ISSN 0711-6659

canadian acoustics
acoustique canadienne

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

SEPTEMBER 2010
Volume 38 -- Number 3

SEPTEMBRE 2010
Volume 38 -- Numéro 3

PROCEEDINGS

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

CAA 2010

COMPTES RENDUS

canadian acoustics

acoustique canadienne

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

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

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

CANADIAN ACOUSTICS is published four times a year - in March, June, September and December. The deadline for submission of material is the first day of the month preceeding the issue month. Copyright on articles is held by the author(s), who should be contacted regarding reproduction. Annual subscription: $35 (student); $80 (individual, institution); $350 (sustaining - see back cover). Back issues (when available) may be obtained from the CAA Secretary - price $15 and 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 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: $35 (étudiant); $80 (individuel, société); $350 (soutien - voir la couverture arrière). D'anciens numéros (non-épuisés) peuvent être obtenus du Secrétaire de l'ACA - prix: $15 (sans affranchissement). 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.



Acoustics Week in Canada
Semaine Canadienne d'Acoustique

Inn at Laurel Point
Victoria BC, 13-15 October 2010
Victoria CB, 13-15 octobre 2010

Organizing Committee / Comité d'organisation

Conference Chair / Président: *Stan Dosso*
Technical Chair / Directeur Technique: *Roberto Racca*
Exhibition / Exposition: *Clair Wakefield*
Registration / Inscription: *Lori Robson*
Treasurer / Trésorier: *Lara Berg*
Website / Site Web: *Brendan Rideout*
Secretary / Secrétaire: *Michael Wilmut*

Sponsors:

JASCO APPLIED SCIENCES

HTI HYDROACOUSTIC TECHNOLOGY, INC.

JADE ACOUSTICS

KINETICS Noise Control

Blachford

Dalimar instruments

Wakefield (Acoustics) Consulting Acoustical Engineers

SVS Sound & Vibration Solutions Canada Inc.

Scantek, Inc.
Sound & Vibration Instrumentation and Engineering
6430 #C Dobbin Rd, Columbia, MD 20145

FDI acoustics

RIM BlackBerry

Exhibitors:

SVS Sound & Vibration Solutions Canada Inc.

keene building products

SPERIAN

Blachford

CUSTOM BUILDING PRODUCTS
LEADING INSTALLATION SOLUTIONS FOR TILE & STONE

Dalimar instruments

KINETICS Noise Control

Brüel & Kjær
Incorporating LDS and Lochard

Pioneer Hill Software

NOVEL DYNAMICS INC.
Sound and Vibration Solutions
DATA ACQUISITION • ANALYSIS • SIMULATION • REPORTING

Scantek, Inc.
Sound & Vibration Instrumentation and Engineering
6430 #C Dobbin Rd, Columbia, MD 20145

ACOUSTIC SOLUTIONS LTD
"Protecting People"

(Page 1 photos: Gary Woodburn, Tourism Victoria; Brendan Rideout; Royal BC Museum; Tyler Ahlgren, Tourism Victoria)

EDITORIAL

On behalf of the Local Organizing Committee, it is my pleasure to invite/welcome you to beautiful Victoria BC for Acoustics Week in Canada 2010, the annual conference of the Canadian Acoustical Association (Wednesday to Friday, Oct. 13-15). Victoria is a great place to visit in October, and we have exciting technical and social programs planned. In keeping with Victoria's oceanic setting and the natural beauty of Vancouver Island, the conference will feature a special focus in the areas of marine and environmental acoustics.

The strong technical program includes 115 talks organized into 20 sessions covering all areas of acoustics, and running from Wednesday morning to Friday noon. (The technical content of the conference is summarized in the pages of this Proceedings issue.) We also have great keynote talks to start each day, including Christine Erbe: The Marine Soundscape, Murray Hodgson: Acoustical Environments in 'Green' Office Buildings, and Garry Heard: Canadian Defence Research in Arctic Acoustics. Thursday will include an all-day Exhibition of acoustical equipment, products, and services, with a dozen companies taking part.

The meeting will be held at the Inn at Laurel Point, an exceptional conference hotel located on Victoria's historic waterfront near to all city centre attractions. The Inn boasts 200 luxurious guest rooms, each with balcony and water views, and state-of-the-art conference facilities.

A Welcome Reception will be held Wednesday evening in the dramatic First Peoples Gallery of the renowned Royal British Columbia Museum, a short harbour-side walk from the conference hotel. The traditional CAA Banquet will be held Thursday evening at the hotel, and will feature outstanding West Coast cuisine and local entertainment.

We look forward to seeing you in Victoria!

Stan Dosso, Conference Chair
University of Victoria

C'est avec plaisir que je vous invite/accueil au nom du Comité Local Organisateur à Victoria CB pour la Semaine d'Acoustique Canadienne 2010, la conférence annuelle de l'Association Canadienne d'Acoustique (Mercredi à Vendredi, 13-15 oct.). Victoria est un endroit magnifique à visiter en octobre et nous avons organisé de formidables activités techniques et sociales. Au regard de la situation océanique et de la beauté naturelle de l'ile de Vancouver, la conférence portera un intérêt tout particulier à l'acoustique marine et environnementale.

Le programme technique, ayant lieu du mercredi matin au vendredi midi, inclut 115 présentations orales réparties dans 20 sessions couvrant tous les domaines de l'acoustique. (Le contenu technique de la conférence est résumé dans les pages de ce numéro d'Actes de conférence.) Nous avons également de formidables présentations d'ouverture pour commencer chaque journée, incluant Christine Erbe: Le Paysage Sonore Marin, Murray Hodgson: Environnement Acoustique dans les Bâtiments 'Verts', et Garry Heard: La Recherche de la Défense Canadienne en Acoustique Arctique. La journée de jeudi inclura également une exposition d'équipement, de produits et de services acoustiques avec une douzaine d'entreprises participantes.

La conférence se déroulera au Laurel Point Inn, un hôtel somptueux situé sur le bord de mer historique de Victoria, près de toutes les activités du centre-ville. L'hôtel a 200 chambres luxueuses avec balcon et vue sur mer et des salles de conférences haut-de-gamme.

Une réception de Bienvenue sera organisée mercredi soir dans la Galerie des Premiers Peuples du très renommé Musée Royal de Colombie Britannique, à seulement quelques pas de l'hôtel de la conférence. Le traditionnel Banquet de l'ACA se tiendra le jeudi soir à l'hôtel, et mettra en vedette l'excellente cuisine de la Cote Ouest, accompagnée de divertissements locaux.

Nous espérons avec plaisir de vous voir à Victoria !

Stan Dosso, Président de Conférence
Université de Victoria

Table of Contents/Tables des matières

Underwater Acoustics

Advanced Audio Applications

Hearing Sciences

Acoustic Standards

Soundscapes

Speech Communication

Architectural Acoustics

Noise Control

Vibration and Transport Vehicle Noise

Musical Acoustics

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

Program an Overview			
DAY ONE	**WEDNESDAY 13 OCT 2010** **Registration Open at 8:15, Marble Lobby**		
8:50–8:55	**Welcome**		
8:55–10:00	**Keynote Talk: Christine Erbe (Room: Spirit AB)** ***The Marine Soundscape and Effects of Noise on Aquatic Animals***		
10:00–10:20	**COFFEE BREAK**		
	ROOM: SPIRIT AB	ROOM: SPIRIT C	ROOM: SPIRIT D
10:20–12:20	*BIOACOUSTICS*	*ADVANCED AUDIO APPLICATIONS*	*SOUNDSCAPES*
12:20–1:20	**LUNCH**		
1:20–3:20	*BIOACOUSTICS / OCEAN ACOUSTIC INVERSON*	*ADVANCED AUDIO / HEARING SCIENCES*	*SPEECH COMMUNICATION I: PROSODY*
3:20–3:40	**COFFEE BREAK**		
3:40–5:00	*OCEAN ACOUSTIC INVERSION*	*ACOUSTIC STANDARDS*	*SPEECH COMMUNICATION II: L1 & L2 ACQUISITION*
5:00–10:00	**Acoustics Standards Committee Meeting (All welcome, Suite 152)**		
6:30-8:00	**WELCOME RECEPTION** **(First Peoples Gallery, Royal BC Museum)**		
DAY TWO	**THURSDAY 14 OCT 2010** **EXHIBITION: Marble Lobby 9:40–5:00**		
8:35–9:40	**Keynote Talk: Murray Hodgson (Spirit AB)** ***Evaluation and Control of Acoustical Environments in 'Green' (Sustainable) Office Buildings***		
9:40–10:20		**COFFEE BREAK**	
	ROOM: SPIRIT AB	**ROOM: SPIRIT C**	ROOM: SPIRIT D
10:20–12:00	*SPEECH COMMUNICATION III: PERCEPTION*	*ARCHITECTURAL ACOUSTICS I*	*ACOUSTIC SEABED MAPPING / UNDERWATER ACOUSTICS I*
12:00–1:00	**LUNCH**		
1:00–3:00	*SPEECH COMMUNICATION IV: PRODUCTION*	*ARCHITECTURAL ACOUSTICS II*	*UNDERWATER ACOUSTICS I*
3:00–3:40	**COFFEE BREAK**		
3:40–5:20	*SPEECH COMMUNICATION V: PROCESSING & PERCEPTION*	*ARCHITECTURAL ACOUSTICS III*	*NOISE CONTROL*
5:30–6:30	**CAA ANNUAL GENERAL MEETING (Spirit AB)**		
7:00–9:30	**AWC 2010 BANQUET & AWARDS (Terrace Ballroom)**		
DAY THREE	**FRIDAY 15 OCT 2010**		
8:35–9:40	**Keynote Talk: Garry Heard (Spirit AB)** ***Canadian Defence Research in Arctic Acoustics***		
9:40–10:00	**COFFEE BREAK**		
	ROOM: SPIRIT AB	ROOM: SPIRIT C	**ROOM: SPIRIT D**
10:00–12:20	*UNDERWATER ACOUSTICS II*	*VIBRATION AND TRANSPORT VEHICLE NOISE*	*NOISE CONTROL / MUSICAL ACOUSTICS*
12:30–1:30	**LUNCH & STUDENT PRESENTATION AWARDS**		

DAY ONE	WEDNESDAY 13 OCT MORNING Registration Open at 8:15, Marble Lobby		
8:50–8:55	Welcome		
8:55–10:00	Keynote Talk: Christine Erbe (Spirit AB) *The Marine Soundscape and Effects of Noise on Aquatic Animals*		
10:00–10:20	COFFEE BREAK		
	BIOACOUSTICS Chair: Christine Erbe Room: Spirit AB	*ADVANCED AUDIO APPLICATIONS* Chair: Jeremie Voix Room: Spirit C	*SOUNDSCAPES* Chair: Frank Russo Room: Spirit D
10:20–10:40	**Jason Wood, Scott Veirs, Val Veirs & Dominic Tollit:** How deep do you call? Depth localization in southern resident killer whales using passive acoustics	**Brady Laska, Chris Forrester & Sean Simmons:** Musical noise limits to speech enhancement	**Frank Russo:** The Toronto Sound Map project
10:40–11:00	**Carolyn Ward & Paul Hines:** Passive classification of marine mammal vocalizations using an automatic aural classifier	**Christian Giguère, Chantai Laroche & Véronique Vaillancourt:** Transmission characteristics of two tactical communication headsets with hearing protection capabilities	**Jennifer Schine:** Movement, memory and the senses in soundscape studies
11:00–11:20	**Mariana Melcón, Amanda Cummins, Sean Wiggins & John Hildebrand:** Assessing the effects of mid-frequency sonar on cetaceans in Southern California	**Antoine Bernier & Jérémie Voix:** Signal characterization of occluded in-ear versus free-air voice pickup on human subjects	**Vincent Andrisani:** Relocating the ear: A cross-cultural exploration of the electrified soundscape
11:20–11:40	**David Bain:** Bias in estimates of numbers of marine mammals affected by underwater noise	**Engbert Wilmink & Pieter van 't Hof:** DST a novel approach for noise dependent hearing protectors	**Hein Schoer:** The Sounding Museum: Two weeks in Alert Bay
11:40–12:00	**Brendan Rideout, Stan Dosso & David Hannay:** Underwater 3D passive acoustic Bayesian tracking of Pacific Walruses	**Hugues Nelisse, Jerome Boutin, Martine Gendron & Tony Leroux:** Measurements of noise exposure during wildfire air firefighting operations	**Milena Droumeva:** Mobile soundscape mapping
12:00–12:20	**Tin Ma, Nicole Collison, James Theriault, Joe Hood & Sean Pecknold:** Detection of precise time events for marine mammal clicks	**Joachim Gossmann:** Temporal metaphorics in auditory strategies of environmental monitoring	**Andrew Czink:** Sound matters: Mediation, mimesis, and embodiment in soundscape music
12:20–1:20	LUNCH		

DAY ONE	WEDNESDAY 13 OCT EARLY AFTERNOON		
	BIOACOUSTICS **Chair: Christine Erbe** **Room: Spirit AB**	***ADVANCED AUDIO APPLICATIONS*** **Chair: Jeremie Voix** **Room: Spirit C**	***SPEECH COMMUNICATION I: PROSODY*** **Chair: Sonya Bird** **Room: Spirit D**
1:20–1:40	**Samuel Johnston, Tracey Steig & David Ouellette:** Fine-scale 3-D tracking of fish behavior in central California using acoustic tags	**Anthony Brammer, Gongqiang Yu, Eric Bernstein, Martin Cherniack & Jennifer Tufts:** Predicting the intelligibility of speech corrupted by nonlinear distortion	**Masako Hirotani:** Use of pitch for processing emotions
1:40–2:00	**Christine Erbe:** Automatic signal detection in noise using entropy	**Eric Bernstein, Anthony Brammer, Gongqiang Yu, Martin Cherniack & Donald Peterson:** Ear cup selection for feedforward active noise reduction hearing protectors	**Zita McRobbie-Utasi:** Constraints on durational increase in boundary positions: A cross-linguistic study of languages with contrastive quantity
		HEARING SCIENCES **Chair: Christian Giguère** **Room: Spirit C**	
2:00–2:20	**Diana Percy:** Acoustic diversity and mating signals in the Psylloidea (Hemiptera)	**Eric de Santis & Mark Gaudet:** Noise levels in pubs and nightclubs in Vancouver	**Barbra Zupan & Madeline Shiah:** Examining vocal affect in natural versus acted expressions of emotion
2:20–2:40	**Melanie Hart, Peter Belton & Gerhard Gries:** The effect of fibril erection on hearing in male Aedes togoi: an open and shut case	**Andrew Dobson & Bill Gastmeier:** Occupational noise exposure in nightclubs	**Janet Leonard:** Acoustic diagnostics of prosodic phrasing in SENĆOŦEN
2:40-3:00	**Omar Falou, Amin Jafari Sojahrood, Carl Kumaradas & Michael Kolios:** Modeling the effect of shell thickness on high frequency ultrasound scattering from ultrasound contrast agents	**Reinhart Frosch:** Analysis of human oto-acoustic emissions	**Yang Zhang, Angela Cooper & Yue Wang:** Processing of speech and non-speech tonal information by native and nonnative tone language speakers: An event-related electrophysiological study
	OCEAN ACOUSTIC INVERSION **Chair: Stan Dosso** **Room: Spirit AB**		
3:00–3:20	**Michael Wilmut & Stan Dosso:** Multiple source localization in an uncertain ocean environment	**Alberto Behar & Willy Wong:** Fit testing of hearing protectors	**Qian Wang:** Perception of stress on accented and unaccented words: A comparison between native and nonnative English speakers
3:20–3:40	**COFFEE BREAK**		

DAY ONE	WEDNESDAY 13 OCT LATE AFTERNOON / EVENING		
	OCEAN ACOUSTIC INVERSION **Chair: Stan Dosso** **Room: Spirit AB**	***ACOUSTIC STANDARDS*** **Chair: Tim Kelsall** **Room: Spirit C**	***SPEECH COMMUNICATION II: L1 & L2 ACQUISITION*** **Chair: Masako Hirotani** **Room: Spirit D**
3:40-4:00	**Jan Dettmer & Stan Dosso:** Bayesian model selection using evidence computations	**Tim Kelsall & Alberto Behar:** Shift work, noise exposure and hearing loss	**Sonya Bird:** An ultrasound investigation of possible covert contrasts in early first language acquisition: the case of /sp ~ sw ~ sm/ > [f] merger
4:00-4:20	**Jan Dettmer, Stan Dosso & Charles Holland:** Transdimensional geoacoustic inversion	**Gordon Reusing & Tim Wiens:** Noise modelling versus reality under worst-case meteorological conditions	**Allison Benner:** Production and perception of laryngeal constriction in the early vocalizations of Bai and English infants
4:20-4:40	**Marie-Noel Matthews:** Estimating geoacoustic parameters of gassy sediment using low-frequency sound in St. Margaret's Bay NS	**Alberto Behar, Tim Kelsall & David Shanahan:** CSA and hearing conservation	**Stephen Winters:** The long-term retention of fine-grained phonetic details: Evidence from a second language voice identification training task
4:40-5:00	**Gavin Steininger, Stan Dosso & Charles Holland:** Determining acoustic scattering properties of marine sediments through Bayesian inversion	**Tim Kelsall & Christian Giguère:** Then new CAA Acoustical Standards Committee	**Ron Thomson & Paris Campagna:** The effect of task type on native speaker judgments of L2 accented speech
5:00-10:00	**Acoustics Standards Committee Meeting (All welcome, Suite 152)**		
6:30-8:00	**WELCOME RECEPTION** **(First Peoples Gallery, Royal BC Museum)**		

DAY TWO	THURSDAY 14 OCT MORNING EXHIBITION: Marble Lobby 9:40–5:00		
8:35–9:40	**Keynote Talk: Murray Hodgson (Spirit AB)** ***Evaluation and Control of Acoustical Environments in 'Green' (Sustainable) Office Buildings***		
9:40–10:20		**COFFEE BREAK**	
	SPEECH COMMUNICATION III: PERCEPTION **Chair: Molly Babel** **Room: Spirit AB**	***ARCHITECTURAL ACOUSTICS I*** **Chair: Murray Hodgson** **Room: Spirit C**	***ACOUSTIC SEABED MAPPING*** **Chair: Stephen Bloomer** **Room: Spirit D**
10:20–10:40	**Santiago Barreda & Terrance Nearey:** How the role of F3 in vowel perception may be influenced by listener expectations	**Ramani Ramakrishnan & Zarko Sopkic:** Noise reduction potential of green roofs	**Stephen Bloomer, Xavier Monteys & Ross Chapman:** Multifrequency classification and characterization of single beam echosounder data offshore Ireland
10:40–11:00	**Mark Scott:** Imagery-induced context effects	**Maureen Connelly & Murray Hodgson:** Measurement of the sound absorption characteristics of vegetative roofs	**Ben Biffard, Steve Bloomer, Ross Chapman & Jon Preston:** The role of echo duration in acoustic seabed classification and characterization
11:00–11:20	**Connor Mayer, Bryan Gick, Tamra Weigel & Douglas Whalen:** Perceptual effects of visual evidence of the airstream	**Bradford Gover & John Bradley:** Measurement of speech privacy of closed rooms using ASTM E2638 and setting criteria in terms of speech privacy class	**Ian Murfit:** Seabed sediment classification and seafloor bathymetry using single beam hydro-acoustic echo backscatter
			UNDERWATER ACOUSTICS I **Chair: Sean Pecknold** **Room: Spirit D**
11:20–11:40	**Terrance Nearey & Benjamin Tucker:** A perceptual study of [liquid + stop] sequences	**Jean-Philippe Migneron & Jean-Gabriel Migneron:** Case study about speech privacy of integrated furniture in an open-plan office	**Sean Pecknold, John Osier & Brendan DeTracey:** A comparison of measured ocean acoustic ambient noise with estimates from RADARSAT remote sensing
11:40–12:00		**Kevin Packer & Clifford Faszer:** Reverberation measurement and prediction in gymnasia with non-uniformly distributed absorption: The importance of diffusion	**Emma Murowinski & Cristina Tollefsen:** Measurements and modelling of atmospheric acoustic propagation over water
12:00–1:00	**LUNCH**		

DAY TWO	THURSDAY 14 OCT EARLY AFTERNOON		
	EXHIBITION: MARBLE LOBBY 9:40–5:00		
	SPEECH COMMUNICATION IV: PRODUCTION **Chair: Bryan Gick** **Room: Spirit AB**	***ARCHITECTURAL ACOUSTICS II*** **Chair: Bradford Gover** **Room: Spirit C**	***UNDERWATER ACOUSTICS I*** **Chair: Sean Pecknold** **Room: Spirit D**
1:00–1:20	**Donald Derrick & Bryan Gick:** Two phonological segments, one motor event: Evidence for speech-motor disparity from English flap production	**Murray Hodgson & Vincent Valeau:** Design and testing of an antenna array for sound-source localization	**Serguei Iakovlev, Adrien Lefieux, Jean-Francois Sigrist & Kyle Williston:** Multiple shock loading on fluid-filled shell structures
1:20–1:40	**Thomas Magnuson:** A look into the plosive characteristics of Japanese /r/ and /d/	**Zohreh Razavi:** Acoustical challenges for a hospital chiller room addition	**Serguei Iakovlev, Kyle Williston & Adrien Lefieux:** Effect of structural enhancement on the acoustic response of a submerged fluid-filled cylindrical shell
1:40–2:00	**Meaghan Delaney, Soraya Savji & Molly Babel:** An acoustic and auditory comparison of implicit and explicit phonetic imitation	**Ahmed Summan, Musarrat Nahid & Murray Hodgson:** Evaluation of the noise-masking system in a community-health-care facility	**Blair Kipple & Chris Gabriele:** Underwater acoustic levels of Southeast Alaska cruise ships
2:00–2:20	**Alexis Black:** Gender differences in automatic phonetic accommodation	**Ramani Ramakrishnan:** Location of horn speakers in a reverberation room	**Cristina Tollefsen & Sean Pecknold:** Comparison of sound speed profile interpolation methods with measured profiles; effects on modelled and measured transmission loss
2:20–2:40	**Noriko Yamane & Bryan Gick:** Speaker-specific place of articulation: Idiosyncratic targets for Japanese coda nasal	**Chris Bibby & Murray Hodgson:** Characterization and improvement of scattering and absorption by architectural surfaces without the use of specialized facilities	**Sean Pecknold, Cristina Tollefsen & John Osier:** Acoustic propagation sensitivity to variability and uncertainty of the ocean environment: a comparison of modeled and measured data
2:40–3:00	**Scott Moisik & John Esling:** An examination of the acoustic contributions of the epilaryngeal tube	**Ben Gaum & Ramani Ramakrishnan:** Acoustical evaluation of temporary performance facilities	**Reinhart Frosch:** Evanescent liquid sound-pressure waves near underwater resonators
3:00–3:40	**COFFEE BREAK**		

DAY TWO	THURSDAY 14 OCT LATE AFTERNOON / EVENING EXHIBITION: MARBLE LOBBY 9:40–5:00		
	SPEECH COMMUNICATION V: PROCESSING & PERCEPTION **Chair: Stephen Winters** **Room: Spirit AB**	*ARCHITECTURAL ACOUSTICS III* **Chair: Ramani Ramikrishan** **Room: Spirit C**	*NOISE CONTROL* **Chair: Clair Wakefield** **Room: Spirit D**
3:40–4:00	**Celia Shahnaz, Wei-Ping Zhu & Omair Ahmad:** Pitch estimation from noisy speech based on residual-temporal information	**Frances King, Stefan Schoenwald & Brad Gover:** Effect of some floor-ceiling construction changes on flanking transmission	**Clair Wakefield:** Analysis and control of bridge expansion joint "croaking" noise
4:00–4:20	**Yu-Tsai Wang:** Voice acoustic analysis of Taiwanese adults with Dysarthria following stroke	**Chip O'Neil:** Providing "good", "better" or "best" acoustical plumbing system proposals to cost sensitive clients	**Andrew Williamson:** Noise survey within patient care wards at the Royal Jubilee Hospital, Victoria BC
4:20–4:40	**Molly Babel & Grant McGuire:** Contrast salience and talker normalization in nonsibilant fricative perception	**Johannes Klein, Berndt Zeitler and Bradford Gover:** Effects of a concrete topping and modified resilient interlayers on sound transmission through a concrete floor	**Marco Berci & Luigi Vigevano:** A nonlinear geometrical acoustic model for sonic boom propagation
4:40–5:00	**Marianne Pelletier, Huiwen Goy, Marco Coletta & Kathy Pichora-Fuller:** Effect of age on lexical decision speed when sentence context is acoustically distorted	**Behrooz Yousefzadeh & Murray Hodgson:** Beam-tracing model for prediction of impulse responses, and effects of surface-reaction modelling and edge diffraction in rooms	**Ramani Ramakrishnan:** Validation of COMSOL multiphysics and acoustical performance of splitter-silencers
5:00–5:20		**Bradford Gover, John Bradley, Trevor Nightingale, Berndt Zeitler & Stefan Schoenwald:** Subjective ranking of low-frequency impact and footstep sounds on lightweight floor-ceiling assemblies	**Andrew Williamson:** Field impact insulation class (FIIC) testing of hardwood flooring on a variety of resilient underlayments in a concrete condominium building
5:30–6:30	**CAA ANNUAL GENERAL MEETING** **(Spirit AB)**		
7:00–9:30	**AWC 2010 BANQUET & AWARDS** **(Terrace Ballroom)**		

DAY THREE	FRIDAY 15 OCT MORNING		
8:35–9:40	**Keynote Talk: Garry Heard (Spirit AB)** ***Canadian Defence Research in Arctic Acoustics***		
9:40–10:00	**COFFEE BREAK**		
	UNDERWATER ACOUSTICS II **Chair: Stefan Murphy** **Room: Spirit AB**	***VIBRATION AND TRANSPORT VEHICLE NOISE*** **Chair: Joana da Rocha** **Room: Spirit C**	***NOISE CONTROL*** **Chair: Clair Wakefield** **Room: Spirit D**
10:00–10:20	**Gary Brooke, Steven Kilistoff, Dale Ellis & David Thomson:** Performance prediction via the Java Acoustic Model Interface	**Rob Jozwiak:** Effect of helicopter noise and vibration on healthcare facilities	**Shira Daltrop, Murray Hodgson & Clair Wakefield:** Experimental investigation of the effects of absorptive surfaces on the acoustical performance of a barrier in an anechoic chamber
10:20–10:40	**Stan Dosso & Michael Wilmut:** Bayesian acoustic source tracking and track prediction with environmental uncertainty	**Joana da Rocha, Afzal Suleman & Fernando Lau:** Prediction of flow-induced noise in aircraft cylindrical cabins	**Werner Richarz & Tony Gambino:** Features of low frequency wind turbine sound
10:40–11:00	**Robert Barton & Kevin Smith:** Characterization of scattered acoustic intensity fields in the resonance region of a motionless rigid sphere	**Ian Matthew, John Emeljanow & Mark Levkoe:** Methods for measuring off-highway vehicle sound emissions and correlation with "near-to-track" and "off-track" sound levels	**Werner Richarz & Harrison Richarz:** Post construction HVAC noise control
11:00–11:20	**Caitlin O'Neill, Graham Warner & David Hannay:** Verification of a bubble curtain model using an impulse response function for a towed source	**Lixue Wu:** A laser position sensing system for the study of vibration shaker tiling	**Tyler Mose & Andrew Faszer:** Air injection vacuum blower noise control
			MUSICAL ACOUSTICS **Chair: Frank Russo** **Room: Spirit D**
11:20–11:40	**Stefan Murphy & Paul Hines:** Temporal robustness of an automatic aural classifier	**Esen Cintosun & Tatjana Stecenko:** Experimental and numerical comparison of viscoelastic material damping to equivalent mass as acoustic treatments to aircraft composite fuselage	**Frank Russo, Lola Cuddy & Alexander Galembo:** The complementary roles of temporal and spectral processing in tonal perception of low-frequency tones
11:40–12:00	**Dale Ellis & Sean Pecknold:** Range-dependent reverberation and target echo calculations using the DRDC Atlantic Clutter Model	**Alexander Serov:** Application of automatic data processing at the systems of condition monitoring of industrial equipment	**Kostas Zolotas & Sonya Bird:** Lingual ultrasound of articulations made with the didgeridoo
12:00–12:20	**Ross Chapman:** Ocean bottom reflection loss from elastic solid materials: Reflections on reflectivity	**Stephan Paul, Andrey da Silva, Erasmo Vergara & Dinara da Paixão:** Brazil's first undergraduate course in Acoustical Engineering	**Lisa Walker:** Whale song or whale music? From a composer's perspective
12:30–1:30	**LUNCH & STUDENT PRESENTATION AWARDS**		

Hand-held Analyzer Type 2270 for sound and vibration professionals

THE COMPLETE TOOLBOX

MEASUREMENT, ANALYSIS AND RECORDING RESULTS IN THE PALM OF YOUR HAND

- Environmental noise assessment
- Noise at work evaluation
- Dual-channel building acoustic measurement
- Product development and quality control
- Machinery vibration measurement with advanced FFT analysis

ONLINE DEMONSTRATIONS

Watch video demonstrations of Type 2270 and its ease-of-use on www.bksv.com/2270

Brüel & Kjær has the world's most comprehensive range of sound and vibration test and measurement systems

BRÜEL & KJÆR | LISTEN... FEEL.. SOLVE.

www.bksv.com/2270

Brüel & Kjær North America Inc. · 2815 A Colonnades Court · Norcross, GA · 30071-1588
Telephone: 770 209 6907 · Fax: 770 448 3246 · www.bkhome.com · bkinfo@bksv.com

HEADQUARTERS: Brüel & Kjær Sound & Vibration Measurement A/S · DK-2850 Nærum · Denmark
Telephone: +45 7741 2000 · Fax: +45 4580 1405 · www.bksv.com · info@bksv.com

Local representatives and service organisations worldwide

Brüel & Kjær
Incorporating LDS and Lochard

THE MARINE SOUNDSCAPE AND THE EFFECTS OF NOISE ON AQUATIC MAMMALS

Christine Erbe
JASCO Applied Sciences Australia Pty Ltd, Brisbane Technology Park, PO Box 4037, Eight Mile Plains, QLD 4113 Australia, Christine.Erbe@jasco.com

1. INTRODUCTION

The marine soundscape is made up of natural ambient sounds (e.g. wind and waves), biological sounds (e.g. animal calls) and anthropogenic sounds (e.g. ship noise). Acoustic ecology studies the relationships—mediated through sound—between organisms and their environment. As ocean water conducts light very poorly yet sound very well, marine mammals rely heavily on acoustics for communication and navigation. Since the onset of the industrial revolution, man-made noise in the ocean has steadily increased. The effects of noise on marine animals can be short-term or long-term, transient or chronic, negligible to biologically significant, where the survival of a population is at risk. This article gives an overview of the components constituting the marine soundscape, of the use of sound by marine mammals and of the effects of noise. The acoustic ecology of animals other than mammals and the effects of noise on animals other than mammals are less understood.

2. THE MARINE SOUNDSCAPE

Natural ambient sound in the ocean is largely related to wind, waves and weather. The actual sources are bubbles generated near the surface. Wenz (1962) published spectra (power versus frequency) typical for various wind and sea state conditions, and other ambient contributors. Rain can dominate locally and temporarily. In the polar regions, generally, ice movement and fracturing dominate.

Anthropogenic contributions to ambient noise include shipping, petroleum and mineral exploration and production, construction, sonar etc. Of these, shipping has become so widespread that it adds to the background din where individual sources cannot be distinguished.

Biological contributors to ambient sound vary with location and time of year. Marine mammals are highly vocal underwater. Fish, in particular in tropical regions, produce night-time choruses creating a distinct peak in the ambient spectrum. Snapping shrimp dominate ambient noise in tropical waters at high frequencies.

The spectral contribution of the different sources to the ambient spectrum depend not only on the spectral output of the sources, but also on the distribution and density of the sources, and on the sound propagation environment (as set by sound speed, bathymetry, seafloor geoacoustics and ocean dynamics), and the receiver depth (Dahl et al. 2007).

3. SOUNDS MADE BY MARINE MAMMALS

Odontocetes (toothed whales and dolphins) produce tonal whistles (with some exceptions), burst-pulse sounds and echolocation clicks. Geographic differences as well as dialects of populations sharing the same geographic region exist. Whistles and burst-pulse sounds serve social functions, while echolocation aids navigation and hunting.

Mysticetes (baleen whales) produce frequency-modulated calls, as well as pulses and clicks, though mostly at a lower frequency than odontocetes. Echolocation has not been proven in mysticetes. Humpback, bowhead, blue and fin whales produce song of complex call patterns lasting over long durations.

Pinnipeds (seals and sea lions) produce tonal sounds as well as pulses and clicks in air and under water. All marine mammals further produce sounds by slapping body parts together or onto the water surface (Richardson et al. 1995).

4. THE EFFECTS OF NOISE ON MARINE MAMMALS

Noise can affect marine mammals in many ways. The effects of noise and the ranges over which they happen depend on the acoustic characteristics of the source (e.g., noise level, duration, duty cycle, rise time, spectrum), the medium (hydro- and geoacoustic parameters of the environment, bathymetry), and the receiver (e.g., age, size, behavioural state, auditory capabilities).

4.1 Audibility

At low levels, noise might be merely detectable. Noise levels decrease with range due to propagation losses. Audibility is limited by the noise dropping either below the animal audiogram or below ambient noise levels. Audiograms, hearing thresholds as a function of frequency, have been measured for only few individuals from about 20 marine mammal species. Indirect information stems from observed responses to sound, from anatomical studies and from the assumption that animals are sensitive to the frequencies of their own vocalizations (Richardson et al. 1995).

4.2 Behavioural Reactions

Southall et al. (2007) reviewed the literature on observed behavioural responses of marine mammals to noise. Such responses include changes in swim direction

and speed, dive and surface duration, respiration rate, changes in acoustic and contextual behaviour etc. Whether an animal reacts to a sound it hears depends on a number of factors including prior exposure (habituation versus sensitization), behavioural state, age, gender and health. To quantify behavioural responses, studies should be multivariate, considering the full range of metrics appropriate for the sound source and the full range of behavioural and contextual variables.

4.3 Masking

Noise can interfere with marine mammal social signals and echolocation, and the sounds of predators, prey and the environment (e.g. surf). Masking depends on the spectral and temporal characteristics of signal and noise. At a low signal-to-noise ratio (SNR), components of a signal might barely be audible. A higher SNR is needed for signal recognition and discrimination and an even higher SNR for comfortable communication. The potential for masking is reduced by good frequency and temporal discrimination, as well as directional hearing abilities of the animal. Masking can be further reduced in some species if the noise is amplitude modulated over a number of frequency bands (comodulation masking release), if the noise has gaps or the signal is repetitive (multiple looks model), and by antimasking strategies such as deliberate increases in call level and repetition, or frequency shifting (Erbe 2008). Models for the masking of complex calls by anthropogenic noise were developed by Erbe (2000; Erbe et al. 1999) based on behavioural experiments (Erbe and Farmer 1998).

4.4 Auditory Threshold Shift

A threshold shift is a loss of hearing sensitivity, which can be recoverable thus temporary (TTS) or permanent (PTS). TTS, but not PTS, has been measured experimentally in a few species of odontocetes and pinnipeds. A review of these studies led to initial noise-exposure criteria aimed at preventing PTS (Southall et al. 2007).

4.5 Non-auditory Physiological Effects

Systems other than the auditory system, which are potentially affected by noise include the vestibular, reproductive, and nervous systems. Noise might cause concussive effects, physical damage to tissues and organs (in particular gas filled), and cavitation (bubble formation), but data for marine mammals do not exist.

Stress is a physiological response intended to help an organism survive in the face of imminent danger; however, chronic stress can negatively affect health in the long run (Wright et al. 2009). The onset of stress might correspond to fairly low noise levels that induce a behavioural disturbance or masking. Stress might be a direct result of noise, e.g., if an unknown noise is detected, or an indirect result of noise causing, e.g., masking.

5. DISCUSSION

The different effects of noise are often connected, e.g. a TTS affects the audibility of a signal and thus alters the typical behavioural response to that signal. While research has historically focused on single animals, single noise sources and single effects, an integration of effects and a better understanding of the more complex soundscape is required. It is quite feasible to model cumulative sound exposure over multiple sources, long durations and large areas (Erbe & King 2009), but the manner in which exposures get accumulated by the animals and the effects of cumulative exposure remain unknown. Furthermore, the interaction of acoustic and non-acoustic environmental stressors needs to be investigated. Regulation would ideally not focus on a single operation limited in space and time but instead consider cumulative impacts experienced by animals over time and space.

REFERENCES

Dahl PH, Miller JH, Cato DH, Andrew RK (2007) Underwater ambient noise. Acoustics Today 3(1): 23-33.

Erbe C (2000) Detection of whale calls in noise: Performance comparison between a beluga whale, human listeners and a neural network. J Acoust Soc Am 108: 297-303.

Erbe C (2008) Critical ratios of beluga whales (Delphinapterus leucas) and masked signal duration. J Acoust Soc Am 124:2216-2223.

Erbe C, Farmer DM (1998) Masked hearing thresholds of a beluga whale (Delphinapterus leucas) in icebreaker noise. Deep-Sea Res II 45:1373-1388.

Erbe C, King AR, Yedlin M, Farmer DM (1999) Computer models for masked hearing experiments with beluga whales (Delphinapterus leucas). J Acoust Soc Am 105:2967-2978.

Erbe C, King AR (2009) Modeling cumulative sound exposure around marine seismic surveys. J Acoust Soc Am 125:2443-2451.

Richardson WJ, Greene CR Jr, Malme CI, Thomson DH (1995) Marine Mammals and Noise. San Diego: Academic Press.

Southall BL, Bowles AE, Ellison WT, Finneran JJ, Gentry RL, Greene CR, Kastak D, Ketten DR, Miller JH, Nachtigall PE, Richardson WJ, Thomas JA, Tyack PL (2007) Marine mammal noise exposure criteria: Initial scientific recommendations. Aquat Mamm 33:412-522.

Wenz GM (1962) Acoustic ambient noise in the ocean: spectra and sources. J Acoust Soc Am 34: 1936-1956.

Wright AJ, Deak T, Parsons ECM (2009) Size matters: Management of stress responses and chronic stress in beaked whales and other marine mammals may require larger exclusion zones. Mar Pollut Bull: 1879-3363.

PASSIVE CLASSIFICATION OF MARINE MAMMAL VOCALIZATIONS USING AN AUTOMATIC AURAL CLASSIFIER

Carolyn M. Ward and Paul C. Hines
Defence R&D Canada – Atlantic, P.O. Box 1012, Dartmouth, NS, Canada, B2Y 3Z7

1. INTRODUCTION

Traditional visual survey methods for detecting marine mammals are both time-consuming and inefficient, because observers can only see them during the short period when they are at the surface, during daylight hours, and when weather does not greatly reduce visibility. In an attempt to overcome these limitations, passive acoustic methods have become increasingly widespread for detection of marine mammals[1]. However, passive sonar systems that are used to localize and track marine mammals by their vocalizations can also be triggered by transient sounds from other sources, leading to a large number of false alarms. Even in the case of successful marine mammal detection, classification of the genus and species is often required, which typically requires expertise in marine mammal vocalization. The shortage of such expert listeners makes their deployment on ships difficult and costly.

Automated classification of marine mammal vocalizations is important from a biological and environmental perspective since marine traffic interferes with marine mammal habitats and populations. For example, the successful detection and classification of cetaceans will allow avoidance maneuvers to be performed by vessels if necessary. The importance of classification will increase as Arctic ice melts and new shipping lanes are established since cetacean populations in these environments are already sensitive[2].

An automatic aural classifier designed at Defence R&D Canada (DRDC) has been used to discriminate between impulsive-source echoes from man-made structures and echoes from clutter. The aural classifier models the human auditory system and uses timbre-based perceptual features that were identified in musical acoustics, to discriminate between acoustic inputs that would sound different to an expert listener[3]. Many of the features used for classification were inspired by research directed at discriminating the timbre of different musical instruments – a passive classification problem – suggesting it may be applicable to classification of marine mammal vocalizations. Thus, the classifier was tested on a set of marine mammal vocalizations to determine its ability to discriminate vocalizations originating from different species.

2. DATA SET

In this paper vocalizations from four cetacean species are examined: the sperm whale, the northern right whale, and the bowhead and humpback whales. The sperm whale was chosen because its 'clicks' are often confused with impulsive anthropogenic transients and hydrophone self-noise. Sperm whales have short broadband (about 80 Hz – 20 kHz[1]) clicks similar to the impulsive-source echoes for which the classifier was originally designed. Northern right whales were chosen because they are a critically endangered species (only about 400 remain[4]) and reside for part of the year in near-shore waters along Canada's east coast. Northern right whales typically vocalize in the 300 – 600 Hz range[4]. The bowhead and humpback species were selected because similarities in duration and bandwidth of their vocalizations often pose problems for automatic classification. Bowheads vocalize in the 20 Hz – 3 kHz range and humpback vocalizations span the 30 Hz – 10 kHz range[1].

Marine mammal vocalizations came from several different sources; data files collected from DRDC's research ship, CFAV Quest, contained many sperm whale clicks that were recorded using an SSQ57B broadband sonobuoy; northern right whale vocalizations were recorded by DRDC Atlantic using a variety of sonobuoy types deployed from a CP140 Maritime Patrol Aircraft; bowhead and humpback vocalizations were obtained from MobySound[5]. Marine mammal experts had previously identified each vocalization in the data set with a cetacean species. Examining sonograms of humpback song lead to the selection of four distinct sound units that were repeated frequently during a song session; units were labelled humpback1-4. However, preliminary classification results showed that humpback1,2&4 were aurally similar. Thus, classification was performed on a total of five classes: right whale, sperm whale, bowhead whale, humpback1,2&4, and humpback3.

3. METHODS

The aural classifier has the ability to perform both binary (two classes) and multiclass (more than two classes) classification. To determine how well the classifier discriminated between all five classes, the multiclass classifier was first run with all vocalizations in the data set. The binary classifier was run twice to determine how well the aural classifier performs on the challenging bowhead and humpback classification task – once with bowhead and all humpback vocalizations and once with bowhead and only the humpback3 vocalizations.

4. RESULTS AND DISCUSSION

The multiclass classifier result for all classes of cetacean vocalizations is shown in Figure 1. Of the 427 vocalizations tested only 61 were misclassified, giving an error rate of 14%. Note how well the sperm whale data points separate from the other classes. The sperm whales were expected to

be an easy case for the classifier since the impulsive characteristics of sperm whale vocalizations make them sound very different than the other types. Right whale data points were also well separated with only a few misclassifications. Bowhead and humpback vocalizations were expected to be a challenging case because of the similarity of their sounds; however, the results show some separation of the data points associated with these vocalizations so that there is minimal overlap between the classes in the principal component analysis (PCA) space.

The binary classification results for bowhead and humpback vocalizations can be viewed in Figure 2a. Only 35 of the 325 vocalizations were incorrectly classified (an error rate of 11%). The two classes separated well with only a few of the 35 incorrectly classified bowhead calls significantly displaced from the correct side of the decision boundary. Since classification of bowhead and humpback vocalizations has proven challenging, these preliminary results are encouraging. When classification of bowhead and humpback3 vocalizations was performed, the classifier correctly identified all 145 vocalizations (Figure 2b). The classifier was expected to more easily discriminate between these two types of vocalizations because the humpback3 unit sounds distinct – a "whoop" sound, rather than the moan-like vocalizations produced by the bowhead whales or the other three humpback units.

The preliminary classification results presented here point to the automatic aural classifier being an effective tool for classification of marine mammal vocalizations. Work continues to expand the data set and include anthropogenic noise sources for analysis of false alarm rates.

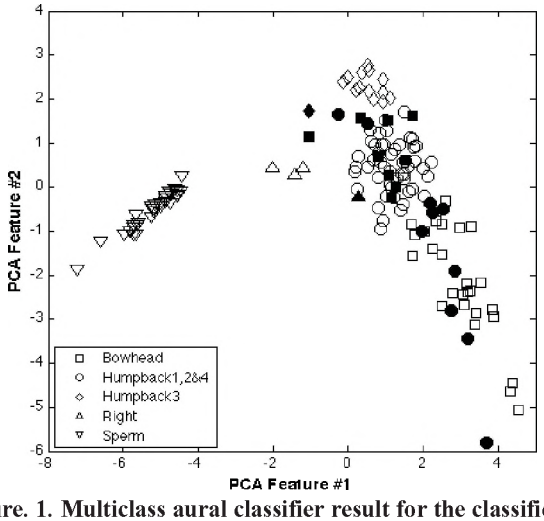


Figure. 1. Multiclass aural classifier result for the classification of all marine mammal types considered. Unfilled data points and filled data points represent correctly and incorrectly classified vocalizations, respectively. For clarity, only a third of the classified vocalizations were plotted. Note that the sperm whale points appear filled because of the density of points, not because they were misclassified.

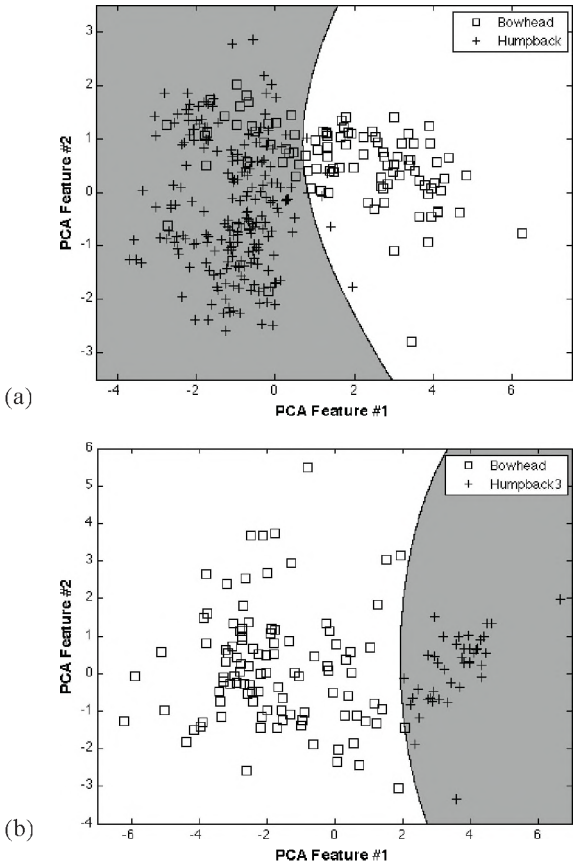


Figure. 2. (a) Results from binary classification of bowhead and all humpback1-4 vocalizations. (b) Results from binary classification of bowhead and humpack3 vocalizations. Correct classification occurs when squares are placed on the white region and crosses on the shaded region. The separation between these two regions is known as the decision boundary.

REFERENCES

[1]D. K. Mellinger, K. M. Stafford, S. E. Moore, R. P. Dziak, and H. Matsumoto, "An overview of fixed passive acoustic observation methods for cetaceans," Oceanography **20** 36-45 (2007).

[2]C. T. Tynan and D. P. DeMaster, "Observations and predictions of Arctic climate change: potential effects on marine mammals," Arctic **50** 308-322 (1997).

[3]V. W. Young and P. C. Hines, "Perception-based automatic classification of impulsive-source active sonar echoes," J. Accoust. Soc. Am. **122** 1502-1517 (2007).

[4]A. S. M. Vanderlaan, A. E. Hay, and C. T. Taggart, "Characterization of North Atlantic right-whale (*Eubalaena glacialis*) sounds in the Bay of Fundy," IEEE J. Ocean. Eng. **28** 164-173 (2003).

[5]S. Heimlich, D. Mellinger, H. Klinck, "MobySound.org," http://www.mobysound.org/index.html (2007).

ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge the advice of Stefan Murphy and Dr. Sean Pecknold of DRDC. The authors wish to acknowledge data preparation completed by Akoostix Inc. and to thank MobySound for access to marine mammal vocalizations.

ASSESSING THE EFFECTS OF MID-FREQUENCY SONAR ON CETACEANS IN SOUTHERN CALIFORNIA

Mariana L. Melcón, Amanda J. Cummins, Sean M. Wiggins and John A. Hildebrand
Scripps Institution of Oceanography, University of California San Diego, 9500 Gilman Dr., CA, USA, 92093 -0205
mmelcon@ucsd.edu

1. INTRODUCTION

Mid-frequency active sonar (MFA) is regularly used during naval exercises to provide an acoustic image of subsurface features, including natural and anthropogenic targets. Because MFA is often operated at high intensities, its sounds can be heard for thousands of square kilometers. MFA signal characteristics can vary considerably over its frequency band, 1-10 kHz, which coincidently happens to be in the audible band for most, if not all, marine mammal species. Over the past decade, correlations have been found between MFA and anomalous mass strandings of beaked whales (Cox et al., 2006). However, the mechanisms by which MFA affects beaked whales are not well understood.

Beaked whales, like most toothed whales, emit echolocation clicks (e.g. Zimmer et al., 2005) which provide them with location and distance information on prey and other objects from returning echoes. These sounds allow beaked whales to be monitored in the same way as MFA using passive acoustic monitoring techniques.

To quantitatively assess the possible impact of the MFA on the beaked whales, we used autonomous passive acoustic monitoring devices at two different deep water sites offshore of southern California for one year to record both beaked whale and MFA sounds. From these recordings, we tested whether or not the presence of beaked whale sounds was less frequent when MFA was detected.

2. METHOD

High-frequency Acoustic Recording Packages (HARPs – Wiggins and Hildebrand, 2007) were deployed at two independent sites off southern California, north and south of San Clemente Island, at depths of ~1000 m. Continuous recordings were made between March 2009 and March 2010. HARP recordings are broad-band with an effective bandwidth between 10 Hz and 100 kHz allowing most sounds produced by marine mammals and anthropogenic sources to be monitored.

Trained analysts evaluated spectrograms using a MATLAB-based software package (*Triton*) to log MFA events and note various marine mammal sound occurrences in the data sets. We also ran an automated detector on the HARP data for beaked whale echolocation clicks.

The detections of both MFA and beaked whales were merged into a sequential time vector, assigned identifying values and plotted to investigate potential patterns. To test whether there were significant differences in the presence of beaked whales when MFA was present (pb(BM)|MFA) or absent (pb(BW)|No MFA), we took 50 random samples of the time vector and calculated the probability of beaked whales with and without MFA. If pb(BW)|No MFA > pb(BM)|MFA, then the pb(BW)|MFA get a "+", and vice versa. After repeating this procedure 1000 times, we counted how many "+" each of the situations had and calculated the associated p value.

3. RESULTS

Recorded MFA events consisted of a wide variety of signals, from constant frequency tones to frequency modulated sweeps, or a combination of both and lasted between a few minutes to days.

MFA and beaked whale sounds occurred more frequently at the southern site N than at site M (Figure 1). The probability of detecting beaked whales when no MFA was present was compared to beaked whale detection probability with MFA present. No significant differences were found ($p=0.47$ for site M and $p=0.26$ for site N) between both MFA conditions.

Time-lag response was investigated by calculating the same probabilities as before, but with one hour delay in beaked whale detections. Again, no significant differences were found for either site ($p=0.46$ for site M and $p=0.26$ for site N) between presence and absence of MFA.

4. DISCUSSION

In this initial attempt to assess the potential effects of MFA on the beaked whales offshore of southern California using passive acoustic monitoring, we did not find significant differences between the presence of beaked whales when MFA was present and when it was not with our current methodology.

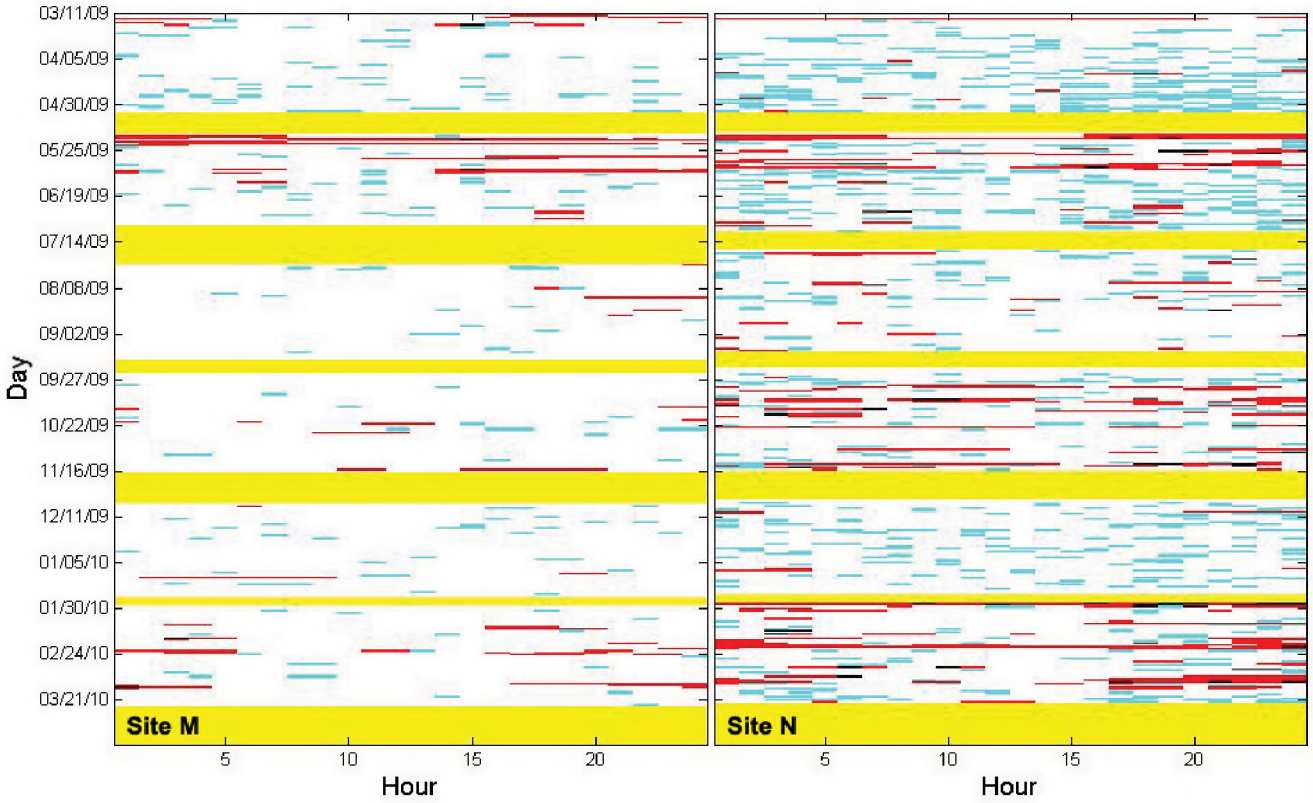


Figure 1. Co-occurrence of beaked whales and MFA in 1-hour bins for over one year of recordings in two different sites offshore of Southern California. Null detections are depicted in white, beaked whales in cyan, MFA in red and occurrence of both in black. No effort is marked in yellow.

Since no predictable periodic pattern was observed in the detections of MFA, it is expected that animals will need some time to react to it –if they show any reaction- instead of avoiding it simultaneously. Thus, the presence of animals was shifted by one hour to investigate if a temporal shift had an effect, but no significant difference in the presence of beaked whales was found when MFA was detected compared to when it was not. However, from this result, we cannot conclude that MFA has no effect in the population of beaked whales because of several issues.

First, the analysis was conducted with 1-h bins, which may be too long for acute responses and too short for long-term responses. Shorter duration bins and recordings over several years should be investigated for responses along with longer time-lags.

Second, all MFA events were considered equally, disregarding type, frequency, duration or intensity. For example, beaked whales may not react in the same way to an intense sound as to one that is barely audible suggesting MFA received sound levels are an important parameter.

Finally, presence or absence constitutes an "all or nothing" measure. However, in nature, responses are often gradual. Counting the number of calls or noting their intensity over time could provide a more progressive measure.

Further work should be conducted to evaluate potential causes and effects to properly assess and judge the impact of anthropogenic noise on these animals.

REFERENCES

Cox,T.M.; Ragen, T.J.; Read, A.J.; et al (2006). Understanding the impacts of anthropogenic sound on beaked whales. *J Cetacean Res Manage* 7(3):177-187

Wiggins, S.M.; Hildebrand, J.A. (2007). "High-frequency Acoustic Recording Package (HARP) for broad-band, long-term marine mammal monitoring" in *International Symposium on Underwater Technology 2007 and International Workshop on Scientific Use of Submarine Cables & Related Technologies 2007* (Institute of Electrical and Electronics Engineers Tokyo, Japan), pp. 551-557.

Zimmer, W.M.X.; Johnson; M.P.; Madsen, P.T.; Tyack, P.L. (2005). Echolocation clicks of free-ranging Cuvier's beaked whales (*Ziphius cavirostris*). *J Acoust Soc Am* 117(6): 3919-3927.

ACKNOWLEDGEMENTS

We would like to thank Dr. F. Stone and E. Young of CNO-N45, Dr. C. Collins of the NPGS and C. Johnson of PACFLT for project support, Dr. S. Baumann-Pickering for the data extracted from the automatic beaked whale detector, Dr. M. McDonald for guidance on beaked whale bioacoustics, G. Campbell for assistance in logging MFA events, and Dr. J. Zwolinski for useful advice on statistical methods.

BIAS IN ESTIMATES OF NUMBERS OF MARINE MAMMALS AFFECTED BY UNDERWATER NOISE

David E. Bain
Friday Harbor Laboratories, University of Washington, Friday Harbor, WA, USA 98250 dbain@u.washington.edu

1. INTRODUCTION

Wade *et al.* (2010) addressed the question of whether scientists' funding sources were correlated with their choice of existing data to consider on how marine mammals are affected by noise. In this paper, I examine whether even straightforward application of existing data is likely to lead to biased conclusions.

The data on effects of noise on marine mammals are limited. As a result, models have been developed to extrapolate existing data to novel contexts. These models include methods for estimating sound propagation efficiency, individual variation in responsiveness within species, inter-specific variation in responsiveness, and factors that can be ignored by the model. Like all empirical data, data on marine mammal responsiveness contain uncertainties due to limited sample sizes and resulting broad confidence intervals. Measurements on the physical sources themselves may include uncertainty due to limitations of measurement equipment. While the data themselves are typically reported as unbiased best estimates within confidence intervals, bias may be generated when best estimates are used in the models without consideration of their uncertainty.

2. METHODS

Published literature, environmental impact statements, and permit applications were reviewed to identify assumptions that could lead to biased conclusions.

3. RESULTS

3.1 Bias from physical measurements

Equipment for making physical measurements of sound is optimized for measurements within a limited range of frequencies and amplitudes. E.g., airguns have been reported to be loud, low-frequency noise sources. The early measurements of these devices were thus made with equipment optimized for loud, low-frequency sound. Reports of the physical properties of these devices reported their low-frequency content, but did not report the level of high-frequency content, as the equipment was not calibrated to measure those frequencies. Consequently, modelers have assumed only low frequency energy is present in airguns.

However, Bain and Williams (in review) found airguns also contain biologically significant levels of high frequency energy. That is, application of the early measurements is biased against predicting effects on high frequency-hearing specialists. As a result, effects on these species tend to be underestimated.

Sound propagation models are also based on physical measurements. The measurements produce a distribution of received levels at various distances from the source, and the mean and confidence intervals can be calculated. In turn, the radius at which effects are likely to cease to be biologically significant is calculated. Then, the area in which effects are likely to occur is calculated, and multiplied by density to estimate the number of individuals likely to be affected.

However, use of the best estimate of the radius does not lead to the best estimate of area when there is uncertainty. This is because received levels tend to decrease with the log of distance and area is proportional to the square of the radius to the threshold. When the radius to the threshold contour is at the upper confidence level, the affected area increases more than it decreases when the radius to the lower confidence level is determined. Note that the bias toward underestimating the affected area increases with increasing uncertainty.

3.2 Bias from biological measurements

As pointed out above, uncertainty increases bias, and individual variation is a source of uncertainty. The US Navy has employed a risk continuum (or "dose-response") function based on the level at which the most sensitive individuals begin to be affected and the level at which 50% of individuals are affected to account for individual variation. However, the parameters plugged into this model can be sources of bias when indices rather than actual values are employed. In field studies, the lower threshold would only be available as an anecdote, and as Wade et al. (2010) pointed out, anecdotes are difficult to publish. The best available approximation is commonly the level at which a statistically significant result is detectable. This, of course, depends on the sample size of the study, which is typically kept small by mitigation protocols. The larger the sample, the less biased a model substituting the significance threshold for the actual minimum would be. The uncertainty in the 50% level will also depend on sample size, and, as mentioned above, uncertainty contributes to bias. A serious methodological source of bias sometimes employed is to use the 50% level for one kind of effect as the minimum level and another type of effect for the 50% level in the model.

Sociality may make the risk model irrelevant. That is, group cohesion depends on all animals in the group responding in the same way to a given level of noise, and may increase the number of individuals affected.

Another source of bias is field studies that are based on the individuals nearest to the noise source. That is, many studies are structured with observers based on the vessel towing the noise source, and thus observations are limited to individuals close to the source. Individuals which move away from and remain too distant from the source to be observed are not included in the study. Therefore, the individuals included in the study are the subset of the population most tolerant of close approaches or least able to avoid the source.

The data of Calambokidis *et al.* (1998) are illustrative of this point. They observed marine mammal behavior from the vessel towing an airgun array; a small vessel that operated at variable distances from the airguns, including distances in excess of 70 km; and aerial platforms. Observations of porpoises near the source vessel consisted primarily of Dall's porpoises. In contrast, observations from the platforms distant from the airguns consisted primarily of harbor porpoises. That is, a larger portion of the harbor porpoise population avoided the airguns at a distance where they could not be observed by the mitigation team than did Dall's porpoises. Conclusions based only on the porpoises near the array would only reflect the responses of a small fraction of harbor porpoises in the study area, and hence be biased. Similarly, bowhead whales have been sighted near noise sources by vessel-based observers, while aerial surveys reveal deflections at long range (LGL *et al.* 1999).

Data from experiments with captive marine mammals are fundamental to many predictions of effects on wild marine mammals. However, there is inter-specific variability in the ability of cetaceans to survive in captivity (Bain 1988, DeMaster and Drevenak 1988). The bulk of data on audition in cetaceans comes from the small number of species that survive relatively well in captivity. To the degree survivorship is correlated to ability to tolerate noise from pumps and filters, there is potential for bias in the availability of data toward noise tolerant species. The exception is harbor porpoises, in which beach-stranded individuals have been extensively studied, but they may prove the rule by being relatively noise-intolerant (Bain and Williams in review).

Inter-specific variation is another source of uncertainty. Attempts to address this have been made by assuming that species with similar anatomy would have similar vulnerability to noise (Southall *et al.* 2007). However, Bain and William (in review) found this was not the case. They found harbor porpoises (*Phocoena phocoena)* were significantly more affected by airguns than the closely related Dall's porpoise (*Phocoenoides dalli*). Similarly, Steller sea lions (*Eumetopias jubatus*) tended to be more strongly affected by noise than California sea lions (*Zalophus californianus*). Stone and Tasker (2006) found similar differences between *Lagenorhynchus* species in the Atlantic.

4. DISCUSSION

These sources of bias have important implications for many issues within the marine mammals and noise arena. Mitigation and monitoring need to be planned over much greater distances than is commonly done currently. Additional habitat-specific risks may need to be addressed in more distant areas. Estimation of cumulative effects will require consideration of longer exposures, more repeated exposures, and more individuals affected overall. New data needed for more accurate estimates of the effects of noise on marine mammals are primarily the level at which 50% of individuals respond to noise in a wide variety of species.

REFERENCES

Bain, D. E. and R. Williams. In review. Responses of marine mammals to airgun noise at long range in coastal waters of British Columbia and Washington State. Mar Ecol Prog Ser

Bain, D. E. 1988. A journey through the NMFS Marine Mammal Inventory. Proc 1987 Int Mar Anim Trainers Assoc Conf 103-130.

Calambokidis, J., D. E. Bain and S. D. Osmek. 1998. Marine mammal research and mitigation in conjunction with air gun operation for the USGS "SHIPS" seismic surveys in 1998. Contract Report submitted to the Minerals Management Service.

DeMaster, D. P. and J. K. Drevenak. 1988. Survivorship patterns in three species of captive cetaceans. Mar Mamm Sci. 4:297-311.

LGL Ltd, Greenridge Sciences Inc., Western Geophysical, National Marine Fisheries Service. September 1999. Marine Mammal and Acoustical Monitoring of Western Geophysical's Open-Water Seismic Program in the Alaskan Beaufort Sea, 1998. LGL Rep. TA2230-3. Rep. from LGL ltd., King City, Ont., and Greeneridge Sciences Inc., Santa Barbara, CA, for Western Geophysical, Houston TX, and Nt. Mar. Fish. Serv., Anchorage, AK, and Silver Spring, MD. 390 pp.

Southall, B.L., A. E. Bowles, W. T. Ellison, J. J. Finneran, R. L. Gentry, C. R. Greene, Jr., D. Kastak, D. R. Ketten, J. H. Miller, P. E. Nachtigall, W. J. Richardson, J. A. Thomas and P. L. Tyack. 2007. Criteria for behavioral disturbance. Aquatic Mammals 33:446-473.

Stone, C. J. and M. L. Tasker. 2006. The effects of seismic airguns on cetaceans in UK waters. J. Cetacean Res Manage 8:225-263.

Wade, L., H. Whitehead and L. Weilgart. 2010. Conflict of interest in research on anthropogenic noise and marine mammals: Does funding bias conclusions? Marine Policy 34:320-327.

UNDERWATER 3-D PASSIVE ACOUSTIC BAYESIAN TRACKING OF PACIFIC WALRUS

Brendan Rideout[1], Stan E. Dosso[1], and David Hannay[2]
[1]School of Earth and Ocean Sciences, University of Victoria, 3800 Finnerty Rd., Victoria, BC, Canada, V8P 5C2,
bprideou@uvic.ca
[2]JASCO Research Ltd., 2101-4464 Markham St., Victoria, BC, Canada, V8Z 7X8

1. INTRODUCTION

This paper describes an MSc thesis research project aimed at three-dimensional (3-D) passive acoustic localization and tracking of vocalizing Pacific Walrus. Passive acoustic localization, in this context, is the process of determining the location of an underwater marine mammal without producing artificial sounds; only the natural sounds of the animal itself are used. Tracking, in this case, is the process of determining a sequence of underwater positions for a sequence of calls. The primary goal of this work is to develop a method for gathering behavioral data on marine mammals without the need to physically attach devices to the animal.

Experimental error is an important part of any scientific endeavor, and must be addressed for the results to be of greatest use. In the context of localization, experimental error is present in the data collection process, both in measuring arrival times for underwater calls and in knowledge of the experiment geometry and characteristics of the acoustic propagation environment (i.e. water depth and sound speed). Quantifying the effect these errors have on the reliability of the location is critical to interpreting the localization results. This work includes uncertainty in the measured data as well as in the sensor locations and the environmental parameters in an attempt to develop a comprehensive picture of localization uncertainty. Without knowing how uncertain the individual call localizations within a track are, it is not possible to say how much more likely the given track is compared with other possible tracks, and therefore how much significance to ascribe to features of the track (such as variations in depth or apparent swim speed of the calling animal).

2. BACKGROUND THEORY

Localization, as implemented in this paper, is similar in principle to triangulation. Triangulation typically only uses energy which travels directly from the unknown source location to the receivers. Because sound bounces off the sea surface and bottom, additional information is available for underwater localization in the form of multipath arrivals (energy which has bounced one or more times off the surface and/or bottom prior to reaching the recorder).

The following derivation describes the approach used to calculate a track for a sequence of calls [1]. To account for all possible sources of error, the set of unknown model parameters $\mathbf{m}$ includes 3-D locations and times for each call in the track, 3-D hydrophone locations, sound speed, water depth and inter-hydrophone time synchronization factors. Arrival times for all direct and multipath arrivals from each hydrophone and call are arranged into the vector $\mathbf{t}$. The relationship between $\mathbf{t}$ and $\mathbf{m}$ is non-linear, but can be linearized about an arbitrary starting model $\mathbf{m}_0$ by retaining the first order term from the Taylor expansion:

$$\mathbf{t} = \mathbf{t}(\mathbf{m}) = \mathbf{t}(\mathbf{m}_0 + \delta\mathbf{m}) \approx \mathbf{t}(\mathbf{m}_0) + \mathbf{J}\delta\mathbf{m} \qquad (1)$$

where $\mathbf{J}$ is the Jacobian matrix composed of partial derivatives evaluated at $\mathbf{m}_0$. Rearranging **(1)** and taking $\mathbf{m}-\mathbf{m}_0=\delta\mathbf{m}$ yields:

$$\mathbf{d} \equiv \mathbf{t} - \mathbf{t}(\mathbf{m}_0) + \mathbf{J}\mathbf{m}_0 = \mathbf{J}\mathbf{m} \qquad (2)$$

This inverse problem is linear but, because hydrophone and source positions are both unknown, is ill-conditioned. Regularization is used to add additional information to the problem in the form of a preference for a smooth track, and prior estimates (with approximate uncertainties) for sound speed, water depth, and hydrophone positions. Specifically, the objective function minimized to fit the data and prior information takes the form:

$$\psi = [\mathbf{d} - \mathbf{J}\mathbf{m}]^T \mathbf{C}_\mathbf{d}^{-1} [\mathbf{d} - \mathbf{J}\mathbf{m}] + \mu_1 \mathbf{m}^T \mathbf{R}\,\mathbf{m} + \mu_2 [\mathbf{m} - \hat{\mathbf{m}}]^T \mathbf{C}_{\hat{\mathbf{m}}}^{-1} [\mathbf{m} - \hat{\mathbf{m}}], \qquad (3)$$

where $\mathbf{C}_\mathbf{d}$ is the data covariance matrix, $\mathbf{R}$ is a 3-D roughening matrix applied to the track locations such that $\mathbf{m}^T\mathbf{R}\mathbf{m}$ represents the L_2 norm of the second derivative (curvature) of the track, $\hat{\mathbf{m}}$ is a vector of prior estimates of model parameters in $\mathbf{m}$, $\mathbf{C}_{\hat{\mathbf{m}}}$ is a matrix of prior uncertainties, and μ_1 and μ_2 are trade-off parameters. Minimizing ψ yields the equation:

$$\mathbf{m} = \left[\mathbf{J}^T \mathbf{C}_\mathbf{d}^{-1} \mathbf{J} + \mu_1 \mathbf{R} + \mu_2 \mathbf{C}_{\hat{\mathbf{m}}}^{-1}\right]^{-1} \left[\mathbf{J}^T \mathbf{C}_\mathbf{d}^{-1} \mathbf{d} + \mu_2 \mathbf{C}_{\hat{\mathbf{m}}}^{-1} \hat{\mathbf{m}}\right] \qquad (4)$$

Due to the linearization step, the solution converges over multiple iterations of (4). Values for the trade-off parameters are selected such that the data and prior estimates are fit to a statistically appropriate level according to the χ^2 criterion [1]. The resulting track is one which maximizes smoothness while still fitting the data to within the data uncertainty. In other words the degree of smoothing, and the closeness of the final solution to the prior estimates, is proportional to the uncertainty in the data. If the data have large uncertainty the resulting track will tend to become smoother and the model parameters become more similar to their prior estimates.

