | 5. Physio / Psycho-acoustique | 9. Acoustique sous-marine |
| 2. Génie acoustique / Contrôle du bruit | 6. Chocs et vibrations | 10. Traitement des signaux / Méthodes numériques |
| 3. Acoustique physique / Ultrasons | 7. Audition | 11. Autre |
| 4. Acoustique musicale / Electro-acoustique | 8. Parole | |

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 70.00 \$
 Membre étudiant(e) 30.00 \$
 Abonnement institutionnel 70.00 \$
Surtaxe d'envoi à l'extérieur du Canada :
 8 \$ vers les États-Unis
 15 \$ tout autre envoi international
 Abonnement de soutien 300.00 \$
(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 :**

Secrétaire exécutif, Association Canadienne d'Acoustique, Casier Postal 74068, Ottawa, K1M 2H9, Canada

The Canadian Acoustical Association l'Association Canadienne d'Acoustique



PRESIDENT PRÉSIDENT

Christian Giguère
Université d'Ottawa
Ottawa, Ontario
V8W 3P6
(613) 562-5800 x4649
cgiguere@uottawa.ca

PAST PRESIDENT PRÉSIDENT SORTANT

Stan Dosso
University of Victoria
Victoria, British Columbia
V8W 3P6
(250) 472-4341
sdosso@uvic.ca

SECRETARY SECRÉTAIRE

David Quirt
P. O. Box 74068
Ottawa, Ontario
K1M 2H9
(613) 993-9746
dave.quirt@nrc-cnrc.gc.ca

TREASURER TRÉSORIER

Dalila Giusti
Jade Acoustics
411 Confederation Parkway, Unit 19
Concord, Ontario
L4K 0A8
(905) 660-2444
dalila@jadeacoustics.com

EDITOR-IN-CHIEF RÉDACTEUR EN CHEF

Ramani Ramakrishnan
Dept. of Architectural Science
Ryerson University
350 Victoria Street
Toronto, Ontario
M5B 2K3
(416) 979-5000 #6508
rramakri@ryerson.ca
ramani@aiolos.com

WORLD WIDE WEB HOME PAGE: <http://www.caa-aca.ca>

Sean Pecknold
(902) 426-3100

ASSISTANT EDITOR RÉDACTEUR ADJOINT

Ralph Baddour
Department of Medical Biophysics
University of Toronto
rbaddour@uhnres.utoronto.ca

DIRECTORS DIRECTEURS

Tim Kelsall
(905) 403-3932
tkelsall@hatch.ca

Richard Peppin
(410) 290-7726
peppinr@scantekinc.com

Jérémie Voix
(514) 932-2674
voix@caa-aca.ca

Vijay Parsa
(519) 661-2111 Ex. 88947
parsa@nca.uwo.ca

Roberto Racca
(250) 483-3300
rob@jasco.com

Clair Wakefield
(250) 370-9302
nonoise@shaw.ca

Sean Pecknold
(902) 426-3100
sean.pecknold@drdc-rddc.gc.ca

Frank Russo
(416) 979-5000 ext. 2647
russo@caa-aca.ca

SUSTAINING SUBSCRIBERS / ABONNES DE SOUTIEN

The Canadian Acoustical Association gratefully acknowledges the financial assistance of the Sustaining Subscribers listed below. Their annual donations (of \$300.00 or more) enable the journal to be distributed to all at a reasonable cost.

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 \$300.00 et plus) rendent le journal accessible à tous nos membres.

ACI Acoustical Consultants Inc.

Mr. Steven Bilawchuk - (780) 414-6373
stevenb@aciacoustical.com - Edmonton, AB

ACOUSTIKALAB Inc.

Jean Laporte - (514) 692-1147
jlaporte@acoustikalab.com - Montral, QC

Bruel & Kjaer North America Inc.

Mr. Andrew Khoury - (514) 695-8225
andrew.khoury@bksv.com
Pointe-Claire, QC

Eckel Industries of Canada Ltd.

- (613) 543-2967
eckel@eckel.ca - Morrisburg, ON

Hatch Associates Ltd.

Mr. Tim Kelsall - (905) 403-3932
tkelsall@hatch.ca - Mississauga, ON

Integral DX Engineering Ltd.

Mr. Greg Clunis - (613) 761-1565
greg@integraldxengineering.ca - Ottawa, ON

Jade Acoustics Inc.

Ms. Dalila Giusti - (905) 660-2444
dalila@jadeacoustics.com - Concord, ON

MJM Conseillers en Acoustique Inc.

MJM Acoustical Consultants Inc.
M. Michel Morin- (514) 737-9811
mmorin@mjm.qc.ca - Montréal, QC

OZA Inspections Ltd.