To calculate uncertainties for the track resulting from (4), the model covariance matrix is calculated for the final solution. The general equation for the model covariance matrix (whose diagonal terms are the variances for the Gaussian distributions which represent the model parameter uncertainties) is:

$$\mathbf{C}_\mathbf{m} = \left\langle \left(\mathbf{m} - \langle \mathbf{m} \rangle\right)\left(\mathbf{m} - \langle \mathbf{m} \rangle\right)^T \right\rangle \qquad (5)$$

By substituting the value for **m** calculated in (4), (5) yields the following formula for C_m for the case of additive data noise:

$$\mathbf{C_m} = G^{-1}\mathbf{J}^T\mathbf{C_d}^{-1}\mathbf{J}\left(G^{-1}\right)^T, where$$
$$G = \mathbf{J}^T\mathbf{C_d}^{-1}\mathbf{J} + \mu_1\mathbf{R} + \mu_2\mathbf{C_{\hat{m}}}^{-1} \quad (6)$$

3. FIELD STUDY

In August 2009, three ocean bottom hydrophone (OBH) recorders were deployed near the Hanna Shoal in the northeastern Chukchi Sea in a roughly equilateral triangle (~400 m on a side). Due to the shallow water (30 m) and flat bottom in this part of the Chukchi, the acoustic propagation environment is well suited for carrying out localization work. Additionally, the Hanna Shoal is known to be a common feeding ground for female and juvenile Pacific Walrus during the summer months. The OBHs recorded with a 16 kHz sampling frequency at 24 bits/sample for 2.5 months. Following retrieval, an automatic marine mammal vocalization detection and classification algorithm (developed by JASCO Research Ltd.) identified a segment of the data containing a high concentration of probable walrus knocks (impulsive calls). The calls in this segment were further processed using a semi-autonomous frequency domain-based edge detector to pick out the direct and multipath arrival times for each call at each OBH. Figure 1 shows time synchronized recordings of a single walrus knock recorded on each hydrophone. Note that the call arrives at each hydrophone at a different time and that the time delays between the various multipath arrivals also differ. These inter-OBH differences represent the information used to locate the point of origin for the knock.

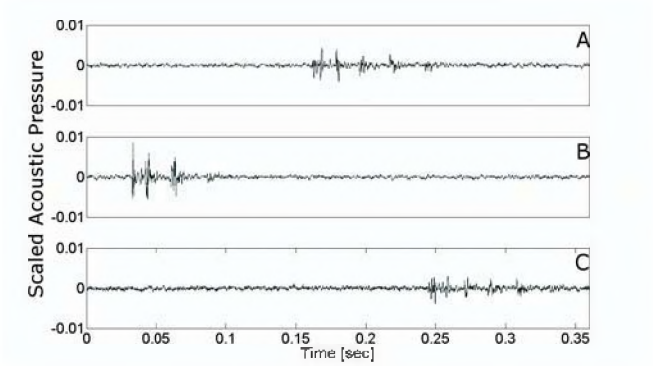


Figure 1 - Time synchronized walrus knock recorded on Hanna Shoal OBH A, B, and C.

4. SYNTHETIC EXAMPLE

Because 3-D localization using real data was not fully implemented at the time of writing, a simulation is used here to illustrate the inversion results. Figure 2 shows a comparison between true source positions and locations calculated using the linearized, regularized tracking algorithm. For each of the true source positions, simulated arrival times were calculated for each of the three hydrophone locations. Zero mean Gaussian noise was added to these arrival times, with standard deviations on the order of 0.0005 s. The measured data is expected to have a similar noise standard deviation. The noisy data were processed using the tracking algorithm. For this plot, only prior estimate regularization was used (no smoothing was done). The top panel shows a plan (*x*-*y*) view of the track. The bottom panel shows source depth (*z*) as a function of *x*. Each call's (x,y,z) position uncertainty is taken to be Gaussian distributed. Standard deviations for these uncertainties are shown in Figure 2 for the *y* and *z* coordinates for each call location. Hydrophone coordinates and uncertainties were also calculated, but are not shown. If data and environmental uncertainties were reduced (indicating that the data is more informative and the environment better understood), call location uncertainties would also be reduced.

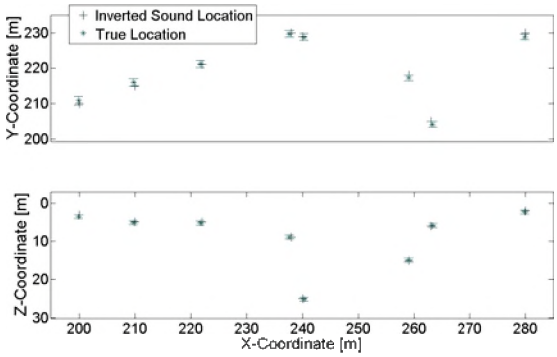


Figure 2 - True and localized 3-D track for synthetic data.

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.

ACKNOWLEDGEMENTS

This work has been partly funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada, JASCO Research Ltd., the Province of British Columbia's Ministry of Advanced Education, and the University of Victoria.

DETECTION OF PRECISE TIME EVENTS FOR MARINE MAMMAL CLICKS

Tin Ma[1], Nicole Collison[2], James Theriault[2], Joe Hood[3], Sean Pecknold[2], and Ben Bougher[3]
[1] Simon Fraser University, thm1@sfu.ca; [2] Defence R&D Canada – Atlantic: nicole.collison, james.theriault, sean.pecknold@drdc-rddc.gc.ca; [3] Akoostix Inc.: jhood, bbougher@akoostix.com

1. INTRODUCTION

Acoustic sonar users, cetacean ecologists, commercial shipping, fisheries, and tourism operators can benefit from improved marine mammal detection and localization. Marine mammals emit various sounds (such as songs, moans, and clicks) for communication, navigation, and foraging purposes. Sensors and techniques for detection, classification, and localization of marine mammal sounds are being developed by numerous institutions. The results presented in this paper benefit from the unique situation of having "truth" information on the source location by emulating marine mammal sounds.

This paper focuses on a localization technique where Time-Differences of Arrival (TDOA) for a signal on multiple receivers is estimated and used as input to a hyperbolic cross-fixing scheme [1]. The algorithm begins by using the discrete Teager-Kaiser energy operator (TKEO) on the dataset [2, 3] to remove phase information and allow event detection using a split-window moving average (SWMA). Detected events are then organized onto "click maps" based on individual animals having unique inter-click interval patterns (ICI) [4], which are then cross-correlated and interpolated to obtain precise measurements for each TDOA. By considering trends in the measurements, a likelihood function can be applied to get a better successive TDOA estimate.

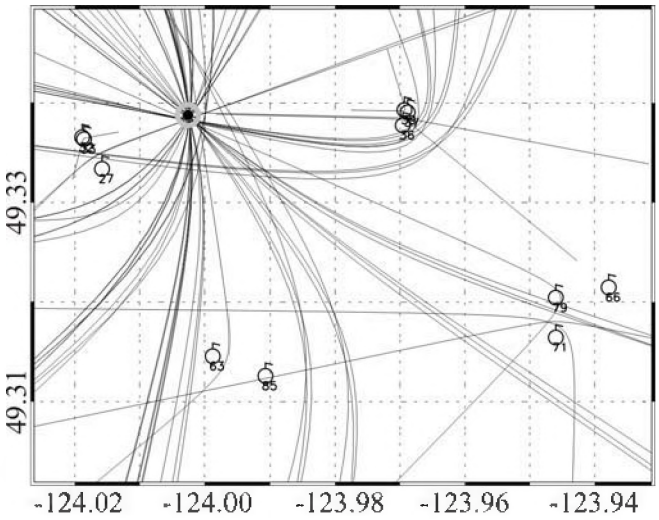


Figure 1: Tactical plot showing trial geometry and an example localization of the source at ~2301 UTC. The estimated source location is within the light grey circle and is indistinguishable from the true location at this scale.

2. SEA TRIAL DATASET

On January 28th 2010, DRDC Atlantic held several sea trials at the Canadian Forces Maritime Experimental and Test Ranges, Nanoose Bay, BC. One dataset was intended to provide clean, known click data that simulated a marine mammal click train. As a first step, frequency-modulated pulses were transmitted, based on a "standard" pulse with a frequency range of 1.0-2.5 kHz and duration of 1.0 s. One hundred pulses, slightly varied from the standard pulse (in amplitude, duration, and frequency range, but not bandwidth), were played as a quasi-periodic pulse train (where the ICI was also slightly varied). The pulse train was repeated to record fifteen minutes of data, starting at 2300 UTC. The recordings were match-filtered with the standard pulse to compress the pulses into simulated marine mammal clicks. This effectively compresses the energy in the 1-sec FM pulses into pulses of approximately 1-ms duration, mimicking the signal characteristics required to test the algorithm (i.e., quasi-periodic ICI, appropriate energy level). We refer to this as the "raw" data in this paper. The signals were received at eleven sonobuoys outfitted with a global positioning system device (GPS) drifting in grouped arrays at four different locations (see Figure 1, where the sonobuoys are denoted by open circles with flags and the source location is within the light grey circle). The speed of sound was assumed constant at 1480 m/s. Analysis of the time period from 23:00:45 to 23:01:15 UTC is presented in this paper's figures.

3. METHODOLOGY

a. Event detection

A series of events (i.e., clicks) are selected from one sonobuoy time series (master channel) for a time period of at least 30 seconds. The time series are later broken up into 15-second time frames with 50% overlap. TDOA values are determined per time frame and remaining sonobuoy time series (slave channels). For each slave channel, the length of data retrieved considers the maximum possible time difference between buoy pairs so that only necessary data is processed. The frame duration determines how many clicks will be cross-correlated at a time. By increasing this parameter, click association becomes more robust at the cost of additional computational time.

An energy-based amplitude envelope is calculated using the TKEO on all selected and retrieved time series data. Figure 2 shows details of the raw data for a single click with the TKEO envelope. The SWMA is used on the TKEO envelope to detect significant events. The SWMA consists of five windows: two noise windows outside of two gap

windows, all of which symmetrically enclose a signal window. The averages of the energies in the signal and noise windows are compared to give a signal-to-noise ratio. Ideally, the signal windows should encompass the majority of the signal energy; the gap windows should be set to prevent the signal tails from being included in the noise estimate. A threshold is set so that significant events are detected and their times listed. Because of the limited sampling resolution, the precise timing for each event is estimated by parabolic interpolation [5] of the TKEO envelope.

b. Event association

After obtaining a list of event times, an event map is generated for each channel. To increase robustness in association, each event is represented by a Gaussian distribution function of unit amplitude centered precisely at each event time. The width of the Gaussian function is set to compensate for small variations in the received click pattern. The resulting event map represents when clicks were detected, whether they are direct- or multi-path, as shown in Figure 2.

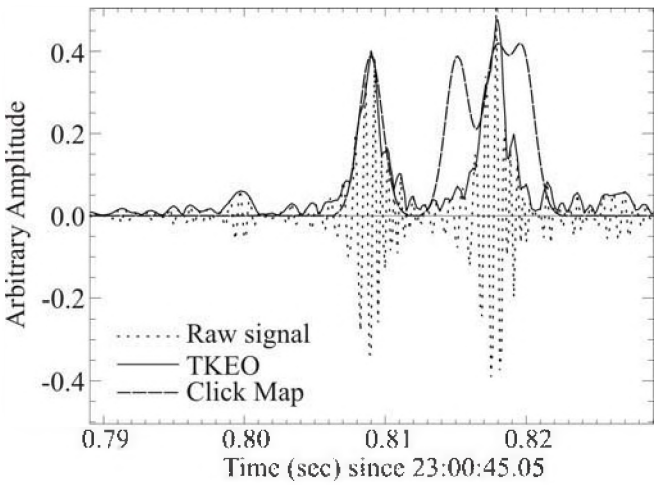


Figure 2: Details of a click. The click map shows spurious peaks because it is triggering on noise peaks.

For each time frame and sensor, event maps are extracted. The event maps from slave channels are cross-correlated with the event maps from the master channel where the output amplitude provides a measure of the match quality. The correlation is subject to noise resulting from partial matches, so candidate TDOA values are found using the SWMA with interpolation on the correlation function.

Candidate TDOA values are then analyzed to select one TDOA estimate for each time frame/sensor combination. The most-likely (i.e., highest-correlated) time is picked for the first time frame. Per subsequent time frame, a Gaussian error model is constructed from prior (selected) TDOA values and used to weigh the correlation of candidate values, so as to find the most consistent (assumed best) delay. This is valid because arrival time will be constrained for a moving source; thus TDOA changes over time will also be limited.

4. RESULTS AND DISCUSSION

The primary result for this paper is shown in Figure 1, where the estimated source location is defined by a black dot within a computed area of probability (denoted by the light grey circle). The drawn hyperbolas are defined by TDOA measurements, and the localization algorithm defines statistically the estimated source location and associated area of probability by clustering the crossing points. The estimated location was 55 m from the source GPS position at a bearing of 200° True. The localization error radius was 115 meters, containing the source.

In Figure 1, the hyperbolas associated with 10 of the 11 sonobuoys provided reasonable inputs to the localization; however, all hyperbolas associated with buoy 71 were found to be inaccurate, possibly due to position error.

In the future, we hope to improve our detection algorithm by implementing an adaptive estimation approach [3]. Different algorithms could be tried such as using a stochastic matched filter [6]. These can all be compared to a ground-truth dataset as we have done here. In addition, this dataset could be used to explore the effect of using a subset of sonobuoys to investigate how the algorithm would react to closely-space buoys. Since this dataset does not include multiple sound-emitting sources, a new dataset would have to be obtained to test this algorithm against this particular case. Testing should also be performed using vocalizing animals to fully test robustness of the algorithm.

REFERENCES

[1] Informal communication, Joe Maksym, DRDC Atlantic.

[2] H. Glotin, F. Caudal, and P. Giraudet, "Whale cocktail party: real-time multiple tracking and signal analyses," *Canadian Acoustics*, vol. 36, no. 1, pp. 141-147, 2008.

[3] V. Kandia, Y. Stylianou, and T. Dutoit, "Improve the accuracy of TDOA measurement using the Teager-Kaiser Energy operator," in *Proceedings of the International Workshop on New Trends for Environmental Monitoring Using Passive Systems (PASSIVE '08)*, IEEE Xplore, Hyeres, France, October 2008.

[4] R.P. Morrissey, J. Ward, N. DiMarzio, S. Jarvis, and D.J. Moretti, "Passive acoustic detection and localization of sperm whales (Physeter macrocephalus) in the tongue of the ocean," *Applied Acoustics*, vol. 67, no. 11-12, pp. 1091-1105, 2006.

[5] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, 1986, Chapter 10.2.

[6] F. Caudal and H. Glotin, "Stochastic Matched Filter outperforms Teager-Kaiser-Mallat for tracking a plurality of sperm whales," in *Proceedings of the International Workshop on New Trends for Environmental Monitoring Using Passive Systems (PASSIVE '08)*, IEEE Xplore, Hyeres, France, October 2008.

AUTOMATIC SIGNAL DETECTION IN NOISE USING ENTROPY

Christine Erbe
JASCO Applied Sciences Australia Pty Ltd, Brisbane Technology Park, PO Box 4037, Eight Mile Plains, QLD 4113 Australia, Christine.Erbe@jasco.com

1. INTRODUCTION

Automatic detection of signals in noise is a common problem in many areas of acoustics. In the field of passive acoustic monitoring of marine mammals, the signals to be detected are vocalizations. The noise originates from natural (e.g. wind, waves, rain) and man-made sources (e.g. shipping, construction, seismic surveys). Signal characteristics vary broadly: frequency ranges from a few Hz to 200 kHz, duration from milliseconds to seconds to hours. Noise characteristics vary by similar orders of magnitude. While specific automatic detectors have been designed to successfully find specific calls in specific environments, the challenge is to find a large variety of calls in a large variety of noise. An exploitable difference between calls and noise is that most noise is a result of stochastic processes (wind, waves, rain, cavitating propellers and seismic airguns generate gas bubbles underwater of varying size and resonance frequency), while many animal signals are a result of deterministic processes (vibrating strings and cavities of predetermined and fixed size). As a result, Shannon entropy (also called information entropy) can be expected to differ between signal and noise. Shannon entropy quantifies the information contained in a data set. The concept was introduced by Claude E Shannon in his 1948 paper "A mathematical theory of communication" (Shannon 1948). The current article investigates whether entropy makes a "good" detector for animal calls in underwater ambient noise.

2. METHOD

Underwater acoustic recordings from the Arctic were used to test and compare automatic signal detectors. The species that were present in the recordings and their common call types are listed in Table 1.

Table 1. Selected arctic species and their call types

Species	Call Types
Bowhead whale	FM tones, song, pulsive calls
Gray whale	FM tones, moans, pulsive calls
Beluga whale	Whistles, pulsed calls, clicks
Walrus	Knocks, bell sounds, grunts
Bearded seal	FM signals
Ringed seal	Barks, yelps

Three different automatic detectors were tested and compared: 1) a broadband peak energy detector, 2) peak energy detection in a set of bandpass filters, and 3) a peak entropy detector. Each detector computed a statistical quantity s(t), a mean $\overline{s}$ and a standard deviation σ. For a given threshold γ, a signal was deemed present if s(t) > $\overline{s}$ + γσ. Two different windows were applied to the time series, an "averaging" window of 1 min length, over which the mean and standard deviation were computed, and a "detection" window of 100 ms length, over which the instantaneous statistic was computed for comparison to the mean. The detection window immediately preceded the averaging window and both were moved through the time series sample by sample.

Given a recorded pressure time series p(t), the broadband energy detector computed $p^2(t)$. The band-passed energy detector split the signal into multiple overlapping pass bands $p_f(t)$ before computing $p_f^2(t)$. This was done by Fourier transforming the time series over 100 ms long windows, and grouping the Fourier coefficients into octave bands covering the recorded bandwidth of 10 kHz. Moving the 100 ms window through the time series sample by sample yielded a time series of Fourier coefficients. Energy was computed in each band, and a signal was deemed present if the energy in any one band surpassed the mean by a preset threshold. A slightly different implementation of the band-passed energy detector and more detail about the entropy detector can be found in Erbe & King (2008).

The entropy detector also Fourier transformed the pressure time series over 100 ms windows and computed the power spectrum |P(f)|. The power spectrum was normalized so that the entropy did not depend on the absolute energy: $\sum P(f) = 1$. Shannon entropy was computed as $-\sum P(f) \cdot \log P(f)$. The 100 ms window was moved through the pressure time series sample by sample, yielding a time series of entropy.

3. RESULTS

Figure 1 shows an example of two marine mammal calls that were detected by the entropy detector. A spectrogram is plotted with entropy (not to scale) overlain as a thin black line.

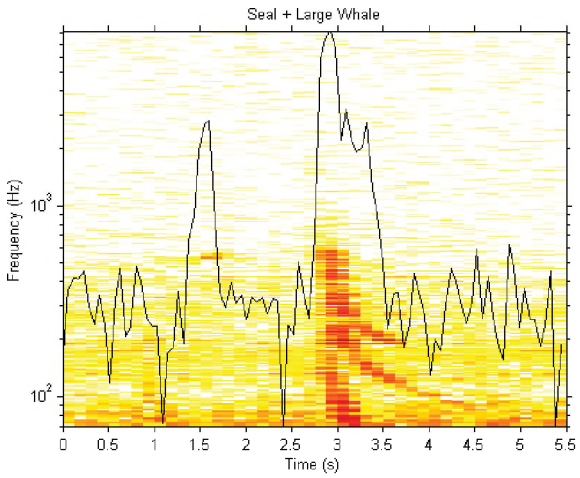


Fig. 1. Spectrogram of a 5.5 s recording showing a faint pinniped call and an FM call of a baleen whale. Entropy is shown as a black line unscaled.

To compare the three detectors, receiver-operating-characteristics (ROC) were computed. An automatic detection task is a binary classification problem with four possible outcomes (hit, miss, false alarm, correct rejection). With P_{FA} as the probability of false alarm and P_{CD} as the probability of correct detection (determined by comparing automatic detections to manual detections), ROC curves are computed by varying the threshold γ. As γ is increased, the number of false alarms decreases at the cost of the number of correct detections, because the number of misses increases. An ideal detector would have a probability of false alarm of 0 and a probability of correct detection of 1. The "best" detector in a comparison of detectors is the one approaching (0|1) most closely, in this case the entropy detector.

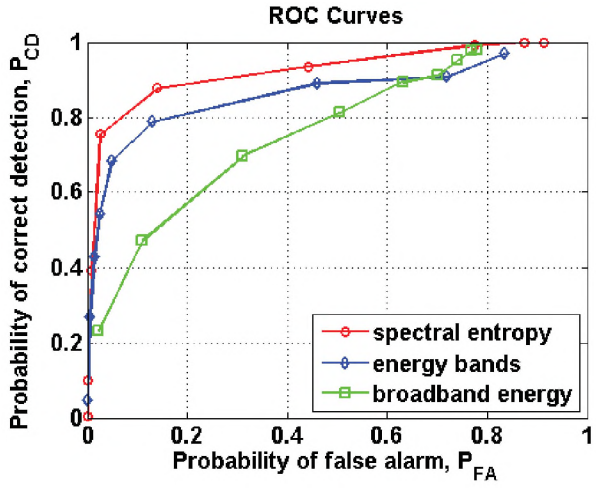


Fig. 2. Comparison of the performances of the three detectors using receiver-operating characteristics (ROC). Each data point corresponds to a set threshold γ.

4. DISCUSSION

All of the animals in Table 1 made calls with tonal components (whistles or frequency modulated (FM) tones). Shannon entropy was significantly higher for these than for ambient noise and these calls were therefore detected very well. The broadband clicks of animals were not detected well by the entropy detector.

Ambient noise can have tonal components (e.g. ice noise and shipping), but these did not cause a significant number of false alarms in the tested data set. The largest number of false alarms was due to ringing bubbles. These were believed to be biological in origin, but the animal making them could not be identified.

For calls of constant frequencies plus harmonics, the band-passed energy detector worked well if the condition was set that energy had to be detected simultaneously in more than two and less than four frequency bands.

Comparing the instantaneous value of the statistic to the median instead of the mean improved performance as ambient noise can have large yet brief (transient) outliers which affect the mean but not the median.

Choosing the window lengths depends on the ultimate goal. If individual calls need to be counted, then the detection window should be short and ideally of the length of typical calls. If a mere species present/absent outcome is desired, window lengths are not critical and can be longer, grabbing more than one call at a time. The length and placement of the averaging window can be made adaptive. E.g. if a group of vocalizing belugas is encountered, the averaging window would ideally remain fixed in time before the vocalizations start rather than moving into the animal sounds and averaging them into ambient noise. Once the vocalizations have stopped, the averaging window can be jumped forward to the end of the vocalizations and continue to move through the data.

Altogether, the entropy detector worked well to find sounds of the target species in their arctic acoustic environment. The entropy detector should be considered a first step in a series of automatic analysis tools. As a second step, all detected signals need to be classified to species, which was not attempted in the current study.

REFERENCES

Erbe, C., and A.R. King (2008). Automatic detection of marine mammals using information entropy. J. Acoust. Soc. Am. 124, 2833-2840.

Shannon, C.E. (1948). A mathematical theory of communication. Bell Syst. Tech. J. 27, 379-423.

ACOUSTIC DIVERSITY AND MATING SIGNALS IN THE PSYLLOIDEA (HEMIPTERA)

Diana M. Percy
Dept. of Botany and Beaty Biodiversity Museum, University of British Columbia
3529-6270 University Boulevard, BC, Canada, V6T 1Z4 diana.percy@ubc.ca

1. INTRODUCTION

Acoustic communication is important in many insect groups (Hill 2008, Čokl & Virant-Doberlet 2003). The mechanisms and diversity of acoustic behavior in psyllids (Psylloidea, Hemiptera) has only been more extensively studied recently (Percy et al. 2006, Tishechkin 2005). Psyllids are small (2-8 mm) plant feeding insects related to aphids, scales, and whiteflies. Their sound making apparatus consists of a "simple" stridulatory system, but there appears to be a much greater diversity of sounds than might be expected from a stridulatory mechanism.

The small size of these insects makes the recording of acoustic behavior particularly challenging. Nevertheless, it has been possible to record the substrate transmitted vibrational signals, and to use playback to elicit species specific acoustic responses.

Whiteflies and aphids, which are both related to psyllids, are known to produce limited acoustic signals, but psyllids are the only members of the Sternorrhyncha clade known to produce and use acoustic signals extensively for mate location and selection.

There are a number of important agricultural pest species in the Psylloidea, the best known being the tomato/potato psyllid, the Asian citrus psyllid, and the carrot psyllid. A number of species also damage horticultural and ornamental plants. With the development of more sophisticated sound processing computer systems and electronic equipment, the possibility of detection and deterrence using acoustic signals is increasingly being investigated. I have recorded examples of male psyllids producing apparent signal jamming noises to deter mate location of rival males, and it therefore seems likely that certain background noises may compromise successful mating. However, it is not clear whether the high densities of individuals often found with pest species would be adversely affected by this approach because acoustic mate location appears to be a strategy to locate mates under conditions of low density.

1.1 Allopatric versus sympatric acoustic signals

Acoustic signals are important in both species recognition and mate selection in psyllids (Percy et al. 2006). Male and female psyllids produce reciprocal acoustic mate signalling, often as highly synchronised duetting. The role of acoustics in psyllid mating and speciation is illustrated with data from allopatric versus sympatric groups and morphologically cryptic species. Two psyllid groups that differ in the amounts of geographical and ecological sympatry were also found to have different correlations between genetic and acoustic diversity depending on whether the species were sympatric or allopatric.

The majority of this acoustic research was undertaken in one of the centers of species diversity for the Psylloidea: Australia.

2. RESULTS

I recorded a high diversity of acoustic signals from 26 species in 12 genera of Australian psyllids. The greatest diversity of Australian psyllids is found in the almost exclusively Myrtaceae-feeding subfamily Spondyliaspidinae (family Psyllidae). I selected a complex of morphologically similar taxa that occur on closely related host plants and used playback experiments to test female receptivity to conspecific and heterospecific male calls from different hosts. Using these methods I have been able to address questions relating to species concepts: specific mate recognition systems (SMRS) and reproductive barriers.

I investigated the divergence in acoustic signals between three closely related, morphologically cryptic species. I found that associations with different host plants coincided with differences in acoustic mating signals, and that the acoustic signals were more divergent in sympatric taxa. There is more phylogenetic information in groups that are relatively recently derived, and groups in which the species are not sympatric.

Among three cryptic taxa in the genus *Cardiaspina*, the results suggest a possibility of unidirectional but not bi-directional gene flow between species occurring on sympatric hosts. As these psyllids cause conspicuous leaf necrosis during feeding, and in outbreak years heavily infested trees may die completely, the reproductive biology and population dynamics of this psyllid group is of broad ecological interest.

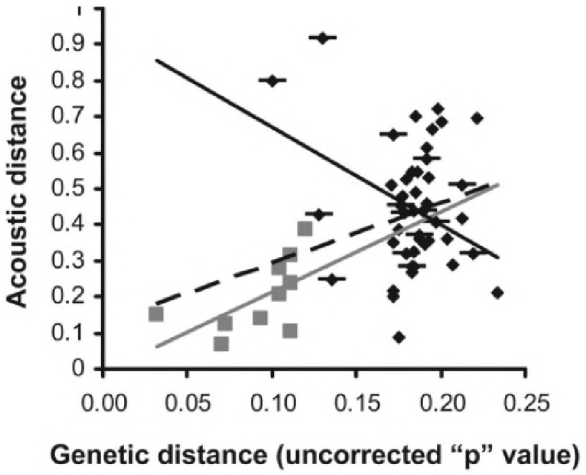


Figure. 1. Comparison of acoustic and genetic distance for 11 triozid psyllid species. There is a significant correlation between genetic and acoustic distance using pairwise comparisons of all 11 taxa (dashed line, $P = 0.005$, $R^2 = 0.1393$). An analysis of trends within the two groups showed that within the *Schedotrioza*-group (grey squares), which are recently derived species that are typically allopatric, the trend also indicates a positive correlation, though non-significant (grey line, $P = 0.07$, $R^2 = 0.3459$), but within the Casuarinaceae-feeding group (black diamonds with horizontal bar), a group of more divergent species that are typically sympatric, the trend is reversed, suggesting a negative correlation (black line, $P = 0.07$, $R^2 = 0.2302$) (from Percy et al. 2006).

3. DISCUSSION

Acoustic data from psyllids in some cases has significant phylogenetic content. However, the divergence in acoustic signals may occur rapidly under certain conditions (e.g., when taxa are sympatric), potentially confounding phylogenetic interpretations. If acoustic characters evolve rapidly through competition or sexual selection in sympatry, and are also prone to convergence due to limited repertoire, then more phylogenetic information may generally be found in acoustic data from recently speciating, allopatric taxa.

Acoustic signals in psyllids are species and gender specific, and therefore the possibility exists to use acoustic identification/detection systems. Vibrational acoustic signals for mate location are likely to be used by insects at low densities such as on colonization of a new host individual, and thus vibrational detection systems could provide an early warning of psyllid colonization on plants. The potential use of disturbance noise to deter psyllid establishment on host plants has some support in the apparent use of these tactics by psyllid males.

REFERENCES

Čokl, A., & Virant-Doberlet, M. (2003). Communication with substrate-borne signals in small plant-dwelling insects. *Annual Review of Entomology* **48**, 29–50.

Hill, P.S.M. (2008) Vibrational Communication in Animals. Harvard Univ. Press.

Percy, D.M., Taylor, G.S. & Kennedy, M. (2006) Psyllid communication: acoustic diversity, mate recognition and phylogenetic signal. *Inververtebrate Systematics* **20:** 431-445.

Tishechkin, D. Yu. (2005). Vibratory communication in Psylloidea (Hemiptera). In *Insect Sounds and Communication* (eds S. Drosopoulos & M. M. Claridge) pp. 357–363. CRC Taylor and Francis: London, UK.

ACKNOWLEDGEMENTS

Funding was provided by the Leverhulme Trust, UK.

THE EFFECT OF FIBRIL ERECTION ON HEARING IN MALE AEDES TOGOI: AN OPEN AND SHUT CASE

Melanie Hart, Peter Belton, and Gerhard Gries
[1]Dept. of Biological Sciences, Simon Fraser University, 8888 University Dr., Burnaby, BC, Canada, V5A 1S6

1. INTRODUCTION

Mosquitoes hear with their Johnston's organs, complex sensory structures in the large second segment of their antennae. These organs detect the direction and amplitude of the movement of the distal part of the antenna as it is displaced by sound waves. In many species, the long fibrils on the male flagellum are erect when the insects swarm at dusk and dawn, and recumbent during the rest of the day (Clements 1999). It has been claimed that when the fibrils are recumbent, males cannot hear (Nijhout and Sheffield, 1979, and references therein). To test this, we used laser vibrometry to examine the movement of the antennae of male *Aedes togoi* with erect and recumbent fibrils to sound stimuli of frequency sweeps and clicks. From the sweeps, we determined the resonant frequency of the antennae. We hypothesized that: a) males with recumbent fibrils would be able to detect sound; b) the resonant frequency of males with open fibrils and their selectivity would be different from those with recumbent fibrils, and c) females would not change their resonant frequency when males had open or recumbent fibrils.

2. METHODS

2.1 Experimental Insects

A colony of *Aedes togoi* was cultured from larvae collected from rock pools in West Vancouver, BC, Canada. Larvae were fed fish flake food, and maintained at 25° C on a 16:8 light:dark photoperiod. Adults were fed sugar water and allowed a bloodmeal. All insects were 3 to 5 days old. Males with recumbent fibrils were tested during the light photoperiod when their fibrils were naturally closed; males were tested in the erect fibril state during the dark photoperiod, when their fibrils were expanded naturally. Females were tested along with males at both times.

2.2 Sound generation

A 10 Hz to 1 kHz logarithmic sweep, 5.1 s long was generated with a Brunelle Instruments model 3050 sweep generator. Sound was kept at 65 dB at the insect antenna by running the signal through a General Radio 1569 Automatic Level Regulator that monitored sound level through a Radio Shack Realistic sound level meter (SLM) 33-1028 or a Brüel and Kjaer 2204 SLM placed directly next to the mosquito. The sound was played from a Sennheiser headphone placed approximately 10 cm from the antenna.

2.3 Laser vibrometry

Laser vibrometer recordings from the base of the flagellum were made with a Polytec OFV-2500 with a Polytec Vib-E 220 data acquisition system. The sampling rate was set at 24 kHz for 10.92 s with 2 mm/s sensitivity. The sweep was recorded concurrently with the antennal response by recording the output of the SLM on the second channel of the Vib-E system. All recordings were made at a temperature of 24-25°C.

2.4 Data analysis

To determine the resonant frequency of the antennae, the laser recordings were analyzed in Adobe Audition 3.0. The time of the phase change between antennal response and the sound played to the antenna was noted, and corresponding frequency determined using Raven Pro 1.4. The directional hypothesis that the erect fibrils of males have a higher resonant frequency than recumbent fibrils was tested with a Student's t-test in JMP 7.0.2, 90% confidence. An ANOVA was also run in JMP 7.0.2 to test for differences between the four tested categories of antennae, 95% confidence.

3. RESULTS

Male antennae with recumbent and erect fibrils and female antennae recorded in the same times as the males all showed a resonant peak between 200 and 400Hz. Examples are shown in Figure 1. Males with erect fibrils had significantly higher resonant frequencies than males with recumbent fibrils when tested with the directional t-test ($p < 0.08$), and were significantly different from the two categories of female antennae when tested in an ANOVA. This test showed no statistically significant difference between the resonant frequencies of female antennae during the active and inactive times (Figure 2).

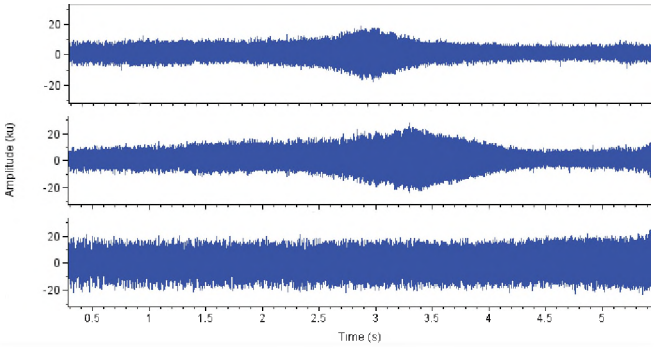


Figure 1. Responses of male *Aedes togoi* with recumbent (top), and erect fibrils (middle) to a flat 65 dB logarithmic sweep, from 10 Hz to 1 kHz (bottom).

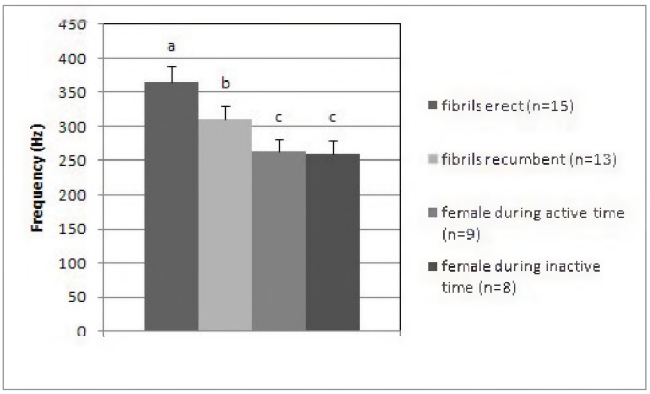


Figure 2. Mean resonant frequency +SE of male antennae with recumbent or erect fibrils (t-test) and of female antennae during active or inactive times (ANOVA, *p* value set at 0.05).

4. DISCUSSION

Laser vibrometry of male antennae with erect or recumbent fibrils and female antennae show very similar resonant responses. The antennae of females have only 10% the surface area of typical 'bottlebrush' male antennae of the same species (Clements 1999), and this makes it unlikely that the physical characteristics of the flagellum play a major role in the sensitivity and tuning of the antenna. Clearly, long fibrils are not essential to detect the wingbeat of nearby flying mosquitoes of either sex. As recently as 1999, Clements in his admirable monograph concluded that "female mosquitoes are believed not to respond behaviourally to sound" but since then, female mosquitoes in four different genera have been shown to change their wingbeat frequency to match harmonics with a nearby flying male (Gibson and Russell, 2006 Cator et al, 2009; Warren et al, 2009; and Pennetier et al, 2010) and one species has been shown to be attracted by the call of its frog food source (Borkent and Belton, 2006). Nevertheless the ultra-sensitive Johnston's organs of male mosquitoes are tuned closely to the wingbeat frequency of a female of the same species, to which they are attracted during swarming. In *Aedes togoi* we measured mean wingbeat frequencies of 523 Hz in males and 306 Hz in females, the latter within the range of resonance we measured in male antennae.

Verticillate (bottlebrush) antennae evidently give male flies that are attracted from swarms by the sound of a female some selective advantage because they have evolved in at least six different families, and in the Ceratopogonid biting midges, not closely related to mosquitoes, the tiny males of most species erect their long antennal fibrils before they swarm and mate.

It is not obvious why the resonant frequency of the male antennae should increase when the fibrils expand because viscous drag would tend to reduce it (Fletcher, 1992). However, our results are similar to those of Pennetier et al. (2010) who found that the antennae were less sensitive but tuned to a higher frequency when the fibrils were extended in *Anopheles gambiae*, another species that erects its fibrils before swarming. Perhaps the contractile properties of the sense cells in Johnston's organ (Göpfert, and Robert, 2001) somehow increase the frequency of its tuning. In *Anopheles*, an alpha-adrenergic transmitter substance is thought to trigger extension of the fibrils and it could possibly also affect the sense cells (Nijhout and Martin, 1978).

We believe ours is the first examination of resonant frequencies in an *Aedes* species known to expand its fibrils, and we show that the antennae of males with recumbent fibrils vibrate in response to sounds much like those with them erect. Female antennae are resonant much like males but their antennae do not have long fibrils and they do not change their resonant frequency before mating.

REFERENCES

Borkent, A. and Belton, P. (2006). Attraction of female *Uranotaenia lowii* (Diptera:Culicidae) to frog calls in Costa Rica. The Canadian Entomologist, 158, 91-94.

Cator, L.J., Arthur, B.J., Harrington, L.C., and Hoy, R.(2009). Harmonic convergence in the love songs of the Dengue vector mosquito. Science, 323, 1077-1079.

Clements A.N. (1999). The biology of mosquitoes Volume 2 Sensory reception and behaviour. CABI Publishing. New York, NY.

Fletcher, N.H. (1992). Acoustic Systems in Biology. Oxford University Press. Oxford

Gibson, G and Russell I.J. (2006). Flying in tune: Sexual recognition im mosquitoes.Current Bilogy 16, 1311-1316.

Göpfert, M.C. and Robert, D. (2001). Active auditory mechanics in mosquitoes. Proceedings: Biological Sciences, 268, 333-339.

Nijhout, H.F., and Martin, S.K. (1978). Alpha-adrenergic activity of isoproterenol in mosquito antennae. Experientia, 36, 758-759.

Nijhout, H.F.and Sheffield, H.G.(1979). Antennal hair erection in male mosquitoes: a new mechanical effector in insects. Science, 206, 595-596.

Pennetier, C., Warren, B., Roch-Dabire, K, Russell, I.J., and Gibson, G. (2010). ''Singing on the Wing'' as a mechanism for species recognition in the malarial mosquito *Anopheles gambiae*. Current Biology, 20, 131-136.

Warren, B, Gibson, G., and Russell I.J. (2009). Sex recognition through midflight mating duets in *Culex* mosquitoes is mediated by acoustic distortion. Current Biology, 19, 485-491.

ACKNOWLEDGEMENTS

Thanks to Ian Gordon and Pilar Cepeda for their help in rearing the mosquitoes and to Mike Hayden for discussing the physics of the system.

MODELLING THE EFFECT OF SHELL THICKNESS ON HIGH FREQUENCY ULTRASOUND SCATTERING FROM ULTRASOUND CONTRAST AGENTS

Omar Falou[1], Amin Jafari Sojahrood[2], J. Carl Kumaradas[2], Michael C. Kolios[2]
[1]Dept. of Electrical and Computer Engineering, Ryerson University, 350 Victoria St., ON, Canada, M5B 2K3
ofalou@ryerson.ca
[2]Dept. of Physics, Ryerson University, 350 Victoria St., Toronto, ON, Canada, M5B 2K3

1. INTRODUCTION

Ultrasound contrast agents (UCAs) are gas-filled, encapsulated bubbles that are administered intravenously to the venous system. They are very small (< 8 microns) which enables them to pass through capillaries. UCAs have a high degree of echogenicity compared to a cell, and therefore they enable contrast between the blood vessels and the surrounding tissue.

Few studies (Moran et al. 2002; Goertz et al. 2005; Ketterling et al. 2007) have looked at the dynamic response and backscatter of UCAs at high frequency ultrasound (HFUS). While some authors reported backscatter from UCAs with no harmonic components when insonicated at a transducer central frequency equal to the resonance frequency of the UCA (Moran et al. 2002), others (Goertz et al. 2005) observed the presence of harmonics in the backscatter of a lipid-shelled UCA, Definity (Bristol-Myers Squibb, North Billerica, MA), at low pressure (0.49 MPa). The results reported by Goertz et al. (2005) are unanticipated since high pressure is required to produce a non-linear response from UCA and generate harmonics in the backscatter. Ketterling et al. (2007) reported the presence of subharmonics in the response of a polymer-shelled UCA, Point (POINT Biomedical, San Carlos, CA), at high pressure (5.9 MPa). However, the experimental subharmonic response occurs at somewhat lower pressure amplitude than the theoretical response.

Analytical solutions to the problem of wave scattering from spherical objects such as ultrasound contrast agents have been studied extensively in the past (De Jong et al. 1993; Church 1995). These solutions are based on the Rayleigh-Plesset equation or variants and can only predict resonance frequencies at which the UCAs undergo radially symmetric oscillations. They cannot easily account for features such as asymmetric bubble oscillations and the interactions of bubbles with their surroundings.

Finite-element analysis (FEA) combined with other numerical techniques, such as the boundary element method, infinite elements, T-matrix method, etc. have been used in the past to model acoustic scattering from various objects submerged in a fluid (Hunt et al. 1975). The scatterer was typically modeled using finite elements while other techniques were used to find the solution in the surrounding medium. Most of these studies concentrated on scattering from rigid objects (Hunt et al. 1975).

In this work, we introduce a 2-D axi-symmetric finite element scattering model that allows asymmetric bubble oscillations and models the interaction of the bubble shell with the incident pressure wave using the constitutive stress-strain relationship coupled to the Helmholtz equation.

2. METHODS

The COMSOL Multiphysics package (COMSOL, Inc., Burlington, MA) was used to develop a 2D axi-symmetric finite element model to study scattering from contrast agents subject to high frequency ultrasound. The UCA was located in the centre of the computational domain with different acoustic properties than those of the surrounding fluid. The scatterer is insonified by a plane wave travelling is the $+z$ direction. Due to the symmetric nature of the problem, the 3-D model can be simplified by a 2-D axi-symmetric, with the z-axis being the axis of the symmetry. This simplification was used since 2-D models require less computational resource and execution times when compared to 3-D models (Falou et al. 2005). The Helmholtz equation was used to describe the propagation of sound waves in the UCA gas core and the surrounding fluid medium. The constitutive equation for the elastic material was used to describe the stress-strain relationship in the UCA shell.

Table 1. Physical properties of the BR14 ultrasound contrast agent (Dollet et al. 2008)

Property	Value
Shell density	1100 Kg/m^3
Shell's Young's modulus	177.6 MPa
Shell's shear modulus	60 MPa
Shell's Poisson's ratio	0.48
Perfluorocarbon density	11.21 Kg/m^3
Perfluorocarbon speed of sound	100 m/s

A perfluorocarbon phospholipid-coated contrast agent known as BR14 (Bracco Research SA, Geneva, Switzerland) surrounded by water was used to validate the finite element model by comparing the resonance frequency predicted by both the finite element model and the analytical Church formulation (Church 1995). This contrast agent was chosen it since it is widely used and has been the subject of research by many investigators (Goertz et al. 2003). Table 1 gives the physical the physical properties of BR14 used in the Church and finite element model. The surface tensions at both the shell-gas and the shell-liquid interfaces were assumed to be negligible. For simplicity, the finite element model does not take into account the viscosity of the shell and the surrounding medium. Initially, the far-field backscatter response of a 5 μm BR-14 having a 3 nm phospholipid shell thickness was considered. Then

scattering from 25 nm, 125 nm, and 250 nm shelled UCAs were studied at 1 – 70 MHz.

3. RESULTS

Table 2 shows a comparison between the resonance frequencies predicted by the Church and the finite element models for 3, 25, 125, and 250 *n*m shell UCAs. Fig 1 illustrates the effect of changing the shell thickness on the backscatter response from the UCAs.

Table 2. Comparison between the resonance frequencies predicted by the Church and the FEA models.

Shell thickness (nm)	Church Model (MHz)	FEA Model (MHz)	% error
3	2.3	2.2	4.4
25	5.6	5.4	3.6
125	12.4	12.1	2.4
250	18	17.6	2.2

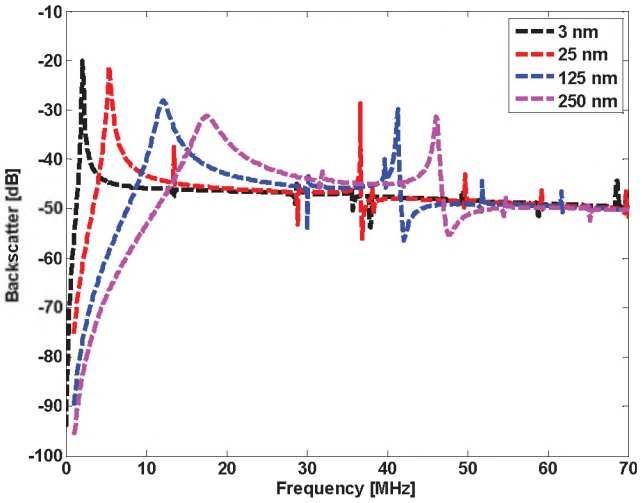


Fig. 1. Backscatter [dB] vs. Frequency [MHz]: effect of changing the shell thickness.

4. DISCUSSION AND CONCLUSIONS

The developed 2-D model has several advantages over conventional 3-D models we have previously developed for studying UCAs behaviour. It requires much less computational resources and execution times and can be used to calculate all quantities of interest, such as stresses and strains at the surface of the UCA, surface modes, etc.

A good agreement (error < 5%) was found between the finite element and analytical solutions (Church model) of the UCA resonance frequencies (the radially symmetric monopole resonance, the first peak in figure 1). Increasing the shell thickness increased the monopole resonance frequency (5.4 MHz, 12.1 MHz, and 17.6 MHz for the 25 nm, 125 nm, and 250 nm shelled UCAs, respectively) and broadened the resonant peaks. The finite element model revealed the presence of a second resonant peak of comparable magnitudes for the BR14 UCA. The frequency of the second peak also increased with shell thickness. Only one dominant resonant peak was found for the 3 nm shell thickness UCA within the studied frequency range. The presence of other resonant peaks in the backscatter from the 25 nm, 125 nm, and 250 shelled UCAs may be used to enhance the effectiveness of the ultrasonic imaging systems at high frequencies. It may also contribute to the generation of harmonics. This may provide an explanation to the presence of harmonics in the backscatter of the Definity UCA at low pressure (Goertz et al. 2005), where non-linear behaviour of the contrast agent is unlikely to occur. This study also shows that more careful design approaches may be taken to maximize the backscatter response from UCAs at high frequencies. This can be achieved by using the developed finite element model to optimize the UCA parameters in order to obtain the desired results. Future work includes the use of the developed model for the optimization of UCAs for high frequency ultrasound imaging.

REFERENCES

Church, C. C. (1995). "The Effects of an Elastic Solid-Surface Layer on the Radial Pulsations of Gas-Bubbles." Journal of the Acoustical Society of America **97**(3): 1510-1521.

De Jong, N. and L. Hoff (1993). "Ultrasound Scattering Properties of Albunex Microspheres." Ultrasonics **31**(3): 175-181.

Dollet, B., S. M. van der Meer, V. Garbin, N. de Jong, D. Lohse and M. Versluis (2008). "Nonspherical Oscillations of Ultrasound Contrast Agent Microbubbles." Ultrasound in Medicine and Biology **34**(9): 1465-1473.

Falou, O., J. C. Kumaradas and M. C. Kolios (2005). A Study of FEMLAB for Modeling High Frequency Ultrasound Scattering by Spherical Objects. FEMLAB User Conference, Boston, MA.

Goertz, D. E., E. Cherin, A. Needles, R. Karshafian, A. S. Brown, P. N. Burns and F. S. Foster (2005). "High frequency nonlinear B-scan imaging of microbubble contrast agents." Ieee Transactions on Ultrasonics Ferroelectrics and Frequency Control **52**(1): 65-79.

Goertz, D. E., M. Frijlink, A. Bouakaz, C. T. Chin, N. De Jong and A. W. F. Van Der Steen (2003). The effect of bubble size on nonlinear scattering from microbubbles at high frequencies. Proceedings of the IEEE Ultrasonics Symposium.

Hunt, J. T., M. R. Knittel, C. S. Nichols and D. Barach (1975). "Finite-Element Approach To Acoustic Scattering From Elastic Structures." Journal of the Acoustical Society of America **57**(2): 287-299.

Ketterling, J. A., J. Mamou, J. S. Allen, O. Aristizabal, R. G. Williamson and D. H. Turnbull (2007). "Excitation of polymer-shelled contrast agents with high-frequency ultrasound." Journal of the Acoustical Society of America **121**(1): EL48-EL53.

Moran, C. M., R. J. Watson, K. A. A. Fox and W. N. McDicken (2002). "In vitro acoustic characterisation of four intravenous ultrasonic contrast agents at 30 MHz." Ultrasound in Medicine and Biology **28**(6): 785-791.

ACKNOWLEDGMENTS

This project was funded by the Canadian Institute of Health Research (grant #79447) and Canada Research Chairs Program awarded to Michael Kolios. Omar Falou is the recipient of an Ontario Graduate Scholarship (OGS).

MULTIPLE SOURCE LOCALIZATION IN AN UNCERTAIN OCEAN ENVIRONMENT

Michael J. Wilmut and Stan E. Dosso
School of Earth and Ocean Sciences, University of Victoria, Victoria BC Canada V8W 3P6, mjwilmut@uvic.ca

1. OVERVIEW

This paper considers simultaneous localization of multiple ocean acoustic sources when properties of the environment (water column and seabed) are poorly known. A Bayesian focalization approach [1, 2] is developed in which the locations and complex strengths (amplitude and phase) of the sources together with uncertain environmental properties and noise variance are all considered random variables comprising the model of unknown parameters. The posterior probability density (PPD) for the model combines information from measured acoustic, formulated in terms of the likelihood function, and by prior information (typically parameter search bounds). The goal then is to maximize the PPD over all parameters to extract the most probable set of source locations.

PPD maximization can be carried out analytically for the source strength and variance parameters by setting partial derivatives of the likelihood to zero. This leads to a linear system of complex equations which is even- or over-determined provided the number of data is greater than or equal to twice the number of sources, and hence is amenable to standard least-squares solution. Maximizing the PPD over the environmental parameters cannot be performed analytically, and is carried out here using a numerical optimization algorithm, adaptive simplex simulated annealing [3], with the analytic solution for source strengths and variance applied for each model realization considered in the optimization process.

2. EXAMPLES

The multiple-source localization procedure outlined above is demonstrated with a synthetic example illustrated in Fig. 1. The 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-column sound speed profile is represented by four unknown sound speeds c_1-c_4 at depths of 0, 10, 50, and D m, where D is the water depth. All of these environmental parameters are considered unknown with prior information consisting of uniform distributions over wide bounds (true parameter values and prior bounds are given in Table 1). Two acoustic sources are present, one at 7-km range and 4-m depth (referred to as the shallow source), and the other at 5.4-km range and 50-m depth (the deep source). Acoustic fields from these two sources are computed at a frequency of 300 Hz at a 24-sensor vertical line array (VLA) using a normal mode propagation model.

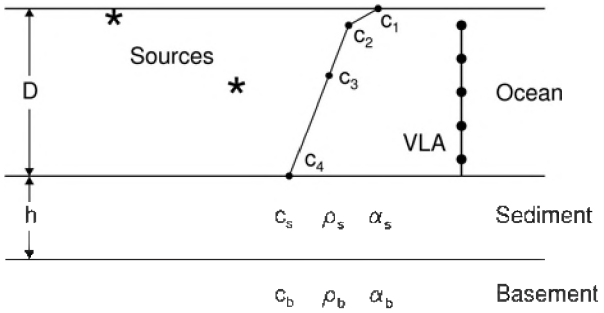


Figure 1. Schematic diagram of the geometry of the two-source localization problem, indicating unknown environmental parameters.

For the study carried out here, acoustic data are considered at five signal-to-noise ratios (SNRs) for the deep source. The SNR of the shallow source is either the same as that of the deep source (Fig. 2) or 12 dB higher than that of the deep source (Fig. 3). For each SNR combination, 50 different noisy data sets were generated and inverted for source locations. To demonstrate the advantage of optimizing over uncertain environmental parameters, focalization results are compared to two cases of localization for fixed environments where the environmental

Parameter & Units	Value	Prior Bounds
SSP:		
D (m)	130	[128, 135]
c_1 (m/s) @ 0 m	1520	[1515, 1525]
c_2 (m/s) @ 10 m	1517	[1510, 1520]
c_3 (m/s) @ 50 m	1515	[1510, 1520]
c_4 (m/s) @ 130 m	1510	[1505, 1515]
Seabed:		
h (m)	9.0	[0, 30]
c_s (m/s)	1494	[1450, 1600]
c_b (m/s)	1529	[1500, 1650]
ρ_s (g/cm^3)	1.38	[1.0, 1.7]
ρ_b (g/cm^3)	1.52	[1.5, 2.2]
α_s (dB/λ)	0.02	[0, 1]
α_b (dB/λ)	0.22	[0, 1]

Table 1. True values and uniform prior bound widths for environmental parameters of the synthetic test case.

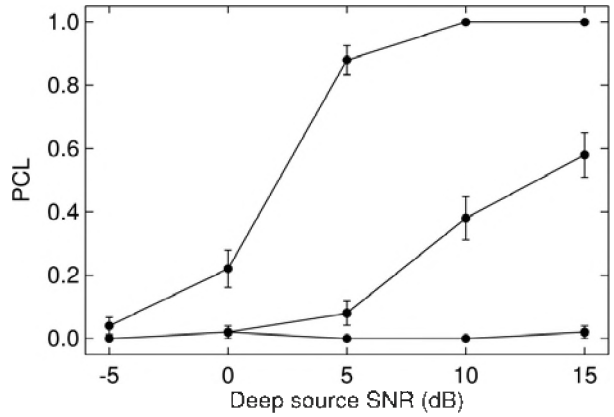


Figure 2. Probability of correct localization for random environmental realizations (lower curve), optimization over wide environmental bounds (middle curve), and exact environment (upper curve) as a function of SNR for the deep source. The shallow source has the same SNR as the deep source in all cases.

parameters correspond to either a random realization from the prior distribution (representing the actual environmental uncertainty) or to the true parameters (i.e., perfect environmental knowledge). The results are quantified in terms of the probability of correct localization (PCL), which is defined as the fraction of localizations achieving mean absolute depth and range errors of less than 10 m and 300 m, respectively, for both sources. One standard deviation binomial uncertainties are indicated as error bars for the PCL values in Figs. 2 and 3.

Figures 2 and 3 show the level of environmental uncertainty assumed in this example essentially precludes localizing the two sources using standard methods, with near-zero PCL values for random environmental realizations at all SNRs.

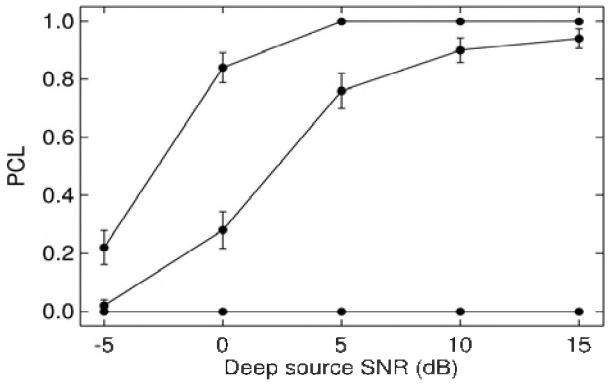


Figure 3. Probability of correct localization for random environmental realizations (lower curve), optimization over wide environmental bounds (middle curve), and exact environment (upper curve) as a function of SNR for the deep source. SNR of shallow source is 12 dB higher than that of the deep source in all cases.

However, optimizing over the unknown environment via focalization provides much better localization results for all but the lowest SNRs, with PCL values approaching those for the true environment for the higher SNR cases in Fig. 3 (localization results are generally better in Fig. 3 than Fig. 2 because of the higher SNR for the shallow source).

Finally, it is interesting to consider how the presence of the shallow source affects the ability to localize the deep source, given focalization over the unknown environment. For instance, to achieve PCL = 0.5 for two source with equal SNR, Fig. 3 shows an SNR of approximately 12.5 dB is required. When the SNR for the shallow source is 12 dB higher than for the deep source, Fig. 3 shows the deep-source SNR required for PCL = 0.5 is about 2 dB. For comparison, single-source localizations with only a deep source present (not shown) required an SNR of 1 dB for PCL = 0.5.

3. SUMMARY

This paper developed a Bayesian focalization approach to multiple source localization in an uncertain environment that made use of analytic solutions for the amplitude and phase of the unknown sources. Synthetic examples considered indicated a substantial improvement in probability of correct localization over localizations with fixed (incorrect) environmental parameters.

REFERENCES

[1] Collins, M. D., and W. A. Kuperman (1991). Focalization: Environmental focusing and source localization, J. Acoust. Soc. Am., vol 90, 1410-1422.

[2] Dosso, S. E. and M. J. Wilmut (2002). Comparison of focalization and marginalization for Bayesian tracking in an uncertain ocean environment, J. Acoust. Soc. Am., vol 125, 717-722.

[3] Dosso, S. E., M. J. Wilmut and A. L. Lapinski (2001). An adaptive hybrid algorithm for geoacoustic inversion, IEEE. J. Ocean. Eng., vol 26, 324-336.

ESTIMATING GEOACOUSTIC PARAMETERS OF GASSY SEDIMENT USING LOW-FREQUENCY SOUND IN ST. MARGARET'S BAY, NOVA SCOTIA.

Marie-Noel R. Matthews[1]
[1]JASCO Applied Sciences, suite 432 - 1496 Lower Water St., Halifax, Nova Scotia, B3J 1R9
marie-noel.r.matthews@jasco.com

1. INTRODUCTION

The geoacoustic properties of gassy sediment not only differ greatly from that of gas-free sediment, but present a strong frequency dependence. These properties are normally measured *in situ* using acoustic sensors embedded in the seabed, or in laboratory by examining pressurized core samples through CT scans, for example (Anderson *et al.*, 1989; Tuffin, 2001). The first objective of this study was to assess the possible use of a Bayesian inversion algorithm to remotely (and cheaply) estimate the geoacoustic parameters of a gassy seabed at frequencies below 500 Hz. The effectiveness of the studied algorithm has already been proven for various environments including shallow-water, multi-layered seabeds. However, its use for gassy sediment is complicated by the frequency-dependence of the geoacoustic parameters and its effectiveness needed to be assessed in this particular environment. This assessment was done through simulations, focusing on environments similar to the central basin of St. Margaret's Bay. The results from these simulations provided basic guidelines describing some limiting factors for which the inversion of gassy geoacoustic parameters may be possible. The main objective of this study was to evaluate the geoacoustic profile of the deep central basin of St. Margaret's Bay, using the inversion algorithm and field data.

2. METHOD

In May 2006, Defence R&D Canada conducted sea trials using a vertical line array of hydrophones in the deep central basin of St. Margaret's Bay, Nova Scotia, Canada. The array, deployed in 63 m of water, comprised 11 hydrophones unevenly spaced in the bottom two thirds of the water column. During one experiment, an acoustic source emitting a signal composed of five tonals below 500 Hz was towed at approximately 26-m depth and horizontal distances from the array of less than 1000 m. The field data used in this project were composed of the acoustic pressure fields recorded by the 11 hydrophones.

The inversion algorithm used in this study is the Adaptive Simplex Simulated Annealing (ASSA) algorithm, developed by *Dosso and Wilmut* (2000). It encompasses a hybrid optimization technique, which was proven very effective in estimating geoacoustic parameters of acoustically complex shallow-water environments (*Dosso*, 2000; *Dosso and Wilmut*, 2000; *Gillard et al.*, 2003). The uncertainties associated with the geoacoustics inversion solution were then estimated using the Fast Gibbs Sampler (FGS) algorithm developed by *Dosso* (2002). The two algorithms were paired with a parabolic equation (PE) propagation model called PE04, and the *implicit* energy function, E_{imp}, based on the Bartlett correlator, $B_f(\mathbf{m})$ (Bartlett, 1947; Tolstoy, 1993):

$$E_{imp}(\mathbf{m}) = N_D \sum_{f=1}^{F} \ln[B_f(\mathbf{m}) \,|\, \mathbf{D}_f^{obs} \,|^2], \tag{1}$$

where, $$B_f(\mathbf{m}) = 1 - \frac{\left|\mathbf{p}_f^{\dagger}(\mathbf{m})\mathbf{D}_f^{obs}\right|^2}{\left|\mathbf{p}_f^{\dagger}(\mathbf{m})\right|^2 \left|\mathbf{D}_f^{obs}\right|^2}. \tag{2}$$

$\mathbf{m}$ is the series of parameters to be inverted, $\mathbf{p}_f(\mathbf{m})$ is the predicted sound field, $\mathbf{D}_f^{obs}(\mathbf{m})$ is the measured field, F is the number of frequencies, N_D is the number of data point, and † represents the conjugate transpose of a matrix.

3. RESULTS

Through the use of simulated data, the effectiveness of the algorithm for a gassy sediment layer was assessed by comparing results in four types of environment: a gassy and a gas-free version of a 2-layer and a 3-layer geoacoustic model. Then, field data were inverted, assuming a 2-layer and a 3-layer environment, in order to estimate the thickness, density, compressional-wave speed and attenuation of each sediment layer found in St. Margaret's Bay.

3.1 Inversion of simulated data

The inversion of the simulated data sets showed that, in a gassy environment, the number of unknown parameters in the inversion problem must be reduced to a minimum in order to achieve convergence of the parameter estimates. Following this limitation, the algorithm distinguished between the gas-free and gassy environments by reproducing the low sound velocity and high attenuation that characterizes the gassy layer. Results also showed that the low sound velocity in gassy sediment causes an important increase in the uncertainty level of the estimated density in the layer. Consequently, in both cases, *i.e.* for the 2-layer and the 3-layer gassy environments, the density of the gassy sediment was irresolvable.

3.2 Inversion of field data

Following the limitations in the number of resolvable parameter, the unknown parameter set was reduced to include the thickness of each layer and the geoacoustic properties of the gassy layer (*i.e.*, the top layer in the 2-layer model; the second layer in the 3-layer model).

Results indicated that the 3-layer geoacoustic profile better simulates the deep central basin seabed than the 2-layer profile. The presence of a thin layer of highly porous clay over the gassy sediment allowed the algorithm to better estimate the low sound velocity and high attenuation expected in the gassy layer. The technique estimated the sound velocity to be up to 1.75 times lower than that of gas-free LaHave Clay, and the attenuation to be up to 350 times that in gas-free Lahave Clay. These values are consistent with the effect predicted by the Anderson and Hampton theory (Anderson and Hampton, 1980a, b).

In addition, the estimated average thickness of this highly porous layer corresponds to the depth at which *in situ* measurements indicated an important decrease in sound velocity (Kepkay, 1977). However, the density and the thickness of the second and third layers were undetermined. It is believed that the use of more accurate values for the density, sound velocity and attenuation in the top and third sediment layers would allow the technique to better determine all parameter values.

4. DISCUSSION

Although many geological parameter describing St. Margaret's Bay sediment properties were unknown, an attempt was made in comparing the estimated sound velocity and attenuation *vs.* frequency to the *Anderson and Hampton* (1980a; 1980b) theory on the acoustic of gassy sediment. This comparison led to an estimated bubble frequency of resonance less than 117 Hz, and an estimated bubble radius between 15 and 25 cm – compared to an averaged maximum observed radius of 10 mm (Tuffin, 2001). Preliminary investigation showed that the effects from bubble size distribution and bubble asphericity does not fully account for this incredibly large bubble size. However, it is possible that as for accumulation of gas bubbles in the water column, at low frequencies (less than 1 kHz), the acoustic effect on gassy sediment may be driven by the resonance of bubble clouds or plumes, instead of the resonance from individual bubbles, via a process in which bubbles pulsate in a collective mode of oscillation (Carey and Browning, 1988; Prosperetti, 1988).

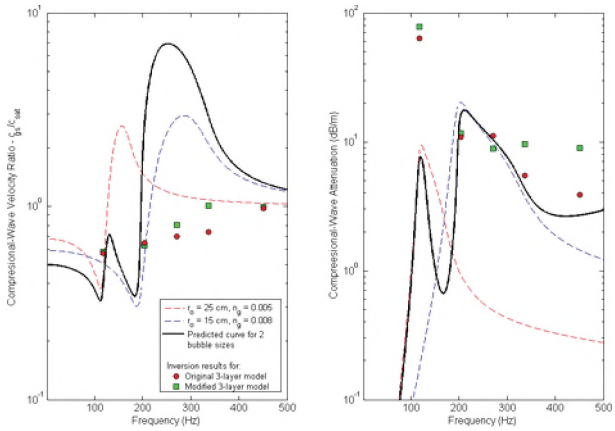


Fig. 1. Example of comparison of theoretical curves and inversion results for the sound velocity and attenuation of gassy sediment in St. Margaret's Bay, NS.

REFERENCES

Anderson, A.L., and D.L. Hampton. (1980a). Acoustics of gas-bearing sediments. I. Background. *JASA*., 67, 1865-1889.

Anderson, A.L., and D.L. Hampton. (1980b). Acoustics of gas-bearing sediments. II. Measurements and models. *JASA*, 67, 1890-1902.

Anderson, A.L., F. Abegg, J.A. Hawkins, M.E. Duncan, and A.P. Lyons (1989). Bubble populations and acoustic interaction with the gassy floor of Eckernförde Bay. *Cont.Shelf Res.*, 18, 1807-1838.

Bartlett, M.S. (1947) Multivariate analysis. *Supplement to J. of the Royal Statistical Society*, 9, 176-197.

Carey, W.M. and D.G. Browning (1988). Low-frequency ocean ambient noise. In *Sea Surface Sound*, Kluwer Acad. Pub., Boston, 361-376.

Dosso, S.E. (2000). Geoacoustic inversion and appraisal. In *Proceedings of Ocean 2000 IEEE*, Providence, RI, 1, 423-428.

Dosso, S.E. (2002). Quantifying uncertainty in geoacoustic inversion. I. A Fast Gibbs sample approach. JASA, 111, 129-142.

Dosso, S.E., and M.J. Wilmut (2000). An adaptive hybrid algorithm for geoacoustic inversion. In *Proceedings of the 5th ECUA*, Lyon, France, 185-190.

Gillard, C.A., D.J. Thomson, and G.J. Heard (2003). Estimating geoacoustic parameters using Marched-Field inversion methods. *EIII JOE*, 28, 394-413.

Matthews, M.N.R., (2009) Estimating geoacoustic parameter of gassy sediment using low-frequency sound in St. Margaret's Bay, Nova Scotia. MSc thesis. Dalhousie Univ., Halifax, NS, Canada.

Prosperetti, A. (1988). Bubble related ambient noise in the ocean. *JASA*. 84(3), 1042-1054.

Tolstoy, A. (1993). *Matched field processing for underwater acoustics*. World Scientific Publishing Co. Pte. Ltd.

Tuffin, M.D.J. (2001) The geoacoustic properties of shallow, gas-bearing sediments. PhD Thesis, University of Southampton, UK.

ACKNOWLEDGEMENTS

The author thanks Defence R&D Canada for providing the field data.

AUTHOR NOTES

This work was conducted as partial fulfillment of the requirements for a degree of Master of Science at Dalhousie University, Halifax, NS, December 2009.

DETERMINING ACOUSTIC SCATTERING PROPERTIES OF MARINE SEDIMENTS THROUGH BAYESIAN INVERSION

Gavin A.M.W. Steininger[1], Stan E. Dosso[1], and Charles W Holland [2]
[1]School of Earth and Ocean Science, University of Victoria, British Columbia, Canada, V8P 5C2
[2]Applied Research Laboratory, Pennsylvania State University, Pennsylvania, USA, 16802

1. INTRODUCTION

Reverberation modeling and sonar performance prediction in shallow waters require good estimates of seabed reflection and scattering properties as well as an understanding of scattering processes in a particular region. This paper describes non-linear Bayesian inversion of synthetic ocean acoustic seabed scattering data for marine sediment parameters, assuming first-order perturbation theory for acoustic scattering. The objective is to determine the necessary angular range to adequately determine the model parameters

2. METHOD

There are three elements in an inversion; these are the data, the forward model (physics), and the inversion scheme. This section gives a brief description of all three of these elements.

2.1 Inversion scheme

Bayesian inversion is based on formulating the posterior probability density (PPD) of the model parameters of interest, which is the product of likelihood and prior information or distribution of the parameters [1,2]. The PPD contains all available information for the parameters; however it can be difficult to interpret directly in an analytic manner. Thus the PPD is approximated using Markov-chain Monte Carlo sampling algorithms. This approximate PPD is then interpreted in terms of its moments, parameter uncertainties (variances, marginal distributions, credibility intervals), and parameter inter-relationships (correlations and joint marginal's). The Bayesian formulation also provides information measures which quantify the evidence provided by observed data to support a particular choice of model parameterization and/or forward (modeling) theory, favoring the simplest choice consistent with the resolving power of the data (Bayesian form of Occam's razor).

2.2 Forward model

In the present application the forward model used describes the backscatter emitted from an insonified rough boundary between two otherwise homogenous half spaces. The configuration is shown in Fig. 1.