(800) 664-8263x25; FAX: (905) 945-3942
oza@ozagroup.com - Grimsby, ON

RWDI AIR Inc.

(519) 823-1311; FAX: (519) 823-1316
peter.vandelden@rwdi.com - Guelph, ON

SILEX Innovations Inc.

Mr. Mehmood Ahmed, (905) 612-4000
mehmooda@silex.com - Mississauga, ON

Spaarg Engineering Ltd.

Dr. Robert Gaspar - (519) 972-0677
gasparr@kelcom.igs.net - Windsor, ON

Tacet Engineering Ltd.

Dr. M.P. Sacks - (416) 782-0298,
mal.sacks@tacet.ca
Toronto, ON

Wakefield Acoustics Ltd.

Mr. Clair Wakefield - (250) 370-9302
clair@wakefieldacoustics.com - Victoria, BC

ACO Pacific Inc.

Mr. Noland Lewis - (650) 595-8588
acopac@acopacific.com - Belmont, CA

AECOM

Frank Babic - (905) 747-7411
frank.babic@aecom.com - Markham, ON

Conestoga-Rovers Associates

Tim Wiens - 519-884-0510 x2352
twiens@croworld.com - Waterloo, ON

ECORE International

Mr. Paul Downey - (416) 440-1094
pcd@ecoreintl.com - Toronto, ON

HGC Engineering Ltd.

Mr. Bill Gastmeier - (905) 826-4044
info@hgcengineering.com - Mississauga, ON

J.E. Coulter Associates Ltd.

Mr. John Coulter - (416) 502-8598
jcoulter@on.aibn.com - Toronto, ON

JASCO Research Ltd.

Mr. Scott Carr - (902) 405-3336
scott@jasco.com - Halifax, NS

Novel Dynamics Test Inc.

Mr. Andy Metelka - (519) 853-4495
ametelka@cogeco.ca - Acton, ON

Peutz & Associés

M. Marc Asselineau +33 1 45230500
m.asselineau@peutz.fr
Paris, FRANCE

Scantek Inc.

(410)-290-7726; FAX: (410) 290-9167
peppinr@scantekinc.com - Columbia, MD

SNC/Lavalin Environment Inc.

M. Jean-Luc Allard - (450) 651-6710
jeanluc.allard@snclavalin.com, Longueuil, QC

State of the Art Acoustik Inc.

Dr. C. Fortier - (613) 745-2003
sota@sota.ca - Ottawa, ON

Valcoustics Canada Ltd.

Dr. Al Lightstone - (905) 764-5223,
solutions@valcoustics.com
Richmond Hill, ON

West Caldwell Calibration Labs

Mr. Stanley Christopher - (905) 595-1107
info@wcll.com - Brampton, ON

Acoustec Inc.

Dr. J.G. Migneron - (418) 834-1414
courrier@acoustec.qc.ca - St-Nicolas, QC

Aercoustics Engineering Ltd

Mr. John O'Keefe - (416) 249-3361
aercoustics@aercoustics.com - Rexdale, ON

Dalimar Instruments Inc.

Mr. Daniel Larose - (514) 424-0033
daniel@dalimar.ca - Vaudreuil-Dorion, QC

H.L. Blachford Ltd.

Mr. Dalton Prince - (905) 823-3200
amsales@blachford.ca - Mississauga, ON

Hydro-Quebec

M. Blaise Gosselin - (514) 840-3000x5134
gosselin.blaise@hydro.qc.ca - Montréal, QC

J.L.Richards & Assoc. Ltd.

Mr. Terry Vivurka, P.Eng. - (613) 728-3571
mail@jlrichards.ca - Ottawa, ON

Mc SQUARED System Design Group

(604) 986-8181; FAX: (604) 929-0642
info@mcsquared.com - North Vancouver, BC

Owens-Corning Canada Inc.

Mr. Salvatore Ciarlo - (800) 988-5269
salvatore.ciarlo@owenscorning.com -
St.Leonard, QC

Pyrok Inc.

(914) 777-7770; FAX: (914) 777-7103
info@pyrokinc.com - Mamaroneck, NY

Sensor Technology Limited

Niru Somayajula - (705) 444-1440
techsupport@sensortech.ca - Collingwood, ON

Soft dB Inc.

M. André L'Espérance - (418) 686-0993
contact@softdb.com - Sillery, QC

Swallow Acoustic Consultants Ltd.

Mr. John Swallow - (905) 271-7888
jswallow@jsal.ca - Mississauga, ON

Vibro-Acoustics

Mr. Tim Charlton - (800) 565-8401
tcharlton@vibro-acoustics.com
Scarborough, ON

Wilrep Ltd.

Mr. Don Wilkinson - (905) 625-8944
info@wilrep.com - Mississauga, ON