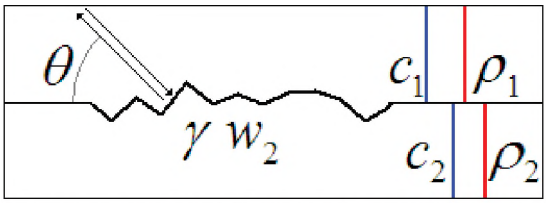


Fig. 1: A schematic of the forward model for acoustic backscatter from a rough interface.

First-order perturbation theory is used to generate and invert the synthetic data [3]. The scattering kernel used (ξ) is given in Equation 1, where k_1 is the wave number of the first medium, and θ is the grazing angle for both thin incident and scatter rays. The scattering exponent is γ, and the spectral strength is w_2.

$$\xi = \frac{4k_1^4 \sin^4(\theta)|R(\theta)|^2 w_2}{K(\theta)^\gamma} \quad (1)$$

The functions R and K are given in Equations 2 and 3, where ρ is the ratio of densities of the second to the first media, and κ is the ratio of wave numbers.

$$R(\theta) = \frac{(\rho-1)^2\cos^2(\theta) + \rho^2 - \kappa^2}{\left(\rho\sin(\theta) + \sqrt{\kappa^2 - \cos(\theta)^2}\right)^2} \quad (2)$$

$$K(\theta) = \sqrt{4k_1^2\cos^2(\theta) + \left(\frac{k_1}{10}\right)^2} \quad (3)$$

2.2 Synthetic data

Nine synthetic data sets were created using Equation 4. This is the decibel representation of Equation 1. A calibration bias term β is also decibel added. This will allow for the inversion of bias data when working with non synthetic data.

$$d_i = 10\log_{10}(\xi_i) + \beta + \varepsilon_i \quad (4)$$

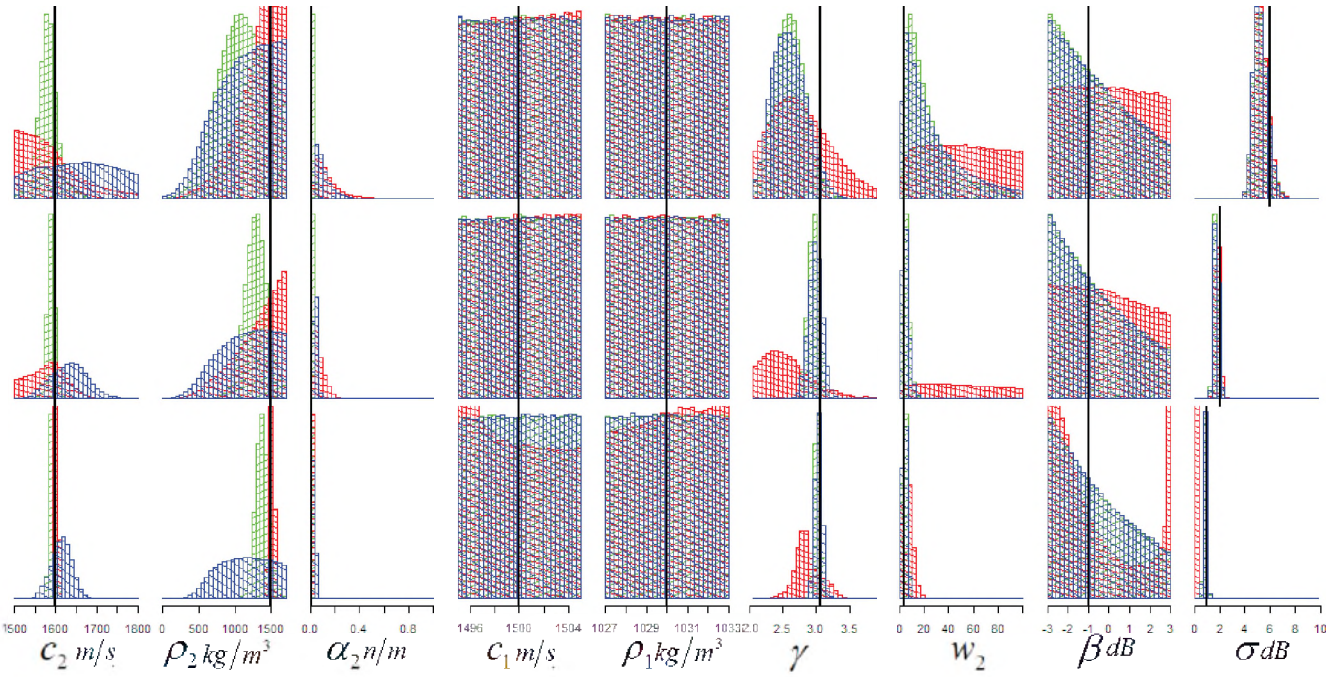


Fig. 2: Marginal distributions for all nine parameters sampled from the PPDs of the data sets. The three rows from top to bottom correspond to data error standard deviations (σ) of 6, 2 and 1dB respectively. The true value is displayed as a black line. for θϵ[1,21], for θϵ [22,89], and for θϵ [1,89].

Each data set contains 45 data points (d_1, d_2, …, d_{45}). These data sets differ in that they consider different ranges of θ and different standard deviations for the Gaussian error terms ε. The ranges of the data are 1-89°, 1-21°, and 22-89°. The standard errors (σ) of the ε s are 6, 2, or 1 dB depending on the data set.

The true values of the nine parameters are $c_2 = 1600$, $\rho_2 = 1500$, $\alpha_2 = 0.02$, $c_1 = 1500$, $\rho_1 = 1030$, $\gamma = 3.06$, $w_2 = 3.6$, $\beta = -1$, and $\sigma = 1,2 \vee 3$.

3. RESULTS

The one-dimensional marginal PPDs for the parameters are shown in Fig. 2. Each row contains nine histograms; one for each of the parameters of interest. The rows are sorted in descending order according to the different values of σ, 6, 2 and 1 dB. Each histogram plot displays three marginal PPDs for the given parameter. They are distinct in that they are estimated from different data sets. The data sets were created using the same true values for the parameters, but with different angular ranges. Thus a total of 81 marginal distributions are presented. The true value of the parameters is displayed as the vertical black line across each histogram.

4. DISCUSSION

As expected, the PPDs have smaller variances (are narrower) as the data errors are reduced. The data sets that consider only θ below the critical angle (21°) are not adequate to estimate the scattering parameters (γ, w_2) because the marginal distributions are too wide; that is, they contain little information. As surface scattering will not dominate above this angle, to estimate these parameters requires more information be added to the inversion. For example, reflection data could be added and a joint inversion performed. An alternative approach would be to consider a more complex scattering kernel that accounts for sub-bottom or volume scattering.

REFERENCES

[1] S. E. Dosso and P. L. Nielsen (2009) "Quantifying uncertainties in geoacoustic inversion II: Application to a broadband shallow-water experiment," J. Acoust. Soc. Am., 111, 143-159.

[2] Dosso, S. E. and C. W. Holland (2006) "Geoacoustic uncertainties from visco-elastic inversion of seabed reflection data," IEEE J. Oceanic Eng., 31, 657-670.

[3] J. Moe and D. Jackson (1994) "First-order perturbation solution for rough surface scattering cross section including the effects of gradients," J. Acoust. Soc. Am., 96, 1748-1754.

MULTIFREQUENCY CLASSIFICATION AND CHARACTERIZATION OF SINGLE BEAM ECHOUSOUNDER DATA COLLECTED FROM OFFSHORE IRELAND

Stephen F. Bloomer[1], Xavier Monteys[2], Ross Chapman[1], and Xavier Garcia[3]
[1]Canadian Marine Acoustic Remote Sensing Facility, University of Victoria, Victoria, BC Canada, V8V 3P1
[2]Geological Survey of Ireland, Dublin, Ireland
[3]Barcelona Center for Subsurface Imaging, Unitat de Tecnologia Marina, CSIC, Barcelona, Spain

1. INTRODUCTION

Since 2000, the Irish National Seabed Survey and INFOMAR (INtegrated mapping FOr the sustainable development of Ireland's MArine Resource) have mapped the area from the outer margins of Ireland's territorial seabed to the shelf as well as nearshore areas. Using several research vessels equipped with a variety of instruments, ocean, seabed and sub-seabed properties have been measured.

During this time, INFOMAR has produced regional seabed classification charts using multibeam backscatter data. Multibeam backscatter returns at high frequencies are generally a good indicator of surface seabed properties, but are not suitable for the subsurface discrimination because of minimal sound penetration. Furthermore, multibeam backscatter returns below nadir are not generally a good seabed discriminator, leaving unclassified gaps in most of these charts. Single beam echoes are precisely imaging most of this region, making the two sonar systems complementary in an integrated approach to seabed characterization.

INFOMAR and the Canadian Marine Acoustic Remote Sensing (C-MARS) Facility partnered in 2009 in a pilot project to examine the utility of using both multibeam (MBES) and single beam echosounder (SBES) data in tandem for improved seafloor and shallow seabed classification and characterization. This paper describes the process developed at C-MARS used to analyze SBES data. The results of one case study are shown to demonstrate the potential for seafloor classification and characterization from the INFOMAR SBES data.

2. DATA

INFOMAR uses a range of vessels for geophysical surveys, governed by the water depth of the area to be surveyed. For surveying in deeper waters, the RV Celtic Explorer is equipped with a Simrad EM1002 (95 kHz) multibeam sonar, a SES Probe 5000 subbottom (3.5 kHz) profiler, and a three frequency Simrad EA600 (12, 38, and 200 kHz) hydrographic echosounder. In shallow waters the RV Celtic Voyager is similarly equipped but uses a dual frequency Simrad EA400 (38 and 200 kHz) hydrographic echosounder and has an additional Simrad EM3002 (300 kHz) multibeam sonar.

In addition, at various study sites, grab samples, vibrocores and video footage have been collected, often guided by the seabed classification map derived from the multibeam data. As well, at a few selected sites, a towed Electromagnetic (EM) system was used to measure the electrical conductivity distribution of around 20m below the seafloor.

3. DATA PROCESSING

At C-MARS, research on improved acoustic seafloor classification has focused on the extraction of attributes from SBES echoes. These attributes, in combination with pre-existing features found in QTC IMPACT™, form a feature set that now includes physically based features such as amplitude. As well, integrated moments of the echo amplitudes, modified from van Walree et al. (2005), were calculated. These new features provide additional information useful for classification and potentially characterization.

The following steps were used to extract features from the SBES echoes:

1. The data were loaded into QTC IMPACT™ to convert the raw Simrad data expressed in decibels to a linear scale. 2. Automatic bottom picking was performed, resulting in less than 2% bad picks, and traces with any further bad bottom picks were eliminated.
3. The echo amplitudes were depth compensated for seawater absorption and imprecise spreading law using a process similar to the depth compensation applied to multibeam backscatter values in QTC MULTIVIEW™ (Preston, 2009).
4. After this gain correction, stacks of 5 echoes were assembled to improve the signal to noise ratio.
5. The data were trimmed to n standard echo lengths (SEL's) before the pick and m SEL's after prior to feature analysis. The standard echo length is defined as the expected duration of an echo from a sand echo for an echosounder with a given frequency and beamwidth. In all cases, n and m were chosen to prevent truncating echoes from soft sediments in the trimming process (for instance, n=0.5 and m=7.5 for the Malin Sea).
6. The trimmed stacked data were processed with the feature algorithms shown in Table 1. The existing features are primarily related to the shape of the echo. Some of the new features are related to seafloor

properties, such as the seafloor acoustic impedance contrast (amplitude) and seafloor roughness (amplitude

Existing Features in QTC IMPACT™
Cumulative Histogram
Histogram
Quantile
FFT
New Features
Modified van Walree moments
Quantile Cumulative Histogram
Amplitude
Amplitude Variability
Fractal

Table 1. Feature families used in analysis of SBES data in the pilot study.

variability).

4. RESULTS

One result of this pilot project is the potential to use SBES features for seafloor characterization. One example of many comes from the analysis of data from a pockmark field in the Malin Sea off the northern coast of Ireland. Figure 1 shows comparisons between the modified van Walree time spread feature derived from 12 kHz SBES data and EM conductivity. From west to east, three regions can be identified.

In region B1, an increase in conductivity is possibly associated with decreases in grain size or compaction, while the time spread feature also increases, consistent with softer or finer sediments and longer echoes. Region B2 is a pockmarked area (as shown in the subbottom profile) with gassy sediments. The EM signature is irregular, with porosity drops and general conductivity values below expected, possibly due to gas related sediment facies.

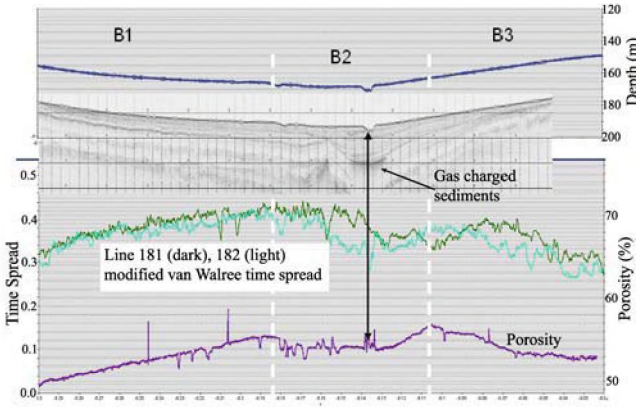


Figure 1. Line 181 and 182 time spread feature (dark and light) correlates well with EM porosity (bottom) in regions B1 and B2, and drops in regions corresponding to gas pockmarks, as shown in the depth and subbottom profiles.

Correspondingly, drops in time spread could be influenced by increased gassy sediment attenuation or gas blanking of the echo return. In region B3, conductivity decreases in a similar way but not reaching the low values of B1. The time spread values show a general negative trend but with more variability and overall lower values than B1, perhaps with coarsening grain size. As with B1, there is fairly good correlation between time spread versus conductivity.

REFERENCES

Preston, J.M. (2009). Automatic acoustic seabed classification of multibeam images of Stanton Banks. Applied Acoustics, 70 (10) 1277-1287.

P. A. van Walree, J. Tegowski, C. Laban, and D. G. Simons (2005). Acoustic seafloor discrimination with echo shape parameters: A comparison with the ground truth. Continental Shelf Research 25, 2273–2293.

ACKNOWLEDGEMENTS

The authors would like to acknowledge funding for this project from an INFOMAR research call.

THE ROLE OF ECHO DURATION IN ACOUSTIC SEABED CLASSIFICATION AND CHARACTERZATION

Ben Biffard[1], Steve Bloomer[1], Ross Chapman[1] and Jon Preston[2]
[1]Canadian Marine Acoustic Remote Sensing Facility (C-MARS), School of Earth and Ocean Science, University of Victoria, Victoria, BC Canada, V8V 3P1 bbiffard@uvic.ca
[2]Quester Tangent Corp., 6582 Bryn Road, Saanichton, BC Canada, V8M 1X6

1. INTRODUCTION

In this paper, the results of two experiments primarily concerning echo duration are summarized. The aim of the experiments is to improve seabed remote sensing using single-beam echosounders by:

- verifying the ray-trace model of echo duration that is used as the basis to remove the affects of depth from the data prior to analysis for classification
- applying a new technique to remove the affects of seabed slope from the seabed echo data
- demonstrating the application of echo duration data for seabed characterization

Current research effort at C-MARS is to develop new and improved seabed classification methods based on the Quester Tangent Corp. method applied in the QTC IMPACT™ software. Seabed classification segments the seabed into areas of similar acoustic character based on statistics. Further, direct seabed characterization is to be explored. To do this, the physics of echosounding are paramount. In general, echoes from the seabed are composed of three types of acoustic response: reflection, volume backscattering and surficial backscattering. The shape of the echo is largely determined by the total backscatter strength as a function of angle of incidence that increases from nadir as the transmit pulse spreads along the seabed. The echo time series from mud, sand, gravel, and rock seabeds differ in echo attributes that may be called shape, overall duration and relative amplitude. The QTC IMPACT seabed classification method is primarily based on measures that respond to the echo shape.

Deeper echoes are stretched out in time due a slower rate of increase of the angle of incidence relative to echoes from shallower water. This depth dependence has to be compensated so that the classification maps are of seabed types, not maps of bathymetry. All non-seabed influences on the echo must also be held constant or compensated for; this topic is explored further in Biffard et. al., 2007. The biggest source of error in single-beam based seabed classification is seabed slope. It acts in a way very similar to depth – it stretches echoes out in time by making the angle of incidence progress slower. It also changes the angle of incidence as well; at slopes greater than the ½ beam width the seabed normal ray is lost severely reducing the coherent reflection from the seabed which changes the echo shape.

The duration of a seabed echo can be modeled with some simple ray tracing, leading to the following:

$$E_d = \frac{2}{c} d \sec\left(\phi + \frac{\theta}{2}\right) + \tau + \frac{2}{c} p + \frac{2}{c} w \quad (1)$$

where d is the water depth, c the sound speed, Φ is the seabed slope, θ is the beam width, τ the duration of the transmit pulse, p the depth of penetration in to the seabed defined by $10/kf$ (k is the seabed attenuation coefficient in dB/m/kHz, f is the echosounder frequency), w is the height of macro-roughness such as large rocks. This equation is valid for $\phi \le \theta/2$. The expression used for the penetration depth is based on 10 dB extinction and does not consider the reflection coefficient or sediment sound speed. Eq. 1 implies plots of measured echo duration versus depth should be linear. The depth compensation algorithm, named *SEL*, used in the QTC View series 5 system is based on Eq. 1 assuming nominal values for beamwidth and attenuation.

2. METHOD

The experiments required varying the water depth and seabed slope while keeping all other variables fixed. To do this, both depth and slope are simulated in the field by raising and lowering the transducer and by tilting the transducer. This was done at three controlled sites – the sites are flat, homogenous and well characterized by video, grab sample and penetrometer data. These sites are part of our Patricia Bay, BC testbed as detailed in Biffard et. al., 2006.

3. THE DEPTH EXPERIMENT

Fig. 1 shows plots of echo duration versus depth. Echo durations were measured using a threshold-based bottom pick and cumulative amplitude 'tail' pick. Linear fits were all are significant, explaining a majority of the variation, and therefore confirming our linear model. The fit produces two parameters: effective beamwidth (slope of the regression line) and effective attenuation (part of the intercept). The effective attenuation values are consistent to those found using the grab samples: 0.6 for sand, 0.4 for gravel, 0.05 for mud. The effective beamwidths were wider than the manufacturer's -3 dB beamwidths, but are consistent between the sands and muds. Beamwidths from the gravels are wider because of off-axis surfical backscatter due to high seabed roughness. Depth compensation works by resampling the data to fit within an analysis window. The effective values tune the depth compensation for better

results, ie the echoes fit in the windows and the variation in echo duration caused by depth is removed.

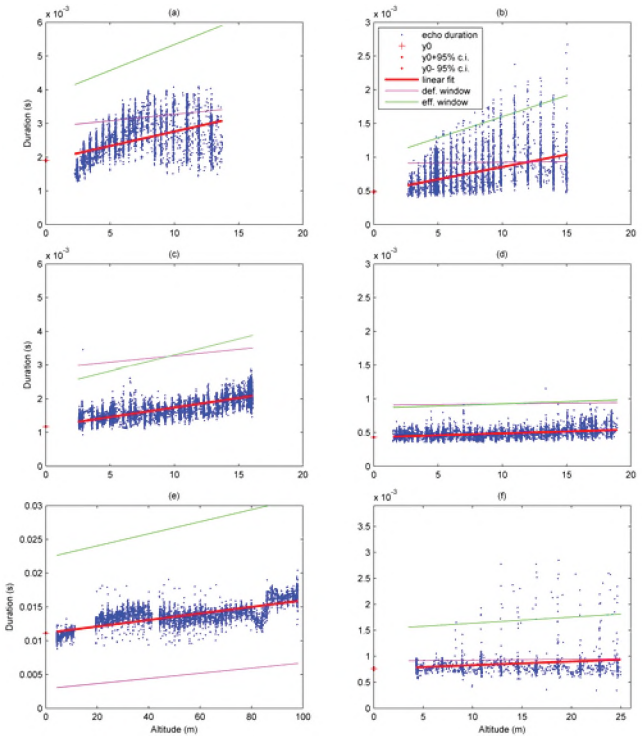


Figure 1. Measured echo duration plotted against the bottom pick of each echo, with linear regressions to 'depth'. The sediments are: gravel site at (a) 24 kHz, (b) 200 kHz; sand site at (c) 24 kHz, (d) 200 kHz; and mud site at (e) 24 kHz, (f) 200 kHz. The red line and dots on the vertical axis indicate the y-intercept and its error. The thick red lines are the regression lines. The magenta lines are the default analysis window, the green lines the analysis calculated from the effective beamwidth and attenuation. The axes ratio is constant for each frequency.

4. THE SLOPE EXPERIMENT

Fig. 2 shows three new statistical measures of echo shape, called 'features', used for seabed classification (Bloomer et. al., 2010). Transducer tilt increases left to right simulating increasing seabed slope. In Fig. 2 (*right*) seabed slope is compensated for by inserting the tilt value into Eq.1 which is then used as if doing depth compensation as before. The result for the gravel site is dramatic. Virtually all of the slope effect is removed. However, this is only effective up to six to eight degrees for the sand and mud sites. This is because the high-roughness surficial backscattering that dominates the gravel site echoes is unaffected by slope, while processes that dominate the sand and mud site echoes are affected. In particular, the specular reflection is lost for slope greater than the ½ beamwidth of 10 degrees. Also, measured echo durations increase with 'slope' as expected.

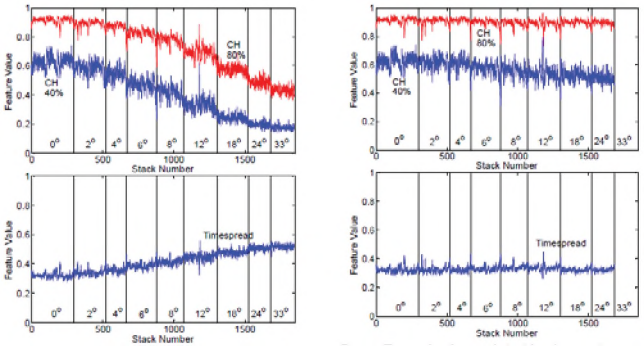


Figure 2. (*top left*) Two cumulative histogram features: 40% (blue) and 80% (red) versus increasing seabed slope (in steps). (*lower left*) The timespread feature. Both panes on the *left* were generated with normal depth compensation, while the panes on the *right* included the slope (ie tilt of the transducer) in Eq. 1 for combined depth and slope compensation. The vertical lines indicate the end of each pass of the 24 kHz echosounder over the gravel site after which the tilt was increased in increments of 2, up to 6, degrees.

5. DISCUSSION

To our knowledge, echo duration data like these have never been published. Perhaps previous attempts failed because of a lack of a controlled testbed site. The setup for the depth experiment included 4-point anchoring and control of the transducer's attitude. Tail picking is also very difficult. The compensation of seabed slope using the echo duration model is a major improvement for single-beam classification.

Echo duration is a direct indicator of seabed type via seabed attenuation. The effective beamwidth also responds to the seabed type. However, this requires a homogenous seabed with some variation in depth. A method that provides such data in general survey conditions and combines the echo duration approach with other characterization methods will be published in the near future (Biffard, 2010).

REFERENCES

B. R. Biffard, J.M. Preston, and N.R. Chapman, "Acoustic classification with single-beam echosounders: processing methods and theory for isolating effects of the seabed on echoes," Proc. MTS/IEEE Oceans Conference, Vancouver, Canada, 2007.

B.R. Biffard, S.F. Bloomer, N.R. Chapman, J.M. Preston, and J.L. Galloway, "Single-beam seabed characterization: A test-bed for controlled experiments," Proc. Eighth European Conference on Underwater Acoustics, Carvoeiro, Portugal, 2006.

B.R. Biffard, Ph.D. thesis (in preparation), U. Victoria, 2010

S.F. Bloomer, X. Monteys and N.R. Chapman, "Multifrequency Classification and Characterization of Single Beam Echosounder data Collected from Offshore Ireland", in this proceedings, 2010

SEABED SEDIMENT CLASSIFICATION AND SEAFLOOR BATHYMETRY USING SINGLE BEAM HYDRO-ACOUSTIC ECHO BACKSCATTER

Ian Murfitt
Fisheries and Oceans Canada, Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo, BC, Canada, V9T 6N7,
Ian.Murfitt@dfo-mpo.gc.ca

1. INTRODUCTION

Marine ecosystems are complex and dynamic; defined by currents, sedimentation, and climatic effects on biological recruitment and abundance. Our knowledge of these sub-tidal benthic systems is limited to anecdotal reference from industry fishers, although often lacking the spatial resolution necessary for stock assessment and management.

In the near-shore, hydro-graphic charts do not provide enough information for benthic biological assessments; therefore, hydro-acoustic surveys are an essential first stage for detailed spatial analysis of shallow sub-tidal marine environments, and single-beam acoustic survey technology is effective and accurate. Hydro-acoustic surveys are a source of supplementary physical habitat information, bathymetry and seabed sediment type, which are used to examine the spatial variability in stock density and biomass distribution.

The geoduck clam, *Panopea generosa*, is a large infaunal bivalve clam that has been harvested in a commercial fishery on the West Coast of Canada, since 1978. Geoducks are harvested by divers from unconsolidated sediments, primarily sand and pea gravel, using a pressurized water pipe to saturate the seabed sediments; so that the clam can be removed undamaged, and for a live export market.

Single-beam echo sounders, onboard fishing vessels, have been the most valuable asset for the exploitation of sub-tidal marine resources; they are used to locate soft bottom sediments for exploratory dives and ultimately charting commercially viable geoduck harvest areas. The geoduck stock assessment program has been collecting hydro-acoustic echo backscatter data, using QTC View and QTC Impact (Quester Tangent Corp., Sydney, BC, Canada) echo digitisation hardware and software, since 2001. To date, more than 10 000 ha of known geoduck harvest areas have been surveyed.

2. METHOD

Currently, a new single frequency shallow water system is being used to collect 50 kHz single-beam (Airmar 9° x 17° transducer) echo backscatter. This transducer frequency penetrates the seabed surface several centimetres, this is preferable in determining sediment consolidation; e.g. the distinction of light grain sand from hard packed sand to gravel or rock. For many past surveys, a 200 kHz frequency transducer that has a smaller beam width of seafloor insonification was also used, which detects more variability in the seabed texture, algal cover and surface sediments.

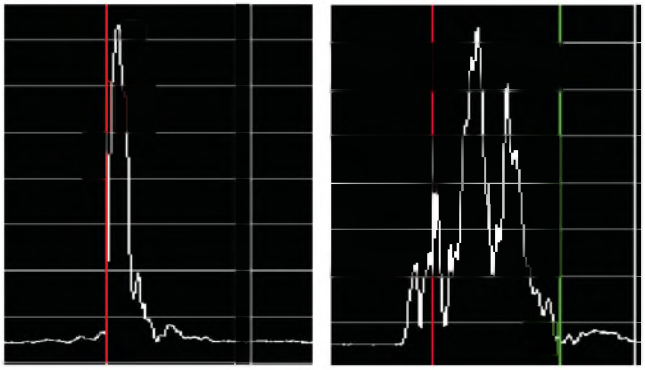


Figure. 1. Echo energy waveforms, less consolidated substrate (left) and more complex/consolidated substrate (right).

The variance in the amplitude and duration of the seabed echo backscatter is relative to the variance in the consolidation and complexity of the seabed surface sediments (Figure 1). The measured energy from single-beam backscatter is analysed to generate statistically descriptive features, which are then reduced to three principal components (Collins et al., 1996, Ellingsen et al., 2002, Murfitt and Hand, 2004). While there are limitations in acoustic technologies, water quality and sea floor slope, single-beam echo sounder systems are an invaluable technology for marine science (Anderson et al., 2008), particularly in refining harvest bed area estimates for geoduck stock assessment.

Most acoustic survey designs follow a 50 m parallel track-line grid in the perpendicular to shore direction, and an along shore 100 m parallel cross track-line. QTC Impact software is used for echo editing and principal component analysis (PCA) (Quester Tangent Corporation, 2004). Post-processing of the acoustical backscatter yields a file of track-line point data. Each file contains survey depth and three principal components, which maximise the variance of the acoustical backscatter. Principal components analysis of the statistical features generated by Quester Tangent echo analysis, and seabed classification as applied by the geoduck stock assessment program are further described in, Murfitt and Hand, 2004.

Depth is first corrected for tide height to chart datum, and then used to generate a continuous bathymetric surface and

depth contours. Idrisi Geographic Information Systems software (Clark Labs, Worchester, MA, USA) is used for geo-statistical interpolation (Kriging method) of the depth and PCA point data into continuous surface images. Interpolated surface images are an additional step in seabed classification, but interpretation and distinction of transition between sediment-types is more obvious and objective than single point track-line data as generated by QTC Impact software. The bathymetric surface is used to calculate area in dive fishery working depth, and the visualization of the distribution of both current stock and stock removed.

The general principle in seabed classification is that the variability of the seabed echo impedance or reflectance can be discretely segmented to represent unique differences in the seabed by the strength and duration of the echo. K-means clustering is an iterative method that minimizes the distance between cluster centres and the data cluster members, grouping similar image cell values from the interpolated principal components (Eastman, 2006). All three PCA images are used to segment the data into a pre-determined number (k) of clusters* that represent distinct changes in seabed consolidation and/or seabed complexity, and are assigned to known seabed-types within the survey area. Acoustic survey data and seabed sediment classification is verified with geoduck harvest locations as reported on fishery dive logs.

3. RESULTS

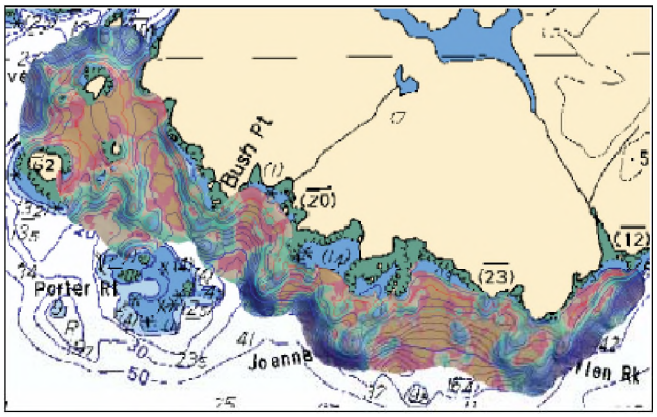


Figure. 2. K-means classification of seabed sediment consolidation, orange-red is a sand substrate.

Figure 2 shows the result of k-mean clustering of the acoustic data, the orange and red classes are the least consolidated sediments and the preferred areas for geoduck harvest effort. The depth range for this survey area is between 2 m and 40 m, the working depth range for divers is between 3 m (min. limit to protect eel grass beds) and 20 m (max. limit for diver safety). There are 3 invertebrate species commercially harvested at this survey area: red sea urchins (*Strongylocentrotus franciscanus*), sea cucumbers (*Parastichopus californicus*), and geoducks. While there is some overlap of species on seabed substrates, cucumbers and urchins are generally less abundant on preferred geoduck harvest sediments. With knowledge of depth and substrate extents, we can better estimate targeted populations and assess the impact of fishing on future stock recruitment.

The primary purpose for these acoustic surveys has been to calculate area of suitable substrate for the geoduck fishery quota calculation. The secondary purpose has been to determine the area in stratified working depths, relative to fishery harvest log depth and dive density survey depth.

4. DISCUSSION

Acoustic technologies have evolved substantially over the past ten years; however single beam is still an established dependable means of collecting seabed backscatter. There would be greater advantages in survey area coverage and resolution with more advanced transducers, particularly, combination lower resolution multi-beam and high resolution side-scan transducers, but with a higher capital expense.

REFERENCES

Anderson, J. T., D. Van Holliday, R. Kloser, D. G. Reid, and Y. Simard. 2008. Acoustic seabed classification: current practice and future directions. ICES Journal of Marine Science, 65:1004-1011.

Collins, W. T., R. Gregory, and J. Anderson. 1996. A digital approach to seabed classification. Sea Technology, 37:83-87.

Eastman, J. R. 2006. IDRISI Andes Guide to GIS and Image Processing. Clark Labs, Clark University, Worcester, MA, USA. 327 pp.

Ellingsen, K. E., J. S. Gray, and E. Bjørnbom. 2002. Acoustic classification of seabed habitats using the QTC VIEWTM system. ICES J. Mar. Sci. 59:825-835

Murfitt, I. and C.M. Hand. 2004. Acoustical substrate classification for the improved estimation of geoduck clam abundance and distribution, p. 289-300. In: Nishida, T., Kailola, P.J., and Hollingworth, C.E. [eds]. 2004. GIS/spatial analyses in fishery and aquatic sciences (Vol. 2). Fishery-aquatic GIS research group, Saitama, Japan.

Quester Tangent Corporation. 2004. QTC Impact Acoustic Classification User Guide, Version 3.4. Quester Tangent Corporation, Sydney, BC, Canada.

ACKNOWLEDGEMENTS

I would like to acknowledge the long and continued support from the Underwater Harvesters Association for the Geoduck stock assessment program.

* Post processing, clusters are referred to as classes.

A COMPARISON OF MEASURED OCEAN ACOUSTIC AMBIENT NOISE WITH ESTIMATES FROM RADARSAT REMOTE SENSING

Sean Pecknold[1], John Osler[1], and Brendan DeTracey[2]
[1] Defence R&D Canada – Atlantic, P.O.Box 1012, Dartmouth, Nova Scotia B2Y 3Z7, sean.pecknold@drdc-rddc.gc.ca
[2]Halifax MetOc, Trinity, CFB Halifax, Halifax, NS, B3K 5X5

1. INTRODUCTION

Ambient noise has been a topic of study in underwater acoustics for over sixty years. In the absence of rain and biological noise sources, wind-generated noise is the dominant source of underwater ambient noise in the 1 kHz to 25 kHz band. The extent to which this noise can be estimated using remotely-sensed data is an area of interest, in particular for the purposes of Rapid Environmental Assessment. Here, synthetic aperture radar (SAR)-derived wind fields measured in littoral areas off the coast of Nova Scotia using Radarsat-1 and Radarsat-2 satellite imagery are compared with data measured using shipboard instruments and by the Gulf of Maine Ocean Observation System (GoMOOS) buoy 'N' in the Northeast Channel. Using an algorithm to predict underwater ambient noise generated due to wind, the modeled ambient noise based on satellite measurements is then compared with in situ measurements recorded on hydrophones.

2. BACKGROUND

2.1 Wind speed dependence of ambient noise

There is an immense amount of literature available concerning ambient noise in the ocean, literally thousands of publications. Urick [1] summarizes much of the early work in this area, including the findings that wind-related processes dominate ambient noise over much of the spectrum below frequencies of 25 kHz. These processes include wind turbulence, surface motion, wave interactions, and spray and cavitation. Shipping can be the dominant noise source for frequencies of tens to hundreds of Hz, but in shallow water without high shipping levels, ambient noise can depend on wind speed at these frequencies as well (e.g., Piggott [2] measured on the Scotian Shelf). Here, a model formulated by Merklinger and Stockhausen (MS) [3], slightly modified [4] is used to calculate expected ambient noise levels given the wind speeds.

2.2 Field trials

Underwater ambient noise levels were collected as part of two field trials off the coast of Nova Scotia. The first trial, designated Q316, took place in September and October of 2008, in the Northeast Channel and Brown's Bank area. The second, designated Q325, took place in Emerald Basin and Emerald Bank area in October and November of 2009. Wind speed data was collected through both trials on ship-borne instruments and deployed sensors. During Q316, hourly averaged wind speeds were obtained from GoMOOS buoy 'N', while during Q325, a meteorology station with anemometer was deployed from October 27, 2009, to November 9, 2009. Their locations are given in Table 1, together with the locations of the ambient noise sensors.

During Q316, three sets of ambient noise data were collected during periods that coincided with the acquisition of satellite SAR imagery, on the 16th, 17th, and 18th of September. The data were acquired using two sub-surface high-fidelity audio recording packages (SHARPs) and one Ocean Bottom Seismometer (OBS). The OBS hydrophone was deployed approximately 1 m above the ocean bottom, in water depth of 107 m. Both SHARP units included two hydrophones, at depths of approximately 55 m and 75 m. SHARP1 was deployed in water depth of 107 m and SHARP2 in water depth of 111 m. All hydrophones were omni-directional with flat frequency responses above 5 Hz.

Table 1. Buoy and sensor locations.

	Latitude	Longitude
GoMOOS buoy	42° 19.014' N	065° 54.010' W
Met station (Q325)	43° 48.348' N	062° 51.534' W
OBS (Q316)	42° 17.488' N	065° 33.123' W
SHARP1 (Q316)	42° 17.637' N	065° 32.987' W
SHARP2 (Q316)	42° 25.412' N	065° 22.287' W
SHARP1 (Q325)	43° 41.555' N	062° 39.391' W
SHARP2 (Q325)	43° 50.823' N	062° 51.161' W

For Q325, ambient noise data were acquired coincidentally with SAR imagery on only one occasion (the 5th of November), although two other sets of ambient noise data were collected as well, one 12 hours after a SAR image in very similar conditions (October 31st), the other set on November 3rd, when the Radarsat overpass was several degrees south of the trial area. Data were acquired on two SHARP units, in water depths of 148 m and 257 m, with hydrophones at approximately 55 m and 75 m depth.

2.3 Radarsat wind speed measurements

Wind speed information can be derived from Synthetic Aperture Radar (SAR) backscatter measurements, using empirical models [5,6]. Both Radarsat-1 and Radarsat-2 use a 5.3 GHz (C-band) SAR sensor in Scan SAR mode to obtain large area wind speeds with 100 m horizontal resolution. Figure 1 shows an example of the wind speed derived from a Radarsat-1 SAR image.

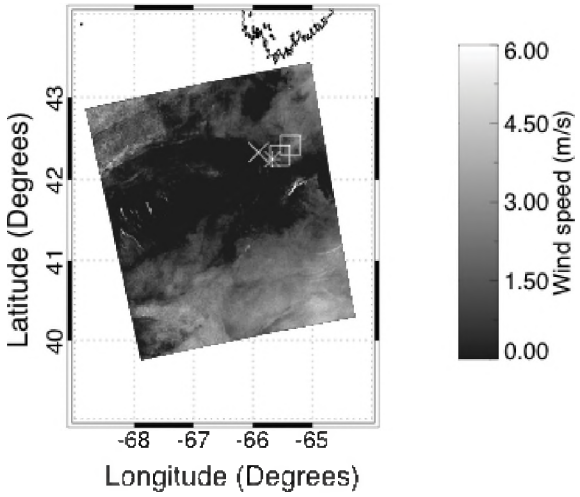


Figure 1. SAR-derived wind speed from Radarsat-1 from September 17, 2008, 2214 UTC. The SHARP recorders are shown as squares, GoMOOS as an X, ship location an asterisk.

3. RESULTS

3.1 Wind speed comparison

Fourteen data points were available for comparison of directly measured wind speed with that obtained using the SAR imagery: three GoMOOS buoy measurements, seven measurements from ship-borne instrumentation, and four readings from the deployed meteorology buoy. An hourly mean and standard deviation of wind speed were calculated for the ship anemometer and meteorology station, the GoMOOS buoy already providing hourly averaged wind speeds. Wind speeds ranged from 0 m/s to 11.2 m/s. Wind speed estimates from the SAR imagery were obtained by calculating the mean and standard deviation of the wind speed in a 10-km radius around the location of each wind sensor and hydrophone. After correcting for sensor height, wind speeds from the Radarsat and other sensors agreed to within one standard deviation for 8 data points, and two standard deviations for 3 data points. The other three data points were all for wind speeds less than 2 m/s, for which estimates made using the Radarsat imagery were too low. It was anticipated that it might be difficult to estimate very low wind speeds from the Radarsat images [7].

3.2 Ambient noise comparison

Figure 2 shows a comparison of ambient noise measured during Q316 on the lower hydrophone of the SHARP2 unit, during the periods when SAR imagery was collected (Radarsat-1 on September 16th, Radarsat-2 on the 17th and 18th). The data was sampled at 22050 Hz. Power spectral densities were computed by averaging 1200 half-second square-windowed periodograms.

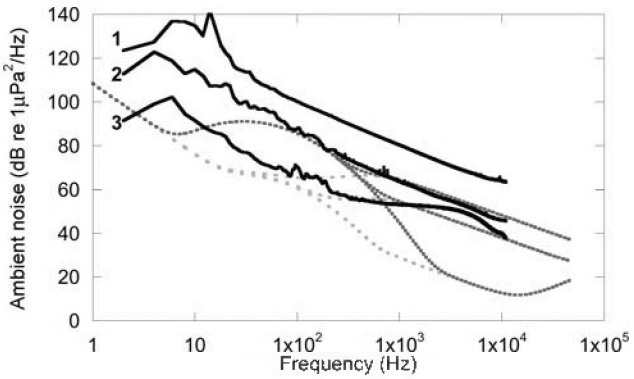


Figure 2. Measured ambient noise from SHARP2, lower hydrophone: (1) Sept. 16, 2008, 6.6 m/s wind speed; (2) Sept 17, 2008, 1.8 m/s wind speed; (3) Sept. 18, 2008, 0.1 m/s wind speed. Dotted lines show MS model results for given SAR-derived wind speeds, together with light and heavy shipping curves (light and heavy dashed lines).

4. DISCUSSION

As seen in Figure 2, there is an obvious dependence of the measured ambient noise on wind speed, demonstrating the potential for ambient noise estimates over a wide area to be derived from Radarsat remotely sensed data. It is evident however that the measured ambient noise is considerably higher than that predicted by the MS model, particularly at the lowest wind speed (0.1 m/s). There is also wind speed dependence at all frequencies, including the frequency regime normally expected to depend on shipping noise. This is consistent with the observations of Piggott [2], however, the noise levels measured here are higher, possibly due to the nature of the propagation in the area. Further analysis of data from both the Q316 and Q325 experiments may help resolve this issue.

REFERENCES

[1] Urick, R. J. (1984). *Ambient Noise in the Sea* (Peninsula, Los Altos, CA).

[2] Piggott, C. L. (1964). "Ambient sea noise at low frequencies in shallow water of the Scotian Shelf," J. Acoust. Soc. Am 36. 2152-2163.

[3] Merklinger, H.M. and Stockhausen, J.H. (1979). ''Formulae for estimation of undersea noise spectra", presented to the 97th Meeting of the Acoustical Society of America, J. Acoust. Soc. Am. 65, S88.

[4] Hazen M. G., Desharnais F. (1997), "The Eastern Canada Shallow Water Ambient Noise experiment," in *Proc. of OCEANS '97* (1), 471-476.

[5] Vachon P, Dobson F. (2000). "Wind retrieval from Radarsat SAR images: Selection of a suitable C-band HH polarization wind retrieval model," Can. J. Remote Sensing 26(4), 306–313.

[6] Danielson R.E., Dowd M., Ritchie H. (2008). "Objective analysis of marine winds with the benefit of the Radarsat-1 synthetic aperture radar: A nonlinear regression framework," J. Geophys. Res. 113, C05019.

[7] Barthelmie R.J., Pryor S.C. (2003). "Can Satellite Sampling of Offshore Wind Speeds Realistically Represent Wind Speed Distributions?" J. Appl. Meteo. 42(1), 83-94.

MEASUREMENTS AND MODELLING OF ATMOSPHERIC ACOUSTIC PROPAGATION OVER WATER

Emma Murowinski and Cristina Tollefsen[1]
Defence Research and Development Canada - Atlantic, P.O. Box 1012, Dartmouth, NS B4Y 3Z7
[1]cristina.tollefsen@drdc-rddc.gc.ca

1. INTRODUCTION

Understanding atmospheric acoustic propagation over water could prove to be a valuable tool for determining the environmental footprint of offshore wind farms or naval gunfire exercises, or for evaluating the effectiveness of acoustic hailing devices used at sea. Atmospheric parameter profiles (temperature, wind speed, humidity, and turbulence) and water surface roughness can dramatically affect the acoustic propagation. Wiener [1] measured acoustic transmission loss in foggy conditions using a fog horn as the acoustic source. Salomons [2] showed that water surface waves can strongly affect transmission loss in long-range, over-water propagation, while Boué [3] showed that cylindrical spreading is an appropriate model up to 700 m range. Bolin and Boué [4] showed that accurate predictions in shadow zones rely on inclusion of atmospheric turbulence in transmission loss models.

In the two experiments presented here, atmospheric acoustic transmission loss over water was measured as a function of range. Simultaneous environmental data acquired included atmospheric parameters and directional wave spectra.

2. METHOD

2.1 November 2009 Experiment

The first experiment was performed at sea on 2 Nov 2009 on board Defence Research and Development (DRDC) Atlantic's research ship, CFAV Quest. Two receivers, a Sony Linear PCM Recorder and an mh acoustics em32 Eigenmike microphone array, were mounted aboard Quest at 7.5 m above the sea surface. The source was a dual-tone Nauticus 3500 horn with nominal frequencies of 530 Hz and 670 Hz and an on-axis source level of 115.7 dB re 20 μPa at 1 m. The horn was mounted aft-facing at 2.1 m height on a rigid hull inflatable boat (RHIB) that was driven towards and away from CFAV Quest at a speed of 7 knots. During the runs, the horn was sounded every 30 s for 10 s. A Brüel & Kjaaer (B & K) sound pressure level (SPL) meter was used to monitor the ambient noise on board Quest. Point measurements were made of temperature and humidity (at 15.2 m height), wind velocity (at 24.7 m height), and significant wave height.

2.2 July 2010 Experiment

The second experiment was performed from 19-23 Jul 2010 in the Bedford Basin, Halifax, Nova Scotia on board DRDC Atlantic's Acoustic Calibration Barge. The receiver was a Core Audio Tetramic mounted above the barge structure 10.2 m above the water surface. The source, the same Nauticus horn used in the 2009 experiment, was mounted aft-facing at 1.25 m height on a Zodiac. The Zodiac was driven towards and away from the barge at speeds of 4 to 8 knots and the horn was sounded every 20 s for 5 s. A Sony Linear PCM Recorder was used in the Zodiac to monitor the source. The ambient noise level and source level were measured several times each day with the B & K SPL meter.

Directional ocean surface wave spectra were measured using a Teledyne RD Instruments Acoustic Doppler Current Profiler (ADCP). Vaisala Radiosondes were launched from Canadian Forces Base Halifax (6 km away) on each day at 0930 and 1230 to record atmospheric parameter profiles. Point measurements were made of temperature, wind velocity, humidity, and air pressure at 9 m height. Parameter profiles from Environment Canada's Global Environmental Multiscale (GEM) model were available at 0900, 1200, and 1500 each day.

3. RESULTS

3.1 November 2009 Experiment

Measured SPL in arbitrary units are plotted as a function of range in Figure 1 for both horn frequencies and two different relative wind directions: crosswind (Figure 1a and 1b) and upwind (Figure 1c and 1d). For all the plots in Figure 1, the wind speed was 10.1 m/s from a direction of 50°. Wind direction was calculated relative to the source-receiver vector; therefore, the relative wind direction is 0° when the receiver is directly downwind of the source. For comparison, each plot shows the SPL resulting from cylindrical and spherical spreading, forced to agree with the measured data at the closest range point. In the crosswind case, the received SPL decreases more quickly with range than predicted by both cylindrical and spherical spreading. In the upwind case, the received SPL shows the same general trend as spherical spreading. The maximum detectable range is greater (700 m) in the crosswind case than the upwind case (350 m).

3.2 July 2010 Experiment

Measured SPL in arbitrary units are plotted as a function of range in Figure 2 for both horn frequencies and two different relative wind directions: approximately crosswind (Figures 2a and 2b), and approximately upwind (Figures 2c and 2d). For the crosswind run, the wind speed was 1.3 m/s from a direction of 81°, while for the upwind run, the wind speed was 1.6 m/s from a direction of 275°. In the crosswind case, the received SPL agrees with cylindrical spreading to a range of 600 m, where it drops to 8 dB below

spherical spreading. In the upwind case, the received SPL agrees with cylindrical spreading to a range of 300 m, and shifts to spherical spreading beyond 300 m. Again, the maximum detectable range is much greater (900 m) in the crosswind case than the upwind case (700 m).

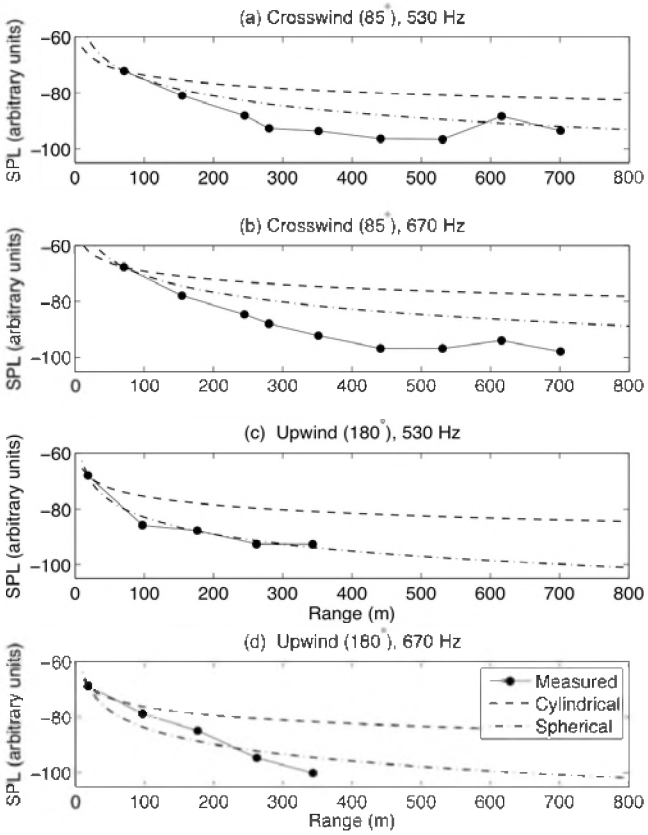


Figure 1 Measured SPL (●), cylindrical spreading (dashed line), and spherical spreading (dash-dot line), in arbitrary units, as a function of range for the Nov 2009 experiment: (a) crosswind, 530 Hz, (b) crosswind, 670 Hz, (c) upwind, 530 Hz, and (d) upwind, 670 Hz.

4. DISCUSSION AND FUTURE WORK

In both experiments, simple spherical or cylindrical spreading did not suffice to describe the range dependence of the measured transmission loss. The maximum detectable distance was greater for the crosswind case than the upwind case in both experiments. Future work will include analyzing the remainder of the data from the Jul 2010 trial, and implementing a propagation model using measured and modelled atmospheric parameters for comparison with the measured transmission loss.

REFERENCES

[1] Wiener, F. M. (1961), Sound Propagation over Ocean Waters in Fog, J. Acoust. Soc. Am. 33 (9), 1200-1205 .

[2] Salomons, E. (2007), Computational Study of Sound Propagation over Undulating Water, In *Proceedings of the 19th International Congress on Acoustics*, Madrid: International Commission for Acoustics.

[3] Boué, M. (2007), Long-Range Sound Propagation over the Sea with Application to Wind Turbine Noise, Vindforsk Report No. V-20, Vindforsk, http://www.vindenergi.org/Vindforskrapporter/V-201_TRANS_webb.pdf, accessed on 30 Jul 2010.

[4] Bolin, K. and Boué, M. (2009), Long range sound propagation over a sea surface, J. Acoust. Soc. Am 126 (5), 2191-2197.

ACKNOWLEDGEMENTS

Thanks to P. Shouldice, D. Graham, P. Anstey, R. Johnson, M. Fotheringham, and LS D. Ratelle for technical support.

AUTHOR NOTES

The work was conducted while Emma Murowinski was working as a co-op student at DRDC Atlantic. The current address is: e.murowinski@gmail.com.

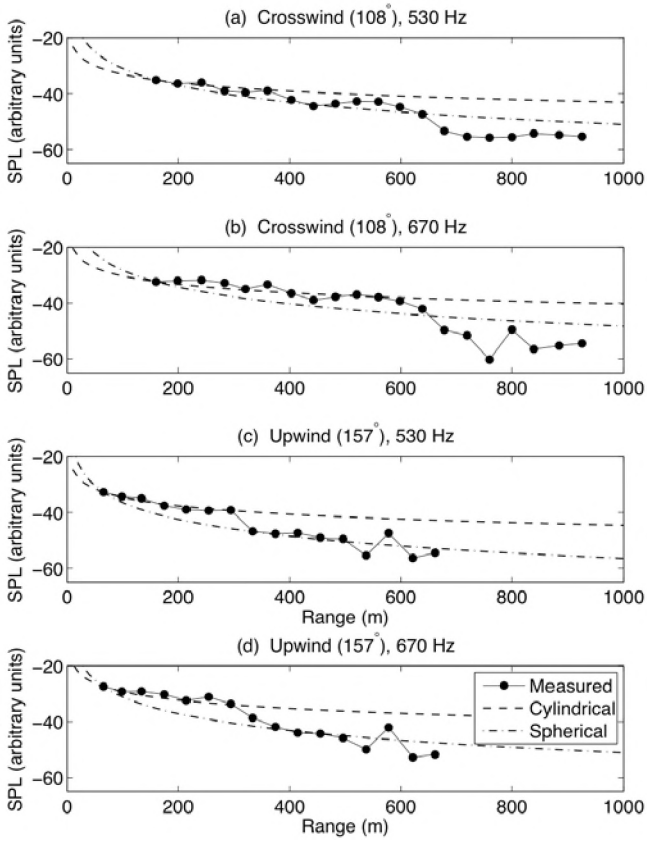


Figure 2 Measured SPL (●), cylindrical spreading (dashed line), and spherical spreading (dash-dot line), in arbitrary units, as a function of range for the Jul 2010 experiment: (a) crosswind, 530 Hz, (b) crosswind, 670 Hz, (c) upwind, 530 Hz, and (d) upwind, 670 Hz.

MULTIPLE SHOCK LOADING ON FLUID-FILLED SHELL STRUCTURES

Serguei Iakovlev [1], Adrien Lefieux[2], Jean-Francois Sigrist[3] and Kyle Williston [1]
[1] Dept. of Engineering Mathematics and Internetworking, Dalhousie University, Halifax, Nova Scotia, Canada
[2] Dept. of Structural Mechanics, University of Pavia, Pavia, Italy
[3] Service Technique et Scientifique, DCNS Propulsion Indret, France

1. INTRODUCTION

The analysis of the scenario where a submerged shell structure is subjected to two or more consecutive shock waves is of definite practical interest, one of the most obvious examples being a complex shock loading generated as a result of the reflection of the original shock wave off one or more reflective surfaces in the proximity of the structure. At the same time, the relevant published studies of such multiple loading scenarios are quite scarce, perhaps owing to the fact that there are a number of higher-priority shock-structure interaction problems that over the years attracted more attention from the research community and more funding from the industry.

More specifically, although the somewhat relevant complex reflection patterns in systems comprised of several components were considered in a number of studies (e.g. [1,2]), it appears that, to the best of our knowledge, the interaction between a submerged fluid-filled shell and a multi-front shock wave has not yet been considered.

Our goal, therefore, is to report some preliminary findings for a fluid-filled submerged cylindrical shell subjected to an external shock wave with two fronts (or, equivalently, to two consecutive shock waves). We are particularly interested in developing an understanding of the structure and evolution of the internal hydrodynamic field induced by such loading since the internal hydrodynamic loading will determine some of the peak values of both the pressure and structural stress.

2. MATHEMATICAL FORMULATION AND SOLUTION METHODOLOGY

We consider a thin elastic circular cylindrical shell filled with and submerged into identical fluids. We assume that the shell is thin enough, and that its deflections are small in comparison to its thickness, so that the linear shell theory can be employed; we further assume that the Love-Kirchhoff hypothesis holds true.

The fluids are assumed to be irrotational, inviscid, and linearly compressible, thus the wave equations are used to model the fluid dynamics. The fluids and the shell are coupled through the dynamic boundary condition on the interface.

The problem is approached with the methodology we developed in our earlier work [3, 4], i.e. we apply the Laplace transform time-wise to the wave equations and then separate the spatial variables in order to arrive at the expressions for the transforms of the internal and external pressure in a form of a series of modified Bessel functions of the first (internal fluid) and second (external fluid) kind. Upon inverting the transforms we obtain the pressure as a Fourier series with time-dependant coefficients which, for the radiation pressure, still depend on the unknown at this stage normal displacements of the shell.

Then, we use the same series form for the shell displacements and, substituting them into the shell equations, we arrive at the systems of the ordinary differential equations for each of the displacement harmonics. The systems are then approached numerically (finite differences) and the resulting normal displacement is used in the pressure series to compute the radiation pressure everywhere around the shell.

The newest addition to the approach is incorporating a shock wave with multiple fronts, more specifically, a shock loading comprised of two consecutive plane shock waves of the same intensity; the equations and relevant discussion for the individual waves can be found in [4].

We note that although this model is quite simple in that any realistic shock wave being reflected off any surface will generate another shock wave with a different location of is `virtual source', with a plane wave being the only exception in some instances, we still think that the results we present here are of value since they establish the basic framework for analysis of multiple shock loading, and since they allow us to highlight the most important phenomena that make the present case so different from its single-wave counterpart.

The methodology we use has been extensively (and successfully) verified for weak shock waves and acoustic pulses in our earlier work [3,4], and since the present loading does not differ from the single-front shock waves used in those studies in terms of its intensity, we are convinced that the methodology can be successfully used to address the problem in question as well.

3. RESULTS AND CONCLUSIONS

A steel shell is considered with the thickness of 0.01 m and radius of 1 m, submerged into and filled with water. The interaction with two identical consecutive plane incident waves is analyzed, with the rate of exponential decay of 0.0001314 s and the initial pressure in the front of 250 kPa. Although the model allows for the consideration of any distance between the wavefronts to be considered (which, of course, corresponds to different delays of the offset of the second wave), we only consider a single value of the dimensionless delay of 0.9 (the dimensionless unity is equal to the time it takes for the shock wave to pass over the radius of the shell). This value means that the second wave arrives at the shell when the first one is just about to reach the axis.

Fig. 1 shows three numerically simulated sequential images of the internal field for the chosen shock wave during the early and middle interaction (the external field is not be shown since we have observed that all the most interesting phenomena due to the presence of the second front are taking place in the internal fluid).

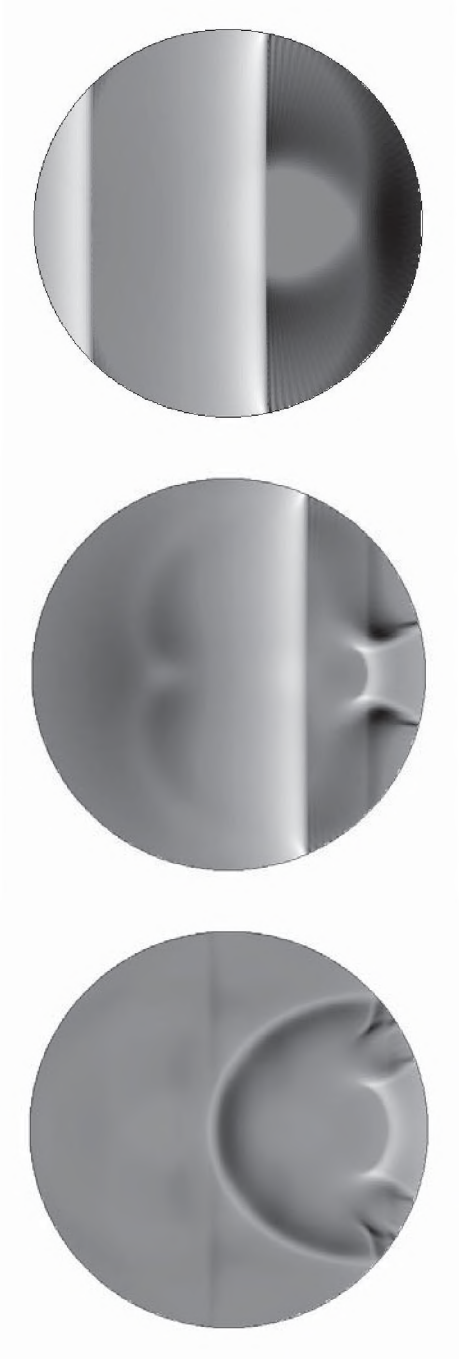


Figure 1. The internal hydrodynamic pattern at the dimensionless time instants of 1.20 (top), 2.30 (middle) and 3.10 (bottom).

The first image shows an instant when both waves has already impacted the shell and originated the respective internal pressure waves propagating downstream (from here on we will refer to the internal pressure waves simply as 'shock waves', minding the remarks on the matter found in [3]). In the second image, the first internal wave has reflected off the tail region and is forming the classical post-reflection pattern, but the second one is still propagating downstream, with its regular reflection pattern developing. In the third image, both waves have reflected and are propagating upstream, resulting in the field allowing one to simultaneously see two different stages of the development of the reflection pattern in one image (an effect somewhat similar to the one seen in one of the experimental images of [5]; in that work, however, the presence of such a double structure was due to the overlaying of images taken at different instances, not due to the existence of an actual double-front wave).

As was expected, the structure of the field for the double-wave case is much more complex than that for the single wave one. Even the three representative illustrations shown here make a strong enough case for a thorough further investigation of the two-wave scenario. Specifically, we expect that the superposition of various features of the internal field will, in some cases, result in new features that are of considerable practical interest.

Investigating such effects is one of the future objectives of our research program, and it will require a more in-depth study of the interaction for a rather wide range of the time intervals between the shock waves. Such an investigation will also facilitate the future studies of systems where multiple surfaces are present. In particular, it will result in a better understanding of what mutual locations of the shell and the reflective surfaces should be given a priority in such studies.

REFERENCES

[1] Jialing, L. Hongli, N. (1997) Study on propagation of blast wave over cylinder in a channel bend, proceedings of the 21st International Symposium on Shock Waves, Great Keppel Island, Australia, July 1997, pp.103–107.

[2] Oakley, J.G., Anderson, M. H., Wang, S., Bonazza, R. (2001) Shock loading of a cylinder bank with imaging and pressure measurements, proceedings of the 23nd International Symposium on Shock Waves, Fort Worth, Texas, USA, July 2001, pp.589–595.

[3] Iakovlev, S. (2006) External shock loading on a submerged fluid-filled cylindrical shell, *Journal of Fluids and Structures* 22, 997-1028.

[4] Iakovlev, S. (2008) Interaction between a submerged evacuated cylindrical shell and a shock wave. Part I: Diffraction-radiation analysis, *Journal of Fluids and Structures* 24, 1077-1097.

[5] Sun, M. and Takayama, K. (1996) A holographic interferometric study of shock wave focusing in a circular reflector, *Shock Waves* 6, 323-336.

ACKNOWLEDGEMENTS

S.I. and K.W. acknowledge the financial support of the Natural Sciences and Engineering Research Council (NSERC) of Canada. S.I. also acknowledges the financial support of the Killam Trusts and the Faculty of Engineering, Dalhousie University.

EFFECT OF STRUCTURAL ENHANCEMENT ON THE ACOUSTIC RESPONSE OF A SUBMERGED FLUID-FILLED CYLINDRICAL SHELL

Serguei Iakovlev [1], Kyle Williston [1], Jean-Francois Sigrist[2] and Adrien Lefieux[3]
[1] Dept. of Engineering Mathematics and Internetworking, Dalhousie University, Halifax, Nova Scotia, Canada
[2] Service Technique et Scientifique, DCNS Propulsion Indret, France
[3] Dept. of Structural Mechanics, University of Pavia, Pavia, Italy

1. INTRODUCTION

Although shell-fluid(s) systems that are more complex than a single shell submerged into (and in some cases filled with) fluids have been studies in the context of the acoustic or shock loading for quite a while (e.g. [1]), it appears that most of the studies, both experimental and numerical, devoted to the visualization of the respective hydrodynamic fields appeared only during the last decade (e.g. [2,3]).

Despite the existence of the research effort mentioned, it appears that the studies where a thorough investigation of the evolution of the hydrodynamic patterns observed in a structurally enhanced system are very scarce, and for some systems simply do not seem to exist. This certainly appears to be the case for a circular cylindrical shell containing a rigid co-axial core, and we present some of our results for such a scenario.

2. MATHEMATICAL FORMULATION AND SOLUTION METHODOLOGY

We consider a thin circular cylindrical shell, assume that the deflections of the shell are small compared to its thickness, and assume that the Love-Kirchhoff hypothesis holds true, thus the respective linear shell theory can be used. We assume that the shell is filled with and submerged into identical fluids, and the fluids are assumed to be linearly compressible, inviscid, and irrotational, thus modeled by the wave equations. We furthermore assume that the shell contains a rigid co-axial cylindrical core.

The methodology we use has been developed in our earlier work [4,5] and it is based on combining the Laplace transform technique with respect to the time with the separation of variables with respect to the spatial coordinates. As a result, the solution is obtained in the form of Fourier series with time-dependent coefficients that are convolution integrals of the hydrodynamic pressure or the shell displacement and the so-called response functions, the latter representing the response of the fluids to the scattering by or radiation of the shell. The methodology has been extensively verified [4,5], and we feel that we can confidently use it for simulating more complex systems for which experimental data does not necessarily exist yet.

3. RESULTS AND CONCLUSIONS

A steel shell is considered with the thickness of 0.01 m and radius of 1 m (r_0), submerged into and filled with water. The interaction with a cylindrical incident wave [5] is analyzed, and the rate of the exponential decay of the wave is assumed to be 0.0001314 s while the initial pressure in the front is 250 kPa. A rigid co-axial core placed inside the shell is assumed to have different radii a, from a small one (a/r_0=0.10) to the one that dominates the internal volume (a/r_0=0.75).

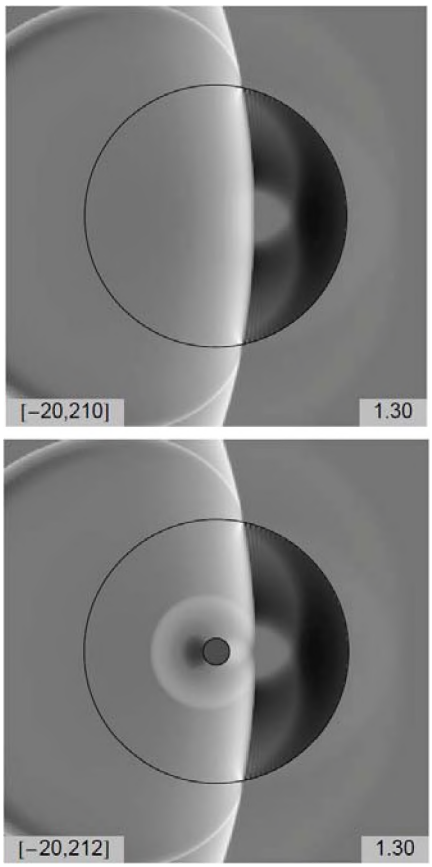


Figure 1. Comparison between the cases when no core is present (top) and when the core has a small radius (bottom).

Fig. 1 shows the comparison between the no-case scenario and the one where a small-radius core (a/r_0=0.10) is added to the system. The single numerical value on the right corresponds to the dimensionless time (with unity corresponding to the time it takes for the incident wave to travel the distance equivalent to one radius of the shell), and the two values on the left correspond to the pressure range shown, the lower value corresponding to the black halftone in the image and the higher value corresponding to the white one. It is apparent that the changes to the front of the internal pressure wave are rather insignificant, although the internal wave pattern is still quite different because of the core-reflected field present. All the main features, however, remain unchanged.

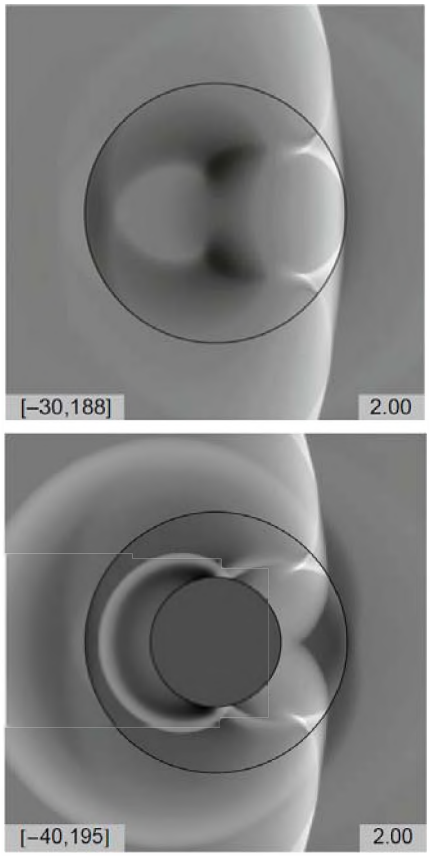


Figure 2. Comparison between the cases when no core is present (top) and when the core has a medium radius (bottom).

Fig. 2 shows a similar comparison for a medium-size core (a/r_0=0.50). In this case, the changes to the internal wavefront are very dramatic, and the influence of the presence of the core is clearly visible. Two core-reflected fronts are visible as well, one that is still inside the shell and one that has propagated into the external fluid.

Fig. 3 shows a snapshot for the case where the core dominates the internal volume (a/r_0=0.75; the respective no-core case is not shown since it is very similar to that seen in Fig. 2). In this case, the pattern is quite a bit more complex than even in the medium-size core case. In particular, three waves are seen in the external fluid in addition to the incident-scattered field, all of which are a result of the transmission of the core-reflected waves outside the shell. The multiple reflection pattern in the layer between the shell and the core is apparent, and it is interesting to point out that the layer seems to exhibit some properties of a waveguide.

The observations made are very interesting and reveal some fundamental facts about the interaction when the structural complexity of the original single-shell system is increased. Namely, it is now clear that the phenomenological nature of the internal field (e.g. the presence of the high-pressure focusing) can be controlled by means of incorporating additional structural elements into the system.

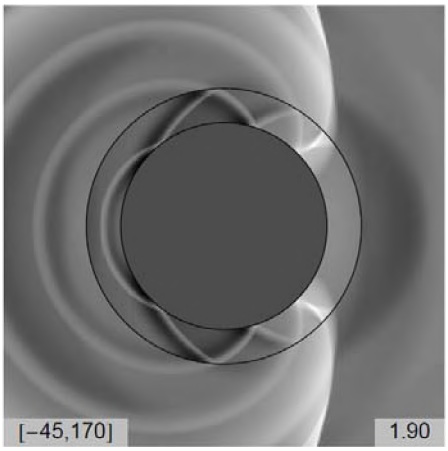


Figure 3. The reflection pattern for the case of a large core.

Furthermore, based on these preliminary observations, it appears that the case of two different fluids would exhibit a number of even more interesting features, especially when the sound speed in the internal fluid is lower than in the external one. Even more interestingly, our preliminary studies for a shell loaded by a shock wave with a double front also indicate that when the present system is subjected to such a loading, one should expect to see dramatically more complex reflection patterns, especially for large core sizes. We intend to investigate these aspects in the future.

REFERENCES

[1] Huang, H. (1979) Transient response of two fluid-coupled cylindrical elastic shells to an incident pressure pulse, *Journal of Applied Mechanics* 46, 513-518.

[2] Wardlaw, A. B., Luton, J. A. Jr. (2000) Fluid-structure interaction mechanisms for close-in explosions, *Shock and Vibration* 7, 265-275.

[3] Derbesse, L., Pernod, P., Latard, V., Merlen, A., Decultot, D., Touraine, N., Maze, G. (2002) Acoustic scattering from complex elastic shells: visualization of S_0, A_0 and A waves, *Ultrasonics* 38, 860-863.

[4] Iakovlev, S. (2006) External shock loading on a submerged fluid-filled cylindrical shell, *Journal of Fluids and Structures* 22 (8), 997-1028.

[5] Iakovlev, S. (2008) Interaction between a submerged evacuated cylindrical shell and a shock wave. Part I: Diffraction-radiation analysis, *Journal of Fluids and Structures* 24, 1077-1097.

ACKNOWLEDGEMENTS

S.I. and K.W. acknowledge the financial support of the Natural Sciences and Engineering Research Council (NSERC) of Canada. S.I. also acknowledges the financial support of the Killam Trusts and the Faculty of Engineering, Dalhousie University.

COMPARISON OF SOUND SPEED PROFILE INTERPOLATION METHODS WITH MEASURED DATA; EFFECTS ON MODELLED AND MEASURED TRANSMISION LOSS

Cristina Tollefsen and Sean Pecknold
Defence Research and Development Canada – Atlantic, P. O. Box 1012, Dartmouth, NS B2Y 3Z7
cristina.tollefsen@drdc-rddc.gc.ca and sean.pecknold@drdc-rddc.gc.ca

1. INTRODUCTION

Range-dependent model predictions of underwater acoustic propagation rely on a series of input sound speed profiles (SSPs) at different ranges from the acoustic source. In order to avoid computational artifacts associated with sudden changes in SSP, some propagation models interpolate between the input SSPs to calculate the SSP at each range step. A variety of SSP interpolation schemes exist, not all of which are suitable for entering measured SSPs into propagation models [1].

2. THEORY

Among the most promising interpolation schemes for direct implementation within propagation models are linear interpolation, triangular interpolation, and trapezoidal interpolation. Detailed formulae can be found in [1]; an overview will be presented here.

In linear interpolation the sound speed is linearly interpolated between the input speeds at different ranges but at the same depth. The risk of using linear interpolation is that features such as sound channel ducts that strongly affect propagation may be smeared out by the interpolation process [1].

Both triangular and trapezoidal interpolation schemes rely on the identification of SSP features. Figure 1a shows two idealized SSPs, each with three features identified: the depth of the sound speed minimum, the depth of sound speed maximum above the minimum, and the depth of the maximum negative gradient between the first two features.

In triangular interpolation, a diagonal line connecting the depths of two features at two different ranges defines two triangles (Figure 1a), and the rate of ascent or descent of the feature along the diagonal line is maintained. For intermediate ranges and depths bounded by the dotted lines and "inside" the triangles (Regions A, B, and C in Figure 1a), the rate of ascent or descent is proportional to the distance from the diagonal line and can be determined by simple algebra. Linear interpolation is used for ranges and depths "outside" the triangles (Regions D, E, and F in Figure 1a).

In trapezoidal interpolation the lines connecting the features divide the interpolation region into trapezoidal areas (Regions A, B, C, and D in Figure 1b). For interpolation at a given depth and range point, the rate of descent or ascent is proportional to where the point lies between the two diagonal lines defining the region.

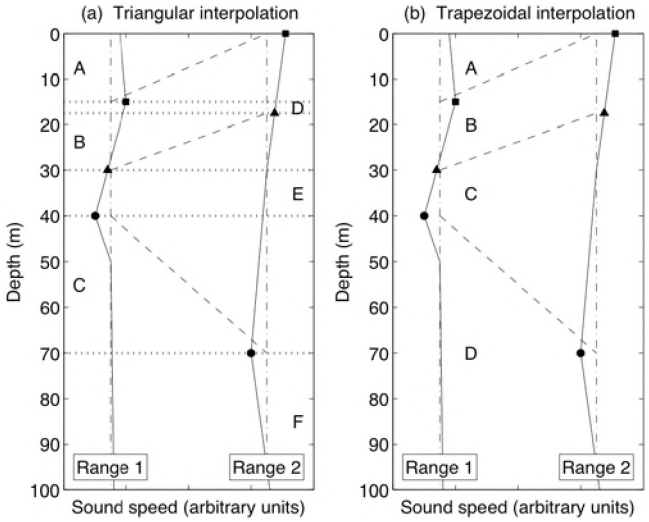


Figure 1 Types of interpolation. The sound speed minimum (●), maximum above the minimum (■), and depth of maximum negative gradient between the first two features (▲) are indicated on each profile. (a) Triangular interpolation: the depths at which the features are found define the diagonals of the triangles. Dotted lines indicate the extent of regions with triangular (A, B, C) and linear (D, E, F) interpolation. (b) Trapezoidal interpolation: the depths at which the features are found define the non-parallel sides of the trapezoidal regions A, B, C, and D.

3. RESULTS and DISCUSSION

During a September 2008 sea trial on the Scotian shelf, SSPs were acquired along a 35-km straight-line track at intervals ranging from 1.1 km to 3.5 km along the track (mean interval = 1.8 km). A subset of the measured SSPs are plotted in Figure 2.

The interpolation was performed between the first and last of the measured profiles in Figure 2, in which the SSP features identified were: the depth of the sound speed minimum (rising from 64 m to 28 m depth), the depth of the sound speed maximum above the minimum (remaining at 4 m depth), and the depth of the maximum gradient between the first two features (rising from 20 m to 18 m depth). Using these features, the sound speed was calculated using linear, triangular, and trapezoidal interpolation, at the same range as each measured profile.

The measured and interpolated sound speeds as a function of depth and profile number are displayed as greyscale images in Figure 3. The measured profiles (Figure 3a) can

be compared with profiles obtained by linear interpolation (Figure 3b), triangular interpolation (Figure 3c), and trapezoidal interpolation (Figure 3d), all performed using profiles #1 and #15 as the interpolation endpoints. The most noticeable difference between the interpolation methods is in the representation of the rising sound speed channel that is first visible in measured profile #11, centred at around 40 m depth, rising to 28 m depth in the last profile. With linear interpolation, the channel remains at the same depth (28 m) but gradually becomes more pronounced between the first and last profiles; however, the triangular and trapezoidal interpolation schemes capture the formation of the channel at 40-60 m depth and its subsequent rise.

Transmission loss estimates were generated using a version of Bellhop ([2] [3]), a Gaussian beam propagation model. Measured and interpolated SSPs from Figure 3 were used with a 100 m deep range-independent environment with sandy bottom. The average and median differences in transmission loss at 28 m depth were calculated (Table 1), comparing the results using measured SSPs to those with interpolated SSPs. The trapezoidal interpolation most closely reproduced the results generated using the measured profiles, with an average difference of -3.1 dB and median difference of 3.0 dB.

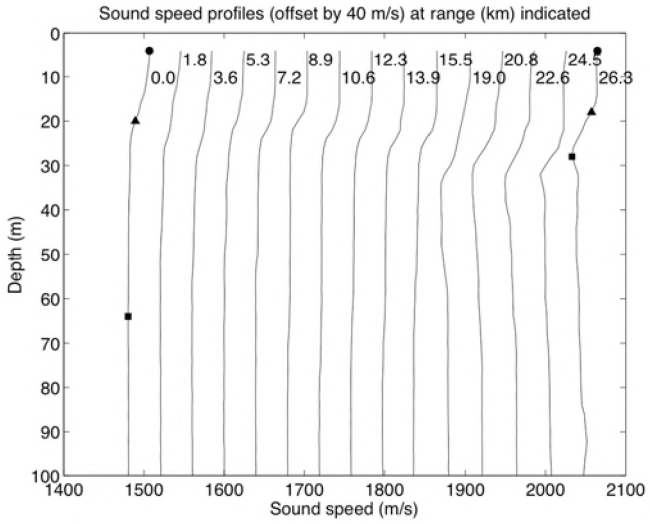


Figure 2 Measured SSPs, offset by 40 m/s, with range from the track start point (km) at the top of each profile. The sound speed minimum (●), maximum above the minimum (■), and depth of maximum negative gradient between the first two features (▲) are marked on the first and last profiles.

Table 1 Summary of average and median difference between modelled transmission loss using measured SSPs and different interpolation schemes, at 28 m depth.

Interpolation Type	Average difference (dB)	Median difference (dB)
Linear	-6.7	-5.5
Triangular	-4.5	-3.9
Trapezoidal	-3.1	-3.0

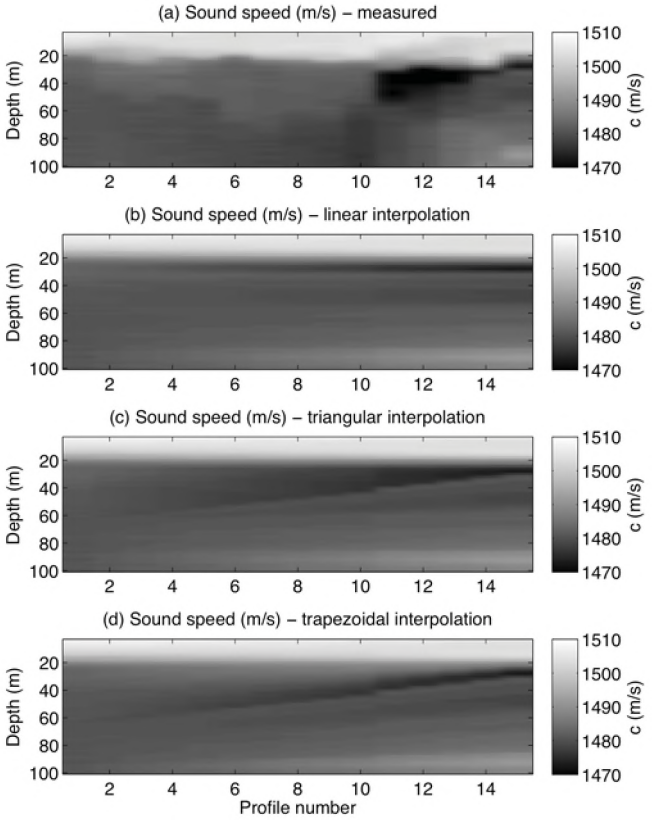


Figure 3 Greyscale representations of sound speed as a function of depth and profile number, for (a) measured data, (b) linear interpolation, (c) triangular interpolation, and (d) trapezoidal interpolation.

4. CONCLUSIONS

Three different methods of interpolating in range between SSPs were investigated for use in range-dependent ocean acoustic propagation modelling. Interpolated SSPs were compared with measured SSPs acquired at the same ranges along a straight-line track. Qualitatively, a sound channel that forms partway along the track is better represented by triangular or trapezoidal interpolation than by linear interpolation. Trapezoidal interpolation resulted in the least difference between transmission loss calculated using measured SSPs and interpolated SSPs.

REFERENCES

[1] McCammon, D. (2009). Researching automatic methods to interpolate between sounds peed profiles at different locations. Defence R&D Canada – Atlantic Contract Report, DRDC Atlantic Report CR 2008-292.

[2] McCammon, D. (2006). Users' guide to BellhopDRDC_V3. Defence R&D Canada – Atlantic Contract Report, DRDC Atlantic Report CR 2006-067.

[3] Porter, M. B., Bucker, H. P. (1987). Gaussian beam tracing for computing ocean acoustic fields. J. Acoust. Soc. Am. 82 (4), 1349-1359.

EVANESCENT LIQUID SOUND-PRESSURE WAVES NEAR UNDERWATER RESONATORS

Reinhart Frosch
Sommerhaldenstrasse 5B, CH-5200 Brugg, Switzerland; reinifrosch@bluewin.ch
PSI (Paul Scherrer Institute), Villigen and ETH (Eidgenoessische Technische Hochschule), Zurich (retired)

1. INTRODUCTION

Evanescent liquid sound-pressure waves (i.e., standing waves, of limited spatial extension, with variable pressure and liquid-particle velocity but negligible density variation) appear to play a fairly important role in the human cochlea, e.g., in the generation process of spontaneous oto-acoustic emissions [Frosch (2010a, 2010b)]. These waves are studied, in the present contribution, with the help of underwater resonators. Mainly because of the kinetic energy in the generated evanescent waves, a tuning fork submerged in water oscillates at a frequency below 440 Hz (typically at ~415 Hz, lower than 440 Hz by about a semitone). In the case of drinking glasses tapped with a spoon, the corresponding frequency reduction is greater than an octave; see Table 1.

Case	Condition	f [Hz]	b [Eq.(11)]
A	empty, on table	~1040	0
B	full, on table	~590	2
C	empty, submerged	~590	2
D	full, submerged	~440	4

Table 1. Frequencies of a wineglass tapped with a spoon.

In Section 2 below, an idealized drinking glass is treated, namely a bottomless hollow cylinder as shown in Fig. 1.

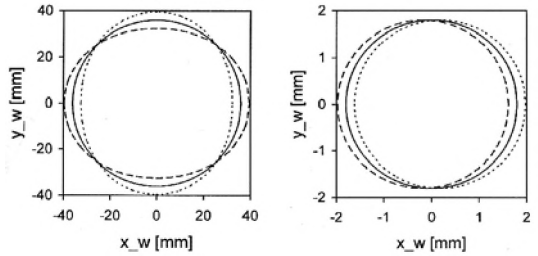


Fig. 1. Left: oscillation of idealized tapped drinking glass. Right: linear oscillation of hollow cylinder or of tuning-fork prong.

2. METHODS

The true oscillation amplitude is much smaller than that shown in Fig. 1. In the corresponding small-displacement approximation [see e.g. Frosch (2010a)] the sound-pressure p and the liquid-particle velocity v in a liquid of density ρ and of negligible compressibility and viscosity obey Newton's second law in the form

$$\rho \cdot (\partial \vec{v} / \partial t) = -\vec{\nabla} p \,, \tag{1}$$

and, in the present two-dimensional case, the Laplace equation,

$$\partial^2 p / \partial x^2 + \partial^2 p / \partial y^2 = 0 \,. \tag{2}$$

A possible (standing-wave) solution for p is:

$$p(x, y, t) = a_p(x, y) \cdot \sin(\omega \cdot t) \,, \tag{3}$$

where the angular frequency $\omega = 2\pi \cdot f$ is assumed to be constant. The real function $a_p(x, y)$ in Eq. (3) must fulfil the Laplace equation (2).

Liquid sound-pressure and streamlines: In case B (see Table 1), a solution compatible with Fig. 1 is obtained if $a_p(x, y)$ on the inside of the hollow cylinder is defined to be proportional to the real part of the analytic function $F(n_c) = n_c^{\,2}$ of the complex number $n_c = x + i \cdot y = r \cdot e^{i \cdot \varphi}$:

$$a_p = (a_{p0} / R^2) \cdot (x^2 - y^2) = a_{p0} \cdot (r / R)^2 \cdot \cos(2\varphi) \,. \tag{4}$$

In Eq. (4), R is the inner radius of the hollow cylinder, a_{p0} is a pressure constant, and r, φ are plane polar coordinates. In Fig. 2 (diagram on the left), lines of constant a_p are shown to be hyperbolae.

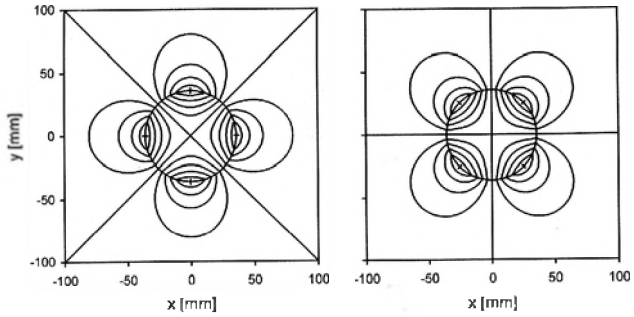


Fig. 2. Left: constant-pressure lines, $a_p / a_{p0} = 0.0, \pm 0.2, \ldots, \pm 1.0$, according to Eqs. (4) and (9). Right: streamlines according to Eqs. (6) and (10), for $N = 5$.

Eqs. (1), (3), (4) and the definitions $v_x = \partial \xi / \partial t$, $v_y = \partial \eta / \partial t$ yield the following equations for the displacements of the liquid particles from their no-wave place x, y:

$$\xi = \frac{2a_{p0} \cdot x}{\omega^2 \cdot \rho \cdot R^2} \cdot \sin(\omega t); \; \eta = -\frac{2a_{p0} \cdot y}{\omega^2 \cdot \rho \cdot R^2} \cdot \sin(\omega t) \,. \tag{5}$$

The *streamlines* of this liquid motion can be found by setting the *imaginary part* of the above-mentioned function $F(n_c) = n_c^{\,2}$ equal to a constant. The equation of streamline number n, where $n = 1, 2, \ldots, N$, is:

$$y = \pm n \cdot R^2 / (2N \cdot x) \,. \tag{6}$$

These streamlines are hyperbolae, too; see Fig. 2 (diagram on the right). A liquid particle touching the wall [no-wave coordinates $x = R \cdot \cos(\varphi)$, $y = R \cdot \sin(\varphi)$] has, according to Eq. (5), the following with-wave coordinates x_w, y_w:

$$x_w = R \cdot \cos(\varphi) \cdot [1 + \varepsilon \cdot \sin(\omega t)]; \; \varepsilon \equiv 2a_{p0} / (\omega^2 \cdot \rho \cdot R^2) \,; \tag{7}$$

$$y_w = R \cdot \sin(\varphi) \cdot [1 - \varepsilon \cdot \sin(\omega t)] . \quad (8)$$

The contours in the x_w-y_w plane defined by Eqs. (7) and (8) at $\omega t = 0, \pi/2, \pi,$ and $3\pi/2$ agree with the hollow-cylinder shapes shown in Fig. 1 (left).

In case C (Table 1), derivations like those just described, but based on the function $F(n_c) = n_c^{-2}$ of the number $n_c = x + i \cdot y = r \cdot e^{i \cdot \varphi}$, yield the following equations:

$$a_p = a_{p0}' \cdot (R'/r)^2 \cdot \cos(2\varphi); \; R' = \text{outer cylinder radius} ; \quad (9)$$

$$\text{streamlines: } r(\varphi) = R' \cdot \sqrt{(N/n) \cdot \sin(2\varphi)} ; \quad (10)$$

see Fig. 2 (regions outside of cylinder).

Prediction of oscillation frequencies: If the potential energy during the oscillation shown in Fig. 1 (left) is assumed to be due to the deviation of the local curvature radii of the half-thickness line from the no-wave radius $R_{av} = (R + R')/2$, then the following theoretical oscillation frequency is obtained [Frosch (2010a)]:

$$f_{th} = [h/(\pi \cdot R_{av}^2)] \cdot \sqrt{(Y \cdot h)/(5h \cdot \rho_w + b \cdot R_{av} \cdot \rho)} ; \quad (11)$$

here, $h = R' - R$ = wall thickness, assumed to be $<< R_{av}$; Y = elasticity modulus of wall; ρ_w = density of wall; b = number depending on the considered case (see Table 1).

Oscillation of hollow cylinder according to Fig. 1 (right): In that case, the functions $F(n_c) = n_c$ and $F(n_c) = n_c^{-1}$ yield straight equidistant lines on the inside and circles on the outside of the hollow cylinder; see Fig. 3 and Eqs (12),(13).

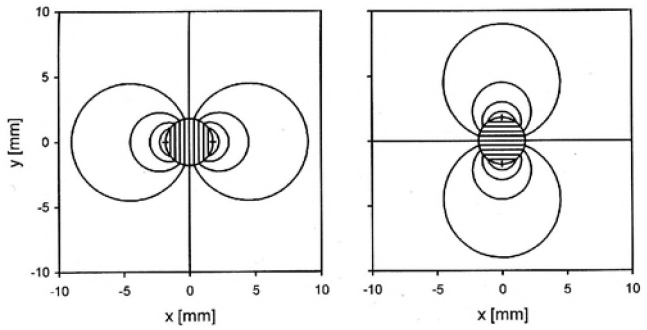


Fig. 3. As Fig. 2, but for hollow-cylinder oscillation as shown in Fig. 1 (right).

Liquid-pressure amplitude (Fig. 3, left):

$$r < R : a_p = a_{p0} \cdot x/R; r > R' : a_p = a_{p0}' \cdot (R'/r) \cdot \cos(\varphi) . \quad (12)$$

Streamlines (Fig. 3, right):

$$r < R : y = \pm R \cdot n/N; r > R' : r = R' \cdot (N/n) \cdot \sin(\varphi) . \quad (13)$$

The hollow cylinder and all liquid particles inside oscillate together in the x-direction.

The oscillation of an idealized tuning-fork prong is also illustrated by Fig. 1 (right) and Fig. 3. During a sinusoidal oscillation of frequency $f = \omega/(2\pi)$ and amplitude $\varepsilon' \cdot R'$, the maximal kinetic energies of the prong (radius R', height H, density ρ_{prong}) and of the surrounding water are [Frosch (2010a)]:

$$E_{prong} = (\pi/2) \cdot R'^4 \cdot H \cdot \rho_{prong} \cdot \varepsilon'^2 \cdot \omega^2 ; \quad (14)$$

$$E_{water} = (\pi/2) \cdot R'^4 \cdot H \cdot \rho \cdot \varepsilon'^2 \cdot \omega^2 . \quad (15)$$

Eqs. (14),(15) yield the following prediction for the tuning-fork frequency ratio $R_f = f(\text{in air})/f(\text{under water})$:

$$R_f = \sqrt{1 + \rho/\rho_{prong}} . \quad (16)$$

3. RESULTS

Insertion of the properties of the wineglass used for Table 1 into Eq. (11) yielded, for case A (i.e., $b = 0$), a prediction of $f_{th} = (0.70 \pm 0.10)$ kHz, lower than the experimental frequency of 1.04 kHz; such a discrepancy is expected, because the glass structure differs strongly from the assumed hollow cylinder (Fig. 1). Eq. (11) correctly predicts equal frequencies in cases B and C. Various hollow metal cylinders yielded good agreement of theory with experiment in case A [Frosch (2010a)]. A typical experimental D/A frequency reduction factor was (0.52 ± 0.01), larger than the factor of (0.42 ± 0.01) predicted by Eq. (11). That discrepancy is attributed to the small height H of the cylinder used $(H/R_{av} = 1.03)$; true streamlines agreed with Fig. 2 near half-height only. Insertion of the steel density $\rho_{prong} = 8.0\,\text{g/cm}^3$ into Eq. (16) yields a predicted frequency ratio of $R_f = 1.061$, corresponding to one semitone, in agreement with observation.

4. CONCLUSION

The experimentally observed oscillation frequency reductions caused by submerging resonators in water are consistent with the hypothesis that these reductions are predominantly due to the kinetic energy of the evanescent (standing) waves generated by the resonators.

REFERENCES

Frosch, R. (2010a). Introduction to Cochlear Waves. vdf, Zurich, pp. 257-279, 301-302, 423-426.

Frosch, R. (2010b). Analysis of Human Oto-Acoustic Emissions. Contribution to this conference.

PERFORMANCE PREDICTION VIA THE JAVA ACOUSTIC MODEL INTERFACE

G. H. Brooke[1], S.J. Kilistoff[2], D.J. Thomson[3], and D. D. Ellis[4]
[1] Brooke Numerical Services, 3440 Seaton St., Victoria, BC V8Z 3V9 gab.bns@gmail.com
[2] 5880 Bear Hill Road, Victoria, BC V9E 2J2 skilistoff@shaw.ca
[3] 733 Lomax Road, Victoria, BC V9C 4A4 drdjt@shaw.ca
[4]DRDC Atlantic, P.O. Box 1012, Dartmouth, NS, Canada, B2Y 3Z7 daledellis@gmail.com

1. INTRODUCTION

Performance prediction provides an important and essential link between the underwater acoustics research community and industry (environmental concerns) or the military (surveillance). Nominally, performance prediction (e.g. CASS-GRAB [Keenan, 2000] or SPADES [Brooke, 2008]) involves the use of an acoustics model to generate field predictions for a particular ocean environment given locations of particular sensors and targets; these data are then combined with some detection criteria and assumed (or computed) noise background to produce and display a measure of the performance of sensor against target [Ellis and Pecknold, 2010]. Currently, no one acoustics model can handle the wide range of frequencies implied by either military or industrial applications. The java Acoustic Model Interface (jAMI) is under development with the intent of providing a framework for a Client-Server approach to performance prediction in which the problem configuration and display reside on a thin Client and the models and other computational engines are centralized on a more powerful backend computer or Server.

In jAMI, the Client is programmed exclusively in Java whereas the backend consists of a combination of C and FORTRAN code designed to take advantage of as many existing and publicly available codes as possible. Currently, jAMI supports SAFARI [Schmidt, 1988], PECan [Brooke et al., 2001], POPP [Ellis, 1985], and BellHop [Porter, 1987] for simple propagation studies and the DRDC Clutter model [Ellis et al., 2008; Brooke et al., 2010] for reverberation and target predictions. A C-interface layer has been written that provides seamless interaction of all three languages. This same interface also provides the links to public domain environmental databases for bathymetry (Gebco), sound speed via temperature and salinity (World Oceanographic Atlas, WOA), and bottom composition (Deck41). Ultimately, the goals are to develop a chart-based application that performs transmission loss, target echo, reverberation, and ambient noise computations for real ocean environments obtained from the databases at specific locations as well as process pathological test cases supplied by the user. Standardized displays for environmental and acoustic parameters are under development, that coupled with the chart and database capabilities should result in a comprehensive and powerful performance prediction application.

In this paper we describe some of the software architecture considerations, the Java parameter interface and display capability, and finally present some transmission loss and reverberation results for select cases.

2. SOFTWARE ARCHITECTURE

A schematic of the jAMI application is shown below in Fig. 1. One of the main design considerations has been the separation of the parameter input and display capability (i.e., the Client) from the computational engine (i.e., the Server or backend). As illustrated, this has been effected using a TCP/IP socket connecting the Client to a C-interface layer on the Server side. This allows the Client to be programmed exclusively in Java; Java is freeware, relatively easy to port between operating systems, and has powerful Graphical User Interface (GUI) building capabilities. As shown, the Client allows the User to input parameter values, choose between models, choose between output field quantities, run the models, display the results, and, finally, using a special 'Browser' component, allows the User access to HTML HELP files (and the Web if desired).

The C-interface layer accepts data from the TCP/IP socket connection and then links up to the appropriate model and if necessary to the environmental databases. Currently, at time of writing, the model suite is limited to the existing FORTRAN codes (listed in the figure) but in future, models written in C could also be accommodated. Similarly, the code currently supporting the manipulation and extraction of environmental information from special, packed ASCII files is written in C but this does not preclude FORTRAN database code in the future. Finally, it is worth mentioning that the jAMI application can be run in various standalone configurations and the total package (including database files) fits comfortably on a modern laptop computer.

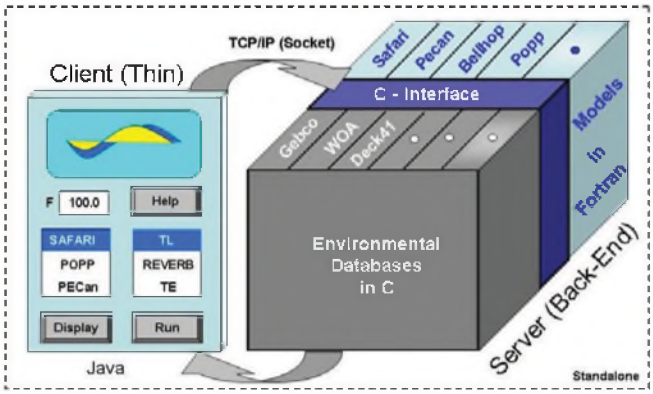


Figure 1. jAMI software architecture and configurations.

3. RESULTS AND EXAMPLES

Active sonar performance prediction involves the calculation of reverberation. jAMI currently employs the DRDC clutter model for such calculations. A typical 'clutter map' display of beam (bottom) reverberation as a function of range and azimuth is presented in Fig. 2 below. This pathological case and involves a 32-element horizontal line array oriented perpendicular to an iso-speed shallow-water wedge waveguide. The figure indicates more reverberation from the directions in which the water depths are shallowest, as expected.

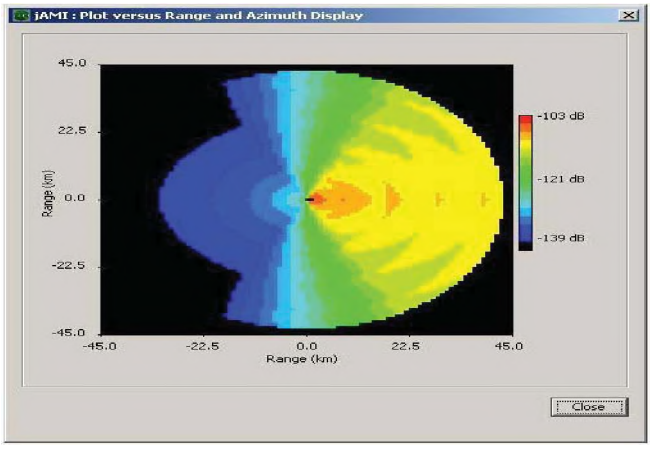


Figure 2. Beam reverberation plotted as a 'clutter map'.

A very important aspect of accurate performance prediction involves the proper use of acoustic models to account for the propagation effects. One of the primary design goals for the jAMI application was to allow access to a number of established propagation models via a common GUI and display. Plots of propagation loss versus range or depth provide useful information when deciphering performance issues. A typical range display is shown in Fig. 3 below where we plot propagation loss versus range for two different models (SAFARI and PECan) applied to the shallow water (100 m), Pekeris waveguide test case known as NORDA 3b [Davis et al., 1982]. The two models agree very well over the entire range as expected. These results were obtained by configuring the models in the Client,

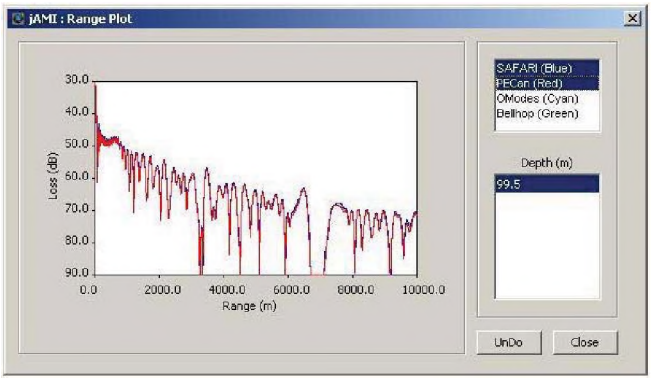


Figure 3. Propagation loss versus range (NORDA 3b test case).

running the models on the Server, and then displaying the results back on the Client.

4. DISCUSSION

The jAMI application is constantly evolving and being refined. One of the strengths of Java is the relative ease by which changes can be made to the GUI and display code. By design, jAMI has been modularized so that other models/databases can simply be 'plugged in'. Future plans include a chart display and environmental viewer/display functionality.

REFERENCES

Keenan R.E., (2000), "An Introduction to GRAB Eigenrays and CASS Reverberation and Signal Excess," in Proceedings of Oceans 2000, **2**, pp.1065-1070.

Brooke G.H., (2008),"An update on the multi-static model, SPADES, and its application to the reverberation workshop test suite," 154th Meeting of the Acoustical Society of America, 27-31 November, New Orleans, LA.

Ellis D.D. and Pecknold S.P., (2010), Range-Dependent Reverberation and Target Echo Calculations Using the DRDC Atlantic Clutter Model , to appear in Canadian Acoustics, (this issue) October.

Schmidt H., (1988), "SAFARI Seismo-Acoustic Fast Field Algorithm for Range-Independent environments," Rep. SR-113, SACLANTCEN ASW Research Centre, San Bartolomeo, Italy.

Brooke G.H., Thomson D.J., and Ebbeson G.R., (2001), "PECAN: A Canadian Parabolic Equation Model for Underwater Sound Propagation," J. Comp. Acoust., **9**, pp. 69-100.

Ellis D.D., (1985), "A two-ended shooting technique for calculating normal modes in underwater sound propagation. Report 85/105, Defence Research Establishment Atlantic, Dartmouth, NS, Canada, September.

Porter M.B. and Bucker H.P., (1987), "Gaussian Beam Tracing for Computing Ocean Acoustic Fields", J. Acoust. Soc. Am., **82**(4), pp.1349-1359.

Ellis D.D, Preston J.R., Hines P.C., and Young V.W., (2008), "Bistatic signal excess calculations over variable bottom topography using adiabatic normal modes," in *International Symposium on Underwater Reverberation and Clutter*, P. L. Nielsen, C. H. Harrison and J.-C. Le Gac, eds., NATO Undersea Research Centre, La Spezia, Italy, pp. 97-104.

Brooke G.H., Kilistoff S.J., and Thomson D.J.,(2010), "Interface Clutter Model with DRDC System Testbed – Phase II," Contractors Report, Contract number: W7707-098249/001.

Davis J.A., White D., and Cavanagh R.C., (1982), "NORDA parabolic equation workshop, 31 March-3April, 1981," Tech. Note 143, Naval Oceanographic Research and Development Activity, NSTL Station (MS).

ACKNOWLEDGEMENTS

Parts of this work were supported under contract to DRDC Atlantic.

BAYESIAN ACOUSTIC SOURCE TRACKING AND TRACK PREDICTION 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. OVERVIEW

This paper considers matched-field tracking and track prediction for a moving ocean acoustic source when 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. The algorithm involves two stages. The first stage (referred to as the tracking stage) consists of probabilistic tracking by inverting acoustic recordings of the source at a sequence of past times. For this problem, a Bayesian formulation is applied in which the posterior probability density (PPD) is integrated over unknown environmental parameters to obtain a time-ordered sequence of joint marginal probability surfaces over source range and depth, referred to as probability ambiguity surfaces (PASs). Due to the strong nonlinearity of the matched-field problem, this inversion is carried out numerically using Markov-Chain Monte Carlo methods. In particular, Metropolis-Hastings sampling is applied to environmental parameters (rotated into principal components) and two-dimensional Gibbs sampling to source locations to take advantage of fast computation of conditional probability distributions over range and depth using normal mode methods. This approach provides a large ensemble of track realizations drawn from the PPD [1, 2].

The second stage (the prediction stage) consists of applying a probabilistic model for source motion to each of the track realizations in the PPD ensemble obtained in the tracking stage, thereby producing a sequence of source range-depth probability distributions for future times. The particular source motion model applied here is based on the assumption of constant source velocity. In this case, the dependence of source range with time, $R(t)$, can be modelled using the law of cosines [3], as illustrated in Fig. 1(a):

$$R(t) = \left[R_0^2 + (v_h t)^2 - 2R_0 v_h t \cos\theta_0\right]^{1/2} \qquad (1)$$

where R_0 and θ_0 are the range and the angle between receiver-to-source radial and direction of motion at an initial time $t_0 = 0$ and v_h is the horizontal velocity. Tracking using acoustic data measured at a vertical line array (VLA) cannot determine horizontal coordinates (x, y), but only the range $R(t)$, as shown in Fig. 1(b). The closest point of approach (CPA) is defined to be the source position that minimizes

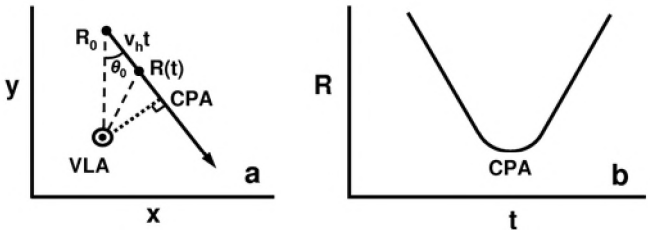


Figure 1. Track geometry in plan view (x-y) and range-time.

$R(t)$, as shown in Fig. 1(a) and (b). Equation (1) can be solved for track-parameter estimates $\hat{R}_0, \hat{v}_h, \hat{\theta}_0$ based on $R(t)$ values obtained by tracking inversion for a series of past times, and the uncertainty in the solution estimated using linearized inverse theory. These track parameter estimates can then be used to predict the source range at a series of future times using Eq. (1). Applying this procedure to every set of past ranges in the PPD ensemble from the tracking stage accounts for the uncertainty in the initial tracking, including the effects of environmental uncertainty. To account for uncertainty in the track prediction model, an ensemble of track predictions is drawn from the track-parameter uncertainty distribution for each set of past ranges. A similar procedure is applied to predict future source depths from past depth estimates (a simpler one-dimensional problem).

The result of the two stages described above is a very large ensemble of source tracks for future times, the variation of which quantifies the uncertainty in both past tracking and track prediction procedures. This ensemble can then be considered in terms of PASs for future times, and the most-probable track estimate/prediction (with uncertainties) can be extracted. Further, for incoming tracks, uncertainty distributions for range and time of CPA can be computed.

2. EXAMPLE

The section considers a simulated example of Bayesian track prediction. The unknown environment and source parameters of this example 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 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,

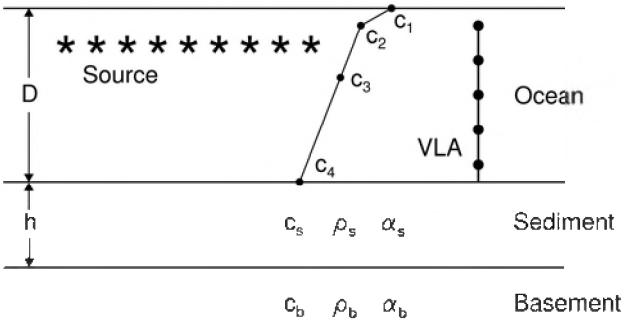


Figure 2. Experiment geometry and model parameters.

and D m. Wide uniform prior distributions (search intervals) are assumed for all parameters. Synthetic acoustic data were computed at a frequency of 100 Hz for 9 source locations at 2-minute intervals along an inbound track with a constant depth of 30 m and ranges defined by R_0=14 km, θ_0=14°, and v_h =8.5 m/s. These ranges and depths are shown in the left column of Fig. 3 (dotted lines). The data were computed at a VLA with 24 receivers at 4-m spacings from 26.120-m depth in 130-m of water. Random Gaussian errors of fixed variances were added to the synthetic data to achieve a mean SNR over the track of -2 dB.

Bayesian source tracking was applied to these data for a source search region of 0-20 km range and 0-129 m depth, with constraints on maximum allowable horizontal and vertical source velocities of 10 and 0.33 m/s, respectively. The resulting PASs are shown in the left column of Fig. 3. The track marginal distributions are multi-modal, with three distinct tracks of relatively high probability and at least one other with lower probability. These tracks differ in range, but are fairly consistent in source depth (near or slightly shallower than the true depth) and velocity. Interestingly, for this noise realization, the track that corresponds most closely to the true track is not the highest probability track, but the second highest. This example illustrates the power of sampling the PPD in the Bayesian approach, since algorithms based on estimating the highest-probability track would miss this important secondary track.

Probabilistic track prediction was subsequently applied for 9 future source locations at 4-minute intervals (i.e., twice the time interval for tracking), producing the sequence of PASs shown in the right column of Fig. 3. The track prediction results initially include peaks for the four distinct tracks estimated from acoustic inversion, but these coalesce within the first three time samples. Beyond this the PAS peaks occur at somewhat longer ranges and shallower depths than the true track, although the true track is encompassed by the high-probability regions.

3. SUMMARY

This paper developed and illustrated a probabilistic approach to the prediction of future locations of a moving ocean acoustic source based on probability distributions for

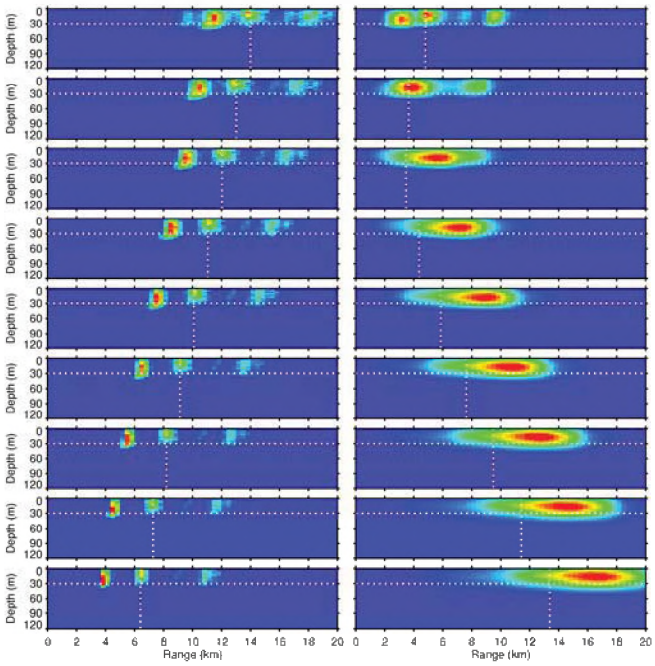


Figure 3. PASs for source tracking (left column) and track prediction (right column) for the synthetic example. Dotted lines indicate the true source range and depths.

past source locations, as estimated by a Bayesian acoustic tracking algorithm which accounts for environmental uncertainty. Markov-chain Monte Carlo methods were employed to sample the posterior probability density over unknown environmental parameters and past source locations, and a probabilistic prediction model for constant-velocity source motion based on the law of cosines was applied to each track-estimate realization in the ensemble to produce probability distributions for future source locations. The results were presented in terms of probability ambiguity surfaces (joint marginal PPDs over source range and depth), which quantify the information content of data and prior for track estimation and prediction.

REFERENCES

[1] Dosso, S.E. and M. J. Wilmut, 2008. Uncertainty estimation in simultaneous Bayesian tracking and environmental inversion, J. Acoust. Soc. Am., vol. 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., vol. 125, 717-722.

[3] M. J. Wilmut, J. M. Ozard, and M. L. Yeremey, 1997. Tracking in range versus time with applications to matched-field processing of vertical line array data, Canadian Acoustics, vol. 25, 21-27.

VERIFICATION OF A BUBBLE CURTAIN MODEL USING AN IMPULSE RESPONSE FUNCTION FOR A TOWED SOURCE

Caitlin O'Neill[1], Graham Warner[1], and David Hannay[1]
[1]JASCO Applied Sciences Ltd., 2101-4464 Markham St., Victoria, BC, Canada, V8Z 7X8
caitlin.oneill@jasco.com, graham.warner@jasco.com, david.hannay@jasco.com

1. INTRODUCTION

The effects of a bubble curtain on acoustic propagation from a towed source was investigated using a bubble curtain wavenumber integral acoustic model. The model was tested using an impulse response function as a source with direct path, surface reflected paths, and curtain reflected paths included.

2. METHODS

Sound propagation and back-scattering in bubbly water has been studied extensively in problems encountered in acoustic oceanography and ultrasonic imaging (Leighton, 1994). Even a very small fractional volume of air bubbles in water can significantly change the sound speed. The primary reason for this effect is that the compressibility of bubbly water is much greater than for non-aerated water. Sound pressure waves incident on the boundaries of bubbly water layers can be reflected strongly due to the large change in acoustic impedance across the boundaries.

The curtains were modelled as uniform layers 4.3 m in thickness, extending 20 m below the surface and separated by 36 m. The source was placed mid-way between the curtains at a depth of 5 m. The fraction of curtain containing air was 15% and each bubble had a 2.8 mm radius. Water sound speed was 1540 m/s.

Attenuation by the bubble curtain was calculated according to the method in Commander and Prosperetti (1989) and was incorporated in the air curtain layer reflection and transmission coefficients by using complex internal sound speeds. The relationship between the attenuation coefficient and frequency is shown in Fig. 1. Bubble layer sound speeds were computed to give frequency dependent curtain layer sound speeds from 105 m/s to 115 m/s, as shown in Fig. 2.

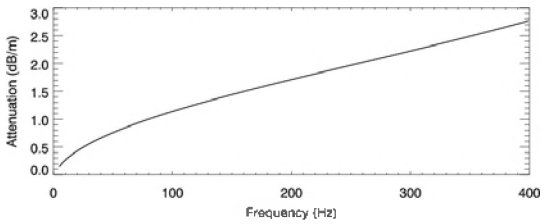


Fig. 1. Attenuation of the bubble curtain due to impedance mismatch, as a function of frequency.

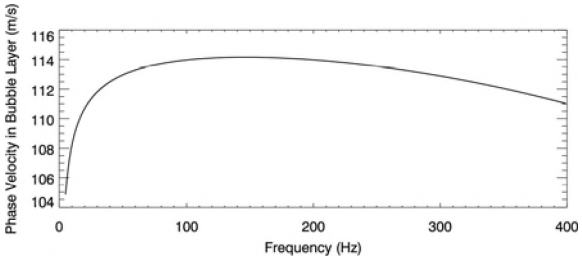


Fig. 2. Phase velocity in the bubble layer as a function of frequency.

Full waveform modelling was carried out from 5 Hz to 400 Hz. Frequencies below approximately 20 Hz are strongly attenuated by destructive interference of surface reflections (ghosts) that are treated here as surface images.

2.1 Wavenumber Integral Model

A wavenumber integral acoustic modelling method was used to compute the pressure fields presented here. Detailed descriptions of the approach are provided by Jensen *et al.* (1993) and Frisk (1994). The wavenumber integral approach is appropriate for the current problem of planar reflectors because it decomposes the spherical pressure field emitted by each of the airguns into a continuum of outward-propagating plane cylindrical waves. Reflections of the plane waves from the planar sea surface and air curtain walls were performed by treating each reflection as a multiplication by the appropriate plane wave reflection coefficient. Plane wave transmission through the air curtain upon each interaction was treated by multiplying by the plane wave transmission coefficient.

The model approach developed here makes two primary assumptions that have not been validated. The first assumption is that reflections from the curtains can be treated as mirror images with appropriate complex reflection coefficients applied. This would be a valid assumption if the curtains were infinite in planar extent, or at least much larger than the Fresnel zone size. The assumption is likely valid at high frequencies (shorter wavelengths) but may have reduced accuracy at lower frequencies. The second assumption is that diffraction around the bottom edges of the curtains is not important.

2.2 Method of Images

Source pressure signatures were modelled by using the method of images to account for multiply-reflected acoustic paths between the air curtains and from the surface. Because a wavenumber integral approach was applied, a solution was computed independently for each wavenumber before reconstruction by the integral. The horizontal wavenumber is preserved through reflection off the vertical reflectors and surface. This method is represented in Fig. 3, where the paths of several parallel rays, corresponding to one wavenumber, are shown (solid lines) and their images (dashed lines). For each reflection, the corresponding image position was calculated and a finite number of images were identified, corresponding only to the specular reflections that led to sound escaping in the direction of measurement. This calculation accounted for the geometry of curtains and was independent of receiver position. Transmission and reflection coefficients of the curtain were accounted for at each image and applied to waves propagating through the curtain and escaping under the curtains.

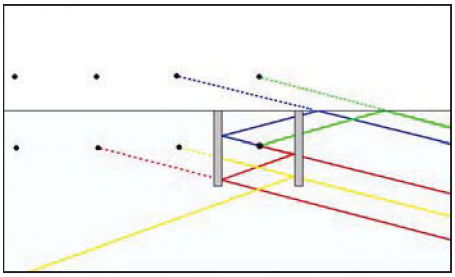


Fig. 3. Source reflection images for horizontal reflections off the bubble curtains and vertical reflections from the surface.

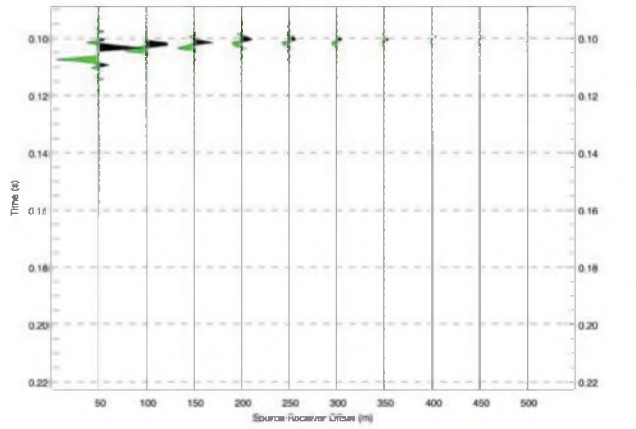


Fig. 4. Impulse response from a source at 6 m depth of direct and surface reflected paths without air curtains at ranges from 50 m to 500 m at a fixed receiver depth of 30 m. Signals have been time-aligned so the direct path signal occurs at approximately 0.1 seconds.

3. RESULTS

Testing of the acoustic model was carried out by examining the impulse response function with only direct path and surface reflected paths included. Fig. 4 shows the impulse response (5 Hz to 400 Hz) for a single towed source with no air curtains and no bottom reflections. The source depth was 6 m and the receiver depth was 30 m. This test shows the direct and surface reflected paths converging with increasing offset as expected. Fig. 5shows a scenario similar to that treated in Fig. 4, but with air curtains present. The individual paths are spaced as expected and with the proper polarity based on the respective numbers of curtain and surface reflections.

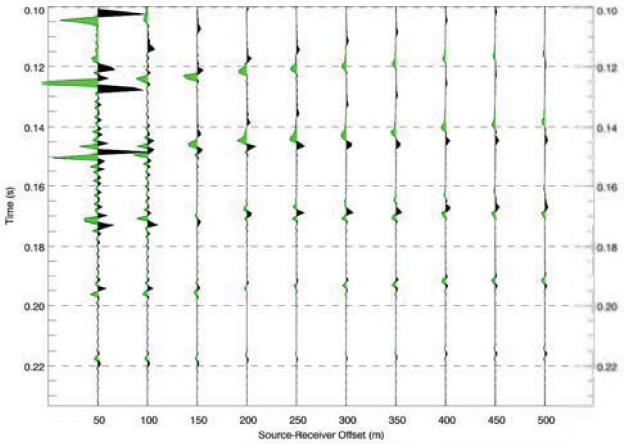


Fig. 5. Impulse response from a source at 6 m depth of direct transmission and multiple internal reflections with bubble curtains at ranges from 50 m to 500 m at a fixed receiver depth of 30 m. Signals have been time-aligned so the direct path signal occurs at approximately 0.1 seconds.

4. DISCUSSION AND CONCLUSIONS

Comparison of Fig. 4 and Fig. 5 show that the shielding effect is observed with the bubble curtain present. Interference between multiple reflections from the bubble curtain influenced the modelled impulse response of the towed source in the horizontal direction. Subsequent investigation will examine the effects of a bubble curtain on a towed airgun array.

REFERENCES

Commander, K.W. and A. Prosperetti (1988). Linear pressure waves in bubbly liquids: Comparison between theory and experiments. J. Acoust. Soc. Am. 85 (2): 732-746.

Frisk, G.V. (1994). Ocean and seabed acoustics: A theory of wave propagation. New Jersey: Prentice Hall.

Jensen, F.B., W.A. Kuperman, M.B. Porter and H. Schmidt (1993). Computational ocean acoustics. New York Springer: Verlag.

Leighton, T.G (1994). The acoustic bubble. San Diego: Academic Press.

ACKNOWLEDGEMENTS

Funding for the project was provided by Shell Offshore Inc.. Stress Engineering Services Inc. collaborated on the project and modelled the hydrodynamics of the bubble curtain. James Ferguson of JASCO Applied Sciences Ltd. contributed to the model development.

TEMPORAL ROBUSTNESS OF AN AUTOMATIC AURAL CLASSIFIER

Stefan Murphy and Paul C. Hines
Defence R&D Canada – Atlantic, P.O. Box 1012, Dartmouth, NS, Canada, B2Y 3Z7
stefan.murphy@drdc-rddc.gc.ca

1. INTRODUCTION

Active sonar systems are used to detect underwater manmade objects of interest (*targets*) that are too quiet to be reliably detected with passive sonar. In coastal waters, the performance of active sonar is often degraded by false alarms caused by echoes returned from geological seabed structures (*clutter*) found in these shallow regions. To reduce false alarms, a method of distinguishing target echoes from clutter echoes is required.

Research has demonstrated that perceptual signal features similar to those employed in the human auditory system can be used to automatically discriminate between target and clutter echoes, thereby improving sonar performance by reducing the number of false alarms [1]. The temporal robustness of this method is tested in this work by classifying recent echoes from 2009 using an automatic aural classifier previously trained with older (2007) echoes. Preliminary dependence on signal-to-noise ratio (SNR) is also presented.

2. EXPERIMENT

An active sonar experiment on the Malta Plateau was conducted during the Clutter'07 sea trial and repeated during the Clutter'09 sea trial. NATO Research Vessel Alliance ran a track to the southeast past Campo Vega Oilfield (ship track published in [2]). Broadband sources were used to transmit linear FM sweeps (600–3400 Hz) and a cardioid towed-array was used as the receiver. The sources and receiver were towed at a depth of 50 m. The original data set consists of over 95,000 pulse-compressed echoes returned from two underwater objects representing targets (an oil rig and a wellhead) and many geological clutter objects.

In order to avoid biasing the classification by SNR, the SNR distributions of target and clutter echoes are matched. After the SNR matching, approximately 25,000 echoes from Clutter'07 are used for *training* the aural classifier, and approximately 10,000 echoes from Clutter'09 are used for *testing* the classifier.

Many environmental factors that affect sonar echoes can change over a 2-year period; however, the aural classifier uses supervised learning and is unable to adapt to new data once trained. The two primary factors for this data set were the sound speed profile in the water column and the sea surface roughness.

Figure 1 shows that the sound speed profiles differed considerably between experiments. In 2007 the sound speed profile was downward refracting; in 2009 the profile was close to isovelocity.

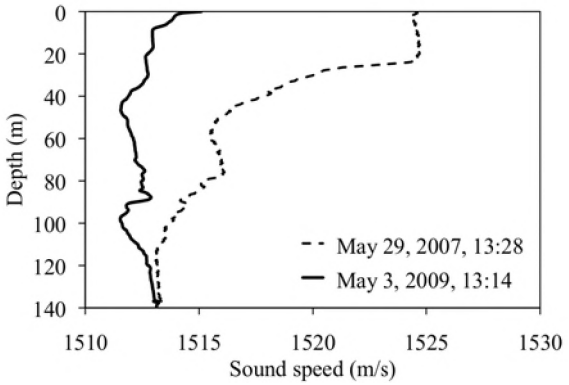


Figure 1 – Sound speed profiles for both experiments calculated from expendable bathythermograph (XBT) data.

Surface conditions also changed significantly between experiments. In 2007, the average relative wind speed of 13.0 knots resulted in Beaufort force 5–6 seas during the experiment. During the experiment in 2009, a lower average relative wind speed of 3.0 knots resulted in near flat seas. (Beaufort force 1).

3. AURAL CLASSIFIER

Details of the automatic aural classifier are published in [1]. The echoes are processed using a human auditory model that quantifies the timbre of each echo using 51 perceptual-based signal features that are not highly correlated ($r^2 < 0.81$) over the training data set.

The features are individually ranked by their ability to discriminate between target and clutter training echoes. Dimensionality is reduced by forming a subset of top ranked features, and further reduction is accomplished by principal component analysis. The number of top features and number of principal components are selected by the user.

A Gaussian classification method is used to calculate a target–clutter decision boundary in the feature space using the training echoes. The representation of the echoes was reduced to 2 dimensions by taking the first 2 principal components of the top 5 features. Figure 2 shows the minimum-error-rate decision boundary (circle) formed using Clutter'07 echoes. A scatter plot of Clutter'09 target and

clutter echoes is overlaid on the decision regions. Since the data set consists of thousands of echoes, the plot is limited to a representative sample of 60 echoes.

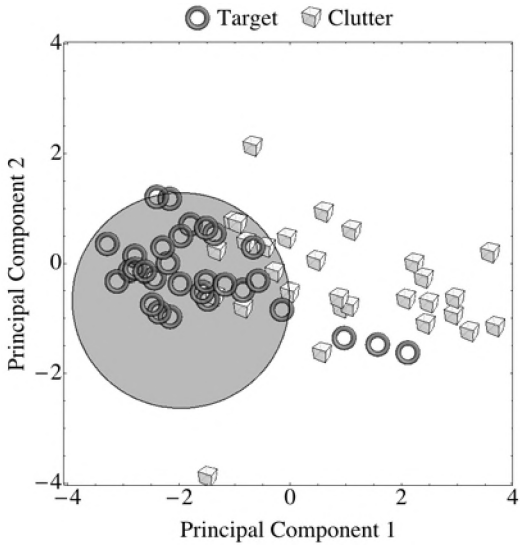


Figure 2 – Scatter plot of echoes in the reduced feature space. The gray circular target region contains 90% of the target echoes (toroids), and the surrounding white clutter region contains 64% of the clutter echoes (cubes).

4. RESULTS

4.1 Classifier performance

The minimum-error-rate operating point (shown in Figure 2) is chosen according to Bayes decision theory, with equal cost of misclassification for both target and clutter classes. In order to take all operating points into account, receiver-operating-characteristic (ROC) curves are used. ROC curves plot probability of detection versus probability of false alarm, and the summary performance metric used is the area under the ROC curve, A_{ROC}. For ideal classification, $A_{\mathrm{ROC}} = 1$, and if classification is performed by random guessing, $A_{\mathrm{ROC}} = 0.5$.

Since the number of features and principal components chosen are user-defined variables, it is possible to adjust them and monitor performance. Figure 3 shows a plot of classifier performance as grayscale intensity versus the number of top features used on the horizontal axis and the number principal components on the vertical axis. Darker color indicates higher performance. Since the number of principal components cannot exceed the number of features, data is constrained below the unit diagonal.

The peak performance ($A_{\mathrm{ROC}} = 0.903$) occurs at 29 features and 3 principal components. Having an A_{ROC} greater than 0.9 indicates a successful classifier, which demonstrates temporal robustness of the aural classifier. As a performance baseline, testing the classifier with the same Clutter'07 echoes it is trained with yields a peak A_{ROC} of 0.943.

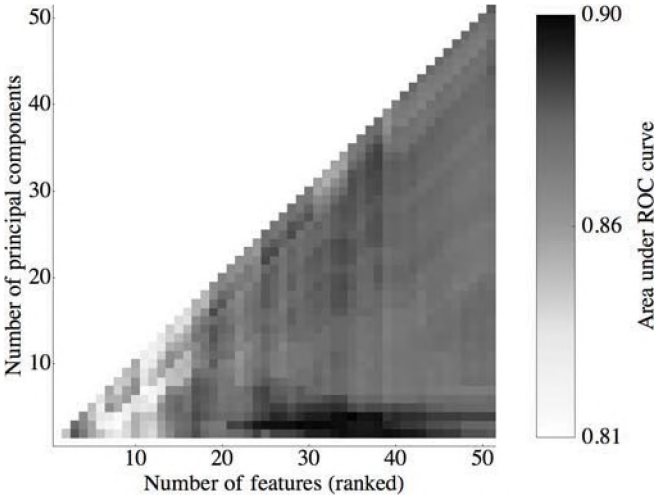


Figure 3 – Testing performance (A_{ROC}) of Clutter'09 echo classification using the aural classifier trained with Clutter'07 echoes.

5. CONCLUSIONS AND FUTURE WORK

The temporal robustness of the automatic aural classifier was demonstrated by classifying echoes using a classifier previously trained with echoes obtained 2 years earlier under different conditions.

Preliminary results show that classifier performance (A_{ROC}) increases with increasing echo SNR (dB); furthermore there is evidence that suggests the relationship is linear. This is an area of continued research.

REFERENCES

[1] Victor W. Young and Paul C. Hines, *J. Acoust. Soc. Am.* **122** (3), 1502 (2007).

[2] Paul C. Hines, Victor W. Young, and Jeff Scrutton, *Proceedings: International Symposium on Underwater Reverberation and Clutter* (ISURC, Lerici, Italy, 2008), pp. 293–301.

ACKNOWLEDGEMENTS

The authors would like to thank Doug Abraham, Charlie Gaumond, and Richard Menis for their contributions to the Clutter'09 sea trial, which was conducted by the NATO Undersea Research Centre (NURC) under Chief Scientist Peter Nielson. Funding provided in part by United States Office of Naval Research.

RANGE-DEPENDENT REVERBERATION AND TARGET ECHO CALCULATIONS USING THE DRDC ATLANTIC CLUTTER MODEL

Dale D. Ellis[1], and Sean P. Pecknold[2]
[1]DRDC Atlantic, P.O. Box 1012, Dartmouth, NS, Canada, B2Y 3Z7 daledellis@gmail.com
[2]DRDC Atlantic, P.O. Box 1012, Dartmouth, NS, Canada, B2Y 3Z7 sean.pecknold@drdc-rddc.gc.ca

1. INTRODUCTION

A shallow water reverberation model based on normal modes has been developed and refined at DRDC Atlantic over the years [Ellis, 1985; 1995]. Originally, the model handled range independent boundary reverberation in monostatic or bistatic geometries, including source and receiver beam patterns for comparison with measured data, and was later extended to handle target echo. The formulation was extended to range-dependent environments using adiabatic normal modes [Ellis et al., 2008]. The computations presented at that time used a Matlab/Fortran hybrid model with modes evaluated on a rectangular grid. While the scattering strength and echo at any point on the grid could be arbitrary, a constant water depth was still required. In 2009 a range-dependent Matlab/Fortran code was developed for monostatic reverberation calculations along a single radial [Kwan and Ellis, 2010]. In 2010, a model (implemented in Fortran 95) was developed to handle both sloping bathymetry, and towed array beam patterns in bistatic geometry on a 2-D grid. The model is computationally efficient and its capabilities are evolving. Work is in progress to implement it into jAMI [Brooke et al., 2010], as well as the DRDC Atlantic System Test Bed and Pleiades System which is sometimes used by the Canadian Forces. Model-data comparisons of towed array clutter data obtained on the Malta Plateau are underway, as well as comparisons with several range-dependent problems from the ONR Reverberation Modeling Workshop [Thorsos and Perkins, 2008], and the UK Institute of Acoustics workshop on Validation of Sonar Performance Assessment Tools [Zampoli et al., 2010].

This paper provides a brief description of the model and a few illustrative calculations of its output.

2. METHOD

The model assumes the environment is defined over a rectangular area. At each (x,y) point the required inputs are water depth, sound speed profile, bottom acoustic properties, scattering strength (surface and/or bottom); for the signal excess calculations an omni-directional target echo strength at specified depth z_T is specified. In addition a number of discrete targets can be specified. A source is at location (x_S, y_S, z_S) and towed array receiver at (x_R, y_R, z_R), heading in direction ϕ_R. The source is simply specified by a center frequency, intensity and pulse length, and can have a vertical beam pattern; the towed array can have beams steered in multiple directions (typically a line array of N hydrophones will have N independent beams at the design frequency). The beam steering angles can be specified, as well as the ambient noise on each beam.

The equations for the formulation in terms of normal modes are given in [Ellis et al., 2008]. The implementation requires the input of the environmental parameters on a grid of points. At each point the sound speed, acoustic properties, and scattering strengths can be different. The user can specify a number of computational parameters, including some normal mode computational controls, and the radials on which the towed array beam time series will be calculated. Two main calculations are done and written to data files for display later: (1) reverberation and target echo at the grid points, assuming ideal "wedgie" receiver beam patterns of uniform response over horizontal beam width ϕ_H, and no sidelobes; (2) beam time series for each towed array steering direction. The first calculation is intended primarily for illustrative purposes and at this point a number of short cuts have been made in the coding; the second is for comparison with data, and the intent is to keep improving the fidelity.

3. RESULTS

Figure 1 shows the signal excess (target echo to reverberation level in dB) on a grid generalized from Fig. 2 in Ellis et al. [2008]. The area is 100 km square with (x, y) coordinates between (-50,0) and (50,100). A 50x50 grid is used for the calculations; [a 51 by 51 might have been nicer], so the centre of the grid points range from approximately (-49,1) to (49,99) with increments of 2 km in each direction. The water depth is 100 m, except for two ridges in the y-direction rising to 60 m, and another ridge rising to 70 m in the x-direction; both have gaps near the middle. There is a single seamount of height 50 m near (-40,85). The bottom has Lambert scattering with a strength of -27 dB, except for a +10 dB enhancement along the line (2,2) to (50,50). Similarly the target (at depth 10 m) has echo strength of 8 dB, except for a 7 dB enhancement along the line (-48,52) to (2,2). The source is at (-10,48) at depth 30 m, and the receiver at (10,48) at depth 50 m.

The basic environment is similar to the ONR 3D problems, and described by Zampolli et al. [2010]. It has isospeed water of 1500 m/s over a sand bottom half space of relative density 2.0, sound speed 1750 m/s and attenuation of 0.5 dB/wavelength; the volume absorption in the water is a version of Thorp's formula.

The source is omni-directional with unit energy (10 dB source level for a duration of 0.1 s). The frequency was 250 Hz. The towed array was chosen to give a horizontal beam width of 3.6°; 39 omnidirectional elements at spacing of 2.5 m with Hann weights were used for the beam time series. The CPU time on a 2GHz computer was only 1.5 s, with modes (usually 16) being calculated each of the 2500 grid points.

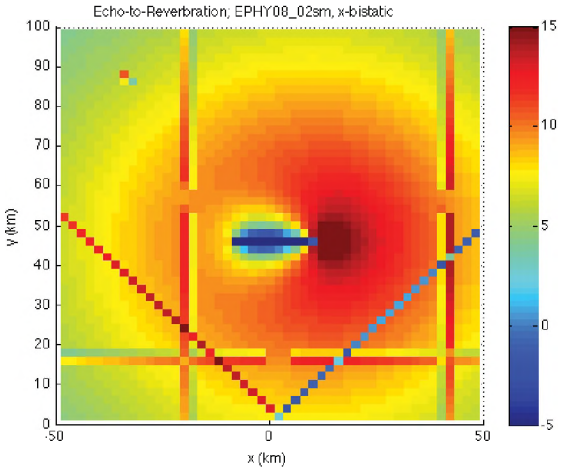


Figure 1. Signal excess on grid using wedgie beam patterns.

In Fig.1 the receiver sees high reverberation (low signal excess) in the direction of the source (west, W); similarly along the high scattering line to the SE. The higher echo along the line to the SW shows up clearly too. Along the 3 ridges, the signal excess first drops (due to the higher reverberation on the up-slope), then increases on the down slope. Beyond the ridges one would expect some shading due to mode cutoff, but perhaps the reverberation and echo are affected similarly. The single point to the northwest affects the signal excess in the adjacent cells.

Figure 2 shows time series with geometry corresponding to Fig. 1, and a selection of beam steering angles relative to the towed array heading of 225°. The predictions have not been checked out thoroughly, but generally seem to make sense. The reverberation on the omni receiver is 10–15 dB above the beam predictions; at short times the reverberation on the 45° and 60° beam are higher since they look in the direction of the source (note the beams have left-right symmetry); the 45° beam also seems to be picking up backscatter from the ridges to the S and W; the 0° and 359° beams are essentially identical; the 90° and 270° beams should be identical, and have the lowest reverberation, except in the region of features; at long ranges the endfire beams (0° and 180°) should have the highest reverberation, and in a uniform environment approach each other.

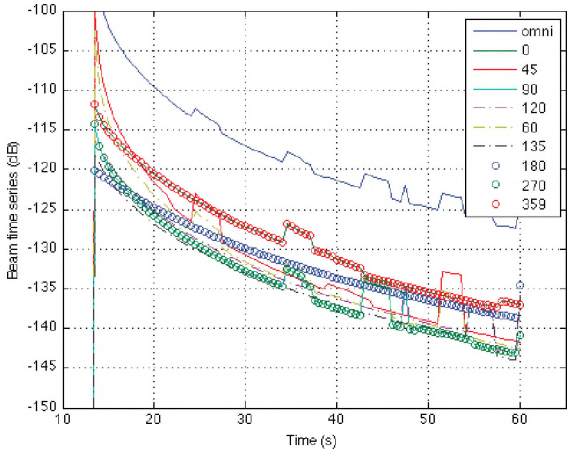


Figure 2. Beam time series of reverb. corresponding to Fig. 1.

4. DISCUSSION

The model continues to evolve. Presently the same sound speed profile and bottom loss are used at all locations. The next step will be to generalize it. For production calculations, it makes sense to pre-calculate the modes on the grid, perhaps interpolating them to a finer scale. Then for other source-receiver geometries and multistatic scenarios they can be re-used for the reverberation and target echo calculations.

REFERENCES

G. Brooke, S. Kilistoff, D. Thomson, and D. Ellis, Performance Prediction via the java Acoustic Model Interface, to appear in Canadian Acoustics, (this issue) October 2010.

D. D. Ellis. A shallow-water normal-mode reverberation model, J. Acoust. Soc. Am. 97:2804–2814, 1995.

D. D. Ellis, J. R. Preston, P.C. Hines, and V. W. Young. Bistatic signal excess calculations over variable bottom topography using adiabatic normal modes, in Neilsen et al., 2008, pp. 97–104.

P. L. Nielsen, C. H. Harrison, and J-C. Le Gac, eds. International Symposium on Underwater Reverberation and Clutter, NATO Undersea Research Centre, La Spezia, Italy, 2008,.

D. D. Ellis. A two-ended shooting technique for calculating normal modes in underwater sound propagation. Report 85/105, Defence Research Establishment Atlantic, Dartmouth, NS, Canada, Sept. 1985.

T. Kwan and D. D. Ellis, Reverberation calculations over sloping ocean bottoms. Technical Memorandum 2009-192, DRDC Atlantic, Dartmouth, NS, Canada, June 2010.

ONR (2006) ftp://ftp.ccs.nrl.navy.mil/pub/ram/RevModWkshp_I/Problem_Definition_documents.

E. I. Thorsos and J. S. Perkins, Overview of the Reverberation Modeling Workshops, in Neilsen et al., 2008, pp. 3–14.

M. Zampolli, M. A. Ainslie, and P. Schippers. Scenarios for benchmarking range-dependent active sonar performance models, to appear in Proc. Institute of Acoustics, 32, Pt. 2, 2010.

ACKNOWLEDGEMENTS

Paul Hines encouraged the development of this clutter model for the DRDC PLEIADES System.

Parts of this work were supported by the US Office of Naval Research, Code 32.

OCEAN BOTTOM REFLECTION LOSS FROM ELASTIC SOLID MATERIALS: REFLECTIONS ON REFLECTIVITY

N. Ross Chapman[1] and Hefeng Dong[2]
[1]Earth and Ocean Sciences, University of Victoria, Victoria, BC, Canada, V8P5C2; chapman@uvic.ca
[2]Norwegian University of Science and Technology, NO-7491 Trondheim, Norway; hefeng@iet.ntnu.no

1. INTRODUCTION

In large regions of the Pacific Ocean the ocean bottom is characterized by a thin layer of sediment over an elastic solid crust. The presence of the solid basement close to the sea floor generates additional losses that affect long range acoustic propagation. This paper describes geoacoustic inversion of reflection loss versus angle data at very low frequencies (~16 Hz) to obtain estimates of the compressional (p) and shear (s) wave velocities in uppermost oceanic crust, Layer 2A. The data were obtained in experiments using small explosive charges and a horizontal line array. The measurements were made at thin sediment sites located at increasing distance from the spreading centre at the Endeavour segment of the Juan de Fuca Ridge of the west coast of British Columbia to determine the effect of crustal age of the basalt on the seismic velocities.

The results for very low frequency data are sensitive to the thickness of the sediment layer, which increases with distance from the spreading ridge, resonances caused by interference of p- and converted shear waves, and possibly Stoneley waves excited at the sediment/basalt interface.

2. EXPERIMENTS

The reflection loss experiments were carried out at three sites along a transect west of the Endeavour segment of the Juan de Fuca Ridge. The ridge axis is aligned approximately 023°, and the general topography to the west consists of a series of valleys and ridges aligned parallel to the axis. The entire region is very thinly sedimented, less than 25 m out to distances of 35-40 km from the spreading centre. The sites were located in the valleys between the ridges: site A in the first valley 3 km west of the ridge crest at a depth of 2250 m; site B in the next valley approximately 18 km west of the crest at a depth of 2440 m; and site C in the third valley approximately 32 km west of the crest at a depth of 2640 m.

The reflectivity measurements were made using two ships according to a broadside experimental geometry, in which the shooting ship opened range on a course of 65° with respect to the array ship's course. For this geometry, the measurement provided an average of the crustal velocity over a distance of ~25 km in the direction of the track. The acoustic paths were perpendicular to the ridge axis.

One ship towed a 40-channel, 1500-m horizontal line array at a depth of 250 m, and the other ship deployed small explosive charges as the vessels opened range along the tracks. The shot spacing was designed to provide data for angles from 10° to 80° at 2° intervals, and the shot depth was ~200 m. Ship-to-ship ranges were measured using a GPS navigation system. The data were processed in a 1/3-octave frequency band at 16 Hz, and spatially filtered using a time-delay beamformer to measure the array response at the specular angle for the first bottom reflection path. The specular beam data were time-windowed to extract the bottom reflected signal. Significant scattering in non-specular paths was observed at higher frequencies at each site, but the impact of this effect was considerably reduced by processing the specular beam for the 16-Hz data.

The bottom reflection loss, BL (in dB), was determined from measurements of the acoustic propagation loss for the signal in the specular beam according to

$$BL(\theta) = (H_m(\theta) - Hc + 6)/n, \quad (1)$$

where H_m is the measured loss in dB for the nth bottom reflection, H_c is the calculated spreading loss, θ is the angle of incidence and 6 dB accounts for the four acoustic paths in each order of bottom reflection. The propagation loss was determined from the received energy level, RL, in the specular beam, using known values of the source levels, SL, of the explosive charge [1],

$$H_m(\theta) = SL - RL(\theta) \text{ (in dB)}. \quad (2)$$

The incidence angles at the bottom were determined from the experimental geometry using ray theory and a measured water velocity profile.

3. GEOACOSUTIC INVERSIONS

An example of the averaged reflection coefficient versus angle data for the 16-Hz band is shown in Figure 1 for site A. The maximum around 38° is associated with the p-wave critical angle in layer 2A, and the low reflectivity at higher angles is related to losses due to s-wave propagation in the sediment and upper crust. The relatively low reflectivity at very low angles (> 75°) suggests that there is no shear wave critical angle in the upper crust.

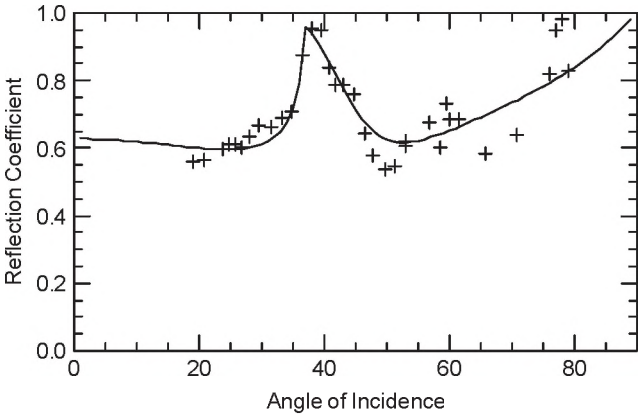


Fig. 1. Reflection coefficient versus angle for site A (~0.1 Ma.). The p-wave critical angle corresponds to the peak at about 38°.

The sound reflected from the ocean bottom carries information about the interaction with a system of interfaces, including the basalt crust and the overlying thin sediment layer. It would be expected that the presence of a thick sediment layer would have to be taken into account for interpreting the acoustic reflectivity, but a thin layer can also cause significant modifications to a simple half-space interpretation of the reflection process. Hovem [2] has shown that the sediment layer generates Stoneley waves at the sediment/basalt interface for certain frequencies and high incidence angles. Also for sufficiently thin elastic solid sediment layers, a resonant condition occurs due to the conversion of p- to s- waves at the sediment-basalt interface [3]. This effect has been shown to cause variations of several decibels in the reflection loss at long ranges or high incidence angles. These resonance effects complicate the interpretation of the experimental data, and it is important to account for them in modeling the reflection loss in the inversion.

Accordingly we use a geoacoustic model consisting of a simple thin sediment layer over an elastic solid basalt half-space. The model was parameterized by the compressional and shear wave velocities, $v_{p,s}$, and attenuations, $\alpha_{p,s}$, the densities, ρ of the sediment and basalt, and the thickness of the sediment layer, H. The parameter values are typical of clayey silt or silty clay. The water layer was given a sound speed of 1500 m/s and density of 1.03 g/cm^3.

In this work, the reflection coefficient versus angle data are inverted using a Bayesian approach to estimate the posterior probability distribution for the parameters of a specific geoacoustic model. In the Bayesian approach, the complete solution of the inverse problem is given in terms of an *a posteriori* probability distribution that specifies the probability of each possible model within the *a priori* limiting values for each model parameter.

The most sensitive parameters in the inversions were the p- and s-wave velocities of the basalt, and the sediment layer thickness. The shear wave speed in the sediment was also sensitive, indicating that the resonance due to converted shear waves was a significant factor in the reflections. The estimated seismic velocities for the uppermost portion of layer 2A and the sediment thickness are listed in Table I for the young crustal sites. The uncertainties are derived from the 1-D marginal distributions for each parameter that were obtained in the inversion. An example of the reflection coefficient calculated using the estimated properties from the inversion is shown in Figure 1 by the solid curve.

Table I. Estimated values of basalt velocities and the thickness of the sediment layer for very young crust.

Site	H (m)	V_p (m/s)	V_s (m/s)
A	6.9 ± 3.8	2547 ± 30	725 ± 178
B	26 ± 3.0	2626 ± 20	1014 ± 35
C	17 ± 2.0	2710 ± 18	1320 ± 46

The relatively low values for the p-wave velocity out to 1.5 Ma suggest that the western portion of the Endeavour segment is open to hydrothermal circulation, and the aging process is continuing at all sites. In contrast to the gradual change in p-wave velocity, there is a significant change in the s-wave velocity over the span of crustal ages observed in this experiment. These results provide new constraints for modeling the porosity in the crustal low velocity layer. The very large value of Poisson's ratio at the youngest site suggests that thin cracks mostly remain open. At the older sites where the sediment cover is thicker, the decrease in Poisson's ratio suggests that the thin cracks are being filled, but the thicker cracks remain open.

REFERENCES

[1] Chapman, N.R. "Source levels of shallow explosive charges", *J. Acoust. Soc Am.*, 84, 697-702, (1988)

[2] Hovem, J.M., and Åge Kristensen, "Reflection loss at a bottom with a fluid sediment layer over a hard solid half-space," *J. Acoust. Soc. Am.*, **92** (1), 335-340 (1992)

[3] S.J. Hughes, D.M.F, Chapman and N.R. Chapman, "The Effect of shear wave attenuation on acoustic bottom loss resonance in marine sediments," *Shear Waves in Marine Sediments* (J. Hovem, ed.), 439-446, Kluwer Academic Publishers, Dordrecht (1991)

MUSICAL NOISE LIMITS TO SPEECH ENHANCEMENT

Brady Laska
Research in Motion Ltd., 295 Phillip St. Waterloo, ON, N2L 3W8 blaska@rim.com

1. INTRODUCTION

Musical noise is a term used to describe short-duration narrowband artifacts present in speech processed by spectral modification noise suppression systems. The phenomenon is most easily understood within the context of the spectral subtraction algorithm [1]. Fig. 1 presents a block diagram of magnitude spectral subtraction. The noisy signal $z[n]$, consisting of the clean speech signal $x[n]$ and the additive noise $v[n]$, is divided into overlapping blocks, then transformed to the frequency domain via the fast Fourier transform (FFT). An estimate of the noise signal FFT magnitude is subtracted from the noisy speech magnitude, and the result is re-combined with the noisy FFT phase to reconstruct the enhanced signal in the time domain. Due to stochastic fluctuations, the actual noise magnitude in a given FFT bin will differ from its estimate. When the true value is lower than the estimate, the noise at that frequency is completely removed; when it is higher, some residual noise will remain. This successive elimination and under-suppression produces isolated peaks in the time-frequency noise representation. This is illustrated in Fig. 2: the top plot shows the spectra of two successive noise-only frames input to a spectral-subtraction system along with the estimated noise spectrum (thick solid); the bottom plot shows the enhanced output spectrum with clearly visible isolated spectral peaks. When converted back to the time domain, these peaks become short-lived tones that randomly vary in frequency. This type of on-off tonal switching is the most well-known manifestation; however any significant random modulation in the noise spectrum will create musical-noise type artifacts. Depending on bandwidth of each FFT bin, the sound of the artifacts may range from a tinkling bell to beeping tones to flowing water. These artifacts are so distracting and un-natural sounding that listeners generally prefer the original noisy signal to the processed speech with musical noise.

The human auditory system is tuned to recognize spectral change. When presented with a tone burst, the discharge rate of auditory nerve fibers will rapidly rise, and then gradually decay to the background level, even as the tone persists [2]. This response to spectral change is beneficial for extracting short-lived consonants and speech onsets in low signal-to-noise ratio (SNR) environments. It may also explain why musical noise is so objectionable: While listeners can adjust to a steady-state noise, a randomly modulated noise constantly reminds the listener of its presence. To achieve a natural sounding output that preserves the character of the residual noise, all noise-only spectral components must receive the same attenuation regardless of their absolute magnitude.

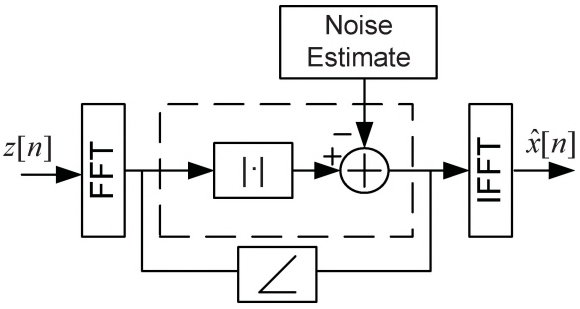


Figure 1. Block diagram of spectral subtraction speech enhancement.

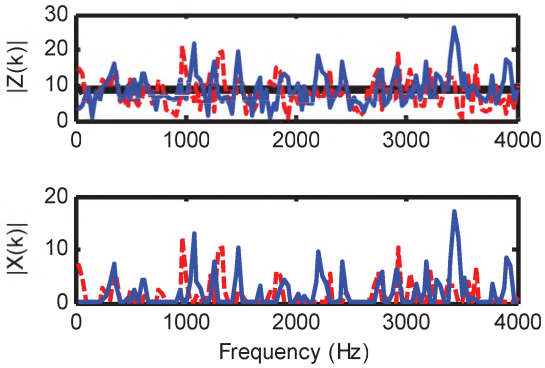


Figure 2. Example of spectral-subtraction processing of noise.

The damaging impact of musical noise on perceived speech quality is well known, and much research has been undertaken to mitigate its impact; however little has been done to understand musical noise on a theoretical basis. In this work we use statistical noise models to demonstrate how musical noise arises in speech enhancement, to characterize the sensitivity of different noise suppression algorithms to artifacts, and to understand the limits to artifact-free noise suppression.

2. GAIN FUNCTION SENSITIVITY

Fig. 3 plots the gain functions of the Wiener filter and the power and magnitude spectral subtraction speech enhancement algorithms as a function of *a posteriori* SNR, $\gamma = 20\log_{10}(|Z(n)| / |V(n)|)$ dB. The plotted attenuation is limited to -25 dB, though the theoretical gain at 0 dB (noise-only) is $-\infty$ dB. At non-zero SNRs, corresponding to intervals where speech is present, the slope of the gain functions is low, so SNR errors have a minor impact. In contrast, around 0 dB the gain functions are very steep, which means that small errors in SNR estimation are amplified to produce large gain variations. Since speech can mask its presence, musical noise is most noticeable in noise-only regions. Thus, the impact of SNR errors is greatest at SNRs where gain fluctuations are most

noticeable. The Wiener gain is the minimum mean-squared error (MMSE) optimal linear estimator of the speech FFT coefficients; however, computing the gain requires knowledge of the SNR, therefore the gain is only optimal insofar as the SNR estimate is exact. However, noise is a random signal; therefore there will always be some uncertainty in the SNR estimate.

2.1 Noise Estimation Error

The FFT coefficients of acoustic noise signals are commonly modeled as complex Gaussian random variables, leading to Rayleigh distributed spectral amplitudes [3]. The Rayleigh distribution is a single parameter distribution with mean and cumulative density function (cdf) given by: $\mu = \sigma\sqrt{\pi/2}$ and $F_X(x) = 1 - e^{-x^2/2\sigma^2}$.

Since the noise is mixed with the desired speech signal, we can only obtain smoothed estimates of the noise statistics. In a stationary environment, a noise estimator will converge to the mean of the noise amplitude distribution. To illustrate the effect of noise fluctuations on a given speech enhancement system, we define the noise to expected-noise ratio as NENR = $20\log_{10}(|V(n)|/\mu)$ dB. During noise-only periods, the NENR is the SNR seen by the speech enhancement system as a result of noise fluctuations. Since 1-$F_X(x)$ is the probability that the random variable X will exceed x, this can be used to give the probability of a given NENR. Fig. 4 plots the gain error as a function of NENR, as well as the probability of the NENR, for Rayleigh distributed noise with parameter $\sigma = 1$.

3. DISCUSSION

Fig. 4 provides some additional insight into the musical noise performance of the enhancement algorithms. For example, there is a 25% probability that the NENR will exceed 2.5 dB. At this NENR the gain error is about 12 dB for magnitude spectral subtraction, 17 dB for the Wiener filter and over 21 dB for power spectral subtraction. This high probability of large fluctuations is why the power spectral subtraction algorithm exhibits the highest levels of musical noise.

Fig. 4 can also be used to explain the fundamental limitations of spectral modification enhancement. The true noise (and NENR) distribution cannot be controlled by the system designer; only the gain function can be adjusted. Classical approaches to musical noise reduction involve a combination of over-subtraction and spectral flooring.

Over-subtraction artificially inflates the noise estimate, which reduces the NENR, thereby shifting the gain functions to the right and reducing the probability of large gain errors. Spectral flooring puts a limit on the amount of attenuation that can be applied, reducing the dynamic range of the possible gain. This has the effect of shifting the gain curves down, so the error for a given NENR is reduced and the maximum gain error is constrained. While these approaches offer some musical noise control, spectral flooring limits the noise attenuation and over-subtraction increases the probability that a low-level speech component will be attenuated or removed.

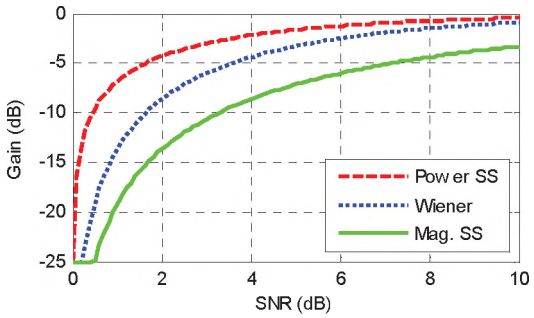


Figure 3. Spectral amplitude enhancement gain functions.

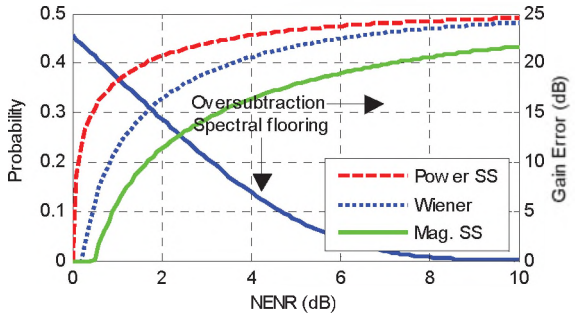


Figure 4. Gain error resulting from noise fluctuations.

4. SUMMARY

Stochastic variations of the noise signal from its expected value prevent us from obtaining exact SNR estimates. In most spectral modification speech enhancement systems, small SNR errors during noise-only periods can produce large fluctuations in the applied gain which modifies the character of the residual noise and resulting in musical noise artifacts. The need to prevent the emergence of these highly objectionable artifacts limits the amount of noise attenuation that can be applied without severely distorting the speech signal.

5. REFERENCES

[1] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoustics Speech Signal Process.*, vol. 27, pp. 113 –120, Apr. 1979.

[2] B. Delgutte, "Auditory neural processing of speech," in *The handbook of phonetic sciences*, pp. 507 – 538, Oxford: Blackwell, 1997.

[3] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," IEEE Trans. Acoustics Speech Signal Process., vol. 32, no. 6, pp. 1109 – 1121, 1984.

SIGNAL CHARACTERIZATION OF OCCLUDED IN-EAR VERSUS FREE-AIR VOICE PICKUP ON HUMAN SUBJECTS

Antoine Bernier, and Jérémie Voix
École de Technologie Supérieure, Université du Québec, 1100 Notre-Dame Ouest, Montréal (QC), Canada, H3C 1K3
antoine.bernier@polymtl.ca jeremie.voix@etsmtl.ca

1. INTRODUCTION

While speaking, a human subject will perceive his own voice mostly through the free-air radiation path. Nevertheless, the vibration of his vocal tracks and mouth cavity are also inducing a bone (skull) conducted vibration that will excite the cochlea. This phenomenon is inherently part of one's perception of his voice.

When the ear canal is occluded with an in-ear device (such as an earplug) providing a good acoustic seal (from the external environment) while leaving the ear canal walls free to vibrate (device shallow inserted), the regenerated sound pressure level inside the occluded ear canal can be very clearly perceived by the human subject (it is called the occlusion effect), while the free-air radiation path will no longer be significant to the resulting perception.

While the occlusion effect is often perceived negatively, it also enables the pickup of one's own voice from within his occluded ear canal. A small microphone placed in a well occluded ear will provide a strong signal, free from external disturbances. However, the resulting voice-pickup will be perceptually different, usually colored and not as "clear" sounding as a free-air propagating voice taken under the same conditions, unless extra signal processing is applied.

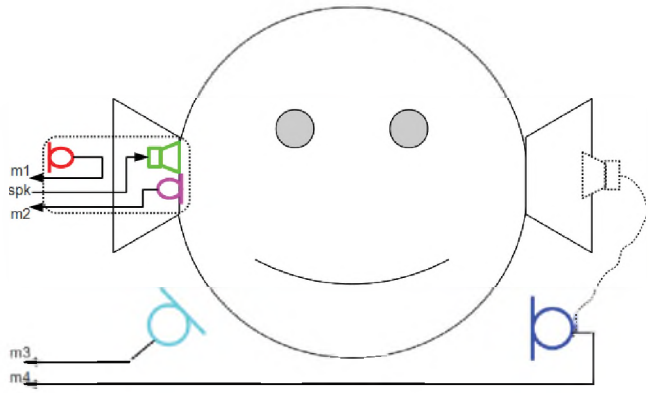
The purpose of the current research study is to characterize the differences in the signal characteristics of the occluded ear voice pickup relatively to the free-air voice pickup, in the form of a transfer function between the two signals: A first step to identify the signal processing parameters that would be required to alter the in-ear voice pickup to have both signal perceptually similar to the human ear.

2. METHOD

2.1 Test procedure and Data Acquisition

Nine subjects (index *s01* to *s09*) of different age, sex and physiological characteristics were guided through to same procedure to acquire the test signals: The subjects were to read seconds a text in their first language (French, denoted *fr* or English, denoted *en*) for 40 seconds while wearing a special in-ear device. Three bilingual subjects were retested in their second language. The in-ear device used is a new communication earpiece developed by Sonomax [1] and designed to instantly custom-fit all ear canals. For this experiment, the same earpiece was reused and, in order to achieve a good seal for each subject, a layer of disposable and malleable wax was added around the curved cone-shaped end of the earpiece.

The original in-ear device, illustrated in Fig. 1, enclosed two microphones: a first microphone (*m1*) inside the earpiece for in-ear voice-pickup through a sound-bore, the second (*m2*) on the outer faceplate of the in-ear device, pointing outward for free-air pick-up. A high fidelity free-field reference microphone (*m3*) is also present. The subjects wore the device on their right ear, and were equipped on the other ear with a commercial hand-free cell phone apparatus composed of an earphone (in dotted line in Fig. 1) with a wire-hanging microphone (*m4*). Although present in the original earpiece, the speaker (*spk*) was not used for this experiment.



All the 4 microphone signals were recorded through National Instruments PXI 4410 data acquisition cards at sample frequency 44.1 kHz with 24 bits resolution. The NI cards were interfaced with Matlab [2] that ran all the post-processing of the signals.

2.2 Post-Processing

To obtain the transfer function between the signals of interest, a dual channel FFT analysis [3, 4, 5] was conducted. To compute the FFT, the 40 seconds long signals were divided into the smallest power of 2 frame length offering experimental stability (from where larger frame length no longer modified the results), and the resulting FFT were used to obtain the Cross-spectra and Auto-spectra. These spectra were then averaged and used to compute the best transfer function estimate. The algorithm was tested and validated using signals altered by known transfer functions. Validity of the transfer function estimate over the speech frequency bandwidth was monitored by calculating the coherence between the signals. Presence of coherence would mean a given speech signal frequency component was picked up by both microphones, meaning it is carried by both the free-air and the bone conducting path, hence permitting a valid transfer function estimation at this point. A lower or no coherence (values lower than 0.1)

would either mean there is no signal to analyze (outside of the speech frequency bandwidth) or that it is absent on one of the microphone (if the bone conducting path only partially transmit voice inducted vibrations, or the microphone is not sensitive enough).

Due to the space contains of this conference proceedings, only the transfer functions between *m1* and *m4* are presented in this paper. Microphones *m4* and *m1* are commercial and pre-commercial MEMS audio microphones with similar sizes both designed to be surface-mounted but produced by different –undisclosed- manufacturers. Their frequency response is similar and almost flat at +0/-6 dB over the 20 Hz – 20 kHz frequency range, but their nominal sensitivity is slightly different, with *m1* at -38 dB (0 dB reference is 1V/Pa)) and *m4* at -40 dB. Neither the 2 dB difference in electrical sensitivity, nor pre-gain used in signal conditioning are accounted for in the data presented here-below.

3. RESULTS

The magnitude of the transfer function between external microphone *m4* and internal microphone *m1* is presented in Figure 2 for the 9 subjects, over the 12 tests performed. The group average and the standard deviation are also computed for every 1/12th octave band: Fig. 2 shows in overlay the mean as well as the mean ± standard deviation. It first can be seen that large inter-subject variations can occur in the magnitude of the transfer function, but that such shift is most often quite constant over a large frequency range. For example, the transfer function magnitude measured for English-speaking subject 5 (mustard color curve with legend *s05-en* in Fig. 2) is consistently larger than the other transfer function (hence shifted upwards), but shows the same frequency pattern than the other transfer function for the other subjects and is hence looking similar to the group average transfer function. It could also be seen, for the bilingual subject tested (*s04*, *s05* and *s09*), that the measured transfer function is not much sensitive to the language used.

The coherence functions are presented in Fig. 3 for the same group of subject, together with the group average. It first can be seen that the coherence value equals or exceeds 0.3 for the 100 Hz – 2 kHz frequency range and that a 60 Hz – 3.5 kHz range can be achieved, allowing for the coherence to be as low as 0.1. Such frequency range is to be compared to the classical 300 Hz – 3.5 kHz used in traditional phone line systems. An abnormal lower coherence value around 1 kHz can also be observed, it may well be related to a know physical/acoustical resonance of the earpiece at that frequency.

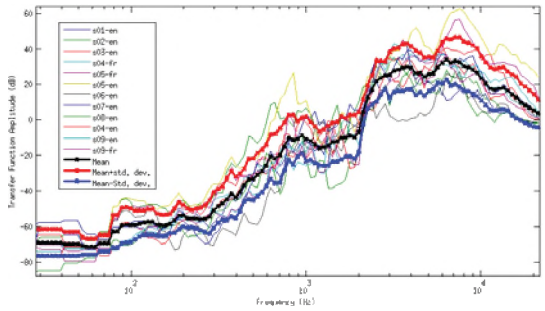


Figure. 2. Measured Transfer Functions

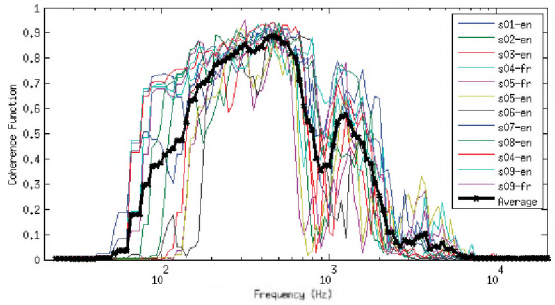


Figure. 3. Measured Coherence Functions

If the average transfer function magnitude is now considered over the meaningful frequency range (i.e where the coherence is higher than 0.1), it can be seen that the typical correction to be applied to an in-ear microphone signal can be quite simple, as it can be in first-approximation considered to be a monotonic rising slope. Such rising pattern can be for example, obtained from an analog high-pass filter used in its cutting-band and is quite straightforward to achieve in the analog or digital domain, depending of the application envisioned.

REFERENCES

[1] Sonomax Technologies (www.sonomax.com)
[2] Mathworks (www.mathworks.com)
[3] Thrane, N. (1979). The Discrete Fourier Transform and FFT Analysers. B&K Technical Review. No 1.
[4] Herlufsen, H. (1984). Dual Channel FFT Analysis (Part I). B&K Technical Review. No 1.
[5] Herlufsen, H. (1984). Dual Channel FFT Analysis (Part II). B&K Technical Review. No 2.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial and logistic support from CRITIAS, the Sonomax-ETS Industrial Research Chair in In-Ear Technologies.

AUTHOR NOTES

Some of the work presented here, was conducted while Antoine Bernier was completing his electrical engineering internship at Sonomax Technologies inc, as a bachelor degree student from École Polytechnique de Montréal.

PREDICTING THE INTELLIGIBILITY OF SPEECH CORRUPTED BY NONLINEAR DISTORTION

A.J. Brammer[1], G. Yu[1], E.R. Bernstein[1], M.G. Cherniack[1], J.B. Tufts[2], and D.R. Peterson[1]
[1]Ergonomic Technology Center, University of Connecticut Health Center, Farmington CT 06030, U.S.A.
[2]Department of Communication Sciences, University of Connecticut, Storrs CT 06269, U.S.A.

1. INTRODUCTION

Common methods for predicting the intelligibility of speech, the speech transmission index (*STI*) and the speech intelligibility index (*SII*), fail when the speech signal is corrupted by nonlinear distortion, e.g., center clipping (Steeneken & Houtgast, 2002). This limitation restricts the applicability of the metrics to many digital communication systems, where peak and center clipping are often unwanted byproducts of the signal processing. For distorted speech, the performance of the *SII* may be improved by calculating the speech signal-to-'noise' (or distortion) ratio from the coherence for three amplitude ranges, and combining the results to assess the intelligibility (Kates and Arehart, 2005).

In this paper we explore modifications to the *STI* to permit estimates of the intelligibility of speech corrupted within a communication system by center clipping. This type of distortion occurs when a signal rapidly changes polarity from a non-zero value, and is associated with quantization errors in digital signal processing. The development of models is first briefly described, followed by a summary of the psychophysical experiment employed to evaluate their performance. Results are presented for word intelligibility in speech-spectrum shaped noise, which is well predicted by the *STI*, and for speech subjected to center-clipping.

2. METHOD

2.1 Speech Transmission Index

The original *STI* was based on determining changes in the temporal envelope modulations of a test signal that substituted for real speech. The analysis involves computing an intensity modulation transfer function for seven octave bands from 125 to 8,000 Hz. A modulation transfer index is constructed for each octave band, MTI_k (k = 1, . .7), from the average intensity modulations within fourteen one-third octave bands with frequencies from 0.63 to 12.5 Hz, $TI_{k,f}$ (f = 1, . .14). The contributions to intelligibility are considered in two ways in the most recent version of the *STI*, the so-called revised *STI*, STI_r (Steeneken & Houtgast, 2002): 1) direct contributions from each octave band, which employ an empirically determined frequency dependent factor, α_k, and 2) indirect, or inter-band, contributions that are termed redundancy corrections, β_k, and arise from the observation that the energy of speech sounds in adjacent frequency bands may be correlated.

An alternative formulation consisting of a combination of the artificial test signal and an octave-band speech probe was necessary for predicting the intelligibility of speech corrupted by peak clipping, and noise. The use of solely speech as the probe signal and, in particular, the need to consider the coherence between speech and noise intensity modulations was discussed by Payton and Braida (1999). Our models build on this work, by employing a speech probe and formally introducing the coherence function between the original and corrupted speech, $\gamma^2_{k,f}$, to generalize the interaction between speech and interfering noise or distortion on the MTI_k, in the following way:

$$MTI_k = \frac{1}{14}\sum_{f=1}^{14} TI_{k,f} * \gamma^2_{k,f} \quad (1)$$

The redundancy between speech information contained in different octave bands is treated formally by introducing the normalized cross-covariance function between intensity modulations in nearby bands, $\rho_{k,j}$ ($j = k$+1, . .7) In this paper we explore the consequence of including interactions between three adjacent octave bands (i.e., $j = k$+1, k+2), which coincides with the non-zero coefficients of the cross-correlation matrix observed for uncorrupted speech. The model for predicting intelligibility is then:

$$STI_{\rho-speech} = \sum_{k=1}^{7} \alpha_k MTI_k + \frac{1}{6}\sum_{k=1}^{6}\sum_{j=k+1}^{k+2}\sqrt{\rho^2_{k,j} * MTI_k * MTI_j} \quad (2)$$

2.2 Modified Rhyme Test

The modified rhyme test (MRT) was used to characterize speech intelligibility (ANSI, 1989). A subject with normal hearing and confirmed understanding of American English was seated in an anechoic chamber. Speech was reproduced by a small, high-fidelity, low distortion loudspeaker located in front of the subject on the ear-nose plane, 2.4 m from the center-head position (Paradigm S1). Center-clipped speech was produced by removing the central 10% to 98% of the long-term histogram of the waveform of the carrier sentences and MRT words (Fig. 1). Speech-spectrum shaped noise was produced by four loudspeaker towers surrounding the subject (JVC SRX715Fs + SRX718Ss), with output processed to produce a flat, pseudo-diffuse field in the horizontal plane at the center-head position (±3 dB).

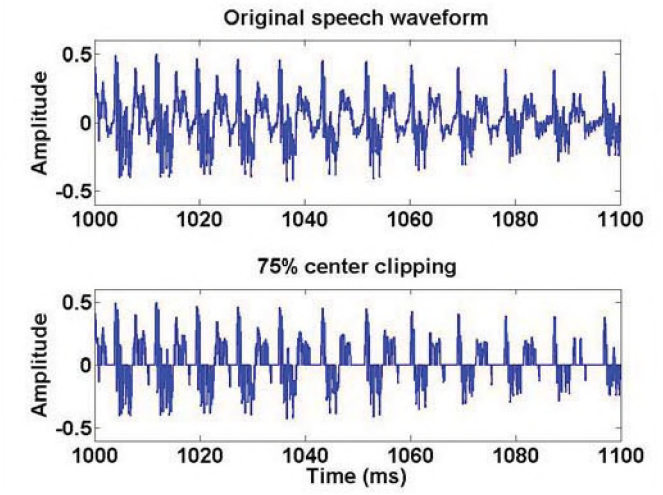


Fig. 1. Example of 75% center clipping shown for a 10 ms duration segment from a speech waveform.

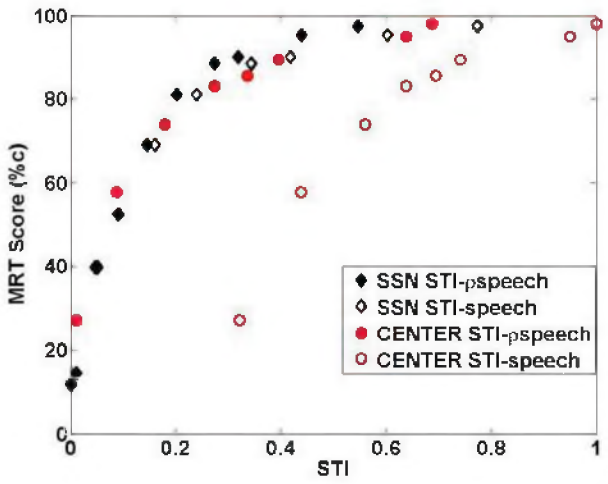


Fig. 2. Mean MRT scores and predictions using two *STI* models for: speech-spectrum shaped noise – diamonds, and; center clipping – circles.

All subjects (3 male, 3 female) gave their informed consent to participate in the study according to the provisions of the University's institutional review board.

3. RESULTS AND DISCUSSION

Mean MRT scores are presented in Fig. 2 for three replications of each experimental condition by each subject, for two *STI* models. The first model, *STI-speech*, consists of a speech probe in which values of the MTI_k are calculated from the transmission indices and the coherence between the original and modified speech according to eqn. 1. The transmission indices are computed following the procedure described in IEC 60628-16, 2003, and the contributions of each octave band to the intelligibility employ the values of α_k and β_k contained in the standard. The MRT scores are shown by unfilled diamonds for speech in speech-spectrum shaped noise, and unfilled circles for speech corrupted by center clipping. The second model, *STIρ-speech*, also employs a speech probe and again introduces the coherence to calculate MTI_k as in the first model. However, the second model replaces the empirically determined redundancy factors (β_k) of Steeneken and Houtgast, and the international standard, by the measured cross-covariance of the intensity modulations between adjacent and the next nearest neighbor octave bands, as given by eqn.2. For this model the MRT scores are shown by filled diamonds for speech in speech-spectrum shaped noise, and filled circles for speech corrupted by center clipping.

Inspection of Fig. 2 reveals that the model employing speech as the probe signal but not accounting for the correlated energy in adjacent (octave) frequency bands (*STI-speech*) yields *STI* values that depend on the competing or distorting sounds interfering with word intelligibility (i.e., the unfilled symbols). In contrast, the model incorporating a measure of the correlation between sounds in adjacent octave bands (*STIρ-speech*) yields *STI* values that are substantially the same irrespective of whether the interfering sounds are speech-spectrum shaped noise or the distortions introduced by center clipping (i.e., compare filled symbols).

It thus appears that a model accounting for the redundancy between energy in adjacent octave bands is required to predict the intelligibility of speech corrupted by center clipping using the *STI*. The formulation proposed here (*STIρ-speech*) is one of a family of models constructed using generalizations of eqn. 2 to introduce varying degrees of interaction between the speech information in different frequency bands. It should be noted that the results for both models described employ only seven adjustable parameters, the values of α_k, in contrast to the 22 adjustable parameters introduced to model speech corrupted by nonlinear distortion using the *SII* (Kates and Arehart, 2005).

REFERENCES

ANSI S3.2-1989 (1989). American National Standard: Methods for the Intelligibility of Speech over Communication Systems (American National Standards Institute, New York).

IEC 60628-16 (2003). Sound System Equipment, Part 16: Objective Rating of Speech Intelligibility by Speech Transmission Index (International Electrotechnical Commission, Geneva).

Kates, J.M., and Arehart, K.H. (2005). Coherence and the speech intelligibility index. J. Acoust. Soc. Am. 117, 2224-2237.

Payton, K.L., and Braida, L.D. (1999). A method to determine the speech transmission index from speech waveforms. J. Acoust. Soc. Am. 106, 3637-3648.

Steeneken, H.J.M., and Houtgast, T. (2002). Validation of the revised STI_r method. Speech Comm., 38, 413-425.

ACKNOWLEDGEMENTS

This work was supported by NIOSH research grant R01 OH008669.

EAR CUP SELECTION FOR FEEDFORWARD ACTIVE NOISE REDUCTION HEARING PROTECTORS

Eric R. Bernstein[1], Anthony J. Brammer[1], Gongqiang Yu[1], Martin G. Cherniack[1], and Donald R. Peterson[1]
[1]Ergonomics Technology Center, Univ. of Connecticut Health Center, 263 Farmington Av., Farmington, CT, U.S.A., 06030

1. INTRODUCTION

Performance of an active noise reduction (ANR) hearing protector device (HPD) is significantly affected by the mechanical design of the ear cup. Two characteristics are of critical importance when evaluating a passive HPD design for potential modification into an active feedforward HPD: passive noise reduction and predicted active noise reduction with active control. Several commercially available passive HPDs have been evaluated using these metrics in order to select the ear cup that would offer the best performance with a feedforward ANR system.

Passive performance of the ear cup serves to augment the electronic control system in an active HPD. A well designed ear cup will compensate for the decreased attenuation provided by the active system at high frequencies as well as provide the maximum possible passive attenuation to improve the total attenuation of the complete system. Shaw and Thiessen (1965) detailed two important features that are necessary to maximize passive attenuation: the cushion must provide a high resistance to air leakage, and the volume underneath the ear cup should be as large as possible.

Active noise reduction systems improve the performance of an active HPD by attenuating low frequency noises whose wavelength approaches or exceeds the size of the ear cup. Simple estimation of the active noise reduction (ANR(ω)) possible with an optimal feedforward controller can be determined from the coherence function ($\gamma_{re}^2(\omega)$) between a reference microphone placed outside the ear cup and an error microphone at the desired point of cancellation inside the ear cup (Nelson and Elliot, 1992), $ANR(\omega)=1-\gamma_{re}^2(\omega)$. This method of estimation of active control system performance is ideal for ear cup selection because it does not require the presence of a secondary source loudspeaker or electronic controller, only the reference and error microphones located where they would ideally be positioned in a final device.

Caution should be exercised when using this equation in practice due to assumptions regarding the disturbance and the controller. First, this equation assumes a stationary disturbance, and thus controller performance could be better than predicted when using an adaptive controller with non-stationary signals. Second, the equation does not require the optimal controller to be causal and therefore may not be physically realizable. For a description of performance prediction with a causally constrained controller, see Elliott (2001).

2. METHOD

Ear cups were evaluated in a reverberation chamber possessing a reverberation time of between 1.5-4 sec over a frequency range of 100-8000 Hz. Tests were conducted using both a Head and Torso Simulator (HATS) (Brüel & Kjær Type 4128C) and human subjects. 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.

HATS passive attenuation values were calculated as the difference between the power spectrums of a white noise signal presented with and without each ear cup recorded at the artificial eardrum microphone (Brüel & Kjær Type 4158C). Coherence was measured from a microphone (Knowles FG 23629-P16) placed on the surface of the ear cup along the axis of the ear canal to the artificial eardrum. Human subject passive attenuation and coherence testing were conducted similar to HATS procedures, however the artificial eardrum microphone was replaced with a custom probe microphone system designed to accurately replicate sound pressures at the subjects eardrum. The system uses custom fitted ear molds to allow repeated placement of the probe tip and has been shown to replicate signals of up to 6 kHz with an accuracy of ±2 dB, (Brammer 2009). Power spectrum and coherence were calculated and averaged over 50 measurement periods using a dynamic signal analyzer (Agilent 35670A). A signal bandwidth from 0-1600 Hz was used as coherence values were negligible above these frequencies and the level of passive attenuation remained approximately constant.

Predicted attenuation was used to select an optimal ear cup for conversion into an active HPD. Total noise reduction (TNR(ω)) was calculated using Eq. 1 by combining both the measured passive noise reduction (PNR(ω)) and active noise reduction predicted using the coherence function. The ear cup selected to provide the best predicted total attenuation was then modified into a feedforward ANR HPD using a floating-point DSP (Texas Instruments TMS320C6713).

$$TNR(\omega) = PNR(\omega) - 10 \cdot \log_{10}(1 - \gamma_{re}^2(\omega)) \qquad (1)$$

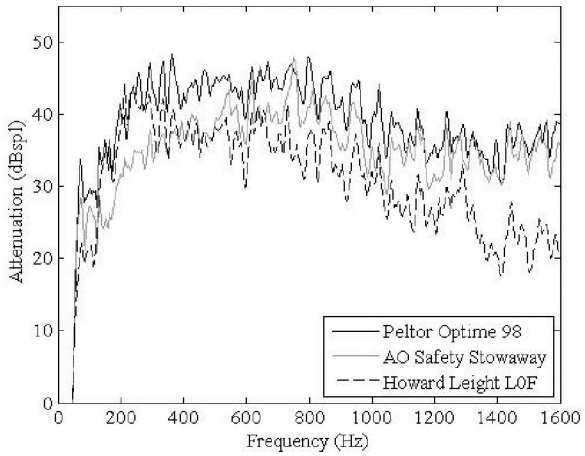


Fig 1. Predicted TNR(ω) of ear cups from human subject tests

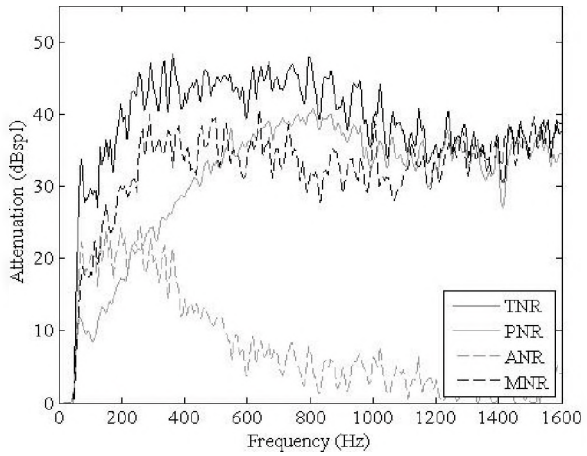


Fig 2. Noise reduction comparison of the Peltor Optime 98 ear muff before and after feedforward ANR modification

3. RESULTS AND DISCUSSION

Representative data from three diverse ear cup designs are included in Fig 1. The Peltor Optime 98 headset is a large volume ear muff intended to provide high levels of noise isolation in an industrial environment. The AO Safety Stowaway ear muff is designed for portability and has a teardrop shape to allow for part of the ear muff to sit underneath a hardhat yet still retain the volume necessary for high passive attenuation. The Howard Leight L0F ear muff is an ultra-slim ear muff that minimizes the storage size and volume underneath the ear cup.

The Peltor Optime 98 ear muff demonstrates exceptional performance over the entire frequency range of interest. Passive attenuation settles to approximately 35 dB at frequencies above 1.6 kHz and high coherence up to 400 Hz demonstrates that a feedforward system could add nearly 20 dB of active attenuation, see Fig 2. While the AO Safety Stowaway ear muff has a passive attenuation comparable to the Optime ear muff at high frequencies, the slight increase to the coherence at low frequencies cannot compensate for the loss of passive attenuation. Conversely, the small distance between the two microphones using the Howard Leight ear muff provides a very high coherence for active control, but the small volume leads to a significantly reduced noise reduction capacity at higher frequencies. Based on these results, the Optime ear muff was selected for modification into an active HPD. It should be noted that while the Peltor Optime 101 ear muff also tested provided slightly better total predicted performance than the Optime 98 ear muff, the Optime 98 ear muff was selected because it provided better coherence for design and testing of active feedforward noise control systems for future research.

When modifying the Peltor Optime 98 ear muff, the external reference microphone was positioned similar to the earlier coherence experiments. However, the error microphone was repositioned for better integration with the active headset. The differences between the TNR(ω) of the constructed system compared with the measured noise reduction (MNR(ω)) are illustrated in Fig 2. While the MNR(ω) closely follows the shape of the TNR(ω), a difference of approximately 10 dB is observed between 100-1000 Hz. This discrepancy can be explained by the modifications involved in installing a secondary source loudspeaker and repositioning the error microphone, which will affect both passive attenuation and coherence respectively. The relative agreement between the TNR(ω) and MNR(ω) curves indicate that this method is suitable for approximation of feedforward ANR HPD performance.

The results of this study suggest several guidelines for characteristics that make a particular ear cup design ideal for feedforward noise control. Comparing both the Peltor and AO Safety ear muffs, it is clear that a large volume underneath the ear cup is necessary for high passive attenuation. However, the coherence results from the AO Safety and Howard Leight ear muffs, not shown, indicate that the best predicted active attenuation is obtained when the distance between the external reference microphone and the error microphone is minimized. Future designs should emphasize an optimal balance between these two considerations to provide the best possible performance from feedforward HPD designs.

REFERENCES

Brammer, A. J. et. al. (2009). Monitoring sound pressure at the eardrum for hearing conservation. Can. Acoustics **37**, 108-109.

Elliott, S. (2001). Signal Processing for Active Control. [Academic Press, London].

Nelson, P. A. and Elliott, S. J. (1992). Active Control of Sound. [Academic Press, London].

Shaw E. and Thiessen G. (1962). Acoustics of circumaural earphones. J. Acoust. Soc. Am. **34**, 1233-1246.

ACKNOWLEDGEMENTS

This work was supported by NIOSH (R01 OH008669).

NOISE LEVELS IN PUBS AND NIGHTCLUBS IN VANCOUVER

Eric M. de Santis and Mark L. Gaudet[1]
[1] BKL Consultants Ltd., #308 – 1200 Lynn Valley Rd., North Vancouver, BC, Canada, V7J 2A2 desantis@bkl.ca

1. INTRODUCTION

As part of the application process for extended hours of operation, the City of Vancouver requires local pubs and nightclubs to demonstrate compliance with its Noise Control Bylaw No. 6555[1]. Since 2007, BKL Consultants Ltd. has provided this service to over 30 entertainment venues in Vancouver.

Part of the method used to establish a venue's contribution to the sound level at noise sensitive locations includes the installation of an overnight noise monitor inside the facility. Compliance with Vancouver's noise control bylaw is assessed both in terms of A-weighted and C-weighted sound levels. Measurement results in this summary paper are presented in terms of A-weighted equivalent sound levels for the purpose of comparing results with generally accepted occupational noise exposure criteria[2].

2. METHOD

Long-term noise monitoring (> 12 hour periods) was performed inside 30 entertainment venues. These venues consisted of nightclubs with live DJ's, pubs with live bands or DJ's and/or pubs playing music through a house sound-system. The primary intention of these measurements was to establish the maximum indoor sound levels received at a fixed microphone location inside a venue under test. Measurements were mostly conducted on Friday or Saturday evenings.

Safety of the measurement microphone (Brüel & Kjær Type 4189) and sound level meter (Brüel & Kjær Type 2250) often dictated the selected monitoring position inside the venues. Where possible, the measurement microphone was suspended from the ceiling above dance floors or near DJ booths.

3. RESULTS

The distribution of equivalent sound levels, L_{eq}'s, measured in all 30 entertainment venues is presented in the boxplot shown in Fig.1. Data is shown in groups of one hour periods between 20:00 – 02:00 hours. The distribution of the 8-hour equivalent sound levels ($L_{eq,8hr}$) measured in the entertainment venues is also shown in the far right group shown in Fig. 1. Median values (horizontal lines inside the boxes) range from 84 - 97 dBA. The lower quartile (Q1-lower horizontal lines) values range from 80 – 98 dBA while the upper quartile (Q3 - upper horizontal lines) values range from 90 – 102 dBA. The median $L_{eq,8hr}$ is 97 dBA. The lower and upper quartile $L_{eq,8hr}$ values range from 90 to 98 dBA, respectively.

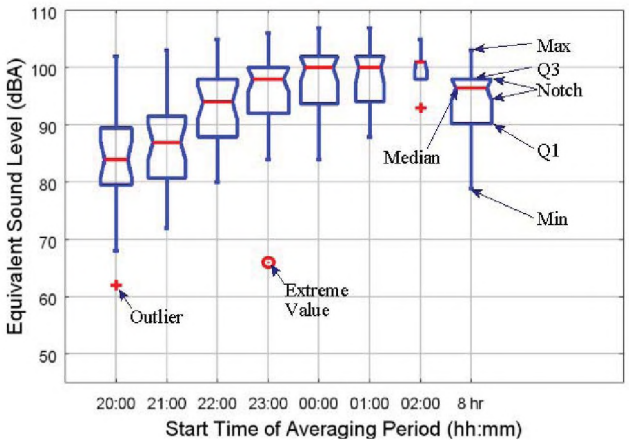


Fig. 1. Boxplot of equivalent sound levels (L_{eq}) by one hour time averaging periods between 20:00 – 02:00 hours and eight hour time averaging period (far right).

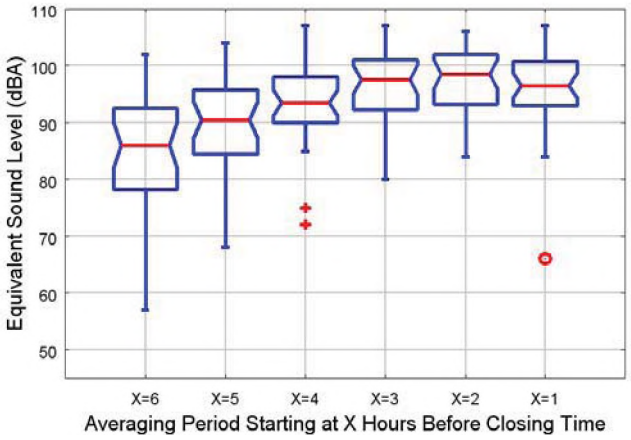


Fig. 2. Boxplot of one-hour equivalent sound levels ($L_{eq,1hr}$) by averaging periods starting at six hours before closing time continuing up to one hour before closing time.

The boxplot shown in Fig. 2 presents the distribution of equivalent sound levels for the complete data set in groups of one hour averaging time periods starting at six hours before closing time continuing up to one hour before closing time. Median values range from 86 – 99 dBA. The lower quartile and upper quartile values range from 78 – 93 dBA and 93 – 102 dBA, respectively.

4. DISCUSSION

The notches (defined by the horizontal V-shaped dips on either side of the boxes) shown in Figures 1 and 2 represent the median confidence interval. They can be used to visually assess whether the median $L_{eq,1hr}$ values between groups are significantly different (with 95% confidence). That is, if the notches between groups are not overlapping, then there is strong evidence that median values are significantly different. The notches of $L_{eq,1hr}$ values grouped at 20:00 - 21:00 hours in Fig. 1 do not overlap with the notches of groups at 22:00 – 02:00 hours. This indicates that median $L_{eq,1hr}$ values between groups at 20:00 and 21:00 hours are significantly different than median $L_{eq,1hr}$ values at later times in the evening. Further, $L_{eq,1hr}$ values grouped at 22:00 hours are significantly different than $L_{eq,1hr}$ values grouped at 00:00 hours. Given that the notches between groups at 23:00, 00:00, 01:00 and 02:00 hours are overlapping , and are barely overlapping with the notch at group 22:00, it is likely that the peak $L_{eq,1hr}$ occurs between 23:00 and 02:00 hours.

The boxplot shown in Fig.2 was presented in order to compare the distribution of measured $L_{eq,1hr}$ values accounting for different hours of operation. In comparison with Fig.1, median values in Fig. 2 are more closely placed in the centre of each box, indicating a more symmetric distribution. The median $L_{eq,1hr}$ values at groups X=6 and X=5 hours before closing time are significantly different than values at times approaching the final hour of operation. Further, $L_{eq,1hr}$ values at groups X=4 and X=2 hours before closing time are significantly different. Given that notches between groups X=3, X=2 and X=1 are overlapping, and are barely overlapping with the notch of group X=4, it is likely that the peak $L_{eq,1hr}$ occurs between one and three hours before closing time.

The boxplots in Figures 1 and 2 would indicate that median $L_{eq,1hr}$ values are tending to increase as the evening progresses. Due to the large amount of variation in this data set, it is difficult to conclude that $L_{eq,1hr}$ values between each group are significantly different from each other. The large variation in sound levels measured in the entertainment venues is likely caused by differences in the measurement microphone's distance to the sound sources, acoustic room influences on sound levels received at the microphone position, the type of audio equipment used and the venues' individual volume preference. The overall variation could have been reduced by categorizing the different types of venues. Due to the wide range of amplified activities going on at the assessed venues and the limited data set, this was not done.

The median $L_{eq,8hr}$ value of 97 dB is an indication that the potential risk for noise-induced hearing damage exists in these entertainment facilities, particularly for those who are in close proximity to DJ booths and/or nightclub dance floors for extended periods of time on a regular basis. The $L_{eq,8hr}$ presented here is not an accurate estimate of an individual's noise exposure as sound levels can vary greatly at different spatial locations inside the facilities.

REFERENCES

[1]City of Vancouver (2010). "Noise Control Bylaw No. 6555." (*vancouver.ca/**bylaws**/**6555**c.PDF)*
[2]ISO 1999:1990."Acoustics -- Determination of occupational noise exposure and estimation of noise-induced hearing impairment." (International Standards Organization, Geneva 1990)

ACKNOWLEDGEMENTS

The author would like to thank all the staff at BKL Consultants Ltd. and the open source developers of GNU Octave. Special thanks also to Ju Helen Wong, University of Manchester.

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

OCCUPATIONAL NOISE EXPOSURE IN NIGHTCLUBS

Andrew Dobson, and William J. Gastmeier
HGC Engineering, 2000 Argentia Rd., Plaza 1, Suite 203, Mississauga, Ontario, Canada L5N 1P7

1. BACKGROUND

This paper presents is summary of sound levels measured inside nightclubs featuring loud music and a study of the hearing damage risk potential to the employees and patrons.

Due to a recently implemented by-law in the City of Toronto, HGC Engineering has had the opportunity to measure the sound levels associated with numerous nightclubs, most of which are situated in Toronto's Entertainment District. The nightclub industry in Toronto is quite intensive, with over one hundred clubs located in the one square kilometre area of the Entertainment District. The high concentration of nightclubs in the area, coupled with high density residential development, has triggered concern amongst residents and city officials, with regard to the environmental (outdoor) noise emissions of the clubs.

A recent City of Toronto by-law requires nightclubs to assess their environmental sound emissions to ensure that the emanating sound does not disturb citizens travelling, or living in the vicinity of the subject club. While the formal purpose of the work conducted was to measure outdoor sound levels and assess compliance with the applicable by-law, measurements were also conducted inside the nightclubs for reference purposes. The results of these measurements are summarized in this paper.

Sound levels were almost always found to be significantly higher than the applicable Ministry of Labour workplace noise exposure regulations at locations where nightclub employees (wait staff, bar tenders, etc) and patrons spend much of their time. Furthermore, staff and patrons were typically not observed to be wearing hearing protective devices (HPD,s).

The regulatory body responsible for defining the applicable sound limits for worker exposure is the Ontario Ministry of Labour. They have developed a regulation intended to protect the hearing of workers exposed to high sound levels in the workplace. It contains the noise exposure limits and expresses the way in which to calculate the potential for employee noise exposure to cause hearing damage.

The content presented herein is not intended to represent a thorough review of the regulation and its application within the industry, but to raise awareness and bring light to the fact that employees in the nightclub industry could be incurring significant damage to their hearing.

2. DESCRIPTION OF THE SUBJECT FACILITIES

Sound level measurements were conducted inside over 60 nightclubs in Toronto, most of which were situated in Toronto's Entertainment District. The areas in nightclubs that are exposed to the high sound levels are generally the areas surrounding the main dance floors. Bars serving beverages are frequently positioned adjacent to, and therefore within the same sound field as the dance floors, resulting in employee exposure levels similar to that of the patrons, but for extended periods of time. Furthermore, walls surrounding these areas tend to be constructed of either brick or painted gypsum board, and these workplaces are therefore reverberant in nature, such that sound levels may not decrease significantly with distance from the source. Sound systems vary in set-up and size, depending on the size and layout of the venue, and tend to range from 10,000 to 40,000 watts in power.

3. WORKPLACE SOUND CRITERIA

As indicated above, noise exposure in the workplace is governed by the Ontario Ministry of Labour, who, in 2007 published "Amendments to Noise Requirements In the Regulation for Industrial Establishments" (Regulation 851, Section 139 of the Occupational Health and Safety Act). This regulation states that "Every employer shall ensure that no worker is exposed to a sound level greater than an equivalent sound exposure level of 85 dBA, $L_{ex,8}$." The $L_{ex,8}$ is a representation of the time-weighted average noise exposure level over an eight hour work shift. Thus, the calculated $L_{ex,8}$ accounts for varying noise levels throughout a shift, and/or shifts that may be shorter or longer than eight hours in length. For reference, Table 1, below, outlines the allowable $L_{ex,8}$ exposure levels, given a specific shift length.

Table 1: Equivalent Noise Exposure Levels

Steady Sound Level [dBA]	Duration
82	16 Hours
85	8 Hours
88	4 Hours
91	2 Hours
94	1 Hour
97	30 Minutes
100	15 Minutes
103	7.5 Minutes
106	3.75 Minutes
109	1.88 Minutes

4. MEASUREMENTS AND DISCUSSION

Sound level measurements were conducted using integrating sound level metres set to record the sound levels as a function of both frequency spectrum in octave bands, and time, on a second-by-second basis. From these measurements, the average (L_{eq}) sound levels in A-weighted decibels were calculated. While the measurement periods did not extend the duration of an entire work shift, they can be considered to be generally representative of the typical exposure levels throughout a shift for an employee working in the vicinity of the sound system and/or dance floor, as indicated by the club managers, and by personal experience.

The overall measured dBA sound levels are included in Table 2 below.

Table 2: Measured Sound Levels

Measurement Date	Measured dBA	Measurement Date	Measured dBA
8-Jun-07	96	27-May-08	96
11-Jul-07	98	8-Jun-08	100
25-Jul-07	103	14-Jun-08	94
1-Aug-07	96	17-Jul-08	105
9-Aug-07	84	3-Aug-08	95
9-Aug-07	84	3-Aug-08	101
16-Aug-07	99	3-Aug-08	96
10-Sep-07	93	3-Aug-08	102
18-Oct-07	94	3-Aug-08	103
23-Oct-07	91	1-Oct-08	84
7-Nov-07	93	7-Oct-08	91
7-Nov-07	94	9-Nov-08	99
17-Dec-07	98	13-Nov-08	100
18-Dec-07	98	22-Jan-09	100
18-Dec-07	91	7-Feb-09	100
18-Dec-07	99	12-Mar-09	97
16-Jan-08	99	6-Apr-09	96
29-Jan-08	95	1-Jul-09	99
30-Jan-08	90	8-Jul-09	92
7-Feb-08	95	19-Aug-09	99
21-Feb-08	94	10-Sep-09	102
22-Feb-08	96	14-Oct-09	100
22-Feb-08	97	21-Oct-09	96
4-Mar-08	96	28-Oct-09	97
6-Mar-08	95	25-Jan-10	93
6-Mar-08	93	1-Mar-10	90
7-Apr-08	101	19-Apr-10	91
		Arithmetic Median (dBA)	**95**
		Standard Deviation (dBA)	**5**
		Arithmetic Average (dBA)	**96**

The overall average sound level in all of the nightclubs was 96 dBA. Depending on the nightclub, work shifts typically do not last a full eight hours. It is safe to assume that many bar tenders and weight staff are exposed to sound levels of this magnitude for a period of four hours, which would result in an $L_{ex,8}$ exposure level of 93 dBA. Even if employees were exposed to this average level for two hours, the resulting $L_{ex,8}$ sound exposure level would be 91 dBA. Both of these situations represent significant excesses over the applicable exposure limit

Interesting information can also be gathered from the frequency spectrum of the measured sound levels, several of which are shown in the following Figure 1.

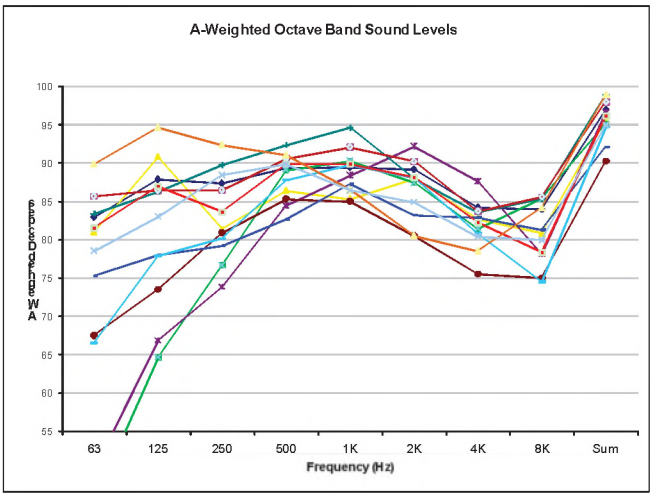


Figure 1: Measured Octave Band Sound Levels

Generally, the human ear is the most sensitive to the higher frequency bands of the sound spectrum (i.e. 1000 Hz to 4000 Hz). Because of this, human hearing is also the most vulnerable to sound in these frequency bands. As can be seen in Figure 1, the A weighted sound levels in these bands tend to be relatively high. It is clear that the measured sound levels and frequency content present a risk of hearing loss.

It is unlikely that voluntarily reducing the interior sound to levels to less than 85 dBA would be considered feasible, as patrons visit the subject establishments often for the purpose of listening to music at high levels. Moreover, the use of hearing protective devices is not generally considered to be viable by staff as frequent verbal communication with customers is vital. Further study of communication ability with HPD's in high noise environments may be helpful with regard to educational possibilities. The installation of absorptive acoustical panelling to the walls of some of these spaces may be beneficial to some degree, by reducing the reverberation of the sound in the upper frequency bands and thereby lowering sound levels at locations remote from the dancefloor. Such treatments would have little effect close to the source.

5. CONCLUSION

The results of these measurements indicate that it is likely that workers employed at nightclubs are frequently exposed to sound levels greater than those outlined in the Ontario Ministry of Labour Occupational Health and Safety Act, Regulation 851, Section 139, and are at risk of hearing damage. Technical solutions short of "turning down the volume" or wearing HPD's are limited.

REFERENCES

[1] Ontario Ministry of Labour, Regulation 851, Section 139 - "Amendments to Noise Requirements In the Regulation for Industrial Establishments"

ANALYSIS OF HUMAN OTO-ACOUSTIC EMISSIONS

Reinhart Frosch
Sommerhaldenstrasse 5B, CH-5200 Brugg, Switzerland; reinifrosch@bluewin.ch
PSI (Paul Scherrer Institute), Villigen and ETH (Eidgenoessische Technische Hochschule), Zurich (retired)

1. INTRODUCTION

Oto-acoustic emissions have been reviewed, e.g., by Kemp (2007). In the present contribution, the TEOAEs (transient-evoked oto-acoustic emissions, also called click-evoked or delayed evoked) in Fig. 1 are analysed in detail.

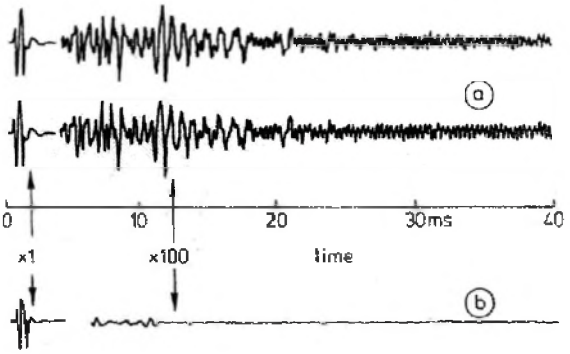


Fig. 1. Reproduction (with permission) of Fig. 16.20 of Fastl and Zwicker (2007), showing click-evoked oto-acoustic emissions from human ears; a: healthy ear of baby; b: human subject with hearing loss.

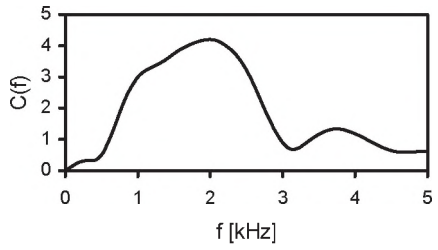


Fig. 2. Frequency spectrum of clicks shown in Fig. 1 at times from 0 to 2 ms. C(f) [in arbitrary units] is the absolute value of the complex Fourier integral defined by Eq. (8.1a) of Hartmann (1998).

In the experiment yielding Fig. 1 there was a sound source (generating the click) and a microphone (generating the three traces shown) in the closed ear canal of the subjects. The frequency spectrum of the click (time = 0 to 2 ms), derived by Fourier transformation [e.g., Eq. (8.1a) of Hartmann (1998)] from the click waveform in Fig. 1, is presented in Fig. 2. The click and the TEOAE (time = 4 to 17 ms; signal magnified, in Fig. 1, by a factor of 100) both contain frequencies from ~1 to ~2.5 kHz. The time-dependence of the instantaneous TEOAE frequency, derived from the waveform in Fig. 1a, is shown by a solid line in Fig. 3. At times later than ~14 ms there is, in Fig. 1a, a "spontaneous" 3-kHz emission, triggered by the click (see Section 2 below).

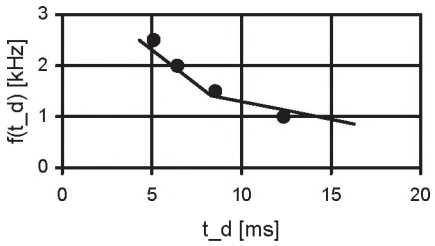


Fig. 3. Instantaneous frequency of emissions shown in Fig. 1 at times from 4 to 17 ms versus delay t_d after click centre [i.e., t_d = time – 0.7 ms]. Solid line: experimental frequency, derived from waveform in Fig. 1a. Filled circles: theoretical frequency, derived from surface-wave formulae according to Section 2 below.

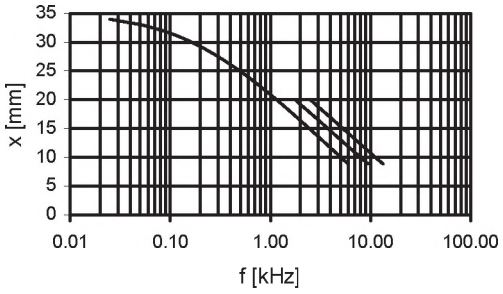


Fig. 4. Preliminary human cochlear maps; *x*: distance from base; *f*: pure-tone frequency; lower curve, 0.025–6 kHz: passive-peak map; same lower curve, 1–6 kHz: internal organ-of-Corti resonator map; middle curve, 2–10 kHz: active-peak map; upper curve, 2.5–13 kHz: basilar membrane resonator map.

2. METHODS

The derivation of the theoretical instantaneous emission frequencies represented by filled circles in Fig. 3 was based on the cochlear maps shown in Fig. 4. These maps are described in Frosch (2009, 2010a). The lowest of the three curves gives the place x [i.e., the distance from the cochlear base, measured along the basilar membrane (BM)] of the "passive peak" [BM oscillation-velocity maximum during the perception of a pure (sinusoidal) tone of frequency f, if the "active" outer hair cells (OHCs) do not function]. For $x < 20$ mm, that same lowest curve in Fig. 1 gives the resonance frequency of the local IOCR (internal organ-of-Corti resonator). The middle curve in Fig. 4 gives the location of the "active peak" (BM oscillation velocity maximum in healthy cochlea during perception of low-level pure tone). The upper curve in Fig. 4 gives, for $x < 20$ mm, the reso-

nance frequency of the local BMR (basilar-membrane resonator).

Interpretation of Fig. 1b: The OHCs in the basal half of the cochlear channel were damaged; cochlear travelling waves generated by the click were "passive", were not significantly reflected, and were extinguished by friction after having passed their passive-peak place.

Interpretation of Fig. 1a: Any strong component of the click frequency spectrum (Fig. 2) caused the OHCs in the corresponding IOCR resonance region to feed "actively" mechanical energy into the travelling wave, to thus give rise to the "active peak" for that component, and also to generate a backward travelling wave, which carried some of the mechanical energy generated by the OHCs back towards the stapes.

The "spontaneous" 3-kHz emission in Fig. 1a is attributed to feedback-generated BMR oscillations involving *evanescent liquid sound-pressure waves* [Frosch (2010a, 2010b)]. Conjectured place of the *submerged* BMR generating the 3-kHz emission: x = 16.6 mm; local BMR frequency according to Fig. 4 (valid for oscillations without cochlear liquid): 4.2 kHz. Resonance frequency of that same local BMR if it is submerged in cochlear liquids and generates evanescent waves: 3.0 kHz.

Formulae for theoretical TEOAE delay $t_d(f)$*:*

$$t_d(f) = 2\tau_{\mathrm{sw}} + \tau_{\mathrm{rise}}, \tag{1}$$

where τ_{sw} = surface-wave group travel time, from $x = 0$ to $x_{\mathrm{IOCR}}(f)$, according to box-model short-wave formula,

$$\tau_{\mathrm{sw}} = \frac{4\rho\cdot\omega}{\alpha\cdot S_0}\cdot\left(\frac{1}{e^{-\alpha\cdot x}-\eta}-\frac{1}{1-\eta}\right); \tag{2}$$

see Chapter 44 of Frosch (2010a). In Eq. (2), ρ is the cochlear-liquid density; the constants S_0 and α define the BM stiffness, $S(x) = S_0\cdot e^{-\alpha\cdot x}$; $\omega \equiv 2\pi\cdot f$ is the angular frequency, and $\eta \equiv M\cdot\omega^2/S_0$ is a dimension-less frequency-dependent quantity; M is the BM surface mass density. The following constants were used:

$$\rho = 10^3\,\frac{\mathrm{kg}}{\mathrm{m}^3};\; S_0 = 10^{10}\,\frac{\mathrm{N}}{\mathrm{m}^3};\; \alpha = 300\,\mathrm{m}^{-1};\; M = 0.1\,\frac{\mathrm{kg}}{\mathrm{m}^2}. \tag{3}$$

The quantity $\tau_{\mathrm{rise}} = Q/(\pi\cdot f) \approx 1.3/f$ in Eq. (1) is the rise time of the forced oscillations of the IOCRs, which have a quality factor of $Q \approx 4$.

Theoretical delay of start of 3-kHz emission in Fig. 1a:

$$t_d = \tau_{\mathrm{sw}}(f_{\max}, x) + \Delta t + \tau_{\mathrm{sw}}(3\mathrm{kHz}, x), \tag{4}$$

where $f_{\max}$ = 2.5 kHz = highest strong component of click spectrum (Fig. 2); the lower-frequency components have shorter τ_{sw}; x = 16.6 mm = place of BMR generating the 3-kHz emission; $\Delta t \approx 1$ ms = time after passage of $f_{\max}$-component at which liquid is sufficiently quiet for generation of evanescent waves [Frosch (2010a), (2010b)].

3. RESULTS

The theoretical TEOAE delays $t_d(f)$ for f =1.0, 1.5. 2.0, and 2.5 kHz derived from Eqs. (1)–(3) are represented by filled circles in Fig. 1 and are seen to agree fairly well with the solid line representing the experimental delays derived from the TEOAE waveform in Fig. 1a.

The theoretical TEOAE delays according to Eqs. (1)–(3) were found to agree not only with Fig. 16.20 of Fastl and Zwicker (2007) [reproduced in our Fig. 1], but also with the many experimental TEOAE delays presented in their Fig. 3.20.

The BMR place x = 16.6 mm inserted into Eq. (4) yields a theoretical delay of t_d = 13 ms (corresponding to time ≈ 14 ms) for the start of the 3-kHz emission, in agreement with Fig. 1a. As mentioned in the text above Eq. (1), the without-liquid BMR frequency at that place is 4.2 kHz.

4. CONCLUSIONS

The measured dependence of the click-evoked otoacoustic-emission (OAE) delay on instantaneous frequency is consistent with the hypothesis that these emissions are generated by the outer hair cells (OHCs) which feed energy into the forward travelling cochlear surface wave.

The place of these OHCs is the place of that internal organ-of-Corti resonator (IOCR) having a resonant frequency equal to the considered click-component frequency; above ~1 kHz, that place is basal of the "characteristic place" (=low-level active-peak place) by ~0.5 octave distance (~2.3 mm), and basal of the corresponding basilar-membrane resonator (BMR) by ~1 octave distance (~4.6 mm).

The "spontaneous" OAEs are hypothesized to be due to BMR oscillations which involve evanescent (standing) liquid sound-pressure waves and so have a frequency ~0.5 octave below the local without-liquid BMR frequency; these oscillations are thought to be generated by feedback from a local IOCR having a resonance frequency region ranging up to exceptionally high frequencies.

REFERENCES

Fastl, H., Zwicker, E. (2007). Psychoacoustics, Facts and Models. Springer, Berlin, pp. 44, 335.

Frosch, R. (2009). Old and New Cochlear Maps. Canadian Acoustics Vol. 37 No.3, 174-175.

Frosch, R. (2010a). Introduction to Cochlear Waves. vdf, Zurich, pp. 257-279, 383-386, 411-426.

Frosch, R. (2010b). Evanescent Liquid Sound-Pressure Waves Near Underwater Resonators. Contribution to this conference.

Hartmann, W. M. (1998). Signals, Sound, and Sensation. Springer, New York, p. 162.

Kemp, D. T. (2007). The Basics, the Science, and the Future Potential of Otoacoustic Emissions. In: M. S. Robinette and T. J. Glattke (Eds.), Otoacoustic Emissions, Clinical Applications, Thieme, New York, pp. 7-42.

FIT TESTING OF HEARING PROTECTORS

Alberto Behar and Willy Wong
IBBME, University of Toronto, 164 College Str., Toronto, ON, M5S 3G9

1. INTRODUCTION

The sound attenuation of a hearing protector, as published by the manufacturer is obtained using statistical calculations from results of measurement performed in a laboratory. Results are therefore, valid for a population but should not be used for individual wearers.

That is precisely the objective of the Field Attenuation Measurement Systems (FAMS) that are intended to measure the attenuation of a hearing protector as worn by the user. Many FAMS are already in the market and are used in the field. Results from the measurements are referred as Personal Attenuation Ratings (PARs) and expressed in dBA. The PAR subtracted for the ambient noise level measured in dBA, is supposed to represent the sound level of the protected ear.

$$SL_{Protected\ ear,\ dBA} = SL_{ambient,\ dBA} - PAR$$

2. FIT TESTING

Fit testing is field verification that a protector is properly worn. In general, it is a simple, fast, qualitative procedure, easy to perform, that does not require specific skills and that allow the potential wearer to test the fit of the protector he is using.

The fit testing principle is well known, especially when applied to respirators: it is a simple test that is performed every time the person is to enter a potentially harmful space such as confined spaces, site on fire, or spaces where the presence of toxic gases or lack of oxygen is suspected.

FAMS, on the other hand, although not intended to be used for every individual who is to be exposed to high noise levels, helps insure that he is wearing the adequate protector in a proper way.

3. FIT TESTING OF HEARING PROTECTORS

In the case of hearing protection devices, there are several reasons for the test to be performed, such as:

a) The wearer can verify the attenuation he is really receiving from wearing the protector under test

b) It helps training the wearer in the proper way of donning the protector.

c) Allows for the selection of a protector that is appropriate for the noise environment he is in.

It has to be pointed out that the procedure is not a panacea, since

a) The test tends to be expensive because of cost of the FAMS and also for the disruption resulting from bringing the worker to the place the test is performed.

b) Results are valid only for the combination wearer/protector and for the particular test, since the same wearer may don his protector in different way in other occasion, resulting in a different PAR.

c) Results cannot be extended to the entire population – no statistics can be developed, unless several measurements are performed on several users, something that detracts the idea of a simple, fast and easy test.

4. FIELD ATTENUATION MEASURING SYSTEMS (FAMS)

Several FAMS that operate under different principles are available in the market. In general, they can be divided into two groups: "objective" or "quantitative" and "subjective" or "qualitative". Results from measurements using different FAMS cannot be compared because of the different principles they apply. An ANSI Working Group is working precisely into the issue of validation and comparison of the different systems[1].

4.1 Objective systems.

Those systems are based on measurements of sound levels outside and under the protector using a two-microphone probe. The subject "lends" his head for the test and no action is required from him. One of the microphones is exposed to the outside (environmental) noise. When testing earplugs, those are replaced by specially prepared, identical to those under test devices, with a probe tube bored through the plug. The second microphone is connected to this tube, receiving the signal that the protected ear is exposed to. In the case of muffs, the inside microphone is located under the muff being tested. A pink noise sound source generates the test signal. Outputs from both microphones (outside and under the protector), are processed using manufacturer provided software that calculates the resulting PAR.

Following are examples of some of the systems.

[1] The ANSI S12/WG 11 Working Group is currently working on the BSR/ASA S12.71-201X the standard Performance Criteria and Uncertainty Determination for Individual Hearing Protector Fit Testing Systems

4.1.1 E-A-Rfit™ VALIDATION SYSTEM by 3M[2]

In this system the two microphones are attached to an eyeglass frame. One of them receives the signal from the environment, while the second is attached to tubing that is introduced through the plug that is under measurement. This special ("surrogate") plug is identical to the one the user is to wear, except for the tube inserted through the body of the plug. In such a way, the second microphone measures the SL of the protected ear. The system also generates a broadband noise through a loudspeaker located in front of the subject. The resulting outputs from both microphones are combined to allow for the calculation of the attenuation in octave bands and as PAR.

4.1.2 QuietDose by SperianProtection/Michael & Associates[3]

As in the previous case, this FAMS requires the use of surrogate ear plugs. The system consists of two dosimeters (contained in one casing) that sample simultaneously sound levels from the environment and behind the protector. The device is worn in the workplace during the entire shift in the usual manner dosimeters are worn. At the end of the shift, the two dosimeters allow for the readings of both $L_{eq,T}$ (outside and under the protector) to be obtained. Their difference is the PAR. The advantage of this system is that the PAR is the real one obtained as a result of a whole-shift operation. The draw back is that it takes a whole shift to measure just one PAR.

4.2 Subjective systems.

They require full participation from the user and are based upon the detection of the hearing threshold with and without the protector. Also, during the test the subject is donning the very protector he is wearing while at work. In such a way the measured PAR appears to be more realistic that the one obtained using an objective system.

Following are examples of some of the systems.

4.2.1 Integra Fit by Workplace Integra[4]

It measures subject's hearing threshold of both ears simultaneously with and without the earplugs in place. Test signal of 500 Hz is provided via specialized deep-dome headset. PAR is calculated as the difference between the thresholds found in both tests.

4.2.2 VeriPro by Howard Leight[5]

Here the measurement of the PAR is performed using a single frequency, loudness balance technique. Signals are conveyed via headset. The test, intended for earplugs, is independent of the background noise. Testing is done in three steps. In each of them the subject is expected to adjust the loudness of the signal to reach a balance between the loudness perceived in both ears. This is done in three steps: a) no protectors, b) only one plug inserted and c) both plugs inserted.

4.2.3 QuickFit by NIOSH[6]

Is a device that generates an octave band of a wide band,1 KHz centered noise. It is contained in a single earmuff-device that has the generator and the controls. The device is applied to the ear that is not protected and the subject adjusts the sound level to his hearing threshold. Then, he inserts the plug and increases the level by15 dBA. If the signal is perceived, then it will indicate that that the resulting PAR is less than 15 dBA. So, the subject is expected to readjust the fitting to increase the PAR to at least 15 dBA.

4.2.4 QuickFit.Web by NIOSH[7]

Is a test performed on line, where a pulsed, wideband test signal is received by the subject through headsets from the web. The test is done in two steps: first, the subject adjusts the signal's level to his threshold while his ears are unclouded. Then he dons his earplugs and listens to the signal that is now increased by 15 dBA. As in the previous case, if the signal is not perceived, it will indicate that the subject has achieved a PAR of at least 15 dBA.

[2] http://www.e-a-rfit.com/

[3] www.howardleight.com/quietdose

[4] http://www.workplaceintegra.com/integrafit/index.html

[5] http://www.howardleight.com/veripro?locale=us

[6] http://www.cdc.gov/niosh/mining/pubs/pubreference/outputid3060.htm

[7] http://www.cdc.gov/niosh/blog/nsb051208_quickfit.html

SHIFT WORK, NOISE EXPOSURE AND HEARING LOSS

Tim Kelsall[1], and Alberto Behar[2]
[1]Noise and Vibration, Hatch, Mississauga, tkelsall@hatch.ca
[2] IBBME, University of Toronto, alberto.behar@utoronto.ca

Non-traditional work shifts such as 12 hours shifts with 3 days on and 2 days off are changing the workplace yet noise regulations use a standard 8 hour shift to evaluate noise. This means that employees with identical long term exposure can have, in the example case, 2 dB different noise exposures. This paper presents the issues both in terms of the difference in measurements but also in terms of the difference in the effect of such shifts on employees' hearing.

More and more industries are abandoning the traditional 8 hour day, 40 hour week in favor of 12 hour shifts, often with 3 days on, 2 days off. Another variation is to work 3 weeks with 1 week off. Although in the long term there is little difference in hours per month or hours per year, the short term effect is a 50% increase in work hours on a particular day.

Ontario and most other jurisdictions now regulate noise exposures based on the exposure during a single day. This means that in order to meet an 8h 85 dBA limit the average sound level over a 12 hour shift must be 83 dBA to allow for the extra 4 hours or exposure. However since we know that hearing loss accumulates over a period of years, their actual exposure according to the equal energy theory would actually be 2 dB lower if their exposure met the 83 dBA limit.

There has been surprisingly little debate on the effect of this extra 2 dB, probably for several reasons. First, 2 dB is a relatively small sound level difference (that is within the measurement error range), and the difference in practice may be hard to measure. Second, we know that people exposed to 8h 85 dBA exposure over 40 years have a 10-15% excess risk of a material hearing impairment and those with an 8h 80 dBA exposure have 0-5 % excess risk[1]. Thus dropping their exposure half way between these two can only reduce the excess risk for those on the longer shifts. Finally, regulators need a practical limit which can be measured relatively quickly and easily. Measuring over a single shift is relatively easy for both regulators and industrial hygienists. Some shift patterns can take weeks or months to repeat, making the effort required to measure longer term exposure much more onerous[2,3].

The picture is even more complicated than this. Two papers (1.2) that appeared in the international journal Noise and Health claim that hearing loss from longer shifts separated by longer periods of rest result in lower threshold shift and, eventually, in reduced risk of hearing loss than those predicted by just using the equal energy theory.

As per the first paper a sample of 218 male workers recruited at a semiconductor factory with no known occupational hazards that affected hearing acuity other than noise worked either in an eight-hour or 12-hour shift. Results from standardized audiometric tests showed that the severity of hearing loss in both ears was significantly lower in subjects who worked a 12-hour shift. In conclusion, working a 12-hour shift followed by a day off is best for workers' hearing.

The second paper also claims that noise-exposed employees working 12 hours a day for two consecutive days followed by two days off, had significantly lower permanent hearing loss than employees working nine-hour shifts from 8 am to 5 pm Monday to Friday.

Obviously, two studies are not sufficient to draw a conclusion. However they might pave the way for more research to determine the use of improved intermittent noise exposure regimes in future design of the noise exposure workday/-week and make future hearing conservation programs more effective.

Meanwhile the debate will continue but it appears that regulating noise exposure on a daily 8h basis is likely providing those working longer shifts with an extra measure of protection.

REFERENCES

1. NIOSH Occupational Noise Exposure Revised Criteria, 1998

2. CSA Standard Z107.56, Procedure for Measurement of Occupational Noise Exposure

3. ISO Standard 9612, Acoustics - Determination of occupational noise exposure - Engineering method

4. Chou, Yu-Fung - Lai, Jim-Shoung - Kuo, Hsien-Wen: Effects of shift work on noise-induced hearing loss, Noise and Health (2009, Vol 11,No45)

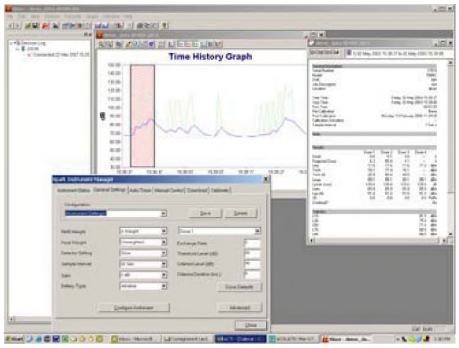
5. Borchgrevink, Hans: Effects of shift work and intermittent noise exposure on hearing: Mechanisms and prophylactic potential, Noise and Health (2009, Vol 11,No45)

Dalimar Instruments Inc.

A voice at your service since 1986

ACOUSTIC SOLUTIONS

Hand held meters



- Noise dosimeters
- Sound level meters (Class 1 or 2)
- Sound level meters with real time filters (1/1 and 1/3 octave)
- Noise exposure analysis software

Permanent and semi-permanent systems

- Sound level meters / real time analyzers
- 1/1 and 1/3 real time analysis with audio recording
- Automatic storage of data
- Stand alone systems with remote communication capabilities

Multi-purpose analyzers

- Frequency analyzer (portable and multi-channel)
- Sound power & acoustic intensity measurements
- Audiometer calibration systems

Rental services

- Rental of a wide range of meters and systems for noise & vibration measurements

Related accessories

Analysis software, microphones, calibrators, cables, etc…

QC: 450-424-0033, ON: 905-707-9000, AB: 403-288-4416
EMAIL:INFO@DALIMAR.CA WWW.DALIMAR.CA

NOISE MODELLING VERSUS REALITY UNDER WORST-CASE METEOROLOGICAL CONDITIONS

Gordon Reusing, and Tim Wiens
Conestoga-Rovers & Associates, 651 Colby Drive, Waterloo, Ontario, Canada, N2V 1C2

1. INTRODUCTION

A constant challenge to acoustic consulting professionals is community sensitivity to noise versus the reality of an adverse noise impact from an industrial neighbour.

Conestoga-Rovers & Associates (CRA) was retained to investigate intermittent noise complaints about an energy Facility. The power generating Facility provides energy to the grid 24-hours per day and is a source of constant and steady state noise emissions. The local government received intermittent noise complaints from adjacent residential neighbors and requested an investigation.

An extensive ambient noise monitoring program was conducted and compared to theoretical noise modelling predictions for the Facility to determine if the perceived adverse impact was reality.

2. METHOD

2.1 Standard ISO 9613-2 Modelling Standard

The Facility's significant environmental noise sources included two landfill gas generator exhausts and two roof mounted radiator units.

A combination of source measurements and manufacturer sound level data were used to develop a standard acoustic model using ISO 9613-2 "Acoustics – Attenuation of Sound During Propagation Outdoors – Part 2: General Method of Calculation." ISO 9613-2 is based on the principle of "predictable worst-case" where downwind propagation is projected for all sources to all receiver locations simultaneously under wind speeds of between 1 to 5 metres per second (m/s) [1].

The off-site noise impacts were evaluated for the nearest surrounding sensitive residential receivers approximately 450 metres (m) (POR1) and 700 m (POR2) from the Facility. The noise impacts predicted at all receivers were below the most stringent government rural nighttime noise limit of 40 dBA.

A far-field audit measurement was conducted on-site for comparison and validation of the acoustic model set-up parameters and to confirm the dominant noise sources. The audit location was approximately 30 m straight-line distance southwest of Generator 2 at an extended microphone height of 4 m above grade, with clear lines of sight to all major noise sources. The modelled noise impact predicted 61 dBA at the audit location, and the audit measurement was 61.5 dBA, which shows good agreement with the modelling work. The exhaust and radiator sources were audible but the exhausts were predominant at the audit measurement location.

2.2 Sound Level Monitoring Program

The Facility is a continuous and steady state noise source, which interrupts operation for temporary maintenance shutdown periods only. Although the noise emissions from the Facility and equipment are constant and do not fluctuate, intermittent noise complaints continued, prompting CRA's investigation. The monitoring was conducted for an extended time period in order to capture a variety of meteorological conditions in an attempt to understand the irregularity of noise complaints received by the government.

Continuous 24-hour environmental sound level monitoring was conducted in March 2010 under late winter conditions (no foliage, minimal snow cover and zero noise influence from wildlife and insects). Type 1 precision sound level monitoring and continuous data logging systems were established approximately 3 to 4 metres (m) from the most exposed façades of the two residences (POR1 and POR2) and the microphones were extended approximately 3 m above grade. A sound recorder was also set-up at one residence (POR1) and audio samples were collected when the environmental noise reached or exceeded the 40 dBA noise limit.

The one-hour Leqs, excluding extraneous background and weather noise influence, representative of noise impact from the Facility generally occur during the quietest nighttime period between the hours of approximately 12 a.m. and 4 a.m. when non-Facility background noise influence was minimal.

3. RESULTS

3.1 Monitoring Program Results

The sound level-monitoring program determined that under select down wind and atmospheric conditions the off-site environmental noise impact from the Facility at POR1 is significantly increased from 44 dBA to 46 dBA. Under downwind conditions, POR2 also experienced increased sound levels from 38 dBA to 42 dBA. The measured sound levels exceed the 40 dBA limit at the nearest residences, which confirmed an adverse noise impact was being experienced at night. However, when

winds were blowing upwind and opposite from the Facility, sound levels at the receivers were measured as low as 33 dBA to 35 dBA.

The ISO 9613 standard modelling protocol was under-predicting the noise impact at POR1 and POR2 by approximately 5 dBA.

3.2 Meteorological Data Analysis

CRA evaluated 5 years of meteorological data specific to the site from 2004 to 2008 and a windrose was generated. The 5-year data analysis was useful in order to determine wind speeds and the overall frequency in which the wind was blowing from the Facility toward each complainant POR based on the critical downwind directions of 345 degrees for POR1 = 3%, and 40 degrees for POR2 = 8%.

The 5-year meteorological data analysis determined an average wind speed of 4 m/s. Wind induced sound is estimated to be approximately 40 dBA.

ISO 9613-2 [1] allows for a meteorological correction, using the following equation:

$$Cmet = C_0 [1 - 10(H_s + H_r) / d_p]$$

Where

h_s is the source height (m)

H_r is the receiver height (m)

d_p is the distance between the source and receiver (m)

C_0 is a factor, in decibels, which depends on local meteorological statistics for wind speed and direction and temperature gradients

The correction accounts for conditions that are unfavorable to propagation as experienced by the off-site receivers POR1 and POR2.

ISO 9613-2 states, "Experience indicates that values of C_0 in practice are limited to the range of zero to approximately +5 dB, and values in excess of 2 dB are exceptional" [1]. The ISO 9613 standard modelling protocol was under-predicting the actual noise impact experienced at POR1 by approximately 5 dBA, which confirmed the practical findings of ISO 9613 and the need for a meteorological correction for the Facility.

The standard model protocol was adjusted to evaluate the true worst-case based on receiver-specific wind speed and wind direction. The frequency of these critical downwind conditions were very low overall, and therefore any noise impact assessment specific to these conditions can be considered extremely conservative, or "exceptional" per ISO 9613, but justified as the cause of noise complaints. A "D - Neutral" atmospheric stability class was used.

The modelled downwind noise impacts are very comparable to the sound level range measured under downwind conditions in March 2010 as follows:

Point-of-Reception	Model Name	Model Result	Monitoring Result Range
POR1	Downwind Conditions at POR1 (4 m/s wind, D stability class)	44.7 dBA	44 to 46 dBA (40 wind + 44.7 Facility = 46 dBA)
POR2	Downwind Conditions at POR2 (4 m/s wind, D stability class)	38.3 dBA	38 to 42 dBA (40 wind + 38.3 Facility = 42 dBA)

4. DISCUSSION

Acoustic professionals must be aware that the unadjusted ISO 9613-2 equations may under-predict the environmental noise impact by up to 5 dBA and critical for sensitive acoustic environments such as rural areas that experience low background noise and have direct line of sight noise exposure.

Meteorological data must be analyzed to determine the appropriate receiver-specific meteorological conditions to evaluate the predictable worst-case.

REFERENCES

ISO 9613-2: 1996(E) Acoustics – Attenuation of Sound During Propagation Outdoors – Part 2: General Method of Calculation.

CSA AND HEARING CONSERVATION

Tim Kelsall[1], Alberto Behar,[2] and David Shanahan[3]
[1]Noise and Vibration, Hatch, Mississauga, tkelsall@hatch.ca
[2]IBMME, University of Toronto, alberto.behar@utoronto.ca
[3]CSA, dave.shanahan@csa.ca

1. INTRODUCTION

CSA has made the decision to disband the Technical Committee Z107 TC on Acoustics and Noise Control. There were several reasons for this action, the main been that there was not sufficient support from users to continue with those areas and also the cost of producing new standards was too high.

Instead, CSA has decided to focus into occupational hearing conservation, including hearing protection and audiometric tests.

There were three groups working in those areas before the change: a subcommittee dealing with hearing measurements, another involved in the measurement of noise exposure (both of them subcommittees in Z107) and a committee dealing with hearing protectors.

The new Technical Committee on Occupational Hearing Conservation is incorporating the above groups plus others from the Z107 committees that are related to hearing conservation.

Standards produced by the Committee are required to be in keeping with existing Canadian standards on Occupational Health and Safety Management Systems and Medical Assessment Practice.

The new TC is one of 43 such committees under the Occupational Health & Safety Standards Program. They all are established under the Strategic Steering Committee on OHS Standards. The Steering Committee, in turn, reports up to the CSA Standard Policy Board. It is the Steering Committee's responsibility to approve TC terms of reference and appointments of TC Chairs.

Five standards from the former Z107 series have been transferred to the new TC. They are: Z94.2, Z107.4, A107.6, Z107.56, and Z107.58.

Meanwhile the Z107.10 standard, Guide for the Use of Acoustical Standards in Canada will be transferred to a new Acoustical Standards Committee formed by the Canadian Acoustical Association. It will also coordinate all non-occupational acoustical standards activities. See Reference 1 for more details.

2. HEARING CONSERVATION

There is a difference between hearing conservation and hearing protection. Hearing conservation refers to a global, management system that deals with noise and vibration in the workplace. As such, the following subjects are focused on:

a) Workplace noise and vibration measurements.
b) Assessment of occupational exposure to noise and vibration.
c) Selection, training and use of hearing protection devices in the workplace.
d) Strategies for reducing the exposure.
e) Noise and vibration control systems in the workplace.
f) Audiometric testing for early detection of occupational hearing loss.

3. THE NEW TECHNICAL COMMITTEE ON OCCUPATIONAL HEARING CONSERVATION.

Following is the composition of the Committee:

Chair: Alberto Behar, University of Toronto,
Vice Chair: Tim Kelsall, Hatch

Subcommittees and Chairs:

SC 1 (S304.3) – Hearing Protection,
SC Chair: Terry Van Volsen, Sperian

SC 2 (S304.4) – Noise Exposure Assessment and Control. SC Chair: Tim Kelsall, Hatch

SC 3 (S304.5) – Hearing Surveillance (Audiometry).
SC Chair: Christian Giguère, Université d'Ottawa

SC 4 (S304.6) – Vibration Exposure Assessment and Control. SC Chair: Tony Brammer, Former NRC

SC 5 (S304.7) – Hearing Conservation Management.
SC Chair: Jeffrey Goldberg, Custom Protect Ear.

REFERENCES

1. Kelsall and Giguere, The New CAA Acoustical Standards Committee, Acoustics Week in Canada 2010

Noise Pollution

The SLARM™ Solution

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

Typical Applications Include:

Community

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

Industrial

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

FEATURES

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

ACO Pacific

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

www.acopacific.com acopac@acopacific.com

ACOustics Begins With ACO™

THE NEW CAA ACOUSTICAL STANDARDS COMMITTEE

Tim Kelsall[1], and Christian Giguere[2]
[1]Noise & Vibration, Hatch, Mississauga, tkelsall@hatch.ca
[2]Audiology & SLP Program, University of Ottawa cgiguere@uottawa.ca

CSA made the decision to disband their Technical Committee Z107 TC on Acoustics and Noise Control which has acted as a clearinghouse for noise standards activity in Canada. They are keeping the occupational noise standards from Z107 and transferring them to a renewed Occupational Hearing Conservation Technical Committee which will also include the Z94.2 committee on hearing protection.

CSA will be transferring their standard Z107.10 to the Canadian Acoustical Association's newly formed Standards Committee. This standard is a listing of acoustical standards from Canada, the US, ISO and elsewhere which are deemed useful for Canadian use. Each listing includes a brief description f the endorsed standard and any recommended changes or concerns for its use in Canada. For example, some ISO standards do not adequately discuss the use of instrumentation in cold weather or the effect of snow on ground absorption and this would be flagged and recommendations made for dealing with these issues.

In addition, the committee is expected to help coordinate the Canadian representation on a variety of ASTM, ISO, IEC and other standards groups. This was one of the original reasons that the CAA was set up and this function has now returned to its original home.

In 2009 the Canadian Standards Association, CSA, made the decision to disband their Technical Committee Z107 on Acoustics and Noise Control which has acted as a clearinghouse for noise standards activity in Canada. CSA are keeping the occupational noise standards from Z107 and transferring them to a renewed occupational Hearing Conservation Technical Committee which will also include the Z94.2 committee on hearing protection.

To fill the gap left in Environmental Acoustics and other fields beyond Occupational Noise, the Canadian Acoustical Association (CAA) board agreed to found an Acoustical Standards Committee, which held its first meeting in May in conjunction with the new CSA committee. The intent is to keep the two committees in close liaison. In addition, the CAA committee will be continuing the Z107 role in coordinating the various Canadian groups working with other standards writing groups, including ISO, IEC, ASTM and ASA, specifically in Building Acoustics, Instrumentation, Acoustics and Noise.

Historically the Canadian Acoustical Association was originally formed to carry out this role of coordinating Canadian acoustical standards activity before it was taken over by CSA Z107. The latter has always held its annual meeting in association with the Canadian Acoustical Association's Acoustics Week in Canada conference and that will of course continue with the new CAA Acoustical Standards Committee.

It is important to realize the difference between CSA and CAA. The CSA is one of only four standards writing groups in Canada authorised by the Standards Council of Canada to prepare National Standards of Canada. The new CAA committee does not have this designation and at present there is no plan to get it. While it can write acoustical standards, they would have no official status in Canada, other than being of use to Canadian Acousticians. By contrast, the US has a much larger group of standards writing organizations and the Acoustical Society of America prepares official US acoustical standards.

The Standards Council of Canada also organizes the ISO and IEC advisory committees in order to review them and recommend Canada's position. These groups have always been harmonised with the Z107 committee to which they have reported regularly on progress and upcoming issues. In addition, Canadian representatives on ASTM and ANSI committees regularly updated Z107. This coordination is also expected to continue with the new CAA committee, which has been structured into Subcommittees encompassing the existing Standards Council Advisory Committees.

CSA has agreed to transfer their standard Z107.10[1] to the new CAA Standards Committee. Z107.10 is a listing of acoustical standards from Canada, the US, ISO, IEC and elsewhere which are deemed useful for Canadian use. This has been the main avenue for the CSA to endorse international standards of importance to Canada. Each listing includes a brief description of the endorsed standard and any recommended changes or concerns for its use in Canada. For example, some ISO standards do not adequately discuss the use of instrumentation in cold weather or the effect of snow on ground absorption and this would be flagged and recommendations made for dealing with these issues.

The intent is to rebrand Z107.10 and continue to update it and augment it. The hope is that this standard will remain the definitive list of acoustical standards for use in Canada. It will be published on the CAA website at little or no cost to the user and it is hoped that this resource will help CAA members and other acousticians in Canada and around the world to understand and appreciate the range of acoustical standards available in Canada and in many other countries,

while at the same time raising the profile of the CAA among Canadian and international acousticians.

CAA is a member of I-INCE and other recognized international acoustics groups. The new Standards Committee will also help provide technical advice to the CAA board on technical issues raised by these groups. Already it has been used to respond to an I-INCE proposal to establish a new Technical Study Group (TSG-9) on Metrics for Environmental Noise Assessment and Control and an INCE/Japan technical questionnaire on Policies on Environmental Vibration in Different Countries. The Standards Committee is available to help answer such technical acoustics questions from abroad and from governments and other organizations within Canada.

Membership at present is from members of the old Z107 committee, however the committee is open to all CAA members with an interest in and expertise on the range of areas which will be discussed. Interested members should contact the current chair, Tim Kelsall, or attend the meeting, which will be held this year on the evening of Wednesday 13 October in Victoria.

REFERENCES

1. CSA Standard Z107.10, Guide for the Use of Acoustical Standards in Canada

MOVEMENT, MEMORY & THE SENSES IN SOUNDSCAPE STUDIES

Jennifer Schine
School of Communication, Simon Fraser University
8888 University Drive, Burnaby, BC, Canada V5A 1S6 jschine@sfu.ca

1. INTRODUCTION

This presentation will explore how the practice of soundwalking can be a tool for memory retrieval. I ask: How are memories created and remembered in the mind and *felt* within the body? What happens to our perception of self, home, and knowing as we move through spaces and places of significance? I aim to explore the subject of memory and movement within the context of soundscape studies; these notions require an understanding of how we "hear" the past and re-evoke our acoustic memories as we move and act through our environment. Traditional methods for recalling the past involve mainly visual cues and focus on materiality—we look to photographs and hold personal objects, etc.—while remaining visual-centered and localized. I suggest that it is the physical act of moving our body through meaningful environments that unifies the senses, places and knowing and that brings together the local past into the present experience (Casey 1987). My main focus is to understand the ways in which people remember—both as individuals and as groups. This presentation explores how the production of memory and act of remembering are evoked during the process of memory walks (or soundwalks) as a way of understanding and engaging with the world.

2. CONTEXT: ACOUSTIC COMMUNICATION

Barry Truax's text, *Acoustic Communication*, describes the concept of an acoustic community—where sound functions positively in the locale to create a unifying relationship with the environment. Acoustic cues and signals are aural reminders and temporal nods to the rhythms of daily life; they help define an area spatially, temporally, socially and culturally. An acoustic community is thus linked and defined by such sounds that signify not only daily and seasonal cycles, but also the shared activities, rituals and dominant institutions of the area. A community with good acoustic definition can easily recognize, identify and derive meaning from the soundscape. In his second chapter, Truax examines the importance of recollecting historical sounds, expanding on the World Soundscape Project's notion of "earwitness accounts". Such accounts reveal how different sounds provide useful information, patterns of association and, perhaps most importantly, which sounds are meaningful enough to be remembered. What Truax is interested in is "the way in which sounds are stored in memory, not separately, but in association with their original context, betray[ing] a fundamental aspect of the listening process" (Truax 2001a). In this vein, it is important to examine cultural "soundmarks" and Schafer's "sound romances".

3. DISCUSSION: MOVEMENT & MEMORY

By walking, we are in a dialogue with the environment; both literally and figuratively, we re-situate ourselves into our story. Anthropologist, Jo Lee, suggests that the act of "walking can be understood as a personal biography: the body moves, in part, due to its links between past, present and future in a life" (Lee 2004). This understanding of place as constituted by movement, memory and biography is a new concept, emerging from such interdisciplinary fields as: soundscape studies (Järviluoma; Hyvarinen; Truax; Kyto; Vikman), ecological anthropology (Ingold; Lee), and ecological psychology (Gibson). In addition, the effects of the practice of walking on memory retrieval have been studied within such disciplines as: psychology, education, social sciences, health and medicine and gerontology (Stones; Dawe). These health studies show that there is strong evidence for an exercise-memory link (Eisner 2004). More importantly, however, they reveal that the body is capable of recovering memory through movement and specifically through the practice of walking.

3.1 Soundwalking

The exploration of soundwalks have been used as an important tool for aural awareness by scholars within the field of acoustic ecology and communication (Schafer; Westerkamp; Smith; Järviluoma; Southworth). Soundwalking is an exploration of sound with the intent of active listening—hearing all environmental sounds while moving in and throughout the environment. I suggest it is not just the act of walking, but the embodiment (Csordas; Feld) of the walk—or the emplacement (Howes; Pink) of the walk—that re-evokes our remembering. Memory is not merely activated through the visual surveillance of landscape, but by our interrelated perceptual understandings and bodily movements—of being in and engaging with the physical and sensory environment—this includes our aural perception and other sensory outlets.

3.2 Nostalgia & the Moment of Recollection

Through the practice of soundwalks, I am interested in how the experience of sound, companioned with the act of walking, affects our sense of time, place and

personhood. What does movement in our environment do during that moment of recollection? To understand this, we must understand that memory is contextual; moreover, the past is constantly being mediated and produced through memory work. The moment of recollection and the concept of nostalgia are thus important to examine.

4. POSSIBLE CONCLUSIONS

This presentation demonstrates that there is a need to examine the relationships between memory, the senses and movement from the viewpoint of embodied sound cognition and the practice of soundwalking. Soundscape studies include an empirical understanding of the senses as interrelated and interconnected; the discipline's theoretical framework has historically explored aural sensibility, perception, and knowledge and has utilized the practice of soundwalks to walk and engage with others. Since the discursive terrain of sensory ethnography (Pink 2009) is part of a new and growing paradigm in ethnographic and scholarly practice, soundscape researchers have an opportunity to voice what other disciplines have previously failed to account for. The question then becomes one of methodology: How do we learn about other people's sensory embodied knowing and remembering? And, in turn, how do researchers situate their own sensory emplacement and memory in relation to the people that they study? What this paper offers, then, is an exploration of walking and listening with others and how the practice of soundwalking can be a method of understanding, itself: "the record of the walk, and of the experience it affords, is just as important—and just as valid a source of field material—as the record of the 'discourse' that might have accompanied it" (Lee & Ingold 2006).

REFERENCES

Casey, E. (1987). Remembering: A Phenomenological Study. Bloomington: Indiana University Press.

Csordas, Thomas J. (1994). *Embodiment and Experience: The existential ground of culture and self* Cambridge: Cambridge University Press.

Eisner, E. W. (2004). *Memory Fitness: A Guide For Successful Aging*. New York: Yale University Press.

Feld, S. (1996). Waterfalls of Song: An Acoustemology of Place Resounding in Bosavi, Papua New Guinea. In Feld, S. and Basso, K. (Ed.), *Senses of Place*. Santa FE, NM: SAR Press.

Gibson, E. and Schmuckler, M.A. (1989). Going Somewhere: An Ecological and Experimental Approach to Development of Mobility. In *Ecological Psychology*. 1(1), 3-25.

Ingold, T. (2004). Culture on the Ground: The World Perceived Through the Feet. In *Journal of Material Culture*. 9: 315-340.

Järviluoma, H. (2009). Lesconil, my home—Memories of listening. In Järviluoma, H., Truax, B., Kyto, M. and Vikman, N. (Ed.), Acoustic Environments in Change. Tampere: University of Joensuu.

Järviluoma, H., Truax, B., Kyto, M. and Vikman, N. (2009). *Acoustic Environments in Change*. Tampere: University of Joensuu.

Howes, D. (2005). Introdution, in Howes, D. (Ed.), *Empire of the Senses: The Sensory Culture Reader*. Oxford: Berg.

Hyvarinen, T. (2009). 'Putt putt' and 'mur'—Old inboard engine and nostalgia. In Järviluoma, H., Truax, B., Kyto, M. and Vikman, N. (Ed.), *Acoustic Environments in Change*. Tampere: University of Joensuu.

Lee, J (2004). Culture from the ground: walking, movement and placemaking. In *Association of Social Anthropologist conference*.

Lee, J. and Ingold, T. 2006. 'Fieldwork on foot: perceiving, routing, socialising'. In S. Coleman and P. Collins (eds), *Locating the Field: Space, Place and Context in Anthropology*. Oxford: Berg.

Pink, S. (2009). *Doing Sensory Ethnography*. London: Sage.

Schafer, R.M. (1977). *The Tuning of the World*. New York: Knopf.

Smith, C. (1993). The Acoustic Experience of Place. In: *Proceedings of the First International Conference for Sound Ecology, The Tuning of the World*. Vol.II, no publisher, no pagination.

Stones, M.J. and Dawe, D. (1993). Acute exercise facilitates semantically cued memory in nursing home residents. In *J Am Geriatc Soc.*, May; 41 (5): 531-4. PubMed PMID: 8486887.

Southwork, M. (1969). The Sonic Environment of Cities. In *Environment and Behaviour*. Vol 1(1).

Truax, B. (2001). *Acoustic Communication 2nd edition*. Westport, CT: Greenwood Press.

Westerkamp, H. (1974). Soundwalking. *Sound Heritage,* III/4, 18-27.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Barry Truax, for his wholehearted support, my colleagues presenting at this conference for their enthusiasm, and Hildegard Westerkamp and the Vancouver Soundwalk Collective for their energy and inspiration. I feel fortunate to be a part of an acoustic community that listens deeply. I would also like to acknowledge my funders: the Social Science and Humanities Research Council (SSHRC), Simon Fraser University and the School of Communication.

AUTHORS NOTES

Jennifer Schine is a graduate student in the School of Communication at Simon Fraser University. Her research investigates concepts of identity, memory and movement within the field of acoustic communication and soundscape studies. She is interested in the relationships between audio heritage, listening practices, sensory experience and other worldviews.

RELOCATING THE EAR: A CROSS-CULTURAL EXPLORATION OF THE ELECTRIFIED SOUNDSCAPE

Vincent Andrisani
School of Communication, Simon Fraser University, 8888 University Dr., British Columbia, V5A 1S6. Email: vincent_andrisani@sfu.ca

1. INTRODUCTION

Throughout the twentieth century, the relationship between the individual and their social environment experienced a profound communicative transformation. The 'acoustic community', regarded as a collection of unique local sounds and their rhythmic interplay (Schafer 1994; World Soundscape Project 1977), has now become characterized by an aura of homogenization (Truax 2008). Both the increased interpersonal reliance upon electroacoustic forms of communication, as well as that which has come at the hands of capitalist discourse (as in the use of Muzak), continues to suppress and mask acoustic attributes that were once definitive of local culture. Simultaneously, we have experienced a shift in the ability of the listener, evidenced by the declining aptitude for soundscape competence, which is defined as "the ability to understand environmental sound as meaningful" (Truax 2001). Once a functional, and even vital component of the communicative process, it has now largely become a neglected means of environmental exchange. The problem we face today is the manner in which we as listeners negotiate not only the surrogate environment created by neo-orality (Ong 1982), but also our own perceptual detachment, so that we may once again begin to engage and comprehend the nature of the relationship between the individual and the environment.

This exploration examines the attributes of the electrified soundscape, and sets them against a backdrop of cross-cultural approaches to soundmaking and listening. In this context, the engaged aural correspondent can positively employ personal perceptual sensitivities, listening habits, and one's own unique history with sound, in order to renegotiate acoustic space with the ability of soundscape competence.

2. NEGOTIATING ACOUSTIC SPACE

The concept of "acoustic space", derived out of the work of Edmund Carpenter and Marshall McLuhan (1960), makes direct reference to the breadth of spatio-temporal information exhibited by the soundscape. It refers not only to the sound of the environment, but perhaps more significantly, to the manner in which relationships are created or neglected through the use of the aural modality.

There are two periods during which the notion of acoustic space has undergone a pronounced adjustment. The first came as a result of the advent of the phonetic alphabet, where the emphasis on the visual modality encouraged the recalibration of the cognitive and therefore perceptive framework (Carpenter and McLuhan 1960; Ong 1982, etc.). The second period of adjustment came at the beginning of the twentieth century with the advent of electroacoustic reproduction. The ability to reproduce sound signaled for the first time its removal from an original context, which thus began the spatio-temporal fragmentation of acoustic space – one that continues to intensify with the use of modern media. It must however be noted that the use and availability of modern extensions, and therefore the dominant mode of communication, varies greatly depending on context. Similarly, the cognitive and perceptive bias is also dependent upon context, as it remains a product of cultural construction (Howes 1991). Thus, by employing a comparative method that is at once sociological and anthropological, it opens up a world of possibility through which the social uses of electroacoustic media can be utilized to reconstruct the dynamics of evolutionary change.

3. CROSS-CULTURAL APPROACHES

By cross-culturally engaging the framework of acoustic communication and soundscape studies, one is offered the opportunity to assess the similarities and differences of the localized acoustic environment. There are three ways in which this is realized. First, the social and cultural "distance" of the aural correspondent offers the privilege of exploring the soundscape from a unique perspective. From this position, one can locate particular features of the acoustic environment that the process of acoustic habituation would otherwise conceal (Schafer 1994; Vikman 2009). Next, is the need for the acoustic researcher to temper the distance between themselves and the functionality of the soundscape. This implies that they are required to make sense of, and derive meaning from, environmental sound. It is at this point that the notion of soundscape competence is once again activated, so as to work toward a heightened understanding of the soundscape in question. And the final service offered by the comparative approach is the revelation of absences. The aural correspondent with a functional comprehension of the communicative process is inclined toward the discovery of difference through absence. Thus, it is only by way of contrast that the opportunity to understand particular acoustical features as unique to the given context can be ascertained.

4. SOCIAL USES OF TECHNOLOGY

Havana, Cuba and Vancouver, Canada offer an example of two distinct environments in which the differences in the social uses of technology are abundant. First, the history of sound in Cuba is one that places a great

deal of communicative significance in the act of music making. This tendency toward acoustical production and reception as a fundamental part of the process of communication is one that extends throughout the continuum of sound, which alongside music also includes speech and the soundscape (Andrisani 2009). Further, the island's political history offers an added twist, where the triumph of the revolution and the economic (and cultural) embargo with the United States has left Cuba in it's own "acoustic bubble" – a socio-cultural environment that is largely absent of the technological forms employed in a metropolis such as Vancouver (e.g. the smart phone, the iPod, and the Internet).

Perhaps the most interesting function of a particular technology in Cuba is that of the shortwave radio. In many ways, it serves as an extension of the public address system, whereby the soundscape is altered on account of the dissemination of the ideology of the Cuban government. In this scenario, acoustic space, both public and private, is reconfigured through the use of modern extensions. These modern technologies help to advance and fortify the success of the revolution, and in so doing, they articulate a social and political hierarchy that is notably different than that which is present in Vancouver. Equally interesting is the manner in which radio functioned as a cultural portal for Cuba during the 1970's. It was through this channel that American bands were introduced to island culture, which subverted the political and economic embargo. As a result, the prevailing influence of African-American music on the local musical form can be traced back to the information offered through shortwave radio emanating from Miami. This in turn illustrates the mutual co-dependence of both the extension and the context, where the two are engaged in constant dialogue, balancing the notion of cultural identity on the axle of this relationship.

5. CONCLUSION

If identity is always about narrative, and producing in the future an account of the past (Hall 1995), then part of the interest in soundscape studies lies in the fact that sound is the medium through which those stories are told. In a manner that echoes the ephemerality of sound, identity too is dynamic, always in flux. It is, as Simon Frith (1996) describes, "not a thing, but a process", one that is most vividly grasped through sound. And while admittedly the soundscape may not offer the equivalent level of organization as the musical form, it is equally a byproduct of cultural construction. Thus, cross-culturally engaging the soundscape offers more than mere insight into the functionality of the communicative process. Rather, the aural correspondent not only experiences what it means to be part of, and identify with, the culture in question; but in so doing, they can also unearth a new manner through which to identify with the culture to which they belong.

REFERENCES

Andrisani, Vincent. "The Soundscape of Havana and the Rhythm of the Clave." *WFAE Sound Megalopolis: Cultural Identity and Sounds in Danger of Extinction.* Mexico City: Fonoteca Nacional, 2009. 257-270.

Carpenter, Edmund, and Marshall McLuhan. "Acoustic Space." In *Explorations in Communication: An Anthology*, 65-70. Boston, MA: Beacon Press, 1960.

Frith, Simon. "Music and Identity." In *Questions of Cultural Identity*, edited by Stuart Hall and Paul DuGuay, 108-125. London: Sage, 1996.

Hall, Stuart. "Negotiating Caribbean Identities." NLR, no. 1/09 (1995): 3-14.

Howes, David. *The Varieties of Sensory Experience: A Sourcebook in the Anthropology of the Senses.* Toronto, ON: University of Toronto Press, 1991.

Innis, Harold. *The Bias of Communication.* Toronto, ON: University of Toronto Press, 1951.

Ong, Walter. *Orality and Literacy: The Technologizing of the Word.* New York, NY: Routledge, 1982.

Schafer, R Murray. *The Soundscape: The Sonic Environment and the Tuning of the World.* Rochester, VT: Destiny Books, 1994.

Truax, Barry. *Acoustic Communication.* 2nd Edition. Westport, CT: Ablex Publishing, 2001.

Truax, Barry. "Soundscape Composition as Global Music: Electroacoustic Music as Soundscape." *Organised Sound* (Cambridge University Press) 13, no. 2 (2008): 103-109.

Vikman, Noora. "The Changing Soundscapes of Cembra Village." In *Acoustic Environments in Change.* Tampereen: TAMK University of Applied Sciences, 2009.

World Soundscape Project. *Five Village Soundscapes.* Edited by R Murray Schafer. Vancouver, BC: A.R.C. Publications, 1977.

ACKNOWLEDGEMENTS

I would like to thank both Professor Barry Truax as well as my colleagues at Simon Fraser University for their continued support and encouragement.

AUTHOR NOTE

This project was conducted during the spring of 2010. It is available in its entirety at www.sensorystudies.org.

THE SOUNDING MUSEUM: TWO WEEKS IN ALERT BAY

Hein Schoer, MA
Research Group Arts in Society, Fontys School of the Arts, Tilburg (NL) / Maastricht University (NL), h.schoer@fontys.nl
for more information: www.fontys.edu/artsinsociety, www.gruenrekorder.de, www.nonam.ch, www.auditorium-mundi.com

EXPOSITION: COMING TO IN ALERT BAY

A raven croaks.

I hear a faint broadband noise, not static but gently varying in volume, centre frequency and Zwicker's parameters. At first there are only short blows that enter my consciousness, but as I regain awareness, it becomes a constant, swooshing sound, that, as I realise after a moment of malorientation, arrives at my ears mainly from the half-open window of my room at the Alert Bay Lodge facing the forest behind it.

It's still dark, so I keep lying and listening. The ocean waves manage to escape the masking grip of the wind in the trees. I can hear them very well now; although their frequency band is not that different from the wind, the internal structure of the waveform allows my brain to separate them from one another. I rub the fatigue out of my eyes, slip into my clothes, grab my gear, and head for the opening in the middle of the forest of Cormorant Island Ecological Park.

The spirits of the skies are with me tonight, I think, as the wind stops to blow and leaves nothing but silence. I had been up here late in the evening, and had made a good quarter of an hour of nice ambience recordings, the dominant sound had been wind, but for the morning atmo I prefer a low noisefloor to be able to capture the awaking bugs and birds. With the whole island still asleep the dying wind leaves me in peaceful quiet (except for a hint of tinnitus).

1ST WORLD/MOVEMENT: THE NATURAL SOUNDSCAPE

I have come to Alert Bay, BC (British Columbia) in the course of my research project "The Sounding Museum", within which I look at (and listen to) the production and presentation of cultural soundscapes.

The crows are the first to come alive, some distant, some closer to my recording position, and all around the place, making it evident that they are the dominant species in the sonic layer of the local ecosystem. Raven may still be the trickster lord of the spirit realm, but he is definitely not alone in the animal kingdom. Next to his croaks a woodpecker knocks, and a number of smaller birds, the names of which I do not know, claim their territory. I wish I had the keenness of the Kaluli of Bosavi who Steven Feld visited, then I could attach names and meaning to their song. Luckily I haven't come here for ornithological studies.

At the inhabited end of the island there also seems to be some early bird, as I can hear a first car engine going off.

As the sun rises above the trees, I turn to the stony beach. The nuthatch that had accompanied me with its sharp high-pitched short yelps now has to share the acoustic space with numerous seabirds, above all seagulls, but also some duck- or gooselike critters.

Now I have to wait for rain, as no BC soundscape can be complete without the constant rain that covers all the land and the sea with a veil of thin grey braids.

2ND WORLD/MOVEMENT: THE ARTIFICIAL SOUNDSCAPE

Despite its small population Alert Bay is one of the cultural centres of the Kwakwawa'wakw, not the least due to the U'mista Cultural Society that is located here, and, as claimed by Bruce Alfred, one of the few who still possess the knowledge and the skill to create bentwood artwork, the heart of woodcarving at the Pacific Northwest Coast.

The natural beach atmo mixes with an Alert Bay ambience. Now the day has really started, the town is getting busy. I hear cars passing, harbour activity, and the starlings and crows in front of the U'mista Cultural Centre. I never get to catch the voice of the bald eagle. I see one sitting on a pole in the water one evening; he would stare at me for an hour, and I would stare back, shotgun ready, but he wouldn't talk.

Sometime around the eagle incident I finally realise that it would be folly to try and create a representational soundscape covering all there is to Northwest Coast indigenous culture and wildlife. What I make instead is a very personal, locally and temporally constricted composition, presenting the impressions of my two-week stay at Alert Bay, featuring the individuals I meet, the events I witness, and the atmospheres I feel in nature and around people.

I had spent some time with Bruce Alfred in the carvers' workshop the other day, made recordings of woodworking tools in action, and a long interview about the difficulty of finding apprentices to continue with the traditional arts, and how the old ways can help the young to find their place in the world. Now I am at Beau Dick's house, recording Beau bucking some block of wood with a handsaw and hollowing the backside of a small totem pole. We go inside, where I just sit for a good hour to talk and listen and watch the carvers at work, while the microphone patiently records.

Being in this busy workshop atmosphere offers a beautiful insight into the many facets and layers of a living culture with such ancient origins. The masks talk of a mythic past and tradition, and some of them will be used in the spirit of that tradition. Others are carved to make a living, exploiting the Man's craving for exoticism in his living room.

3RD WORLD/MOVEMENT: THE HUMAN SOUNDSCAPE

At the T'lisalagi'lakw Native School great efforts are made to counter the effects of the many decades of cultural deracination. Teachers like Sandi Willie, Lina Nichols, and Vera Newman not only teach the children how to read, write, count, and about the different parts of the nervous system, there are also daily classes on traditional song and dance, and Kwakwala language class.

I present them as I see (hear) them, not focussing exclusively on the "exotic" parts of the Alert Bay Soundscape, but rather trying to provide the whole picture of my stay, which included science class and a steak sandwich dinner (didn't record that, though) at the Nimpkish Pub just as much as a Hamatsa initiation (didn't include that, because it's not approved for public presentation) at a potlatch.

The children are obviously having fun, but also a hard time catching up. To me it seems doubtful if they will ever fully master this language, as they are growing up with English at home and everywhere else; Kwakwala is not spoken in everyday conversation anywhere, the residential school system has taken care of that.

For us the girls dance in full regalia, with their family crests on their button blankets. We witness four traditional dances: Ladies' Professional, Salmon Dance, Paddle Dance, and a dance with red shawls. I decide to keep the hum of the ventilation system in the mix, I won't try to remove it with my denoising tools; unlike the hardware noise resulting from a poor signal-to-noise ratio, it is part of the original soundscape, annoying maybe, but genuine.

4TH WORLD/MOVEMENT: THE CULTURAL SOUNDSCAPE & THE SPIRIT WORLD

Potlatch is the most important ceremonial festivity in Northwest Coast culture with deep religious and ritual meaning; dances, songs, masks, and treasures are presented to the audience that is provided for with vast amounts of food and presents. Fiercely challenged by the arrival of Europeans, the culture of the Northwest Coast people had been almost wiped out, potlatch as central expression of this culture was prohibited in 1884. With the ban officially lifted around 50 years ago, the Natives continue to bring back the original spirit of potlatching to life in a contemporary form.

Wa (William Wasden jr.) has invited me to join him at Chief Robert Duncan's potlatch at Campbell River, so the next morning sees us at the Bighouse at nine o'clock. Mourning songs have to be sung before noon, otherwise they will invite the spirits to come and prey on more people.

I will not attempt to try and describe what's happening, one has to be there, to watch and listen, and even then one can only get a vague idea of what it all means to the people.

The feast will go on until past midnight, countless dances and songs, a Hamatsa initiation, a clown's act, the chief will open his box of treasures, we will all eat and drink and have the stripes of red cedar bark around our heads, many speeches will be held, Wa will spread the eagle down… I feel I am being sucked into a wall of sound that soon is all around me. I begin to really understand what Beau Dick meant, when he explained how you get caught up in the music and the sound and "something takes over."

And indeed it does. Engulfed by the hypnotising power of voice and beat, I wander from the dim confinement of the bighouse, the music begins to blur and before I know it, I am in the Spirit World. From the mists of my limbic forest I hear a rejoicing fun dance cheer, the chief has shown his greatness, his people are glad. But they sink into the mists and I realise I am in the animal kingdom. Before I know it, mighty Thunderbird flies over my head and takes me even farther into the mist. I am thinking of using a couple of drum beats from the recording to mimic the thunder of his wings by means of some convolution reverb plug-ins.

The animals have stayed behind, the curtain is lifted to show the spirits lined up to pay me a visit (or do I visit them? After all, it is their world, and I am the intruder). There is tame Bukwas, there is Nutslala the firedancer, and Nutslama, screaming "Wi!Wi!Wi!", and he's running around, he's laughing, and he's throwing snot at me. There is Xwixwi, his tongue is sticking out of his mouth and his eyes are bulgy, and of course Dzunuqwa, coming out of the bushes with her "hu-hu" to scoop me into her basket. As I hear Dzunuqwi coming from the sea with his bubbling voice – his throat is filled with water, and he's got seaweed all over him – I notice that the soundscape becomes clearer again.

REPRISE: DOWN TO EARTH

The day ends in the cosy atmosphere of Woss Lake cabin, with a fine share of grilled, smoked, sweetened and teriyaki sockeye, halibut, prawns and eulachon (no, that will be at the seafood dinner with the totem poll restoration workshop guys, but the dinner at Woss Lake is no less opulent, just instead of the halibut we have hamburgers and potato salad), Wa singing children songs and the family talking.

U'melth, the Raven, croaks one last time for me, spreads his wings and flies away. Tomorrow I will be on my plane home.

It finally starts to rain.

ACKNOWLEDGEMENTS

I wish to thank the people of Alert Bay and the U'mista for sharing their world with me: William Wasden jr., Andrea Sanborn, Pewi Alfred, Marcus Alfred, Bruce Alfred, Beau Dick, Sandi Willie, Lina Nichols, Vera Newman, Sean Whonnock and many others, my endorsers and collaborators at fontys, NONAM, Maastricht University and elsewhere. Dedicated to Richard Schuckmann and Andrea Sanborn.

MOBILE SOUNDSCAPE MAPPING

Milena Droumeva
Faculty of Education, Simon Fraser University, Burnaby, BC, mvdroume@sfu.ca

1. INTRODUCTION

Implicit in the notion of Acoustic Ecology as a discipline, is a deep sense of understanding the surrounding acoustic environment, the nature of its balance, and the subtlety of its sonic components and patterns of our interactions with it. One of the forms of understanding, developed and promoted by Schafer, is a trained aural awareness – the cultivation of a keener ear, a re-sensitizing of our hearing to the sounds of our environment. Other ways of engaging into a meaningful understanding of acoustic ecology are acoustic design on one hand, and soundscape composition on the other. As Shafer notes himself, "Acoustic design should never become design control from above. It is rather a matter of the retrieval of a significant aural culture, and that is a task for everyone." (Schafer, 1977). This point is further contextualized by acoustic communication theorist, Barry Truax who points out that, "the necessity of the ecological concept springs from the context of loss, or at least from the present threat to survival. The question for us now is whether a new balance can be regained. Can we – with consciousness – be part of a new eco-system?" (1992). In light of soundscape design, we have to be informed by the past, and maintain an ecological balance of sound components. To start, we have to understand design as a system that "comprises the knowledge and the techniques that we understand and can put into practice," (1992) and that it involves everyone as listeners and soundmakers, not just the designer/composer.

The changing conditions of our media and technology-saturated world may be in part the cause of the 'dulling' of our ears, however, they also offer surprising and exciting new possibilities for soundscpae and acoustic deisgn, sharing of listening experiences and building of online/media communities around soundscape mapping. Mobile technologies and the web - particularly social networking - have in the last several years "democratized" the collection and sharing of soundscape recordings. While audio recorders are now a standard feature in many personal and mobile computing devices, social networking sites like facebook, twitter, and audio-specific sharing sites such as soundcloud, audioboo and woices, not only make it possible, but encourage the "mapping" of soundscapes in relation to geographic locations, images of the surrounding and built environment, as well as sharing all of these mappings with a larger online community - thus building a "world soundscape project" of a different kind. For owners of smartphones there are now accessible and reasonably accurate sound level meters and related real-time analysis applications. In this report I will trace some of the available tools for soundscape analysis, mapping, recording and sharing, and present the beginnings of a study that aims at better understanding listening practices and the role of urban soundscapes in people's lives through mobilizing these tools in the purpouseful mapping of the soundscape and reflection of sounds and listening.

2. ONLINE SOUNDSCAPE MAPPING

There seem to be several types of online communities for the sharing, recording, uploading and re-constructing of soundscapes and soundwalk recordings. Most use a type of geo-tagging or the aide of GPS-based geographic online typology, based on Google's already established project to allow audio google mapping.

2.1 Community-based online geo-tagging

Below are two screenshots – one from a dedicated online soundscape/soundwalking community called woices.org (world voices – Fig.1) and another from a project of radio Aporee (Fig. 2) initiated as a commemorative initiative for World Listening Day 2010.

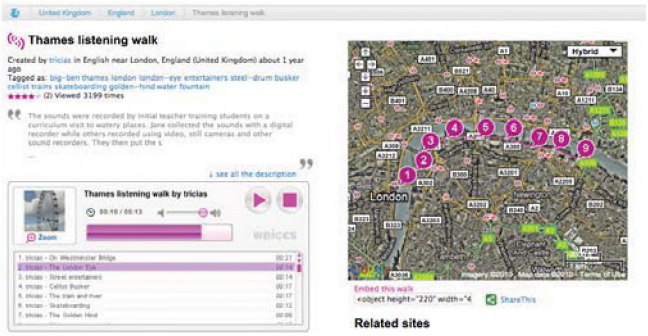


Fig. 1. An entry from woices.org – soundwalk around the Thames River in London presented as a series of nine recordings, contributed from several participants in the soundwalk.

Both are modeled after the World Soundscape project, and attempt, through a community-based contributions, to populate geo-web locations around the world, thus not only preserving sound environments, local soundmarks and personal experience, but also, in a more general way, promoting attentive listening and the active, creative, designer-ly engagement of people and soundscapes, an idea that has enjoyed a growing popularity among community artist groups and academics alike (Agoyard & Torgue 2005, Labelle 2010, McGregor et al 2002).

Fig. 2. A screenshot from radio aporee's contribution interface for World Listening day - utilizing Google Earth's API.

2.2 Generative geographic soundscape re-synthesis

A notable mention in surveying online community-based soundscape mapping spaces goes to Soundwalks.org (Fink et al. 2010) following Valle, Shirosa and Lombardo's work (2009). Soundwalks is a tool which collects user-uploaded sounds, organizes them according to an acoustic, semantic and social ontology and is then capable of re-synthesizing a desired soundwalk (as drawn in Fig. 3 below).

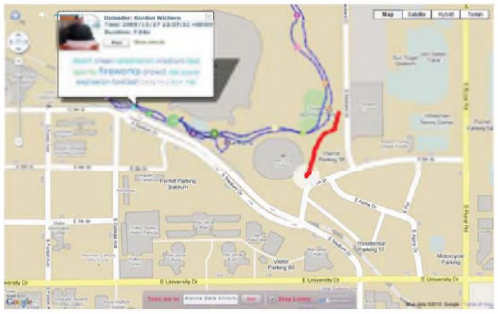


Fig. 3. Screen from Soundwalks.org – a tool that automatically generates a re-synthesis of a soundscapes, in the form of a soundwalk, from a user-uploaded sound file database for given locations.

3. MOBILE MAPPING / ANNOTATION

What I'd like to suggest now is that one element missing from the examples discussed above – an element that may be crucial to soundscape research, rather than just community soundscape mapping – is a way of qualifying, annotating and otherwise analyzing the uploaded sound files. As it stands, contributors upload a variety of audio and recording quality, not to mention, the subjective selection of sound environments and events goes unacknowledged. What I would like to present is a different way of annotating and presenting sound recordings where the sonic environment or acoustic community is of research interest. Using only an iPhone with a Recorder app and Faber Acoustical's dB app, I've collected, visually organized and described a number of sound locales into a sonic inventory of 'aural postcards' (Tonkiss, 2003) Posted on my blog (natuaural.com) I hope to build an ontology of archetypal urban sound environments organized within an analytical framework combining visual, narrative and measurement information. What I hope to demonstrate is that this way of presenting sonic maps offers a valuable addition to other geo-tag-based soundscape communities online, and a potential combination of the two could enrich the field of acoustic ecology research by combining descriptive, geographical, sonic and analytical accounts.

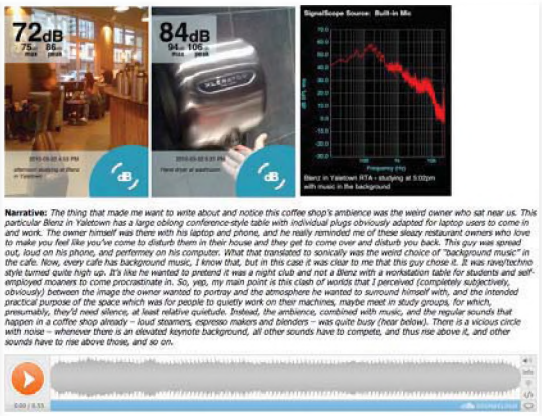


Fig. 4. An example of a blog entry with pictures of the space, overlaid decibel SPL readings, a screenshot of a real-time RTA graph, the sound file (using Soundcloud.org player) and a personal narrative/description.

REFERENCES

Augoyard, J.F. & H. Torgue (2005) *Sonic Experience: A Guide to Everyday Sounds.* Montreal, CA: McGill Queen's University Press.

Fink, A. et al. (2010) Re-sonification of geographic sound activity using acoustic, semantic, and social information. In *Proceedings International Conference on Auditory Display, 2010*, pp. 305-312.

Labelle, B. (2010) *Acoustic Territories:Sound Culture and Everyday Life*. London, UK: Continuum.

McGregor, I., A. Crerar, D. Benyon, and C. Macaulay, (2002) Soundfields and soundscapes: Reifying auditory communities, in *Proceedings of the 2002 International Conference on Auditory Display*.

Schafer, R. M. (1977) *The Tuning of the World*. New York: Knopf. Reprinted as Our Sonic Environment and the Soundscape: The Tuning of the World. Rochester, VT: Inner Traditions International, 1993.

Tonkiss, F. (2003) Aural Postcards: Sound, Memory and the City. The Auditory Culture Reader. In Eds. Michael Bull & Les Black, Oxford: Berg. pp.303-310.

Truax, Barry, (1992) Electroacoustic Music: The Inner and Outer World, In P.Seymour & T.Howell Eds., *A Companion to Contemporary Musical Thought* (pp. 363-386). London: Ruthledge.

Truax, B. (2001) Acoustic Communication. 2nd Ed. Norwood, NJ: Ablex Publishing.

Valle, A. and M. Schirosa, and V. Lombardo, (2009) A framework for soundscape analysis and re-synthesis, *in Proceedings of the Sound and Music Computing Conference*, Porto, Portugal, 2009.

SOUND MATTERS: MEDIATION, MIMESIS, AND EMBODIMENT IN SOUNDSCAPE MUSIC

Andrew Czink
Graduate Liberal Studies, Simon Fraser University, 8888 University Drive., BC, Canada, V5A-1S6 aczink@sfu.ca

1. INTRODUCTION

Audio technologies offer affordances for forms of representation in musical practices unavailable before their advent. The recording 'chain,' from microphone to storage medium to amplified reproduction allows the incorporation of 'real-world' sounds into musical, or organized sound, contexts. These electroacoustic technologies frame environmental sounds in seemingly neutral or naturalized ways, although they are by no means neutral. It is important to remember that environmental sound recordings are transduced, mediated representations of environmental sound(s), and that both human and technological agency are always implicated in these processes of mimesis. The temporal and spatial displacement of environmental sound recordings allows the creation of "…surrogate environments…" (Truax, 2008, p. 104) both in terms of our mediated everyday soundscape, as well as in the potentialities afforded to composers.

Mimetic electroacoustic practices will be considered in terms of embodied forms of transduction and prosthesis. The environment will be considered as a form of music, and music, as a kind of niche building, will be considered as a form of environment made possible through audio technologies. Finally, both listening and sound-making will be considered as technologies of the self and forms of awareness practice affording us a more nuanced understanding of our world.

2. RECORDING

2.1 The Role of the Microphone

The recording process begins, of course, with the microphone transducing acoustical into electrical energy, allowing composers to record sounds from the environment and to incorporate them into their work. While a seemingly neutral documentation process, the microphone, like the camera lens, requires the recordist to choose what to record and from what position, thus framing the recording in a particular way. The resulting recording is an index of the recordist's "composed listening." (Norman, 2004)

While the recordist clearly has intention and agency in making a recording the microphone itself exhibits a form of agency by requiring the recordist to move in particular ways, and by mediating his or her choices through the polar response pattern and frequency response of the device itself. The microphone, in effect, defines the space and sound of the resulting recording in conjunction with the intentions, interest, and goals of the recordist. As soundscape composer Hildegard Weterkamp, says: "the microphone can impart an intense glamour… [and] listening is a silent intelligence that directs us to what we think matters." (Norman, 2004, pp. 86, 77)

The definition of microphone as transducer may be extended to that of a 'transducing prosthesis', since we may consider the recordist to be actively transducing his or her composed listening into the material form of a recording and composition through its use, allowing the recordist to reach, or in this case, listen, further or more closely than otherwise possible. The model of transduction is developed to included transduction of acoustical energy into electrical energy which is further transduced into feelings, thoughts, and emotions in order to finally be transduced into compositions and other sound artifacts.

2.2 The Recording

By allowing one to store sounds, recording has fundamentally altered how we listen. As alluded to above, sound recordings are more than merely passive documents that give us access to some objective reality. The sound recording as document changes our understanding and interrelationship with the soundscape immediately in the act of creating it.

Not only is sound displaced from its original time and place of occurrence, but traces of the space within which the original sound took place are inscribed into the recorded document. We can hear the "aural architecture" (Blesser & Salter, 2007) of the space to some degree in every recording. Even close microphone techniques, used to minimize the sound of the acoustic space, impose their own particularly intimate and eroticized perspectives.

2.3 Amplification and Reproduction

Amplification and reproduction allow us to experience sound in previously unheard of ways. Amplification may work like a sonic microscope giving us access to sounds that would otherwise either be difficult to hear, or would be impossible to hear at all. Westerkamp's *Kits Beach Sound Walk* lets us hear the sound of barnacles, something normally impossible, particularly within an urban environment. David Dunn in *The Sound of Light in Trees*

using specially developed microphone technology allows us to hear the movements of beetles under the bark of trees.

The medium of recording, through its mediation of our listening experiences, and by allowing new relationships to be discovered and created, changes our relationship to sound, the environment, music, ourselves, and each other.

3. ENVIRONMENT AND MIMESIS

The environment, through recordings and soundscape compositions, becomes music. We can hear our environment musically and compose our listening in accordance. By composing with real-world sounds we are engaging in mimesis: we are imitating aspects of our world. Audio technologies, including recording as well as computer based audio technologies, allow us unique opportunities to engage in mimesis.

The use of environmental recordings in soundscape compositions is one obvious way of using mimesis. There are also numerous examples of computer modeling of natural processes that are clearly mimetic. Algorithms have been developed over the years to model the behaviour of complex natural events. Stochastic distributions are used to model the sound of raindrops for example, and algorithms have been developed to model the behaviour of swarms, herds, and flocks.

Adorno contrasted mimesis with rationality, (Windsor, 1996, p 192) and considered mimesis as a threat to the autonomy of the artwork: a process he associated with 'primitive' forms of art and culture. Adorno characterized art "as the product of enlightened rationality." (ibid) It may perhaps be the conscious pursuit of the 'primitive' or the pre-rational that fuels soundscape composers to pursue mimesis in their work.

Michael Taussig, in his book *Mimesis and Alterity,* considers mimesis as fundamental to our understanding of the world we inhabit. To Taussig, mimesis is positive in its opposition to Adorno's notion of a universal and context-free rationality and valorization of the abstract.

4. EMBODIMENT AND PROSTHESIS

Auditory experience and audio technology afford unique forms of embodiment. Sound is experienced as an enveloping and immersive medium. This is true whether we are hearing acoustic sound or technologically mediated sound reproduced through loudspeakers. It is important to make a distinction between the origin of a sound and the source of a sound. The source of electroacoustic sound is the loudspeaker, whereas the origin of the sounds heard will refer back to the now displaced original context of the sound recorded. This distinction clears up the problem of considering recorded sound as 'disembodied.' (Chanan, 2000) We can consider that the origin may be disembodied or displaced, but the experience of the reproduced sound is located in the situated space of the listener. Embodied experience is the only kind of experience we have, and are able to have, whether it occurs in a mediated environment or not.

Auditory experience, which encompasses the experience of music in any form, may be seen as a form of awareness practice that utilizes music as a prosthetic device to 'reach' further into our world and to enhance our understanding. Music allows us to occupy and be part of the world in the environment or territory within which we are situated. Mimetic forms of electroacoustic music "…enhance our understanding of the world, and its influence carries over into everyday perceptual habits." (Truax, 2008, p. 106) Music may be considered as prosthesis precisely because it enables us to explore aspects of our world in ways in which we would be unable to accomplish otherwise.

If we consider music as an environment, a territory, we can understand it as a form of niche-building, where we can build virtual yet material worlds for which we are perfectly adapted, in the case of our own creations, and perfectly adaptable through the transformative experience of engaging in the other's constructed niche.

Listening and soundmaking constitute 'technologies of the self' insofar as auditory experience affords us the opportunity to transduce the vibratory forces of our world into thoughts, beliefs, desires, and actions and through our actions transduce these into material prostheses with which we may sing the world into tangible form and partake of its power to engage.

REFERENCES

Blesser, B., & Salter, L. (2007). Spaces Speak, Are You Listening?: Experiencing Aural Architecture. Cambridge, MA: MIT Press.

Chanan, M. (2000). Repeated Takes: A Short History of Recording and Its Effects on Music. London: Verso.

Norman, K. (2004). Sounding Art: Eight Literary Excursions Through Electronic Music. Aldershot, Hants, England: Ashgate.

Taussig, M. T. (1993). Mimesis and Alterity: A Particular History of the Senses. New York: Routlege.

Truax, B. (2008). Soundscape Composition as Global Music: Electroacoustic Music as Soundscape. *Organised Sound*, *13*(02), 103-109.

Windsor, W. L. (1996). Autonomy, Mimesis and Mechanical Reproduction in Contemporary Music. *Contemporary Music Review*, *15*(1), 139.

USE OF PITCH FOR PROCESSING EMOTIONS

Masako Hirotani[1, 2]
[1]School of Linguistics and Language Studies, and Institute of Cognitive Science, Carleton University,
1125 Colonel By Drive, Ottawa, Ontario, Canada, K1S 5B6
[2]Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04103 Leipzig, Germany

1. INTRODUCTION

Previous studies in both linguistics and psych- and neurolinguistics have shown that various prosodic factors contribute to listeners' processing of emotion (e.g., Davidson et al. 2003; Schirmer & Kotz 2006). The present study examined how listeners use one prosodic parameter, pitch, and its role in emotional processing. Specifically, it examined 1) whether fine distinctions in pitch height influence listeners' processing of emotion and 2) to what degree pitch information is utilized when different emotions are processed (i.e., happy vs. sad). Three off-line listing experiments were conducted to address these questions. Overall, the results indicated that pitch expansion has an effect on listeners' perception of emotion, in particular in processing happy emotion.

2. EXPERIMENT 1

Experiment 1 tested whether pitch (pitch expansion vs. compression at three different heights) of a sentence influences listeners' processing of happy and sad emotions.

2.1. Method

Participants. 60 undergraduate students from Carleton University (36 females, mean age 19.83 years old, SD 2.04) participated in the experiment. All participants were right handed, native speakers of Canadian English, with normal or corrected to normal vision, and none of the participants reported any history of neurological or hearing disorders.

Stimuli. A total of 480 English sentences were constructed. They were in a wide variety of syntactic structures but were all similar in length (3.08 s). In order to choose the best 120 sentences with neutral content, the content of the sentences was normed on a seven-point scale (1 being sad, 4 neutral, and 7 happy) by 32 Canadian English undergraduate students (19 females, mean age 22.5 years old, SD 2.78), separate from those that participated in the listening experiments (See Table 1 for example sentences).

Table. 1. Example sentences of neutral content.

1) Alice cleaned the house before she picked up her son.
2) Mick used the coat rack at the restaurant to hang his jacket.

The 120 sentences rated as having "neutral" content in the norming study were recorded at a monoral 16-bit/44.1-kHz sampling rate by a female Canadian English speaker. The experiment had a 2 x 3 design, with the first factor being whether pitch of the sentence was expanded or compressed (pitch range) and the second factor being three different heights in pitch (pitch height). For each sentence, F0 values were extracted every .01 second and converted using two equations, $[x + (x/40)*2]$ for expanded conditions and $[x - \sqrt{x}\,]$ for compressed conditions. To create three different pitch height conditions (i.e., low, mid, and high), 0 Hz, 5 Hz or 10 Hz was added to or subtracted from each of the sentences. Table 2 presents acoustic data for the stimuli.

Table. 2. Acoustic analysis of stimuli.

Condition	Mean F0 (Hz)	Max F0 (Hz)	Mean Intensity (dB)	Max Intensity (dB)
Expanded Low	173.67	328.27	65.96	81.04
Expanded Mid	177.34	331.40	65.97	81.04
Expanded High	180.94	337.74	65.99	81.04
Compressed Low	147.85	287.46	66.21	81.37
Compressed Mid	150.70	289.67	66.17	81.30
Compressed High	154.61	294.87	66.12	81.23

Procedures. Participants were tested individually in half hour sessions. After each sentence had finished playing, the participants indicated which emotion, either happy or sad, the sentence portrayed by choosing the corresponding key on the key pad.

2.2. Results

Figures 1 presents mean percent choice of "happy" for each condition. ANOVAs were conducted using error terms based on participant (F_1) and item variability (F_2). There was a significant effect of pitch range ($F_1(1,59)=72.41$, $p<.0001$, $F_2(1,119)=242.75$, $p<.0001$), which suggests that the pitch expanded conditions were associated with happy emotion more often than the pitch compressed conditions. In addition, a significant interaction of pitch range and pitch height was found ($F_1(2,118)=3.10$, $p<.01$, $F_2(2,238)=4.16$, $p<.05$). This interaction is probably attributed to the difference between the two extreme conditions in pitch height (high vs. low) for the compressed conditions.

Experiment 1 showed that sentences with a large pitch range elicit happy emotion more often than those with a small pitch range. However, as the reader may recall, Experiment 1 used a forced choice task, in which participants were asked to choose between the two options provided (either "happy" or "sad"). This may be unnatural as a task. To overcome this problem, Experiment 2 used seven options.

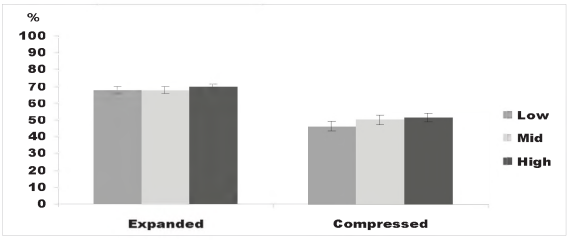


Figure. 1. Mean percent choice for "happy".

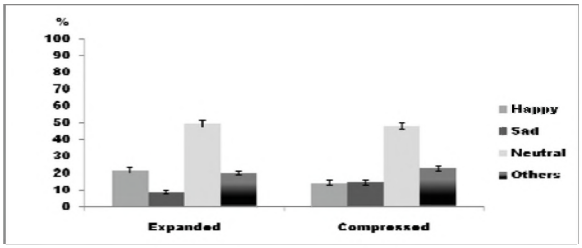


Figure. 2. Mean percent choice for emotion categories.

3. EXPERIMENT 2

3.1. Method

Participants. 60 participants (41 females, mean age 19.86 years old, SD 2.27) took part in the experiment. Participants were recruited from the same participant pool, using the same criteria, as in Experiment 1.

Stimuli and Procedures. The stimuli and procedures for the experiment were the same as those for Experiment 1 except that in this experiment, participants were provided seven emotion categories, i.e., six basic emotions (Ekman 1992) and neutral, from which they were asked to choose the emotion best associated with the sentence they heard.

3.2. Results

The results of experiment are presented in Figures 2. Data were averaged over three different conditions of pitch height (low, mid, and high) for each type of pitch range conditions (pitch expansion and compression). In addition, results for "surprise", "disgust", "fear", and "anger" were clustered into the category "others".

As shown in Figure 1, "neutral" was chosen most often for both pitch expansion and compressed conditions (about 50% of all responses). This is probably due to the fact that the stimuli were created based on the sentences with neutral content (see the stimuli section for Experiment 1). More importantly, "happiness" was chosen more often than "sadness" for the pitch expansion conditions (21.67% happiness vs. 7.75% sadness), whereas such a data pattern was not found for the pitch compression conditions. For the pitch compression conditions, "happiness" and "sadness" were chosen equally often (14.20 % happiness and 14.65% sadness). The statistical analyses (omitted due to space limitations) support the description of data presented above.

The results of Experiment 2 are consistent with those of Experiment 1, suggesting that a large pitch range is associated with happy emotion more frequently than sad.

4. EXPERIMENT 3

As a follow-up to the previous experiments, Experiment 3 tested listeners' sensitivity to subtle differences in pitch. This was to examine a relation between the listeners' choice of emotion and pitch perception.

60 undergraduate students from Carleton University (31 females, mean age 19.90 years old, SD 2.02) participated in the experiment. The participants were asked to listen to the sentence stimuli used for Experiments 1 and 2 and evaluate pitch height by using a 1-7 scale (1 being lowest and 7 highest). The results are presented in Figure 3.

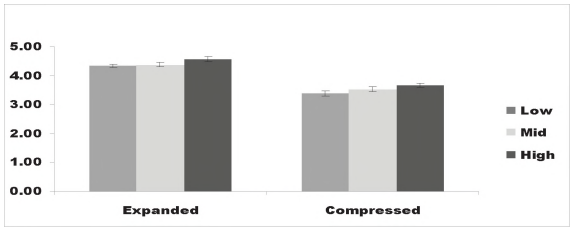


Figure. 3. Mean rating of pitch height.

The data showed a significant effect of pitch range ($F_1(1,59)=87.85$, $p<.0001$, $F_2(1,119)=903.55$, $p<.0001$) as well as pitch height ($F_1(2,118)=24.28$, $p<.0001$, $F_2(2,238)=23.62$, $p<.0001$). These results indicate that listeners were capable of perceiving the small differences in pitch associated with the present stimuli.

5. CONCLUSION

This study investigated the role of pitch in emotion processing. While listeners are sensitive to subtle pitch differences, a pitch range (pitch expansion vs. compression) plays a crucial role in associating sentences of neutral content with specific emotion categories (e.g., happy, sad).

REFERENCES

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*, 169-200.

Davidson, R.J., Scherer, K.R., & Goldsmith, H.H. (2003). Vocal expression of emotion. *Handbook of affective sciences* (pp. 433-453). US: Oxford University Press.

Schirmer, A., & Kotz, S.A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences, 10*(1), 24-30.

ACKNOWLEDGEMENTS

The present study was supported by the research grant, Social Sciences and Humanities Research Council of Canada, Standard Research Grant # 410-2008-0987.

CONSTRAINTS ON DURATIONAL INCREASE IN BOUNDARY POSITIONS : A CROSS-LINGUISTIC STUDY OF LANGUAGES WITH CONTRASTIVE QUANTITY

Zita McRobbie-Utasi
Department of Linguistics of Simon Fraser University, Burnaby, British Columbia, Canada. V5A 1S6.
mcrobbie@sfu.ca

1. INTRODUCTION

The objective of this study is to explore boundary signaling in languages with contrastive quantity. In particular, the role of duration is examined, for segmental quantity tends to constrain the degree of durational increase in boundary positions. This constraint on lengthening coincides with a lesser degree of durational variation within the grammatical unit in comparison with languages where duration is not contrastive.

The phenomenon of preboundary lengthening has been the focus of extensive research for over three decades; a recent summary is found in Rao, 2010. It has been observed in a great variety of languages and the assumption that final segment lengthening is universal (Vaissiére, 1983) continues to be the focus of ongoing research. However, while the lengthening phenomenon may be universal, a wide variety associated with its realization has to be acknowledged.

It has been widely accepted that while the lengthening in boundary positions is phonological, the phonological constituents are not independent from the morphology and syntax of the language on the one hand and rhythmic factors on the other, resulting in different degrees of lengthening Gussenhover & Rietveld, 1992. The authors argue that "prosody driven durational effects… should be kept distinct from preboundary lengthening". In the present study, the analysis of durational variations in relation to preboundary lengthening provides support to this assumption.

2. METHOD

Recordings of six paragraphs by 24 speakers of eight languages were acoustically analyzed. 398 paragraphs were analyzed in this study. The experiment was designed to enable comparison between languages with contrastive vs. non-contrastive quantity; Finnish, Hungarian, Hindi and Latvian representing the former type, Brazilian Portuguese, English, Cantonese and Russian the latter. Each paragraph consists of three sentences (hereafter A, B and C); the sentences in the six paragraphs were the same, except their ordering. The paragraphs were constructed in such a way that all sentence orderings were possible. It is assumed that paragraphs have a linguistically determined structure and their suprasegmental characteristics are strongly bound to the realization of this structure Lehiste, 1975.

3. DISCUSSION

The mean durations of paragraphs cluster around similar values implying a tendency to maintain a durational target Measurements of sentence and pause and segment durations were obtained in order to identify durational variations in relation to their position in the paragraph. Further, paragraph-final syllable durations were examined.

3.1. Sentence and Pause Durations

Sentence durations regardless of their ordering were compared with durations in different positions. The tendencies observed point to greater durational variations in languages with non-contrastive quantity. In first position the maximum durational difference from the mean sentence duration is 0.11s in languages with contrastive durations (0.09s in Hungarian, 0.07s in Finnish, 0.06 in Hindi and 0.11s in Latvian); the maximum durational difference from the mean sentence duration is 0.17s in languages with non-contrastive duration (0.2s in Brazilian Portuguese, 0.16s in English, 0.12s in Cantonese and 0.17s in Russian. In the second position the maximum durational difference from the mean sentence duration is 0.12s in languages with contrastive durations (0.09s in Hungarian, 0.11s in Finnish, 0.08 in Hindi and 0.12s in Latvian); the maximum durational difference from the mean sentence duration is 0.18s (0.17s in Brazilian Portuguese, 0.18s in English, 0.9s in Cantonese and 0.12s in Russian) . In the third position the maximum durational difference from the mean sentence duration is 0.14s in languages with contrastive durations (0.08s in Hungarian, 0.13s in Finnish, 0.07 in Hindi and 0.14s in Latvian); the maximum durational difference from the mean sentence duration is 0.21s (0.21s in Brazilian Portuguese, 0.16s in English, 0.14s in Cantonese and 0.19s in Russian).

The durational patterns emerging from the measurements as presented above evidence a tendency for durational variations to be greater in languages with where quantity is non-distinctive. Durational variations appear to be greatest for sentences in third position in both language types. Measurements of paragraph durations, however, show little effect of sentence duration variations – a fact pointing to a tendency on the part of the speakers to maintain a durational target. In examining sentence duration variations relating to pause durations it may be hypothesized that intersentential pauses play a role in the keeping to a target duration associated with the grammatical unit. The tendency is summarized in Figure 1.

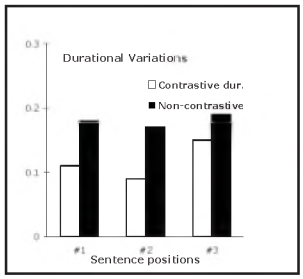


Fig. 1. Sentence-level duration variations

3.2 Paragraph-final syllable durations

Syllable durations in paragraph-final positions were examined in order to establish the degree of preboundary lengthening (while acknowledging the relevance of lengthening before the final syllable (Swerts *et al*. (1994), Hakohari *et al*. 2007), it is assumed that the greatest increase occurs in the last syllable). The hypothesis tested here argues that, while durational variations occur in all position at the word and syllable levels, it is the paragraph-final syllabic position that durational variations occur as increases. It is further hypothesized that although durational increases occur in both language types the increase is significantly greater in languages with non-contrastive quantity. This pattern is presented in Figure 2.

3.3 Segment durations

In examining segmental durations the hypothesis tested was that there is a constraint in the degree of variation languages with contrastive quantity stemming from the necessity of maintaining the phonemic role of length distinctions. In accordance with this hypothesis it was further assumed that durational variations at the segmental level would be of lesser degree in comparison with non-quantity languages. The analysis of vowel duration variation supports this assumption (Fig. 3).

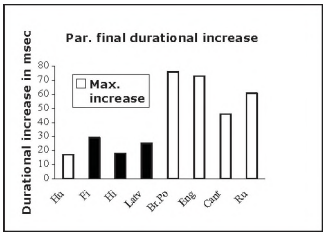


Fig. 2. Durational increase in the final syllable

In connection with durational variations at the segmental level in languages with contrastive duration, the following tendencies could be identified: (i) long stressed vowels occur with the greatest degree of durational variation, and (ii) the smallest degree of variation occurs in connection with short unstressed vowels.

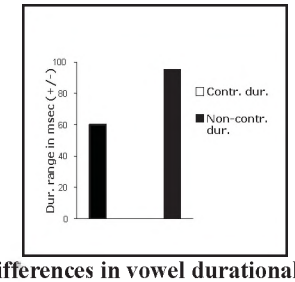


Fig. 3. Differences in vowel durational variations

4. SUMMARY

The results presented above are comparable to those obtained in a pilot study working with a smaller data base (McRobbie-Utasi, 2003) confirming the two findings reported there: (i) durational variations associated with the temporal structure of the paragraph and durations associated with boundary signaling represent two separate mechanisms, (ii) the realizations of the two tendencies are dependent on language types, i.e., there is a constraint on the degree of variation in languages with contrastive quantity.

REFERENCES

Lehiste, I. (1975). The phonetic structure of the paragraph. In Cohen, A. & Nooteboom, S.G. (Ed.s.). *Structure and process in speech perception*. (pp. 195-203). Berlin/Heidelberg/New York: Springer Verlag.

Rao, R. (2010). Final lengthening and pause duration in three dialects in Spanish. In Ortega-Liebari, M.. (Ed.). *Selected Proceedings of the 4th Conference on Laboratory approaches to Spanish phonology*. (pp. 69-82). Somerville, MA: Cascadilla Proceedings Project.

Vaissiére, J. (1983). Language-independent prosodic features. In Cutler, A. & Ladd, R. (Eds*.). Prosody: Models and measurements*. (pp. 55-66). Hamburg: Springer-Verlag.

McRobbie-Utasi, Z. (2003). Phonetic realizations of the New Zealand *KIT* vowel in relation to two social variables, in *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 1201-1205).

Swerts, M. Collier, R. & Terken, J. (1994). Prosodic predictions of discourse finality in spontaneous monologues. *Speech Communications* 15, 79-90.

Hakohari, J., Saarni, T., Salakoski, T., Isoaho, J. & Aaltonen, O. (2007). Measuring relative articulation rate in Finnish utterances. In *Proceedings of the 16th International Congress of Phonetic Sciences*. (pp. 1105-1108.).

EXAMINING VOCAL AFFECT IN NATURAL VERSUS ACTED EXPRESSIONS OF EMOTION

Barbra Zupan[1] and Madeline Shiah[2]
[1]Department of Applied Linguistics, Brock University, 500 Glenridge Avenue, St. Catharines, Ontario, Canada L2S 3A1
[2]Department of Communication Disorders and Sciences, University at Buffalo, 122 Cary Hall, Buffalo, NY 14214

1. INTRODUCTION

Vocal Affect (VA) is one of the most significant characteristics by which human emotion is identified. This non-linguistic component of speech includes prosody, intonation, speaking rate, and vocal effort; all of which aid in conveying emotional information beyond the semantic content of a message.[1] Although listeners pay attention to both non-linguistic and linguistic properties of the speaker's utterance to derive meaning from the message, a large amount of information can be drawn from vocal affect information alone.[1,2] The average listener is generally proficient in detecting a communicative partner's emotional state based on the associated acoustic properties of a happy, sad, or angry tone of voice. These acoustic properties are important since the intention of a speaker's message may change significantly depending on which emotional tone of voice is used. Thus, a listener must combine what is said with how it is spoken to accurately gauge the meaning of the speaker's message.

During a typical face-to-face conversational exchange, a listener is able to use numerous sources of information to discern the intended meaning of the speaker including facial, gestural and vocal cues. In fact, successful social interactions rely on the integration and interpretation of facial and vocal affect cues.[2] However, some real-life situations, such as having a telephone conversation or in the case of visual deficits, force one to rely on only VA. This type of scenario was of particular interest to us. Thus, the aim of the current investigation was to examine a person's ability to recognize emotion in a vocal message in the absence of facial and gestural cues. It was hypothesized that listeners would be able to identify natural expressions of emotion more accurately than acted portrayals of emotion due to the physiological changes said to underlie the vocal qualities of the speakers in natural emotional expressions. In addition, we would assume that natural expressions of emotion are more typically heard, experienced, or encountered, therefore should be more easily identifiable.

2. METHOD

2.1 Participants

Phase One included seven Brock University students as speakers (4 female, 3 male), ranging in age from 21-25.

Phase Two included 31 Brock University students as listeners (27 female, 4 male), with a mean age of 21.8 years. All participants spoke English as their primary language, had no hearing or uncorrected visual deficits, and reported no current or previous communication disorders.

2.2 Procedure

Phase One: The examiners and a research assistant viewed DVDs of reality television shows ("Punk'd", "Amazing Race", & "The Real World") for the purpose of selecting clips of characters speaking semantically neutral phrases using the emotions fear, anger, and happiness (5 clips of each emotion). Only clips identified by the examiners as being significantly characteristic of the tone of voice associated with the selected emotions were selected. To create the acted stimuli, speakers were asked to say the same phrases as those selected from the reality shows, in the same emotional tone(s) of voice. Speakers were recorded using an Olympus digital audio recorder with consistent mouth to microphone distance.

Phase Two: Listeners were administered the vocal affect portion of the Diagnostic Analysis of Nonverbal Affect-2 (DANVA-2). This test requires participants to listen to 24 trials of a sentence and identify the speaker's tone of voice (Happy, Sad, Angry, Fearful). The vocal affect portion of the DANVA-2 served as a baseline for participants' general ability to identify emotional tones of voice. Following completion of the DANVA-2, participants were seated comfortably in front of a computer and instructed to listen to the test stimuli (natural and acted audio clips) and identify as accurately as possible which of the three emotions was being portrayed in each of the vocal clips. Participants were also asked to indicate whether they thought the expression was a natural or acted (pretended) portrayal. Stimuli were randomized for each participant using Cedrus SuperLab software. Responses were provided using a 6 button response pad with the following options: Happy; Happy Acted; Angry; Angry Acted; Fearful; and Fearful Acted. Listeners received no feedback regarding whether the stimuli were from the natural or acted conditions or if the selected emotion was correct.

3. RESULTS

A repeated measures ANOVA revealed a significant interaction between acted versus natural emotion showing better identification of acted than natural expressions of emotion [$F(2,29)=18.277$ $p<.01$]. Follow-up paired samples *t* test were conducted with alpha set at .01 to control for Type 1 error, to evaluate participants' ability to correctly identify natural and acted expressions within emotion type (see Figure 1). The results showed that both Happy Acted

and Fearful Acted stimuli were identified with significantly higher accuracy better than their Natural counterparts. Results of the identification of Angry however were the opposite, with natural expressions of Angry being significantly better identified by participants than acted stimuli.

Overall, participants were better able to identify acted versus natural expressions of emotion. Since the DANVA-2 includes acted vocal expressions, we computed a correlation analysis between participants' DANVA-2 score and their performance in identifying acted expressions of the three emotion types. No significant correlation was found, $r(29) = .32$, $p = <.085$.

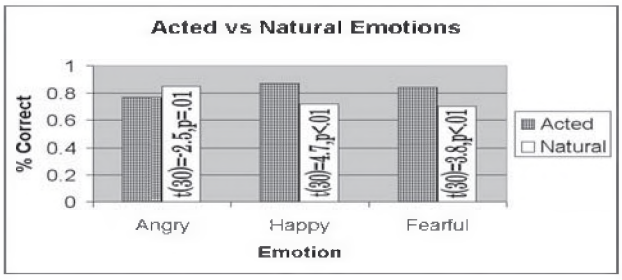


Figure 1. Accuracy of identification of acted vs. natural expressions.

4. DISCUSSION

The current study aimed to examine the ability of listeners to identify vocal expressions of emotion (Happy, Angry, Fearful) under vocal-only conditions in both natural and acted conditions. It was hypothesized participants would more accurately identify natural expressions of emotions since these expressions have been associated with underlying physiological changes that affect the acoustic pattern of the voice.[3] This assumption is an important one since equal accuracy in both categories (natural and acted) would indicate that these physiological changes are not a necessity in the identification of vocal expressions. Our hypothesis was confirmed for Angry vocal expressions, but not for Happy and Fearful vocal expressions. Interestingly, the vocal expressions portrayed in the acted portrayals were more exaggerated expressions of these two emotions while the natural expressions of Anger contained more prominent acoustic cues. Thus, it appears that participants may perceive more prominent acoustic cues as acted because they are more exaggerated than the emotion expressions they encounter in typical daily interactions. Participants in this study were aware that there were examples of both acted and natural expressions and therefore, may have used exaggerated versus subtle expressions in their decision making process. This was reflected in the type of response participants selected since they tended to label more obvious expressions as acted on the response pad. However, these results need to be interpreted cautiously since the number of acted stimuli in comparison to natural stimuli in the current study was unbalanced (105 acted clips; 15 natural clips). The minimal number of natural clips reflects the challenge of finding semantically neutral phrases that portray one of the three emotions used in this study within reality television shows. The use of unequal acted versus natural expressions may have masked inaccuracies of participants simply due to the fact that they had a better chance overall at correctly identifying an acted clip since there were more of them. In addition, participants may have been expecting a more equal balance which may have influenced their choice. This was certainly a limitation of the current study.

A lack of correlation between responses to the DANVA-2 and acted vocal expressions was surprising, since both sets of stimuli focus on acted portrayals. However, the current study included only portrayals of angry, fearful, and happiness whereas the DANVA-2 also includes portrayals of sadness. Due to the increased jitter and breathiness in both fearful and sadness, listeners often confuse these two emotions. Since the current study did not include sadness, the lack of correlation may simply reflect this difference.

Future work in this area should also examine the acoustic cues of the acted versus natural emotion expressions. Although the use of natural clips from reality television shows provided us with stimuli more likely to represent the acoustic patterns associated with physiological changes for each emotion category, we were unable to analyze the stimuli due to the presence of background noise and music present in the majority of clips. Since we could not remove the music, we resolved the issue for the current study by adding similar background music to the acted stimuli to ensure that responses were not being influenced by this factor. However, spectrogram analysis of these two types of stimuli remains important to examine whether the acoustic cues associated with each emotion are influenced by this condition.

REFERENCES

1. Nygaard, L.C. & Queen, J.S. (2008). Communicating emotion: linking affective prosody and word meaning. *Journal of Experimental Psychology: Human Perception and Performance, 34*(4), 1017-1039.

2. Zupan, B., Neumann, D., Babbage, D.R., & Willer, B. (2009). The importance of vocal affect to bimodal processing of emotion: implications for individuals with traumatic brain injury. *Journal of Communication Disorders, 42*(1), 1-17.

3. Johnstone, T., & Scherer, K.R. (2003). Vocal communication of emotion. In M.L. Lewis & J.M. Haviland-Jones (Eds.), *Handbook of Emotions* (2nd ed., pp 220-235). New York: The Guilford Press.

PROCESSING OF SPEECH AND NON-SPEECH TONAL INFORMATION BY NATIVE AND NONNATIVE TONE LANGUAGE SPEAKERS: AN EVENT-RELATED ELECTROPHYSIOLOGICAL STUDY

Yang Zhang[1], Angela Cooper[2], and Yue Wang[2]

[1]Dept. of Speech-Language-Hearing Sciences, University of Minnesota, Minneapolis, MN, 55455, USA
zhang470@umn.edu

[2]Dept. of Linguistics, Simon Fraser University, Burnaby, BC V5A 1S6, Canada; yuew@sfu.ca

INTRODUCTION

One long-deliberated question in speech perception is the extent to which it involves language-specific mechanisms. Some research findings support the independence of speech and nonspeech processing (*1-3*). Others indicate general auditory mechanisms shared by speech and nonspeech processing (*4-7*). Lexical tones present an interesting case at the suprasegmental level to study brain organization of auditory and linguistic processing as a function of learning experience. Imaging data for lexical tones have shown right-hemisphere dominance for auditory processing and left-hemisphere dominance for linguistic processing (*8-10*). However, the dynamic interplay between the two processes and the temporal courses have not been fully addressed (*6*). The present investigation employed high-density ERPs to examine two basic issues: (a) whether auditory processing of pitch patterns fundamentally differs from lexical tone processing, and (b) whether there is automatic transfer of learning for tonal pattern processing from speech to nonspeech in tone language speakers but not in non-tone language users.

METHODS

The participants included 14 native Mandarin Chinese listeners, and 16 native English listeners with no tone language experience. They were all right-handed healthy young adults. Five participants were excluded (2 native and 3 nonnative) due to incomplete data or excessive noise in EEG recording.

The stimuli for the speech condition included two Mandarin monosyllables (*ju, ci*) produced with two tones (rising, falling) by a female native Mandarin speaker. For the nonspeech condition, the same speaker hummed these two syllables with rising and falling tones.

A modified oddball paradigm was employed *(3)*. The oddball sequences contained both syllables (*ju, ci*) in alternation. Each condition consisted of four different sequences, varying in syllable order (*ju-ci* or *ci-ju*) and deviant syllable (*ju* or *ci*), producing a total of eight different blocks. For each condition, 800 stimuli were presented (20% deviant, 80% standard). A 128-electrode light mesh (Geodesic net) was applied, with the vertex (Cz) electrode used as a reference with the ground at the nasion. The stimuli were presented via speakers.

The raw EEG data for each participant were analyzed and averaged with common average reference in BESA software (Brain Electrical Source Analysis) with a passband of 0.5~40 Hz. Two windows, 100~300 ms and 300~ 600 ms, were selected after inspection of the grand mean data to search mismatch negativity (MMN) peak in the subtracted waveform (deviant – standard) (*11*). MMN quantification used the mean of sample values in a window of 40 ms centered around the MMN peak *(6)*. Global field power for the subtracted waveforms was calculated for each participant and each stimulus condition. Furthermore, point-to-point t-tests were performed to see the temporal evolution of significant differences for key factors of interest (*12*). Topographical maps were calculated for selected peak points. Finally, repeated measures MANOVA was performed for the MMNs at F3 site (electrodes 19, 23, 24) and F4 site (electrodes 3, 4, 124) *(6, 11)*.

RESULTS

In the speech condition (Fig. 1), natives showed one early MMN (corresponding to the consonant portion of the syllable) and one late MMN (corresponding to the vowel portion) in both hemispheres. Nonnatives did not show point-to-point significances in the left site, and the right hemisphere site had a later MMN response than the first MMN of the native group. MANOVA results confirmed the significant MMN differences between the native and nonnative listeners for the early MMN [$F(1,23) = 5.02$, $p < 0.05$].

In the nonspeech condition (Fig. 2), both native and nonnative groups showed an early MMN response in both hemispheres followed by P3a response. The native group showed earlier and larger MMN and P3a than the nonnative speakers ($p < 0.05$; MANOVA results).

To avoid channel-selection bias, GFPs *(12* for all the 129 electrodes*)* were calculated for the MMN waveforms (Fig. 3). The results confirmed the MMN results at F3 and F4 sites, showing that native listeners had earlier and stronger MMN responses for both speech and nonspeech conditions. Interestingly, the nonspeech condition elicited larger MMN responses than the speech condition in the native listeners ($p < 0.05$). The MMNs for the speech and nonspeech conditions were comparable for the nonnative listeners.

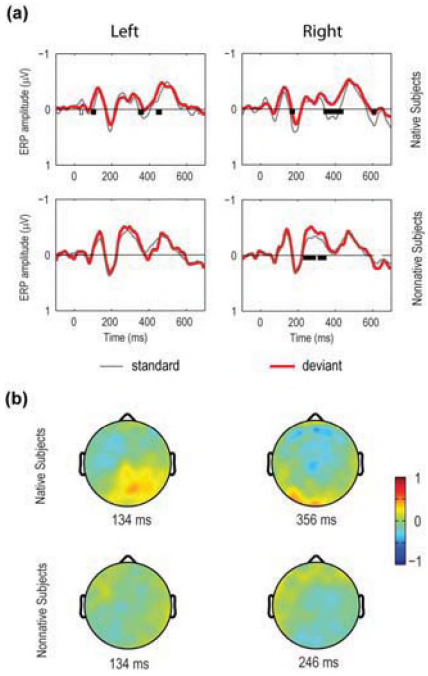


Fig. 1. ERPs for the speech stimuli in the two participant groups. (a) Grand average ERPs for the standard and deviant stimuli from F3 (electrodes 19, 23, 24) and F4 sites (electrodes 3, 4, 124). Point-to-point significant differences were indicated by the black/white bars on x-axis ($p < 0.01$). (b) Topographical maps for the peaks in MMN waveforms.

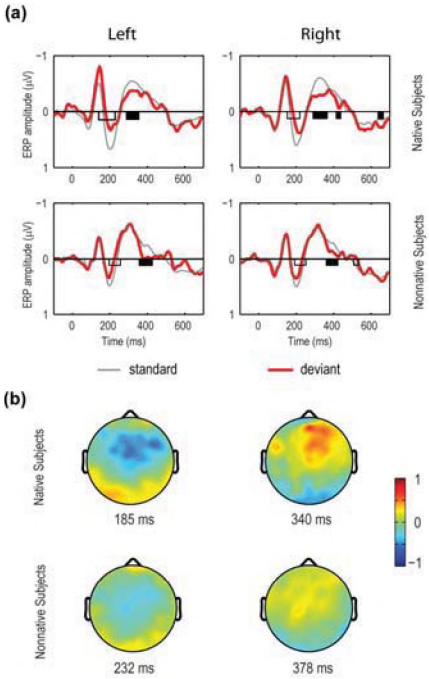


Fig. 2. ERPs for the nonspeech stimuli in the two participant groups. (a) Grand average ERPs for the standard and deviant stimuli from F3 and F4 sites. Point-to-point significant differences are indicated by the black/white bars on x-axis ($p < 0.01$). (b) Topographical maps for the peaks in MMN waveforms.

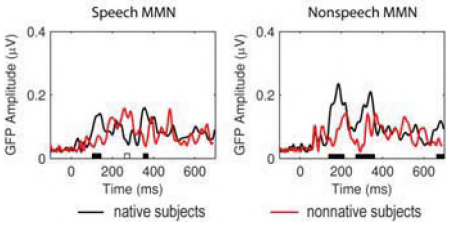


Fig. 3. Global field power of MMN waveforms in all the 129 electrodes for the speech and nonspeech conditions. Significant point-to-point differences between the native and nonnative groups are indicated by the black/white bars on x-axis ($p < 0.01$).

DISCUSSION

Despite different patterns of MMN elicitation in the speech and nonspeech conditions, the MMNs were present for both native and nonnative listeners, but the magnitude was much smaller in the nonnative listeners for the speech stimuli. These data are consistent with previous studies *(3, 6, 7,11)*. Unlike a previous report (3), the native listeners showed earlier MMN responses than nonnative listeners in both speech and nonspeech conditions. Furthermore, the MMNs for nonspeech were larger than those of the speech condition only in native listeners, suggesting that the effects of language experience may automatically transfer to non-speech tonal pattern extraction *(6)*. The results indicate a high degree of interconnectivity for speech and nonspeech in terms of neural sensitivity to pitch categories.

ACKNOWLEDGMENTS

This research is supported by a Discovery Grant to YW from the Natural Sciences and Engineering Research Council of Canada (312457-2006). YZ is supported by two Brain Imaging Research Awards from UMN to work on the project. The authors acknowledge assistance from Mathieu Dovan, Anish Jain, Amy Lee, David Potter, Rebecca Simms, and Marcus Watson.

REFERENCES

1. A. M. Liberman, D. H. Whalen, *Trends Cogn. Sci.* **4**, 187 (2000).
2. K. Miyawaki *et al.*, *P&P* **18**, 331 (1975).
3. Y. Zhang, P. K. Kuhl, T. Imada, M. Kotani, Y. Tohkura, *Neuroimage* **26**, 703 (Jul 1, 2005).
4. A. L. Francis, V. Ciocca, B. K. Ng, *Percept. Psychophys.* **65**, 1029 (Oct, 2003).
5. L. L. Holt, A. J. Lotto, R. L. Diehl, *J. Acoust. Soc. Am.* **116**, 1763 (Sep, 2004).
6. J. Xi, L. J. Zhang, H. Shu, Y. Zhang, P. Li, *Neuroscience*, (In press).
7. E. Sussman *et al.*, *Hear. Res.* **190**, 128 (Apr, 2004).
8. R. J. Zatorre, J. T. Gandour, *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**, 1087 (2008).
9. Y. Wang, J. A. Sereno, A. Jongman, J. Hirsch, *J. Cogn. Neurosci.* **15**, 1019 (Oct 1, 2003).
10. P. C. Wong *et al.*, *Cereb. Cortex* **18**, 828 (Apr, 2008).
11. E. Kaan, R. Wayland, M. Bao, C. M. Barkley, *Brain Res.* **1148**, 113 (May 7, 2007).
12. Y. Zhang *et al.*, *Dev. Sci*, (In press).

AN ULTRASOUND INVESTIGATION OF POSSIBLE COVERT CONTRASTS IN FIRST LANGUAGE ACQUISITION: THE CASE OF /sp ~ sw ~ sm/ > [f] MERGERS

Sonya Bird
Dept. of Linguistics, University of Victoria, PO Box 3045, Victoria BC, V8T3K5 sbird@uvic.a

1. INTRODUCTION

Child language often involves consonant cluster simplification. In the case of clusters beginning with /s/, e.g. /sp/, the most common strategy is to delete /s/, resulting in forms like [pun] for *spoon* (Ingram 1991). An alternative strategy attested in some children is to merge the two consonants, e.g. /sp/ > [f], resulting in [fun] for *spoon*. (Locke 1983). Mergers provide interesting insights into phonological representation in children: they suggest that children are able to decompose sounds into their composite features. Indeed, in the example above, [f] is not a compromise between /s/ and /p/ along any single phonetic continuum; rather it is a merger of the manner of articulation of /s/ with the (near-) place of articulation of /p/.

To fully understand the implications of mergers for phonological representation in children, it is essential to consider the phonetic detail of these mergers. Specifically, is the result of merger truly identical to the underlying version of the sound? In the above example, is the perceived [f] in *spoon* identical to the /f/ in *food*, or does it retain some trace of the underlying segments /sp/? This question has been difficult to answer because the relevant data from child language is based primarily on auditory analyses (in the form of transcriptions), which reflect what the listener hears rather than what the speaker is doing. As a result it is not clear whether perceived mergers are truly complete from the speaker's perspective (Scobbie et al. 2000). As a step towards overcoming this problem, this paper reports on an preliminary articulatory study conducted with a single, normally developing, child whose pronunciation of /s/ + labial consonant sequences (e.g. /sp/) is not auditorily distinct from her pronunciation of /f/. Results suggest that in her case, the mergers are indeed complete.

2. METHOD

The experiment was conducted in two sessions, separated by approximately one month. The participant for the study was a single, normally developing girl – AC – who was 2;8 years at the first session and 2;9 years at the second (her pronunciations did not change across sessions). In both sessions AC was recorded using lingual ultrasound (GE Logic E portable ultrasound machine with 8C-RS convex transducer) pronouncing randomly ordered words from sets like *sponge* – *sun* – *fun*. Table 1 provides the stimuli list, which included only words familiar to AC (*-ed words were elicited in the second session only).

During the recordings, AC was seated on her mother's lap, with the ultrasound machine to the left. AC's mother held AC's head stable against her chest with one hand; she held the transducer under AC's chin with the other hand, monitoring the screen of the ultrasound machine to ensure the probe remained immobilized. AC's father sat to her right, and elicited the target words by asking her to name pictures he presented to her one at a time. The sessions were also audio-taped, the first session using a Countryman AKG C429 head mounted microphone resting on the stand of ultrasound machine (being too big to place on AC's head); the second session using a Senheiser ME-55 directional microphone (with a K6 capsule). Ultrasound and audio recordings were mixed through a Mackie 1402-VLZ3 mixer and captured on an external PC computer using Sony Vegas Pro 8. Because of technical difficulties in the first session, the recording was repeated 3 times, such that the first two repetitions acted as practice sessions. This did not appear to affect the pronunciation of the target words. In the first session, only the midsagittal view of the tongue was recorded; in the second session, the recording was done twice, to get a cross-sectional as well as a midsagittal view of the tongue (by turning the probe 90 degrees).

Table 1. Experimental stimuli.

V	/sw~sp~sm/	/s/	/f/
[I]	swimming; swing; swing set	sing; singing	finger
[u]	spoon; smoothy	soup *	food
[ɛ]	sweater *	sesame *	fender *
[□]	Sponge Bob	suns; sunglasses	fun

Figure 1 illustrates the experimental set-up. Although it did not provide the kind of controlled stabilization required for rigorous, quantitative analysis of the data, it was deemed sufficient to gain a qualitative understanding of tongue position and movement during the target sounds. Certainly a more complete follow-up study should include more reliable head and probe stabilization, to the extent that this is possible with such a young participant.

Fig. 1. Experimental set-up.

3. RESULTS

In total, 40 tokens were elicited in the two sessions, based on a midsagittal (from the side) view of the tongue (AC repeated some of the tokens; these were kept in the analysis): 12 of /s/, 8 of /f/, 11 of /sw/, 6 of /sp/ and 3 of /sm/. Using the audio signal as a reference, short ultrasound (video) clips corresponding to the target sounds were inspected visually to see how /sw~sp~sm/ compared to /s/ vs. /f/ in terms of tongue position. Because the tongue contour was not always clear, and because the probe angle was not consistent within and across sessions, judgments on tongue position were not always easy to make. However in general, results indicate that /sw~sp~sm/ are identical to /f/: they were consistently transcribed as [f], with the exception of two tokens of /sw/, transcribed as [f^{w}]. Articulatorily, the most reliable difference between /sp~sw~sm~f/ and /f/ vs. /s/ was tongue tip (TT) position: /sw~sp~sm/ and /f/ never showed clear TT raising; while tongue contours for /s/ were often unclear, they suggested TT raising in 9/12 tokens.

Figure 2 provides illustrations of this overall pattern. In each frame, the white line represents the tongue contour, with the tongue tip on the right. Panel 2a is of *swing set*; it shows the beginning and end frames of the /sw/ cluster followed by frames from /I/ and the initial /s/ of 'set'. These frames show that /sw/ is relatively stable throughout its duration (1st and 2nd frames); it does not exhibit a raised TT like the following /s/ (4th frame), nor does it exhibit tongue back raising as one might expect for /w/ (tongue position is stable from /sw/ to /I/ - 2nd and 3rd frames). Panel 2b is from *finger* and shows that /f/ is also stable throughout its articulation, and without TT raising. Panel 2c is from *a sun* and shows TT raising for the initial /s/ (2nd and 3rd frames).

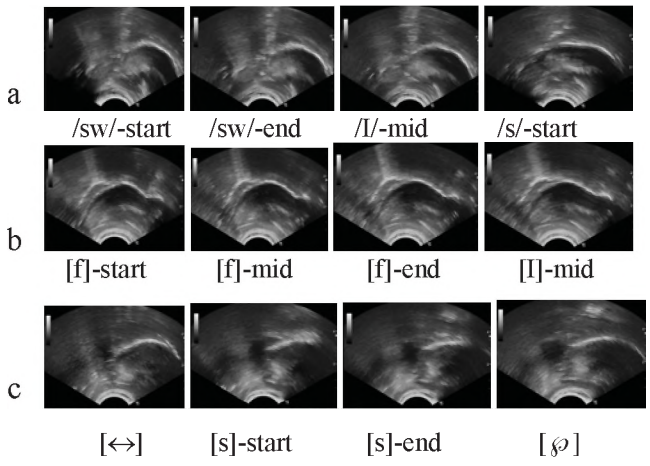


Fig. 2. Illustrative examples – midsagittal view. Panel a: *swingset;* panel b: *finger*; panel c: *a sun*.

During the second session target words were elicited using a cross-sectional (from the front) view of the tongue as well as a midsagittal view, to investigate tongue grooving. In total, 17 tokens were elicited this way: 5 of /s/, 5 of /f/, 3 of /sw/, 2 of /sp/ and 2 of /sm/. In addition, sustained, isolated /s/ and /f/ tokens were elicited. Unfortunately, none of the 5 /s/ tokens elicited in real words had clear tongue contours, making it impossible to compare initial consonants across sets of words. Again, results suggest that /sw~sp~sm/ are identical to /f/: aside from a single frame in a single /sp/ token that exhibited /s/-like grooving (see Fig. 3b, 1st frame), all /sw~sp~sm/ tokens looked like /f/: they had either no tongue grooving or a shallow wide grove, in contrast to the narrower, deeper groove of the sustained /s/. Figure 3 provides typical examples: Panel 3a is of sustained /s/ and two versions of /f/, with and without grooving; panel 3b is of /sp/ in *Sponge Bob* (/s/-like), sweater (/f/-like, with groove) and *spoon* (/f/-like, no groove).

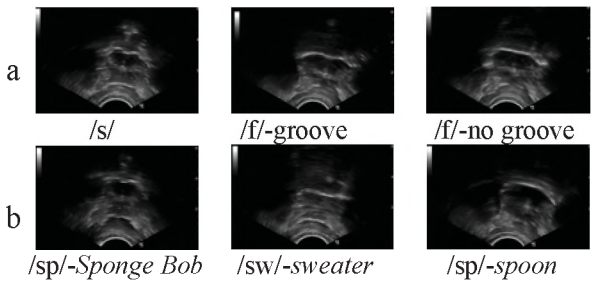


Fig. 2. Illustrative examples – cross-sectional view. Panel a: sustained /s/ vs. /f/; panel b: /sp/ in 3 different tokens.

4. DISCUSSION

If some trace remained of the two underlying segments in /sw~sp~sm/ clusters, one might argue that the observed merger is simply a phonetic effect, a result of AC's developing motor skills. The results tentatively indicate that this is not the case: the /sw~sp~sm/ > [f] merger is complete. For this reason, it is proposed that the merger is a phonological effect, because it can only occur if AC has access to some level of abstraction. Indeed, it is only in abstraction that [f] consists of a compromise between /s/ and /p/, incorporating the manner of articulation of /s/ (fricative) with the place of labial consonants. Crucially, [f] does not represent a compromise between /s/ and /w~p~m/ along any single phonetic continuum. In terms of place for instance, the primary articulators involved are different for /s/ vs. /w~p~m/: the tongue vs. the lips. Therefore, without the ability to abstractly decompose sounds into their place and manner features, there would be no reason to come up with [f] as a merged version of / sw~sp~sm/.

REFERENCES

Ingram, D. (1991) *First language acquisition: method, description and explanation*. Cambridge: Cambridge University Press.

Locke, J. (1983) *Phonological acquisition and change*. New York: Academic Press.

Scobbie J., F. Gibbon, W. Hardcastle & P. Fletcher (2000). Covert contrasts as a stage in the acquisition of phonetics and phonology. In M.B. Broe & J.B. Pierrehumbert (Eds) *Papers in Laboratory Phonology V: Acquisition and the lexicon*. Cambridge: Cambridge University Press.

PRODUCTION AND PERCEPTION OF LARYNGEAL CONSTRICTION IN THE EARLY VOCALIZATIONS OF BAI AND ENGLISH INFANTS

Allison Benner[1]
[1]Dept. of Linguistics, University of Victoria, British Columbia, Canada, abenner@uvic.ca

1. INTRODUCTION

This exploratory study investigates the relationship between infants' production of laryngeal voice quality in the first year of life and adult perceptions of these features in infant vocalizations, building on the work of the Infant Speech Acquisition (InSpA) team at the UVic (Benner et al., 2007; Esling et al., 2006) and the phonetic model provided by Esling (2005). In exploring this relationship, the study focuses on two languages that differ in their use of laryngeal voice quality: English, a Germanic language that uses voice quality non-contrastively in paralinguistic expression (Laver, 1980); and Bai, a Tibeto-Burman language that employs laryngeal voice quality distinctively at the syllabic level in its register tone system (Esling & Edmondson, 2002).

2. METHOD

2.1 Infant Production

Four Bai infants (2 female, 2 male) and four English infants (3 female, 1 male) were recorded monthly in spontaneous interaction with a primary caregiver using a Sony DCR-HC26 digital video camera recorder with an integrated microphone at 16-bit and 44,000 samples/second. Recordings were segmented into vocalizations using SoundForge 8.0. A vocalization was defined as any non-crying sound of at least 500 msec separated from other sounds by at least 2 seconds of silence. A total of 2400 utterances (1200 for each language group; 300 for each of the age groups 0-3 months; 4-6 months; 7-9 months; and 10-12 months) were analyzed using auditory coding and wide-band spectrograms. Utterances were classified by utterance type (babbled, non-babbled, or mixed) and laryngeal voice quality (constricted, unconstricted, or dynamic). Constricted utterances were produced with harsh, creaky, or whispery voice; unconstricted utterances with modal, breathy, or falsetto voice; and dynamic utterances alternated between constricted and unconstricted laryngeal settings.

2.2 Adult Perception

Forty adult native speakers of Bai (24 female, 16 male, aged 20 to 75) and 40 adult native speakers of English (19 female, 21 male, aged 22 to 77) participated in the listening task, which was conducted on a laptop computer with headphones, using E-prime 2.0. Participants rated 36 randomly ordered infant vocalizations on a 5-point scale according to how important they judged the utterances to be in learning to speak their native language (English or Bai). Following the listening task, participants were asked to discuss factors that contributed to their ratings.

The 36 infant vocalizations (18 Bai, 18 English) were excerpted from the infant utterances included in the production study. All sounds were produced by infants aged 10-12 months and were selected to equally represent the laryngeal voice quality and utterance type categories used in the production component of the study (see 2.1 above).

3. RESULTS

3.1 Infant Production of Laryngeal Voice Quality

As shown in Tables 1 and 2, for Bai and English infants, constricted vocalizations decline and unconstricted utterances increase over time, but constricted sounds decrease less markedly for Bai infants and unconstricted sounds increase more for English infants. Production of dynamic utterances increases in months 4-6 for all infants, but Bai infants' production of these sounds steadily increases as the year progresses, compared to English infants'. Chi-square analysis of the relationship between laryngeal voice quality and age revealed a significant association between age and laryngeal voice quality for Bai infants (χ^2 = 129.423, p < .001) and English infants (χ^2 = 239.015 , p < .001). Differences in the development of laryngeal voice quality between the two language groups become significant in months 10-12 (χ^2 = 19.362 , p < .001).

Table 1. Laryngeal Voice Quality: Bai Infants.

Age	Constricted	Dynamic	Unconstricted	Total
0-3	257 (86%)	22 (7%)	21 (7%)	300
4-6	186 (62%)	85 (28%)	29 (10%)	300
7-9	162 (54%)	89 (30%)	49 (16%)	300
10-12	129 (43%)	113 (38%)	58 (19%)	300
Total	734 (61%)	309 (26%)	157 (13%)	1200

Table 2. Laryngeal Voice Quality: English Infants.

Age	Constricted	Dynamic	Unconstricted	Total
0-3	248 (83%)	45 (15%)	7 (2%)	300
4-6	158 (53%)	127 (42%)	15 (5%)	300
7-9	145 (48%)	89 (30%)	66 (22%)	300
10-12	105 (35%)	89 (30%)	106 (35%)	300
Total	656 (55%)	350 (29%)	194 (16%)	1200

3.2 Infant Production of Utterance Type

Production of utterance types is shown in Tables 3 and 4 for Bai and English infants, respectively. The most notable

result is that while Bai infants produce babbling at least as early as English infants, babbling develops more slowly, only increasing significantly in months 10-12. In months 7-9, the two language groups differ significantly in the development of utterance type ($\chi^2 = 18.419$, $p < .001$). These differences remain significant in months 10-12, but the differences are less marked ($\chi^2 = 8.743$, $p < .005$).

Table 3. Utterance Type: Bai Infants.

Age	Non-babbled	Mixed	Babbled	Total
0-3	234 (78%)	59 (20%)	7 (2%)	300
4-6	227 (76%)	60 (20%)	13 (4%)	300
7-9	234 (78%)	51 (17%)	15 (5%)	300
10-12	159 (53%)	47 (16%)	94 (31%)	300
Total	854 (72%)	217 (18%)	129 (11%)	1200

Table 4. Utterance Type: English Infants.

Age	Non-babbled	Mixed	Babbled	Total
0-3	277 (92%)	23 (8%)	0 (0%)	300
4-6	197 (66%)	89 (30%)	14 (5%)	300
7-9	192 (64%)	61 (20%)	47 (16%)	300
10-12	100 (33%)	71 (24%)	129 (43%)	300
Total	766 (64%)	244 (20%)	190 (16%)	1200

Concerning the relationship between laryngeal voice quality and utterance type, for both language groups, non-babbled utterances are most likely to be constricted (for Bai: 89%, 69%, 62%, and 48% in months 0-3, 4-6, 7-9 and 10-12, respectively; for English: 84%, 64%, 60%, and 58%). Beyond months 0-3, mixed utterances are most likely to be dynamic (for Bai: 48%, 49%, and 47% in months 4-6, 7-9, and 10-12, respectively; for English: 60%, 41%, and 52%). Babbling only occurs frequently for both language groups in months 10-12, by which time laryngeal voice quality begins to differ significantly between the two language groups. Among Bai infants, 36% of babbled utterances are constricted, 39% are dynamic, and 24% are unconstricted. The corresponding figures for English infants are 20%, 27%, and 53%, illustrating the tendency for English infants' babbling to be unconstricted. Chi-square analysis of the relationship between utterance type and laryngeal voice quality by language showed no significant language differences for non-babbled and mixed utterances, but highly significant differences for syllabic utterances in months 10-12 ($\chi^2 = 14.199$, $p < .001$).

3.2. Adult Perception

Results for adult perception, reported only briefly here, suggest that Bai and English infants' production patterns mirror, to some extent, Bai and English adults' auditory preferences. Bai listeners show no obvious preference for any of the laryngeal voice quality categories in the listening task, all of which are exploited in Bai tonal distinctions. By contrast, laryngeal voice quality significantly affects English listeners' ratings ($F(2, 476) = 27.941$, $p < 0.001$, $r = .205$). English listeners prefer unconstricted utterances to dynamic utterances, which they in turn prefer to constricted utterances.

In terms of utterance type, Bai listeners significantly prefer babbling to non-babbling ($F(1, 238) = 207.204$, $p > .001$, $r = .327$), as reflected in overall mean ratings of 3.94 and 3.16 for babbling and non-babbling, respectively. English listeners' mean ratings reflect the same significant preference ($F(1, 238) = 866.605$, $p > .001$, $r = .749$), but the preference for babbling is stronger, as reflected in a mean overall rating of 3.98 for babbled utterances versus 2.44 for non-babbled vocalizations. Given the importance of laryngeal voice quality in Bai, it is possible that Bai listeners judge the ability to produce a range of laryngeal voice qualities (as tends to happen in infants' non-babbled utterances) to be nearly as important as their ability to produce syllables, as in babbling. By contrast, English listeners may regard the production of various laryngeal voice qualities as less important to language development, reflecting emotional expression rather than linguistic or phonetic development. Results of the interviews conducted with participants following the listening task, while not reported here, bear out this speculation.

4. DISCUSSION

The results of this exploratory study suggest that adults are attuned to laryngeal voice quality in infants. In turn, the results of the production study suggest that infants become attuned to the use of laryngeal voice quality features in their ambient language early in development, and that these features are reflected in their evolving patterns of production in the first year of life.

REFERENCES

Benner, A., Grenon, I., & Esling, J. H. (2007). Infants' phonetic acquisition of voice quality parameters in the first year of life. In J. Trouvain & W.J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, vol. 3* (pp. 2073-2076). Saarbrücken: Universität des Saarlandes

Esling, J. H. (2005). There are no back vowels: The laryngeal articulator model. *Canadian Journal of Linguistics*, *50*, 13-44.

Esling, J. H., Benner, A., & Grenon, I. (2006). Phonetic development in the first year of life: A comparison of English, Moroccan Arabic, and Bai infants. Paper presented at the conference of the British Association of Academic Phoneticians, Edinburgh, April 2006.

Esling, J. H., & Edmondson, J. A. (2002). The laryngeal sphincter as an articulator: Tenseness, tongue root and phonation in Yi and Bai. In A. Braun & H. Masthoff (Eds.), *Phonetics and its Applications: Festschrift for Jens-Peter Koster on the Occasion of his 60th Birthday* (pp. 38-51). Stuttgart: Franz Steiner Verlag.

THE EFFECT OF TASK TYPE ON NATIVE SPEAKER JUDGMENTS OF L2 ACCENTED SPEECH

Ron I. Thomson and Paris Campagna
Dept. of Applied Linguistics, Brock University, St. Catharines, ON, Canada, L2S 3A1, rthomson@brocku.ca

1. INTRODUCTION

Research investigating adult second language (L2) accent often relies on human listeners to assess speech. While human judgments can measure intelligibility in absolute terms, they might be less precise for measuring fine-grained speaker-dependent variables. For example, there appears to be a relationship between lexical frequency [1] and familiarity [2], and L2 pronunciation accuracy. Some argue that these lexical effects do not reflect real differences in the speakers' articulation of more vs. less frequent words. Rather, judges may perceive the more frequent or familiar words to be more intelligible, because they are more frequent or familiar to the judges themselves [3]. We test this claim by measuring the effect of lexical familiarity on the intelligibility of L2 vowel productions judged in words, and as isolated vowels extracted from those words, thereby masking lexical context.

2. METHOD

2.1 Speakers

19 native speakers (NSs) of Canadian English (*M* age = 22.8, range = 19-49; 2 male, 17 female) provided baseline data. Beginner L2 English learners included 19 who spoke Standard Mandarin (*M* age = 40.1, range = 29-49; 4 male, 15 female) and 19 who spoke a Slavic L1 (*M* age = 38.6, range = 29-49; 12 female, 7 male), mostly Russian. The L2 speakers' mean length of residence in Canada was 15.6 months (range = 4 - 40).

2.2 Stimuli

Auditory elicitation stimuli were produced by a male speaker of Canadian English. They comprised 30 high frequency monosyllabic English verbs - three for each of ten Canadian English vowels: /i/, /□/, /e/, /□/, /æ/, /a/, /□/, /o/, /□/ and /u/ (e.g., 'keep', 'feed' and 'beat' for the vowel /i/).

2.3 Speaking task

In a quiet room, participants were given a written list of prompts, and heard each item played through headphones. They were recorded repeating each word immediately after hearing the auditory prompt, using a high quality digital recorder. Immediately after the speaking task, participants evaluated how familiar they were with each word they had just produced, using a 4-point scale 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 Judgments

Using a sound-editing program, each recorded word was extracted from the original recording, normalized, and saved as a separate sound file. These files were then used to create a second set of files, in which the vocalic portion of each word was extracted, from the first glottal pulse after the initial consonant to the last glottal pulse before the following consonant. Perceptual screening was used to confirm that the selections were as accurate as possible.

Using *Praat* (www.praat.org), five phonetically trained judges assessed the intelligibility of each vowel production in three listening conditions. In Condition 1, all 5160 isolated vowel recordings were randomly presented to listeners, who were asked to identify the vowel they heard by clicking on one of ten buttons that represented the target categories. In Condition 2, the same procedure was used, except that listeners heard the entire word rather than the isolated vowel. The first 2 conditions were conducted in 26 equally sized blocks over several weeks. In Condition 3, recordings of the whole word were again presented over several weeks. However, items were blocked by word (e.g., all productions of 'keep' were presented before moving on to all productions of another word), and judges were asked to click on one of two buttons indicating that the target vowel was produced either correctly or incorrectly. For each condition, vowels were considered intelligible if 3 out of 5 judges agreed that it was a member of the intended category.

3. RESULTS

Words for each vowel category were assigned to one of three groups based on their mean familiarity scores, from 1, "most familiar" to 3, "least familiar". In one case, a tie in mean familiarity scores was broken through reference to its frequency in the British National Corpus. The NSs were perfectly familiar with all of the words. Hence, lexical familiarity rankings reflect only L2 learner ratings.

Among the 19 NSs who provided baseline production data, identification rates for isolated vowels varied from 60% to 100% (*M*=86%). In addition, 97% of the vowels extracted from the elicitation prompts were intelligible (one /u/ production was not). This confirms that it is possible for NS judges to identify vowels played in isolation.

Overall results by L1, listening condition and lexical familiarity ranks are shown in Figure 1.

Native speaker baseline intelligibility: A repeated measures ANOVA revealed significant main effects for listening

condition [$F(2,36) = 44.847$, $p < .000$, $\eta^2 = .714$] and lexical familiarity [$F(2,36) = 5.739$, $p = .007$, $\eta^2 = .242$], but not vowel. A significant interaction between listening condition and lexical familiarity was also found [$F(4,72) = 5.594$, $p = .001$, $\eta^2 = .237$]. Post-hoc Bonferroni adjusted *t*-tests indicated that mean intelligibility scores were significantly lower in the isolated vowel condition than in the other two listening conditions [$t(18) = -6.815$, $p < .000$; and $t(18) = -6.649$, $p < .000$ respectively]. Furthermore, mean intelligibility scores were significantly lower for the most familiar words, versus those that were least familiar [$t(18) = -3.172$, $p = .005$]. Three repeated measures ANOVAs, one for each listening condition, were conducted to investigate the interaction between listening condition and lexical familiarity. These revealed that the significant effect of lexical familiarity was only detectable in the isolated vowel condition [$F(2,36) = 6.187$, $p = .005$, $\eta^2 = .256$].

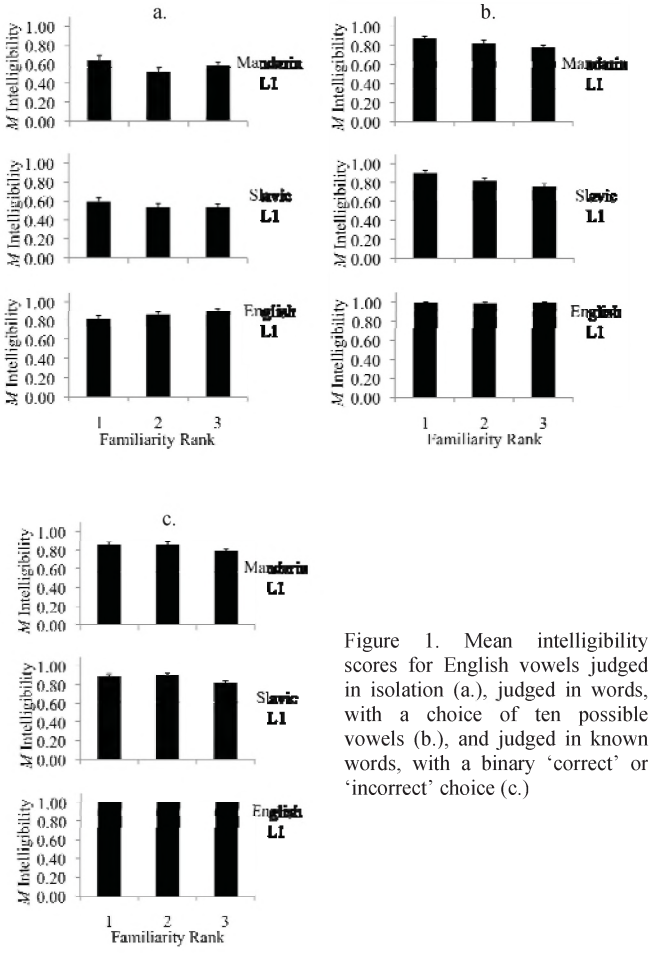


Figure 1. Mean intelligibility scores for English vowels judged in isolation (a.), judged in words, with a choice of ten possible vowels (b.), and judged in known words, with a binary 'correct' or 'incorrect' choice (c.)

L2 speaker intelligibility: A repeated measures ANOVA revealed significant main effects for listening condition [$F(2,72) = 169.997$, $p < .000$, $\eta^2 = .825$], lexical familiarity [$F(2,72) = 11.040$, $p < .000$, $\eta^2 = .235$], and vowel [$F(9,324) = 4.971$, $p < .000$, $\eta^2 = .121$]. In addition, significant interactions between listening condition and lexical familiarity [$F(4,144) = 5.124$, $p = .001$, $\eta^2 = .125$]; listening condition and vowel [$F(18,648) = 9.763$, $p < .000$, $\eta^2 = .213$]; and lexical familiarity and vowel [$F(18,648) = 5.059$, $p < .000$, $\eta^2 = .123$] were also found. There was no significant effect of L1. Post-hoc Bonferroni adjusted *t*-tests indicated that mean intelligibility scores were significantly lower in the isolated vowel condition than in the other listening conditions [$t(37) = -14.099$, $p < .000$; $t(37) = -13.172$, $p < .000$]. Condition 2 also resulted in lower scores than Condition 3 [$t(37) = -2.743$, $p = .009$]. In contrast to the results for NS productions, mean intelligibility scores were significantly higher for the most familiar words versus those that were second most and least familiar [$t(37) = 2.956$, $p = .005$; and $t(37) = 4.371$, $p < .000$ respectively]. There was no significant difference between vowel intelligibility in the second most and least familiar words. Three repeated measures ANOVA, one for each listening condition, revealed that the significant difference for lexical familiarity was detectable in all three listening conditions [$F(2,36) = 4.220$, $p = .018$, $\eta^2 = .105$; $F(2,36) = 15.80$, $p < .000$, $\eta^2 = .306$; and $F(2,36) = 9.691$, $p < .000$, $\eta^2 = .212$ respectively].

4. DISCUSSION

The results of this study indicate that lexical effects on the intelligibility of L2 vowel productions are present in the speech signal, and are not merely an artifact of listener bias towards evaluating vowels found in more familiar words as more intelligible. The intelligibility of NS vowel productions was also found to vary as a function of lexical context. However, while L2 vowels were most intelligible in the more familiar words, NS vowels were most intelligible in the less familiar words. This has implications for L2 speech learning, as it seems to indicate that the words that L2 learners are exposed to the most contain the poorest examples of the categories to be learned. Besides affecting speakers, lexical context also affects listeners. When they know what the intended category should be, listeners are more likely to evaluate the production as intelligible. This resulted in a ceiling effect for the NS productions, where real differences in the intelligibility of vowels, detected in the isolated vowel condition, were undetectable in listening conditions where lexical information was available.

REFERENCES

[1] 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.

[2] Thomson, R. I., & Isaacs, T. (2009). Within-category variation in L2 English vowel learning. *Canadian Acoustics, 37,* 138-139.

[3] Levi, S. V., Winters, S. J., & Pisoni, D. B. (2007). Speaker-independent factors affecting the perception of foreign accent in a second language. *Journal of the Acoustical Society of America, 121,* 2327-2338.

ACKNOWLEDGEMENTS

SSHRC and a Brock University, Dept. of Applied Linguistics Research Grant provided funding for this study.

HOW THE ROLE OF F3 IN VOWEL PERCEPTION MAY BE INFLUENCED BY LISTENER EXPECTATIONS

Santiago Barreda, and Terrance M. Nearey
Dept. of Linguistics, University of Alberta, 4-32 Assiniboia Hall, Edmonton, AB T6G 2E7, Canada sbarreda@ualberta.ca

1. INTRODUCTION

The relationships among apparent speaker identity, a vowel's physical properties and its phonetic quality are not fully known. In a previous experiment, Experiment 1, we studied the relationship between f0, apparent speaker identity and vowel quality by asking participants to make simultaneous vowel and apparent speaker size and gender judgments. We considered three possibilities: that f0 might be directly related to vowel quality in the same way that the formants are, that f0 mainly affects vowel quality indirectly by affecting apparent speaker characteristics, and that f0 is not related to vowel quality at all. Results indicated that f0 affects vowel quality mainly indirectly, via its effects on apparent speaker characteristics. However, some listeners find judgments of size ('large' to 'small') difficult to make consistently and listeners may adapt different size criteria within and between genders. Experiment 2 was similar to Experiment 1, but listeners were instructed that all the synthetic voices were males and that they should report speaker's apparent age. We compare the results of the two experiments that involve the same stimuli and same phonetic responses, but different instructions and 'size-related' responses. There are some surprising results that suggest listeners' use of F3 and higher formants in forming phonetic judgment varies with differences in simultaneous non-phonetic tasks.

2. METHOD

2.1 Participants and stimuli

Participants were 25 students from the University of Alberta. All participants were students taking an introductory level, undergraduate linguistics course. Participants were drawn from a participant pool in which undergraduate students take part in experiments in exchange for partial course credit.

A continuum was designed that spanned roughly from the average F1-F2 frequencies of the /□/ of a male to those of the average /æ/ produced by a female in seven equally spaced steps. The fourth point of this continuum had F1-F2 frequencies appropriate for either an /æ/ produced by an adult male or an /□/ produced by an adult female. Production data collected at the Alberta Phonetics Lab indicated that F3 distributions were nearly identical for the two vowels, implying that F3 could carry little to no direct phonetic information. It was therefore expected that F3 could be manipulated without directly affecting the phonetic quality of the vowels. This seven-step F1-F2 continuum was combined with three different higher formant conditions (where higher formants F4 through F10 varied proportionally with F3) and three different f0 conditions. This yielded a total of 63 distinct vowel stimuli. The frequencies of all of the continuum points and f0 and F3 levels used are presented in Table 1.

	f0 Levels			F3 Levels		
	L	M	H	L	M	H
μ	108	153	215	2475	2755	3068

#	1	2	3	4	5	6	7
F1	684	735	789	848	911	978	1051
F2	1354	1455	1563	1679	1803	1937	2081

Table 1. Formant frequencies and f0 levels for the stimulus vowels used.

2.3 Procedure

2.3.1 Instructions and judgments.

The same stimuli were used in two experiments. In both experiments, participants were instructed that they would be hearing a human-like, 'robotic' voice producing vowels intended to be either /□/ or /æ/. Participants were asked to listen to the vowel and decide which of the two vowel categories the vowel sounded most like.

In Experiment 1, 19 listeners were told that the speakers were either males or females of varying sizes. In addition to the vowel responses, they were required report the gender and relative size of the apparent speaker, by clicking on appropriately marked response areas. (Barreda & Nearey 2010). In the new Experiment 2, listeners were told that these speakers were all male but varied in age from 5 to 25 years of age. Listeners were asked to indicate vowel quality the apparent age of the speaker. Typically, listeners responded to three repetitions of the stimulus list (189 responses), followed by a short break, after which the participant performed another three repetitions of the same list. A total of 9,450 responses were collected across all 25 participants.

2.3.2 Comparison of the two experiments

Experiments 1 and 2 were identical except for (a) the different instructions given to the participants about the

synthetic voices and (b) the non-phonetic response variables. We focus here on comparing only those instances from Experiment 1 in which participants indicated that they thought they were listening to a male speaker. This included 4,548 responses used out of a total of 6,921 responses overall. Since the vowel stimuli were identical in both experiments and only trials where participants had heard a male speaker were considered, any significant differences across both experiments must be attributable to the different instructions given or to differences induced by the differences in the non-phonetic tasks.

3. RESULTS

Initially, the objective of this second experiment was to see if apparent speaker age explained more of the variance in vowel quality than apparent speaker size. Results indicated that age in Experiment 2 was a worse response variable than speaker size in Experiment 1 in terms of its correlation with vowel quality. To further investigate the differences between the two experiments, logistic regression was carried out within-participant using vowel category as a response variable and F1, F3 and f0 as the predictor variables. The estimated coefficients were collected across all participants and compared across both experiments.

	F1	F3	f0
Experiment 1	40.22	-16.15	-2.72
Experiment 2	34.26	-4.25	-2.94

Table 2. Means of the estimated within-participant coefficients across both experiments.

Two sample t-tests were carried out on the within-participant coefficients across both experiments to see if the relationships between vowel quality and each of the three predictors (F1, F3, f0) differed between the two experiments. For example, the mean effect for F1 was 40.22 in the first experiment and 34.26 in the second experiment but this difference was not statistically significant ($t = 1.58$, $df = 41.26$, $p = 0.123$). Nor was there a significant difference for the f0 coefficients ($t = 0.3$, $df = 30.48$, $p = 0.767$). However the difference between the estimated F3 coefficients is highly significant ($t = -5.28$, $df = 22.13$, $p < .0001$), with Experiment I showing a substantially larger effect of F3.

4. DISCUSSION

Our results indicate that F3 has a substantially weaker effect on vowel decisions in Experiment 2 than Experiment 1. Since the stimuli in the two experiments were the same, the only differences between the two experiments are in the instructions given to participants and in the responses used to gauge apparent speaker characteristics.

The most striking result in this experiment, the differential weight of F3, was a surprise, for which we have no compelling explanation. Instead we will speculate on two possible factors that we hope to explore in future experiments.

One factor is that the variety of speakers contemplated by listeners was greater in Experiment 1. Furthermore, the gender distinction is a categorical one that seems to be very salient for listeners. It is possible that this induced listeners to pay more attention to F3 as a potential cue to speaker identity. Even though we are considering only the vowel categorization for voices judged as male, the additional attentiveness to F3 may have carried over to within-gender comparisons.

The second factor is that in Experiment 2, voices were described as male speakers only and the only size-related judgment, age, was a continuous one. Listeners reported some difficulty in estimating ages. Perhaps because of this uncertainty, they were generally less likely to make size-related voice distinctions of the kind we hoped to induce by varying F3.

To further investigate these ideas, we will need to conduct additional experiments, such as combining age and gender judgments, or asking for judgments of size without gender. In any event, the differential use of F3 in judgment of vowel quality is a phenomenon that deserves more thorough study. Our results also indicate that there are interactions between F3 and f0 which make the effect of F3 on vowel quality weaker as f0 increases.

Most of the theories of vowel perception we have encountered do not readily accommodate such variable cue weighting. Additional experiments and theory refinement are clearly required.

REFERENCES

Barreda, S. & T.M. Nearey. (2010). The relationship between fundamental frequency and vowel quality. *JASA, vol. 127, issue 3, p. 2019*

ACKNOWLEDGEMENTS

This work was supported by funding from SSHRC and carried out at the facilities at the Alberta Phonetics Lab and the Center for Comparative Psycholinguistics at the University of Alberta.

PERCEPTUAL EFFECTS OF VISUAL EVIDENCE OF THE AIRSTREAM

Connor Mayer[1], Bryan Gick[1,2], Tamra Weigel[1], and Douglas H. Whalen[2]
[1]Dept. of Linguistics, University of British Columbia, 2613 West Mall, British Columbia, Canada, V6T 1Z4
[2]Haskins Laboratories, 300 George Street, Suite 900, New Haven, Connecticut, USA, 06511

1. INTRODUCTION

It has been long understood that perceivers of speech integrate visual and auditory information from articulator movements, resulting in both interference (e.g., McGurk and MacDonald 1976) and enhancement (e.g., Sumby and Pollack 1954) of auditory perception. Few studies have investigated integration of other types of information in speech perception. Gick and Derrick (2009) found that during auditory speech perception, perceivers integrated tactile information in the form of light air puffs. These puffs, delivered cutaneously on the hand or neck, were designed to resemble speech aspiration. When puffs were present, aspirated stops were more often correctly identified as being aspirated, and unaspirated stops were more often misidentified as aspirated, showing that listeners integrate tactile information in auditory perception in much the same way as visual information.

The goal of the present study was to examine whether speech aspiration can influence perception when it is not felt on the skin, but is rather recoverable indirectly from the visual signal. We predict that when visual evidence of aspiration is present, it will be integrated in speech perception in much the same way as first hand tactile information is. An additional question that arises is whether perceivers automatically make use of this kind of ambient information, or whether they need to be consciously attentive to it.

2. METHODS

Participants were seated in a sound proof room and shown short video clips of a speaker producing the sequence "pom" and "bomb" in a noisy bar setting with multi-talker babble. The babble was set to such a volume that correct auditory-only identification of the sounds was about 70%. There were a total of nine conditions in the experiment: conditions 1 and 2 had a candle placed in front of the speaker: in 1, the speaker said "pom", visibly perturbing the candle by the aspiration of the /p/, while in 2 the speaker said "bomb", and the candle was not perturbed because of the lack of aspiration of /b/. Conditions 3 and 4 were identical to 1 and 2 except that the candle was placed to the side of the speaker, and thus was not perturbed. Conditions 5 and 6 used the same video as conditions 1 and 2, but with mismatched audio: in condition 5, perceivers saw a video "pom" accompanied by an auditory "bomb", while in condition 6 they saw the opposite. Conditions 7 and 8 used the video from conditions 3 and 4, but with ambiguous audio between "pom" and "bomb" created by morphing audio of the two words from conditions 3 and 4 using the program STRAIGHT (Kawahara 2003). Because morphing resulted in half the original sound files, both the "pom" videos and "bomb" videos in this condition used the same audio. This condition was intended to factor out the possibility of facial cues disambiguating the sounds. Condition 9 featured the candle to the side as in conditions 3 and 4, but with perturbation of the candle flame occurring at times not corresponding to the effects of the airstream. This condition was designed primarily for training purposes: perceivers were shown 10 tokens of it at the beginning of the experiment to downplay the significance of the flickering candle. Aside from condition nine, all conditions had 20 repetitions, resulting in a total of 170 tokens. Subjects were given a forced-choice task to identify whether they heard "pom" or "bomb" in each video clip by pressing the left and right arrows on a keyboard. Half the subjects pushed left for "pom", the other half pushed right. Stimuli were presented and input recorded using Psyscope B53 on an iMac. When the experiment was completed, subjects were asked what they had observed, whether they had been consciously aware of the candle flickering and whether they had used it in any conscious strategy to disambiguate the sounds. Subjects' data was separated into those who reported being consciously aware of the candle as offering perceptual cues and those who were not. A total of 19 native English speakers participated: 6 reported being aware of the candle while 13 did not. No subjects had any training in linguistics. All statistical analysis was done using R 2.9.1. Repeated measures analyses of variance were conducted with three audio conditions (aspirated, unaspirated, and ambiguous) and four video conditions ("pom" with candle, "pom" without candle, "bomb" with candle, and "bomb" without candle) for both groups of subjects with post-hoc Tukey's Honestly Significant Difference tests. Data from the training condition was not included in the analysis.

3. RESULTS

Both groups of subjects displayed significant differences in responses between "pom" with and without candle and "bomb" with and without candle, indicating that they were generally able to perceive a difference between aspirated and unaspirated stops. Overall results for subjects who reported not being aware of the candle showed significant main effects only for audio [$F(2, 82) = 8.0845$; $p < 0.001$]. There was no significant effect for video [$F(3,82) = 0.7703$, $p = 0.51$], nor any interaction between audio and video [$F(2,82) = 0.0203$, $p = 0.89$]. Post-hoc tests showed that tokens with an auditory "pom" were more likely to be

identified as such. There were no significant differences between any video factors. For interactions, there was no significant difference in response between "pom" with and without candle (conditions 1 and 3; $p = 1$), nor between "bomb" with and without candle (conditions 2 and 4; $p = 1$). Conditions with mismatched audio and video (conditions 4 and 5), both resulted in lower percentages of correct answers than in conditions with natural audio, but these differences were not significant. Most interestingly, condition 5, auditory "bomb" with an accompanying candle flicker, did not show significant differences from conditions with natural audio "bomb" with candle (condition 2; $p = 0.10$) and without candle (condition 4; $p = 0.17$). There were also no significant differences between conditions where videos for "pom" and "bomb" were paired with identical ambiguous audio (conditions 7 and 8; $p = 1$).

Overall results for subjects who reported being consciously aware of the candle flickering showed a significant main effect of video [$F(3, 36) = 27.0888$; $p < 0.0001$]. There was no significant effect for audio [$F(2, 36) = 0.6778$, $p = 0.68$] nor any interaction between audio and video [$F(2, 36) = 1.3413$, $p = 0.25$]. Post-hoc tests showed that tokens with a candle flicker were more likely to be identified as "pom". However, there was no significant difference between video "pom" with candle to the side and video "pom" with candle in front ($p = 0.99$). As well, there were significant differences between audio "pom" and "bomb", with "pom" being more likely to be identified as such ($p < 0.05$): the lack of a main effect likely comes from the lack of a difference between the ambiguous and "pom" audio ($p = 0.7$). For interactions, there was no significant difference between "pom" with candle and "pom" without candle (conditions 1 and 3; $p = 0.85$). There was also no significant difference in responses between "bomb" with and without candle (conditions 2 and 4; $p = 0.99$). "Bomb" audio with "pom" video (condition 5) did show significant differences with "bomb" with candle (condition 2; $p < 0.05$) but not with "bomb" without candle (condition 4; $p = 0.28$). "Pom" audio with "bomb" video (condition 6) did not show a significant difference for "pom" with an accompanying candle flicker (condition 1; $p = 0.14$), nor did it show a difference with "pom" without candle (condition 2; $p = 0.97$), although in both cases the tokens without a flicker were more often identified as "bomb" . Subjects who noticed the candle did not show any difference in responses between conditions where videos for "pom" and "bomb" were paired with ambiguous audio (conditions 7 and 8; $p = 0.99$).

4. DISCUSSION

Depending on whether subjects reported being consciously aware of it, the presence or absence of the candle flickering had different effects on their responses. The subjects who were not consciously aware of the candle did not show evidence of any integration or interference effects: an accompanying perturbation of the candle had no bearing on correct identifications of "pom", and a flicker accompanying auditory "bomb" did not produce interference effects. Rather, these subjects appeared to make their choices based solely on the audio.

Subjects who were aware of the candle, however, clearly showed effects of it on their responses. When the flame was perturbed, regardless of the accompanying audio, "pom" responses increased. Significant differences in responses between 'pom' and 'bomb' audio, however, suggest that audio did play a role in disambiguation for this group, mainly in cases where the candle was absent. Neither group showed a difference between visual "pom" and "bomb" coupled with identical ambiguous audio, suggesting that subjects were not able to use facial cues in differentiation.

Although both direct and indirect consequences of articulation, whether auditory, visual, or tactile, clearly influence perception, these results indicate that for certain types of information it is not enough that they are merely present: to be used in perception they require listeners' active attention. It is difficult to say where the boundary lies between information that can be unconsciously integrated and information that cannot. Tactile stimuli as in Gick and Derrick (2009) are a relatively indirect consequence of speech articulation and one with which speakers presumably have less experience (eg. puffs on the back of the neck), yet these can be unconsciously integrated. When these stimuli are present only in the visual modality, however, they can only influence perception when perceivers are consciously aware of them: this suggests that there is a wide range of ambient information that can be used in speech perception, but not all of it can be unconsciously integrated.

REFERENCES

Gick, B., & Derrick, D. (2009). Aero-tactile integration in speech perception. Nature, 462, 502-504.

Kawahara, H., & Matsui, H. (2003). Auditory morphing based on an elastic perceptual distance metric in an interference-free time-frequency representation. ICASSP, V.1, Hong Kong. 256–259.

McGurk, H., & MacDonald, J.W. (1976). Hearing lips and seeing voices. Nature, 264, 746-748.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. J. Acoust. Soc. Am. 26, 212-215.

ACKNOWLEDGEMENTS

Thanks to Donald Derrick, Mark Scott, Molly Babel, and the members of the UBC ISRL. This project was funded by an NSERC Discovery Grant to the second author and by NIH Grant DC-02717 to Haskins Laboratories.

TWO PHONOLOGICAL SEGMENTS, ONE MOTOR EVENT: EVIDENCE FOR SPEECH-MOTOR DISPARITY FROM ENGLISH FLAP PRODUCTION

Donald Derrick[1], Bryan Gick[1, 2], and Ian Stavness[3]
[1]Dept. of Linguistics, University of British Columbia, 2613 West Mall, Vancouver, BC, Canada, V6T 1Z4
[2]Haskins Laboratories, 300 George St., 9th Fl., New Haven, CT 16511, USA
[3]Dept. of Electrical and Computer Engineering, University of British Columbia, 5500 - 2332 Main Mall, Vancouver, BC, Canada, V6T 1Z4

1. INTRODUCTION

A pervasive assumption in speech motor behaviour is that units of speech production correspond to preplanned motor routines. In a possible counterexample, Browman and Goldstein [1] showed that, kinematically, a single glottal opening event can span two adjacent segments in s + stop clusters. However, because the segments are adjacent, it has not been clear whether this is a case of preplanned motor efficiency or simply a local concatenation of two separate motor events.

Here we present an argument that in the English word `Saturday', the two flaps may be generated as one articulatory gesture at the onset of the first (upward) flap, while the rhotic and the second (downward) flap occur automatically as a result of gravity and elasticity.

We know that in English there are four subphonemic categorical kinematic variations of flaps, identifiable via B/M ultrasound [2,3]. The first is an alveolar tap (AT). The tongue (tip) moves from below the alveolar ridge, makes contact and moves back. The second is a down-flap (DF). The tongue moves from above the alveolar ridge, makes contact, and continues downwards. The third is an up-flap (UF). The tongue moves from below the alveolar ridge, makes contact, and continues. The fourth is a postalveolar tap (PAT). The tongue moves from above the alveolar ridge, makes contact, and moves back.

There is also a strong relationship between the type of flap selected and the tongue tip position before and after the flap. In the case of a non-rhotic vowel preceding and a tongue tip-up rhotic vowel following the flap, there is a higher likelihood of UF production. In the case of a tip-up rhotic vowel, there is a higher likelihood of a DF [3].

Previous work [4] found an unexplained preference for a UF-DF movement pattern in sequences of flap allophones. Also, DFs are unique among the four variants in that they take advantage of gravity and elasticity: Gravity because, while a speaker is upright, the tongue tip moves from a high position to a low position, and elasticity because the tongue moves from a retroflex to a tongue rest position vowel (see [5]).

The human nervous system does not completely compensate for the effects of gravitational load on jaw motion [6]. Jaw motion during speech differs based on whether a speaker is prone or supine, and tongue motion does not entirely compensate in place of jaw motion.

Similarly, [7] has provided experimental 2D simulation based evidence that much of the forward looping pattern of velar stop production in VCV sequences is based on the anatomical structure of the tongue such that planning may be based on target sequence as much as trajectory motion.

If our hypothesis about flaps in 'Saturday' is correct, we expect more tip-up rhotics in 'Saturday' than in the no-flap control 'peppermint'. We don't expect such differences with 'herded her', a phrase where we expect DF-UF sequences, vs. 'herbifer'. We also expect shorter duration between flaps, and less variability in the flap sequences for 'Saturday' than for the phrase 'herded her'. A production experiment tests these predictions.

To identify whether gravity and myoelasticity alone can, in principle, complete a DF closure, we will also use biomechanical simulation. Biomechanical simulation is well suited for this study as it characterizes how forces within the system interact in order to generate observed movements. ArtiSynth is a biomechanics simulation toolkit, targeted toward modeling and simulation of the human vocal tract [8]. Recently, a model of coupled jaw-tongue-hyoid dynamics has been developed within the ArtiSynth framework [9] that includes bone structures, a deformable tongue model, muscle forces, dynamic coupling (tongue muscle forces act on the jaw and vice versa), and contact (tongue-jaw and tongue-palate).

2. METHODS

Eighteen native speakers of NAE participated. Participants were seated in an adjustable chair with headrest. Microphone and B/M ultrasound recordings were taken of them producing 12 repetitions of 38 sentences, 17 control sentences, and 19 sentences with single and double flap sequences. We focused on four words within the study, 'Saturday', 'herded her', 'peppermint', and 'herbifer'. We measured flap, and rhotic type and duration between flaps.

The reference tongue model [10] uses the Finite Element Method (FEM) to represent the non-linear, large deformation tissue properties of the tongue. Muscle fibers are embedded to represent the muscle structure of the tongue. The reference jaw model [11] is composed of rigid body components for the skeletal structures (cranium, mandible, hyoid bone) connected with point-to-point Hill-type muscles and has been used to analyze forces during unilateral chewing. Simulations reported for the coupled jaw-tongue-hyoid model have shown plausible speech and chewing motions.

3. RESULTS

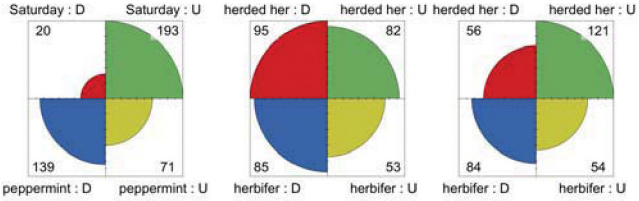


Fig. 1. Rhotic variants by subject by flap phrase (top) vs. control (bottom): U = tip-up, D = tip-down. Left = 'Saturday' vs 'peppermint', mid = initial rhotic 'herded her' vs. 'herbifer', right = final rhotic 'herded her' vs. 'herbifer'.

The difference in rhotic orientation is significant for 'Saturday' vs. 'peppermint' (AIC = 405, c = 2.94, z = 10.6, *p < 0.001), but not for the initial or final rhotic in 'herded her' vs. 'herbifer'.

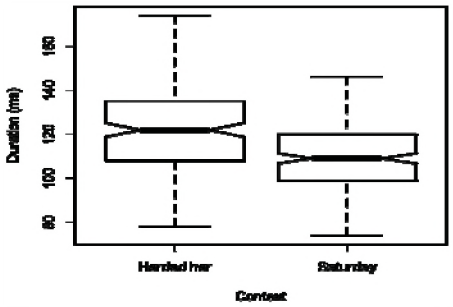


Fig. 2. Duration of 'Saturday' vs. 'herded her'.

The duration between flaps is shorter for 'Saturday' vs. 'herded her' [glmer (REML) t = -8.62 (AIC = -2091)].

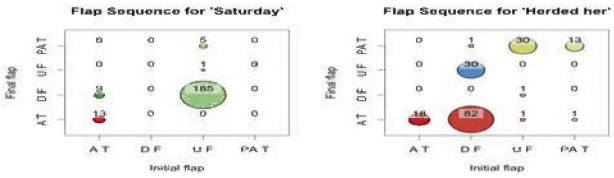


Fig. 3. Flap sequences in 'Saturday' vs. 'herded her'.

Flap sequence variability is much lower in 'Saturday' (mostly UF-DF) than 'herded her' (some DF-AT, and four other sequences prominent), as shown in by Fig. 3.

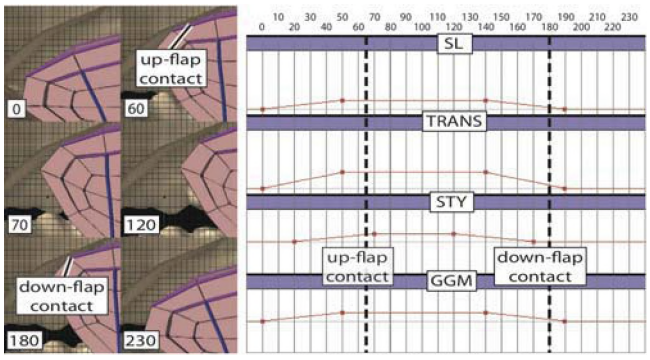


Fig. 4. Frame-by-Frame illustration of UF-DF sequence (left) in relation to muscle activation (right).

In the simulation, contracting the superior-longitudinal (SL), transverse (TRANS) and mid-genioglossus (GGM) lifts the tongue to the alveolar ridge, and the styloglossus (SG) pulls the tongue tip farther up into a retroflex. Upon relaxation, the tongue tip lowers, contacting the alveolar ridge en route. Duration between flaps is 115 ms, similar to Fig. 2.

4. DISCUSSION

The results of the experiment support all three hypotheses, indicating that an UF-DF sequence in the word 'Saturday' is preferred, and produced more quickly than flap sequences in 'herded her'. The simulation supports the hypothesis that the same cluster of muscle contractions can produce UFs and retroflex rhotics, and relaxing the muscles will produce DFs. The results demonstrate disparity between phonology and motor behaviour in that one motor event can encompass the production of two segments spanning a syllable boundary.

REFERENCES

[1] Browman, C. P. and Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3:219–252.

[2] Derrick, D. and Gick. B. (2008). Quantitative analysis of subphonemic flap/tap variation in NAE. *Proceedings of the Acoustics Week in Canada*, 36-3:162-163.

[3] Derrick, D. and Gick, B. (2010). There are no fixed motor programs in speech: Evidence from categorical kinematic variants of English flaps and taps, *Under review*.

[4] Gick, B. (2006). Errors and strategy shifts in speech production indicate multiple levels of representation. *Journal of the Acoustical Society of America*, 119 (5):3302.

[5] Gick, B., Wilson, I., Koch, K., and Cook, C. (2004). Language-specific articulatory settings: Evidence from inter-utterance rest position. *Phonetica*, 61:220–233.

[6] Shiller, D. M., Ostry, D. J., and Gribble, P. L. (1999). Effects of Gravitational Load on Jaw Movements in Speech. *Journal of Neuroscience*, 19(20):9073–9080.

[7] Perrier, P., Payan, Y., Zandipour, M., and Perkell, J. (2003). Influences of tongue biomechanics on speech movements during the production of velar stop consonants: A modeling study. *Journal of the Acoustical Society of America*, 114(3):1582–1599.

[8] Fels, S., Stavness, I., Hannam, A., Lloyd, J. E., Anderson, P., Batty, C., Chen, H., Combe, C., Pang, T., Mandal, T., Teixeira, B., Greena, S., Bridson, R., Lowe, A., Almeida, F., Fleetham, J., and Abugharbieh, R. (2009). Advanced tools for biomechanical modeling of the oral, pharyngeal, and laryngeal complex. In *International Symposium on Biomechanics Healthcare and Information Science*, page online.

[9] Stavness, I., Lloyd, J., Payan, Y., and Fels, S. (2010). Coupled hard-soft tissue simulation with contact and constraints applied to jaw-tongue-hyoid dynamics. *Under review*.

[10] Buchaillard, S., Perrier, P., and Payan, Y. (2009). A biomechanical model of cardinal vowel production: Muscle activations and the impact of gravity on tongue positioning. *Journal of the Acoustical Society of America*, 126(4):2033–2051.

[11] Hannam, A., Stavness, I., Lloyd, J. E., and Fels, S. (2008). A dynamic model of jaw and hyoid biomechanics during chewing. *Journal of Biomechanics*, 41(5):1069–1076.

ACKNOWLEDGMENTS

This research was funded by (NSERC) to the second author, and by National Institutes of Health (NIH) Grant DC-02717 to Haskins Laboratories. The authors thank Alan Hannam for providing the jaw model and his expertise, Gipsa-Lab, Grenoble for providing CT image data (P. Badin) and the tongue model geometry (S. Buchaillard, J.M. Gérard, P. Perrier, Y. Payan), Sid Fels and the ArtiSynth team.

A LOOK INTO THE PLOSIVE CHARACTERISTICS OF JAPANESE /r/ AND /d/

Thomas Magnuson
Dept. of Linguistics, University of Victoria,
PO Box 3045 Victoria B.C., Canada, V8W 3P4 thomasm@uvic.ca

1. INTRODUCTION

At first glance, two phonemes such as /r/ and /d/ do not seem likely candidates to share a great deal of phonetic similarity. In near-natural speech where articulatory reduction is expected, though, /d/ realized as [ɾ] (a flap/tap) is commonplace in a number of languages including Japanese. Japanese /r/, or more properly /ɾ/, is largely considered to be an alveolar or post-alveolar tap subject to a wide degree of phonetic variability (e.g. Vance 1987; Akamatsu 1997) . Alongside canonical [ɾ] realizations, variants of /r/ in Japanese can include the lateral flap [ɺ], lateral- and rhotic approximants [ɹ, l] (Okada 1999; Amanuma et al. 2004) as well as what Hattori (1951) calls a 'weak [d]' (or 'weak plosive' in Kawakami 1977; see also Vance, 1987). The weak plosive variant of /r/ is generally discussed as a positional variant, occurring word-initially or following a pause. This author's earlier auditory-phonetic work with a corpus of extemporaneous (Kansai) Japanese dialogue suggested (Magnuson 2008) that variation may also in part be individual. That is, in one conversational dyad each speaker's pattern of phonetic realization of /r/ differed widely. One speaker ('JFB') produced taps which included transients akin to the release bursts of stops (i.e., 'weak plosives'), while the other speaker ('JFA') produced a number of these alongside a variety of taps as well as lateral- and rhotic approximants.

This paper acoustically re-examines the same dataset to explore the hypothesis that the speaker who produced more plosive-like /r/s (JFB) would lengthen her phonetic realizations of /d/ so as to avoid neutralizing her /r, d/ contrast. The conclusion arrived at is that this speaker does use duration to augment the contrast, but she does so by reducing the length of /r/ as opposed to lengthening /d/.

2. METHOD

The speech data analyzed here consist of one approx. 30-minute telephonic conversation held between two female speakers of Kansai Japanese, JFA and JFB. This conversation is the last in a series of 10 such conversations between the same two speakers, recorded in separate acoustically-controlled environments as a subset of the JST/ATR ESP-C corpus of unscripted Japanese conversation (Campbell 2004, 2007).

Each speaker's side of the conversation (separate digital recordings) were analyzed using Praat. Tokens of /d/ and /r/ were segmented and labeled based on their acoustic properties as were visible in the sound spectrogram and waveform (see Fig. 1, overleaf). Plosive realizations of /d/ and /r/ were identified as those that featured a release burst followed by a short span of aperiodicity leading into a subsequent vowel. Although this sort of release phase is a defining characteristic of [d], it is not something normally associated with flaps/taps. Nonetheless, each speaker produced both plosive-like tokens (with release-like bursts) and flap-like tokens (with no burst but a general decrease in amplitude across a broad range of frequencies) as realizations of both phonemes. The start of the release phase of plosive-like realizations was identified as a brief spike in the waveform prior to the onset of the following vowel.

Only tokens of /r/ and /d/ which were directly comparable to one another were used in the present analysis of duration. That is, post-nasal and post-pausal tokens were excluded as their start-points could not be reliably determined acoustically; also, /r/ does not frequently occur in these positions in Japanese. Approximant realizations of /r/ and /d/ were also excluded as their precise durations could not be reliably determined. Thus, the tokens investigated here consist of both speaker's (phonetically) intervocalic /r/s and /d/s produced either as plosives or as flaps. Table 1 summarizes the tokens used in the acoustical analysis. Where appropriate, the statistical significance of differences in duration among the realizations (within and across speakers) were determined via single-sample t-tests using the mean value of the contrasting group as the test value.

Table 1. Plosive (p) & flap (f) realizations of /d, r/ by speaker.

	JFA	**JFB**	Total
	(p) 79	(p) 136	215
/d/	(f) 82	(f) 150	232
	161	**286**	**447**
	(p) 22	(p) 105	127
/r/	(f) 84	(f) 206	290
	206	**311**	**517**
Total	**367**	**597**	**964**

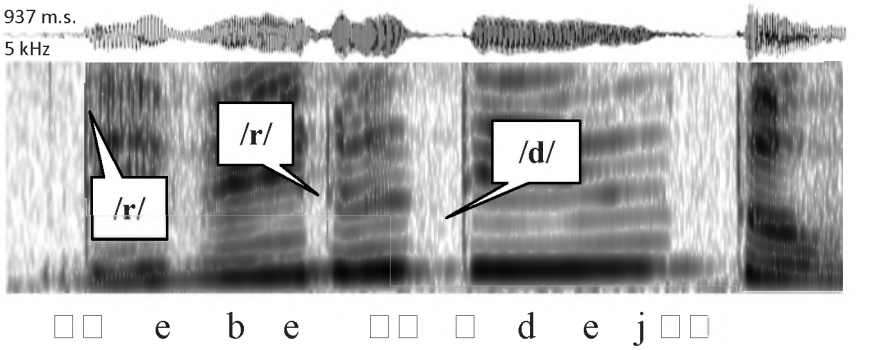


Fig. 1. A spectrogram and narrow transcription of *reberu-de iu-to* ('in terms of level'), spoken by JFB. A burst-like element is apparent in the 1st post-pausal /r/ (not included in the analysis). Note the comparative durations and presence of bursts among the 2nd /r/ and /d/. Both /r/s are transcribed as raised flaps ([□□]) to reflect a robust articulatory closure.

3. RESULTS

It is worth mentioning that, with respect to /r/, the dataset analyzed here reflects only a subset of the phonetic variety produced by the two speakers during their conversation. Both produced lateral- and rhotic approximants with varying degrees of articulatory reduction, in addition to lateral flaps (which have been grouped together here with non-lateral flaps/taps). That said, the tokens of /r/ analyzed here account for 68.7% of JFA's 300 total /r/s and 83.6% of JFB's 372 total /r/s. JFB produced substantially more plosive-like realizations for /r/, which comprised 28.2% (N= 105) of her total /r/s as compared to JFA, for whom plosive-/r/s constituted 7.3% (N= 22) of her total productions.

The hypothesis tested here is that JFB, for whom plosive-like realizations of /r/ were more frequent than JFA, would lengthen her /d/s so as to avoid confusion with her /r/s. This hypothesis would be supported if JFB's /d/s were significantly longer than JFA's. As it happened, the hypothesis was not supported; however, there were indications that JFB phonetically differentiated her /r/s from her /d/s via a different strategy. Specifically, JFB's /r/s (plosives and flaps alike) were significantly shorter than JFA's. These results are summarized in Fig. 2 below.

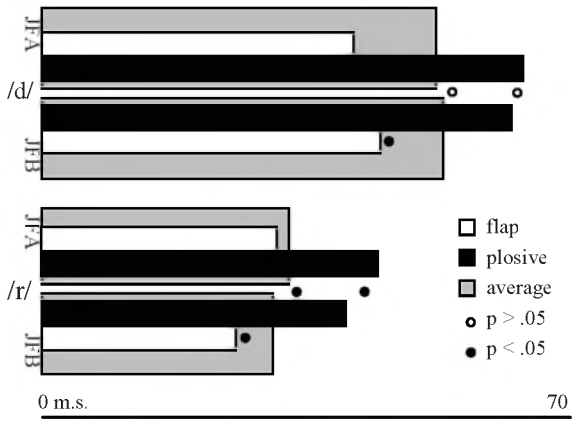


Fig. 2. Duration of both speakers' plosive and flap realizations of /d/ (upper) and /r/ (lower). Averaged flap (white) and plosive (black) duration is depicted in grey. Significance based on one-sample, two-tailed t-tests.

Either speaker's overall duration for /d/ (JFA: 52.2 m.s., JFB: 53.1 m.s.) and for plosive-/d/ (63.7 and 62.2 m.s.) did not differ significantly although JFB's flapped-/d/ duration was longer (44.8 versus 41.2 m.s.; $t = 3.64$, $p < .001$, mean difference = 3.65 m.s.). JFB's duration for both plosive- and flapped-/r/ were significantly shorter than JFA's: 40.2 and 25.7 m.s. versus JFA's 45.0 and 31.1 m.s. (Plosives: $t = -4.57$, $p < .001$; mean diff. = -4.71 m.s. Flaps: $t = -8.52$, $p < .001$; mean diff. = 5.41 m.s.). JFB's combined mean duration for /r/ was also shorter than JFA's: 30.6 versus 32.6 m.s. ($t = -2.95$, $p = .003$; mean diff. = -1.98 m.s.).

4. DISCUSSION

Taken together, these results suggest that speaker JFB exploited the duration of her /r/s in such a way as to maintain the phonological contrast with /d/. Since a duration distinction was also apparent between her flapped-/d/s and /r/s ($t = 7.59$, $p < .001$; mean diff. = 11.06 m.s.), I interpret JFB's use of duration for /r/ as a strategy for maintaining or enhancing her /r, d/ contrast. Much more study is needed to ascertain whether this is a pervasive influence on how /r/ is realized by speakers of Japanese.

REFERENCES

Akamatsu,T. (1997). Japanese Phonetics: Theory and Practice. München, Newcastle: Lincom Europa.

Amanuma,Y., Ohtsubo, K., Mizutani, O. (2004). Nihongo Onseigaku [Japanese Phonetics]. 3rd Ed. Tokyo: Kuroshiro.

Campbell, N. (2007). Resources for interacting with large-corpus speech data. Oriental COCOSDA, 4-7 Dec., Hanoi, Vietnam.

Campbell, N. (2004). Speech & expression: The value of a longitudinal corpus. Proc. 4th International Conference on Language Resources and Evaluation, 26-28 May, Lisbon.

Hattori, S. (1951). Onseigaku [Phonetics]. Tokyo: Iwanami Shoten.

Magnuson, T. (2008). It's not just a liquid, it's a stew: Phonetic variation in Kansai Japanese /r/. Poster presented at the Annual Conference of the Canadian Linguistic Assn., 31 May - 2 Jun., 2008, Vancouver, B.C.

Okada, H. (1999). Japanese. In: Handbook of the IPA: A Guide to the Use of the International Phonetic Alphabet. Cambridge: Cambridge University Press, 117-119.

Vance, T.J. (1987). An Introduction to Japanese Phonology. New York: State University of New York Press.

AN ACOUSTIC AND AUDITORY COMPARISON OF IMPLICIT AND EXPLICIT PHONETIC IMITATION

Meaghan Delaney, Soraya Savji, and Molly Babel
Dept. of Linguistics, University of British Columbia, 2613 West Mall, Vancouver, BC, Canada, V6T 1Z4

1. INTRODUCTION

Talkers have been found to spontaneously accommodate to a model talker in single word shadowing tasks (Babel, 2009, 2010; Goldinger, 1998; Namy et al., 2002; Nielsen, 2008; Shockley et al., 2004). This behaviour has been termed *implicit* phonetic imitation, as talkers have no awareness of having modified their speech. In this study, we compare implicit phonetic imitation to explicit phonetic imitation using an acoustic measure of vocalic phonetic distance and a holistic AXB perceptual similarity measure.

2. METHOD

2.1 Speech Production

Two groups of participants (n = 20, male = 4) completed an auditory naming task which included 50 monosyllabic words with the vowels /i æ ɑ o u/ produced by a male model talker. With the exception of task instructions, the procedure was identical for both groups. In the task instructions the EXPLICIT group was told to explicitly imitate the model talker. The other participants, termed the IMPLICIT group, were instructed to repeat the words naturally. For both groups, the task consisted of a Pre-task reading of the wordlist, followed by a series of three shadowing blocks and a final Post-task reading of the wordlist. This created a total of 800 productions per participant.

2.2 Acoustic Analysis

First and second formants were extracted from the single word productions from a series of Gaussian windows spanning the middle 50% of the vowel with a 2.5 ms step size. Vowel formants for the participants and model talker were normalized using the Lobanov (1971) method.

The Euclidean distance was calculated from each word production to that of the model talker. To calculate the amount of phonetic change in vowel production as a result of exposure to the model talker, the original baseline distance of each word was subtracted from the distance for each following instance of that word for each talker. This created a difference in distance (DID) value. A negative DID value indicates the phonetic distance between the participant and the model talker shrank. A positive value indicates an increase in phonetic distance and phonetic divergence. A value of 0 indicates no change as a result of auditory exposure to the model talker.

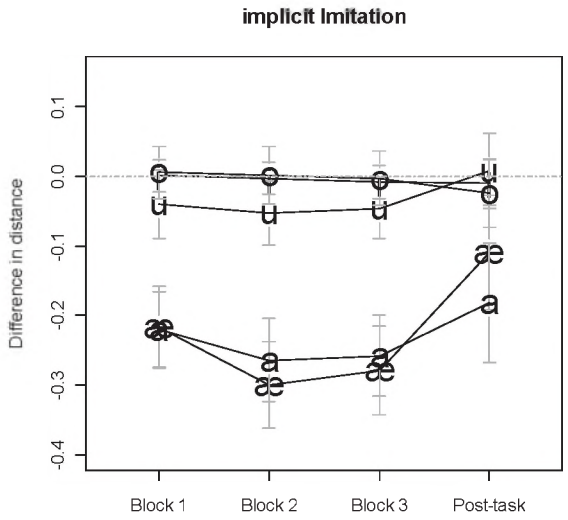


Figure 1. Implicit phonetic imitation by Vowel and Block. The Difference in Distance measure indicates the amount of phonetic imitation. A value of zero shows no change, a negative value indicates imitation and a positive value indicates divergence.

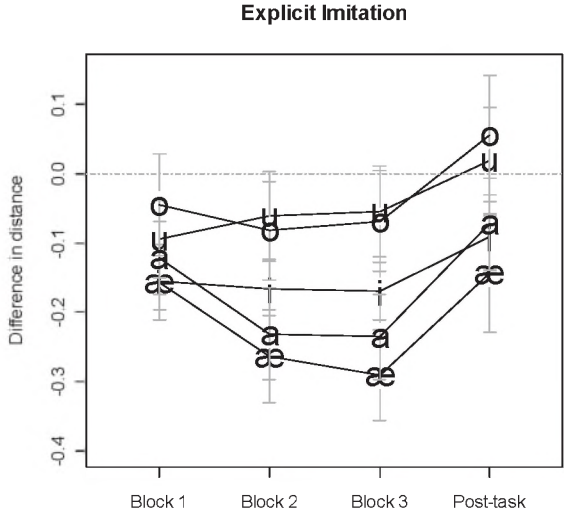


Figure 2. Explicit phonetic imitation by Vowel and Block. The Difference in Distance measure indicates the amount of phonetic imitation. A value of zero shows no change, a negative value indicates imitation and a positive value indicates divergence.

2.3 AXB Perceptual Similarity Task

Participants (n = 13, male = 5) completed an AXB task which included 27 of the monosyllabic words from the

original auditory naming task ending in stop consonants. In the task instructions, participants were told to determine whether the first or third word in a sequence of three identical words sounded more like the middle word. The stimuli consisted of productions from the pre-task, original model talker and the third shadowed block from the auditory naming task. The interstimuli interval was 200ms. The stimuli were blocked by talker and randomized. Counterbalancing was used, whereby participants heard both possible series; pre-task, model talker, shadowed, and shadowed, model talker, pre-task. These responses were analyzed for the proportion of shadowed productions judged more similar to that of the model talker.

3. RESULTS

3.1 Acoustic Analysis

A mixed effects linear model with Subject as a random effect and Block, Vowel, Condition , and Talker Gender as fixed effects was fit to the data using DID as the dependent measure. Figures 1 and 2 highlight some of the main results from the acoustic analysis. In short, the effects of the vowels and blocks across the two conditions were similar. The key exception to this was a main effect of Condition for the /i/ vowel ($\beta = 0.12$, $t = 2.1$), where /i/ was found to be imitated more in the Explicit Condition than in the Implicit.

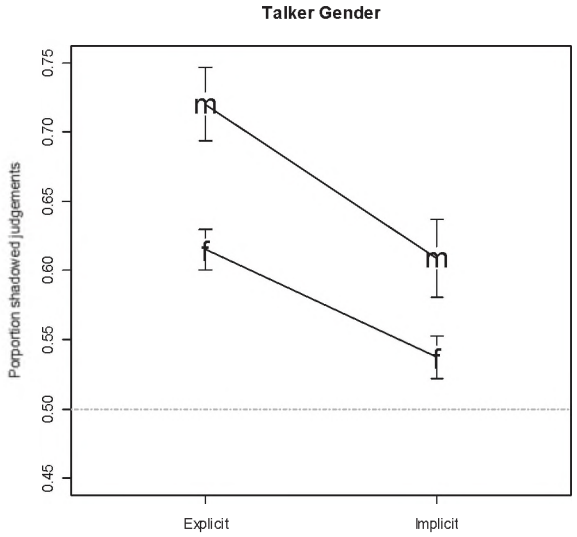


Figure 3. This figure illustrates the main effects of Condition and Talker Gender (male or female). Listeners judged shadowed tokens as more similar to the model talker for talkers in the Explicit condition. Listeners also judged male talkers' shadowed productions as more similar to the model.

3.2 Perceptual Analysis

The proportion of shadowed productions judged as more similar-sounding to the model talker were entered into a mixed effects model with Condition, Talker Gender, and Vowel as fixed effects; Subject was entered as a random effect. In the interest of space, we will report a small selection of the results. There was a main effect of Condition [β =-0.106, t = -5.74, $p < 0.001$] and Talker Gender [$\beta = 0.11$, $t = 3.91$, $p < 0.001$]. These results are shown in Figure 3. There were several significant interactions between Condition and Vowel environment; these are presented in Figure 4.

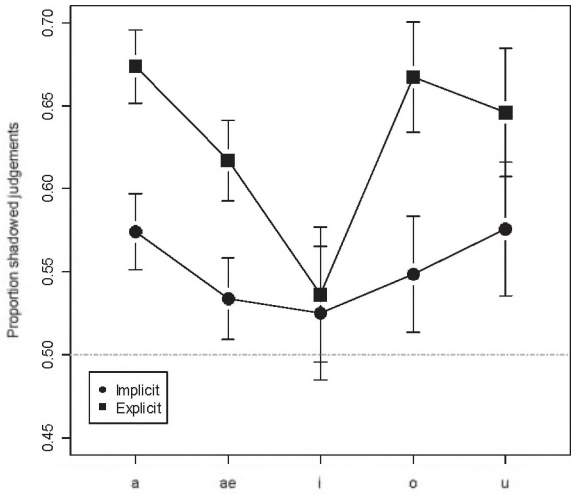


Figure 4. The Condition by Vowel interactions.

4. DISCUSSION

The similarity between explicit and implicit vowel imitation suggest similar cognitive biases mediate the two processes. The similarity judgement task, however, revealed that listeners are more sensitive to explicit imitation. Considering what we found for vowel imitation, this suggests that talkers are imitating more than just formant frequencies and that listeners are sensitive to these differences.

REFERENCES

Babel, M. (2009). *Phonetic and Social Selectivity in Speech Accommodation*. Ph.D. Dissertation, UC Berkeley.

Babel, M. (In press). Dialect divergence and convergence in New Zealand English. *Language in Society,* 39, 1-20.

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 &SocPsych.*, 21:422–432.

Nielsen, K. (2008). *The specificity of allophonic variability and its implications for accounts of speech perception*. Doctoral Dissertation, UCLA.

Shockley, K., L. Sabadini, & C.A. Fowler. (2004). Imitation in shadowing words. *P&P*, 66:422–429.

ACKNOWLEDGEMENTS

Order of the first two authors is alphabetical. Thanks to Keith Johnson for collecting the explicit imitation data at UC Berkeley, Grant McGuire and his lab for help in the data coding, and to UC Berkeley and UBC for financial support. A special thank-you goes out to our research participants as well.

GENDER DIFFERENCES IN AUTOMATIC PHONETIC ACCOMMODATION

Alexis Black
Dept. of Linguistics, University of British Columbia, 2613 West Mall, B.C., Canada, V6T 1Z4, akblack2g@gmail.com

1. INTRODUCTION

Automatic accommodation studies have shown that subtle acoustic features of a person's speech shift in the direction of an interlocutor during and following exposure to that person's speech (Goldinger, 1996; Namy, 2002; Nielsen, 2008; Babel, 2009). These temporary effects have been demonstrated in a variety of paradigms, and are frequently interpreted as support for exemplar-based theories. Research, however, has been inconclusive regarding the effects of socio-demographic characteristics on degree of imitation[1], and has not yet examined the importance of feature hierarchies (e.g. preferential imitation of VOT vs pitch).

Gender effects have been demonstrated in numerous studies on phonetic accommodation, yet no conclusive pattern has yet emerged. Namy et al. (2002) concluded from an immediate shadowing task that women show a higher degree of imitation, that both genders preferentially imitate men, and that women were more sensitive to the imitated acoustic qualities of other speakers' post-exposure productions. Pardo (2006), however, found that men demonstrated higher levels of accommodation in a conversational task. Babel (2009), on the other hand, found equal amounts of imitation across men and women.

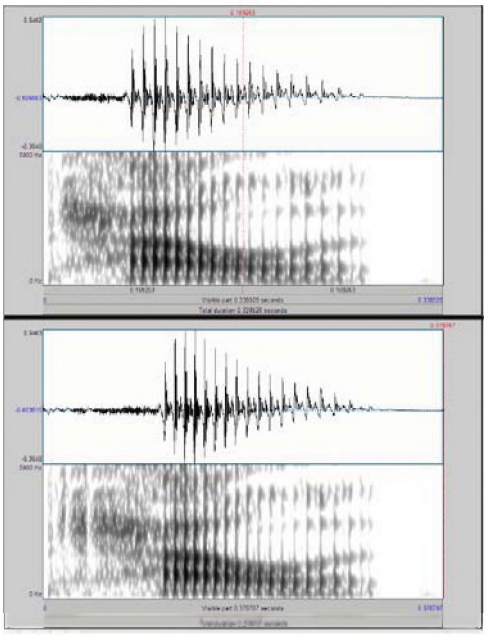
As in Pardo, in Nielsen's spontaneous imitation task with exposure phase (2008), men showed a greater degree of imitation than women. However, the female participants had longer baseline voice onset times (VOTs) than their male counterparts. Nielsen proposed two hypotheses to account for this pattern. Women may have accommodated less because the modified VOT stimuli were less salient (i.e. too close to their own baseline productions). Alternatively, perhaps this pattern is a reflection of a universal tendency to converge more towards the same gender -- as only a male model voice was presented during the experiment, female participants showed lower levels of imitation.

The present study examines the relationship of accommodation to participant and model talker gender. The experimental design allows accommodation to be indexed via multiple phonetic features, including VOT, vowel quality, and pitch.

2. METHOD

Following Nielsen (2008), the experiment involved a blocked-shadowing task that consisted of two production phases and one exposure phase. Stimuli include a total of 187 words: 86 /p-/ initial, 20 /k-/ initial, and 82 resonant or fricative initial. All words were controlled for lexical frequency (CELEX), and were selected based on vowel, initial consonant type, or as filler. A male and female model talker were asked to produce the /p-/, resonant, and fricative-initial stimuli. A novel aspect of the experiment design was that both the male and female voices were presented during the exposure phase; however, one of the two voices had been modified. Subjects participated in one of two conditions: in one, the male model voice had been modified to extend initial /p-/ VOTs by 40 ms.; in the other, the female VOTs were modified. Example stimuli can be seen in Figure 1 below.

Figure 1
Extension of Voice Onset Time in stimulus token "palate"



The original production of "palate" is shown in the upper window. The version in the bottom window has been extended by 40 msec through multiple splices of the original aspiration.

Stimuli were presented through E-prime Experimental Software 2.0 (Schneider et al., 2007). In phase one of the experiment, participants were asked to produce all 170 words as they were presented on a computer screen. In phase two, auditory stimuli (a subset of the total word list) were played in a continuous stream, exposing the participants to four repetitions of each word (two identical male tokens;

[1] The terms imitation and accommodation are used interchangeably throughout this paper.

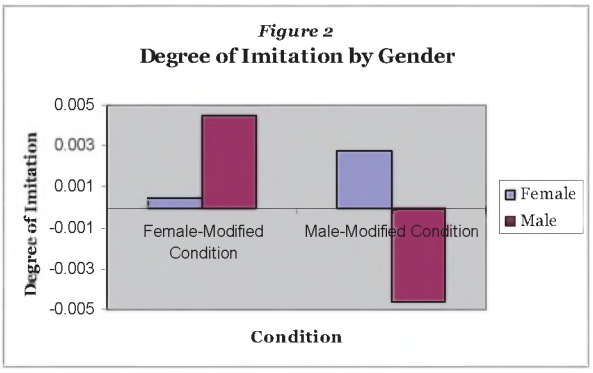
two identical female tokens). Finally, in phase three participants reproduced the entire word list. Stimuli were randomized in every phase. Follow-up questionnaires and informal interviews were conducted post-experiment to assess participants' implicit/explicit emotional stance towards the vocal stimuli.

Twelve participants in total were recorded and analyzed; four female and two male subjects participated in each condition. All sound was recorded using a head-mounted AKG C520 microphone in a sound-insulated booth at the University of British Columbia. Subjects were given ten dollars CAD for their participation.

3. RESULTS

Only VOT differences have been processed and analyzed to date. Degree of imitation was determined by calculating the mean of the difference scores for each participant's pre- and post-exposure /p-/ token productions. These imitation scores were then compared across condition and gender. As demonstrated in Figure 2, females showed a higher degree of accommodation to the male voice (Cohen's *d*: .60), and males displayed the reverse tendency (Cohen's *d*: 1.07).

Degree of imitation was also analyzed in relation to vocal preference and emotional reactivity (Nock et al., 2008). Using the same methods as above, there was an association between vocal preference and degree of imitation (Cohen's *d*: .54); however, it was less robust than the gender-imitation relationship. To determine the relationship between emotional profile and degree of imitation a Spearman correlation was calculated between the difference scores and Nock questionnaire scores, yielding a modest positive correlation (rho = .26).



Males showed higher degrees of imitation to the female-modified voice, while women showed higher degrees of imitation o to the male modified voice

Overall presence of imitation was calculated for each participant across conditions through a paired-samples *t*-test. A participant was considered an imitator if their post-exposure VOT duration was significantly longer than pre-exposure ($p < .05$). By this criterion, five of the twelve participants imitated the lengthened VOT.

4. DISCUSSION

The results of this experiment suggest that imitation is influenced by gender. Specifically, speakers are more likely to imitate the acoustic characteristics of a talker of the opposite gender. We might hypothesize that this is a reflection of heightened sensitivity/attention to speakers' that are perceived as more different from the listener. Alternatively, perhaps there is an evolutionary motivation, in that listeners are more attuned to potential reproductive partners. Continuing analyses will seek to determine whether participants showed differential degrees of accommodation to vowel quality and pitch.

Overall levels of accommodation to the lengthened VOT were quite low. Though VOT is known to be highly variable, previous studies on VOT accommodation still achieved a greater than 50% rate of accommodation. It is probable that the low levels found in this study derive from the novel paradigm design. Listeners were presented with two model talkers – only one of which had lengthened VOTs. These low levels of imitation suggest that the mere presence of a novel auditory stimulus in the speech stream is insufficient to provoke automatic accommodation; rather, external factors such as gender, voice preference, and emotional stance to the speaker modulate the process.

REFERENCES

Baayen, R.H., R. Piepenbrock, and H. van Rijn. 1993. The CELEX lexical database (CDROM). Linguistic Data Consortium, University of Pennsylvania.

Babel, M. (2009). Phonetic and Social Selectivity in Speech Accommodation. Doctoral Disseration, University of California, Berkeley.

Goldinger, Stephen D. 1996. Words and voices: Episodic traces in spoken word identification and recognition memory. Journal of Experimental Psychology 22:116–1183.

Namy, Laura L., Lynne C. Nygaard, and Denise Sauterberg. 2002. Gender differences in vocal accommodation: the role of perception. Journal of Language and Social Psychology 21:422–432.

Nielsen, Kuniko. 2008. The specificity of allophonic variability and its implications for accounts of speech perception. Doctoral Dissertation, University of California, Los Angeles.

Pardo, Jennifer S. 2006. On phonetic convergence during conversational interaction. Journal of the Acoustical Society of America 119:2382–2393.

Schneider, Walter, Amy Eschman, and Anthony Zuccolotto. 2007. E-Prime: User's Guide, version 1.0. Psychology Software Tools.

ACKNOWLEDGEMENTS

Many thanks to Dr. Molly Babel for her guidance and support.

SPEAKER-SPECIFIC PLACE OF ARTICULATION: IDIOSYNCRATIC TARGETS FOR JAPANESE CODA NASAL

Noriko Yamane[1] and Bryan Gick[2]
[1]Dept. of Linguistics, University of British Columbia, 2613 West Mall, Vancouver, BC, Canada, V6T 1Z4
[2]Haskins Laboratories, 300 George Street, Suite 900, New Haven, Connecticut, USA, 06511

1. INTRODUCTION

The long-standing question of whether place targets are necessary for speech gestures has focused primarily on the purportedly "targetless" English schwa [1, 7, 11]. Recent work suggests that English schwa is not placeless or "targetless" after all [3, 4], but rather that schwa may simply be particularly susceptible to coarticulation [2].

A sound also famously purported to be "targetless" is the Japanese coda nasal (henceforth N). Like English schwa, N is susceptible to coarticulation with a following consonant (e.g., aNka, aNta, aNpa; [12, 8]), and has often been called "placeless" ([5, 10]). One x-ray study [6] describes N as having a general dorsal place ('velar or uvular'). However, previous articulatory studies of Japanese N have been inconclusive, being limited to a single subject and lacking quantitative measurement.

A lingual ultrasound study was conducted to test whether N shows evidence of a stable place of articulation.

2. METHOD

2.1 Participants

Seven native speakers of standard Japanese (5 females and 2 males, ranging from early 20s to early 40s) participated in the experiment. The purpose of this experiment was not told. One participant's data was omitted, as her speech was affected by previous temporomandibular joint surgery.

2.2 Materials

All stimuli - *aNa, ahha, akka* - were pseudo words, phonologically and morphologically controlled. They are presented in katakana orthography as in アンア, アッハ, アッカ. The tokens consist of randomized word lists in which each token was repeated 14 times. 10-12 examples of each token were used for the current study.

2.3 Procedure

Recording was conducted in Interdisciplinary Speech Research Laboratory at University of British Columbia. Participants were trained to read all tokens with initial-accent (e.g., a'N.a) at a natural rate.

An Aloka SSD-5000 ultrasound machine was used with a UST-9118 and 180° electronic convex EV probe. Movie clips were recorded into iMovie, and converted to DV files. Still images were extracted at midpoints of consonants and vowels, using Final Cut Express ver. 1.01. Midsagittal tongue contours were produced using EdgeTrak software [9].

2.4 Design

The constriction degree (i.e., peak tongue height; y-axis) and location (i.e., peak tongue position; x-axis) were recorded, in order to compare intervocalic N with (a) flanking /a/ vowels, and (b) the *velar* consonant /k/ and the guttural consonant /h/.

3. RESULTS

3.1 Constriction Degree: N vs. Local Vowel Context

For each participant, N showed a constriction degree greater than that of surrounding /a/ vowels, as in Fig. 1.

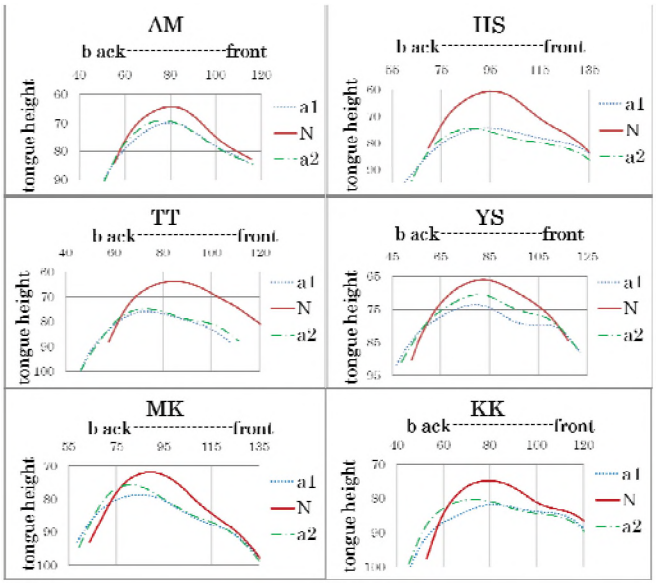


Fig. 1. Tongue Contours of *aNa*. The scale at x/y-axis is in pixels (1 pixel of x is around 0.89mm, 1 pixel of y is around 0.98mm).

3.2 Constriction Location: N vs. /k, h/

The *locations of these tightly constrained constrictions for N varied dramatically across speakers*. As in Fig. 2, the peak location of N for subjects AM and HS is further forward than both /h/ and /k/. TT and YS's N peak location is identical to /k/. MK's peak location of N is between /k/ and /h/. KK's peak location of N is identical to /h/.

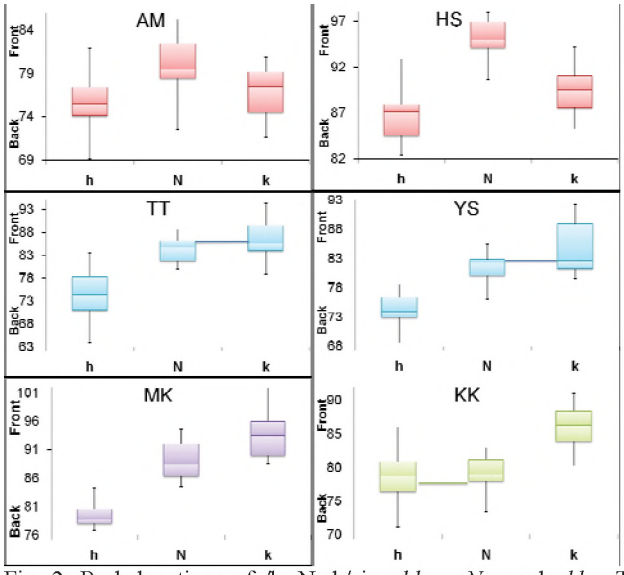


Fig. 2. Peak locations of /h, N, k/ in *ahha*, *aNa*, and *akka*. The scale at y-axis is in pixels (1 pixel is around 0.89mm).

The location of /N/ vs. /k/ or /h/ is confirmed using t-tests, as summarized in Table 1.

Table 1. T-test results (unpaired)

	AM	HS	TT	YS	MK	KK
N vs *h*	**<.01	**<0.01	***<.001	***<.001	***<.001	0.73
N vs *k*	**<.01	**<.01	0.24	0.066	**<.01	***<.001

3.3 Variability of N

N's Standard Deviation (SD) of x values at peak constriction locations was compared with those of /h, k/. As Table 2 shows, SD of N was, on average, less than that of /h/ or /k/.

Table 2. Standard Deviations of consonant location

	AM	HS	TT	YS	MK	KK	AVG
h SD	3.12	5.62	5.29	3.14	2.52	4.19	3.98
N SD	3.84	2.3	3.02	2.84	3.42	3.09	**3.09**
k SD	2.95	2.62	4.6	4.77	4.23	4.6	3.96

The result of f-tests also shows that the variance of N was comparable to that of /k/ for all speakers.

4. DISCUSSION

The above results suggest that N is not targetless, but has a constriction location target that is as stable (within speaker) as /k/, a sound with undisputed consonantal place. Further, places of articulation for N varied individually, ranging from palatal (significantly anterior to /k/) to velar (identical to /k/) to postvelar (significantly between /k/ and /h/) to uvular/upper pharyngeal (identical to /h/). This is illustrated in Table 3.

This high degree of idiosyncrasy in place of articulation accounts for the variability observed across previous single-subject studies. These highly speaker-specific findings for N's constriction location speak in favor of the ontogenetic emergence of place features within individual phonetic/phonological systems.

Table 3. Categorization of peak location

Type of N	Pharyngeal <--> Palatal			Participants
palatal	h	k	**N**	AM, HS
velar	h	k, **N**		TT, YS
postvelar	h	**N**	k	MK
Uvular/ pharyngeal	h, **N**		k	KK

REFERENCES

[1] Browman, C. P. and Goldstein, L. (1992). 'Targetless' schwa: an articulatory analysis. In G. J. Docherty, & D. R. Ladd, *Papers in laboratory phonology II: Gesture, segment, prosody* (pp. 26–56). Cambridge: Cambridge University Press.

[2] Flemming, E. The phonetics of schwa vowels. To appear in Donka Minkova (ed.) Weak Segments in English.

[3] Gick, B. (2002). An X-ray investigation of pharyngeal constriction in American English schwa. *Phonetica* 59, 38-48.

[4] Gick, B. and I. Wilson. (2006). Excrescent schwa and vowel laxing: Cross-linguistic responses to conflicting articulatory targets. In L. Goldstein, D. H. Whalen & C. T. Best (eds.) *Papers in Laboratory Phonology VIII: Varieties of Phonological Competence*. Berlin, New York: Mouton de Gruyter. 635-660.

[5] Itô, J. (1986). *Syllable theory in prosodic phonology*. Doctoral dissertation, University of Massachusetts, Amherst. [Published by Garland, New York, 1988.]

[6] Kokuritsu Kokugo Kenkyūjo (Japan). (1990). *Nihongo no boin, shiin, onsetsu: chōon undō no jikken onseigakuteki kenkyū*. Tokyo: Shūei Shuppan.

[7] Kondo, Y. (1994). Targetless schwa: is that how we get the impression of stress timing in English? *Proceedings of the Edinburgh Linguistics Department Conference '94*, 63-76.

[8] Nakajima, Y. (2003). Kounai-koudou: hassei-kikan no doutai-bunseki ni okeru chou-ompa imeejingu no yuuyou-sei [phonetics and speech technology] *Journal of the Phonetic Society of Japan* 7(3), 55-66, 2003-12.

[9] Stone, M. (2005). A guide to analyzing tongue motion from ultrasound images. *Clinical Linguistics and Phonetics*, Sept-Nov, 2005 Volume (19) 6-7, 455-502.

[10] Trigo, R. L. (1988). The phonological behavior and derivation of nasal glides. Ph.D. Dissertation, MIT.

[11] Van Bergem, D.R. (1994). A model of coarticulatory effects on the schwa. *Speech Communication* 14, 143-162.

[12] Vance, T. J. (1987). *An introduction to Japanese phonology*. Albany, N.Y.: State University of New York Press.

ACKNOWLEDGEMENTS

The authors thank colleagues at UBC and Haskins for helpful comments, and acknowledge funding from an NSERC Discovery Grant to the second author.

EXAMINING THE ACOUSTIC CONTRIBUTIONS OF THE EPILARYNGEAL TUBE TO THE VOICE SOURCE AND VOCAL TRACT RESONANCE

Scott R. Moisik[1] and John H. Esling[1]
[1]Dept. of Linguistics, University of Victoria, BC, Canada, srmoisik@uvic.ca

1. INTRODUCTION

In this study, we provide observations from videofluoroscopic and high-speed laryngoscopic techniques, combined with spectral analysis, to increase our understanding of the acoustic nature of aryepiglottic trilling and the epilaryngeal tube in general. The underlying goal is to work towards a complete biomechanical model of the laryngeal framework that acknowledges the role of the supraglottal structure of the larynx in shaping the nature of the voice source and of vocal tract resonance.

The epilaryngeal tube (sometimes referred to as the laryngeal vestibule; see Fig. 1) constitutes an important component of the larynx that can modulate the voice source in several ways and partly determines the resonances of the lower vocal tract. Under static configurations, Titze (2008) demonstrates how epilaryngeal tube constriction changes the acoustic coupling between the sub- and supralaryngeal vocal tract sections, giving rise to changes in vocal fold dynamics of a non-linear nature (where the filter is influencing the source). With sufficient airflow, various parts of the epilaryngeal tube can also be set into oscillation. This means that the acoustic coupling effect of the epilarygeal tube can vary as a function of time.

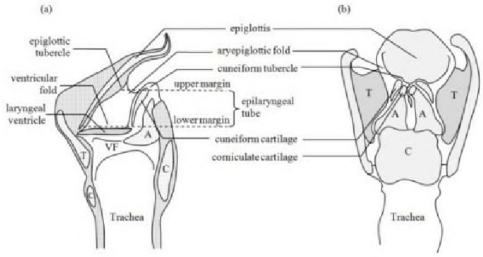


Fig. 1. Epilaryngeal tube anatomy. Mid-sagittal (a); Posterior (b).

There are three main components of the epilaryngeal tube that self-oscillate: the ventricular folds, the aryepiglottic folds, and the epiglottis. Ventricular fold oscillation (or ventricular phonation) in non-pathological cases occurs simultaneously with vocal fold oscillation (Fuks et al. 1998; Lindestad et al. 2002). It is most familiar in the ethnic singing styles of East Asia (e.g. Mongolian, Tibetan, and Tuvan throat singing). Notably, the ventricular folds undergo an irregular transient phase before reaching a highly stable steady state of oscillation where they become entrained with the vocal folds (at 1/2, 1/3, or equal to the glottal F0). Aryepiglottic (AE) oscillation (or trilling), impressionistcally labeled 'growling', is less thoroughly documented (although see Sakakibara et al. 2004). AE trilling is commonly heard in pop music singing styles (most famously by Louis Armstrong), but can be heard in throat clearing and coughing, and also occurs as a phonetic variation of pharyngeal segments in languages such as Iraqi Arabic and in a subset of phonatory registers, most notably in !Xóõ (Painter 1986; Esling & Harris 2005). When the epilaryngeal tube is heavily constricted, the aryepiglottic folds run perpendicular to the longitudinal dimension of the vocal and ventricular folds, and, depending on individual anatomical configuration, there is the possiblity for several epilaryngeal tube apertures or air-channels to exist. The region from the lateral margin of the epiglottis to the cuneiform tubercles defines the left and right apertures associated with the left and right AE folds. It is also common to see a medial aperture/channel formed by the medial surfaces of the cuneiforms. Furthermore, the arytenoid-corniculate complex can also form an aperture. Sakakibara et al. (2004) estimate that the lumen of the epilaryngeal tube contributes a pole in the vocal tract transfer function around 1.5 kHz but they note that there also acoustic losses due to the laryngeal ventricle. Finally, we have observed epiglottal oscillation (or epiglottal trilling; estimated to be at ~30-40 Hz; see Fig. 2) during the production of voiceless pharyngeals in Somali. The speaker in Fig. 2 possesses a flat epiglottis, which likely facilitates the oscillation. It remains an open question how common the linguistic use of epiglottal trilling is compared with aryepiglottic trilling. Even in this example, the aryepiglottic folds can be seen to oscillate.

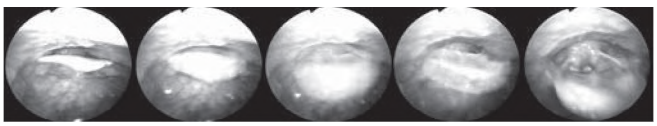


Fig. 2. An epiglottal trill (234 ms) in Somali (/ħaniin/ 'testicle').

2. METHOD

Videofluoroscopic (VIF) and high-speed laryngoscopic (HSL; 500 Hz) data of voiced aryepiglottic trills (produced at a glottal F0 of 100 Hz in an [a] context) were obtained. An estimate of epilaryngeal tube compliance is made based

on the VIF data using ImageJ to obtain pixel areas of the epilaryngeal tube lumen which were then converted to metric by using an estimated scaling factor. A laryngographic (EGG) and acoustic signal were simultaneously collected with the HSL data, and these are analyzed in conjunction with AE aperture information obtained from analysis of the HSL using kymographic and simple image analysis techniques (see Moisik et al. 2010 for more details based on a similar analytical approach).

3. RESULTS & DISCUSSION

The lumen of the epilaryngeal tube expands as airflow from the glottis impinges upon the surface of the tube immediately below the tubercle of the epiglottis (Fig. 3). This compliance is suspected to play an important role in self-sustained oscillation of the AE folds by inducing a phase shift between the volume velocity and the AE aperture area, effectively breaking the symmetry in energy exchange between the AE folds and the air flow.

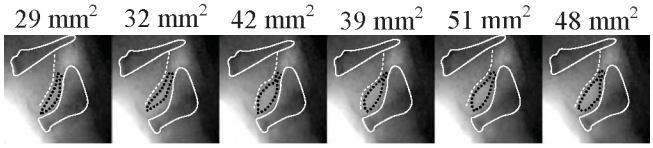


Fig. 3. Estimated epilaryngeal tube sagittal cross-sectional area using VIF. Black dotted line = measured area; white outlines = hyoid bone and cricoid-arytenoid cartilages; white dashed line = epiglottis surface contour.

Although the AE trill is highly unstable/irregular in its vibratory pattern (see Moisik et al. 2010), there are moments when entrainment between the vocal folds and the AE folds stabilizes; such a moment is presented in Fig. 4. In this figure, the right AE aperture (the left is statically patent during this sequence) is compared with the glottal aperture (indicated by the inverted EGG signal which shows moments of glottal opening as peaks). During the AE opening phase (markers 1, 3, and 5 in Fig. 4) vocal fold closure velocity is monotonic and rapid; the acoustic state becomes resonant and is characterized by strong spectral density up to the resonances associated with the F1 and F2 at which point the energy in the signal dips until around 1.5-1.7 kHz, where another resonance is encountered (F3; although this value is close to the resonance said to be associated with the epilaryngeal tube by Sakakibara et al. (2004)). During the AE closure phase (markers 2 and 4) there is an increase in the high frequency components as evident in the acoustic signal and spectral envelopes. This is likely attributeable to the airflow during the glottal open phase impacting with the inferior surface of the right AE fold. Vocal fold movement during the closure phase of the AE cycle is delayed, which may be a result of increasing back pressure on the vocal folds due to the ever narrowing epilaryngeal tube air channel. The spectral analysis reveals that at this point in the cycle a zero is introduced into the vocal tract filter in the region of 400-500 Hz. This acoustic loss may reflect the geometry of the epilaryngeal tube, which, in a way similar to lateral sounds, has a medial obstruction to airflow due to the cuneiform tubercles.

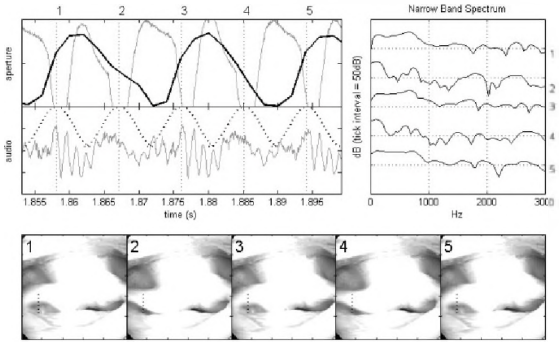


Fig. 4. Right AE aperture vs. inverted EGG (to show opening pattern). Dotted lines in audio plot show 10 ms Hamming windows used to calculate the FFT for spectral moments corresponding to different phases of the trill at glottal closure. Dotted black line in HSL = kymographic line used for estimating right AE aperture.

4. CONCLUSION

The epilaryngeal tube is a unique structure in the vocal tract in that it can simultaneously modulate and contribute to the vocal source as well as shape the acoustic coupling between the source and filter and yield its own resonances. This work represents a continual effort to expand the knowledge we have of this structure with the hopes of incorporating this knowledge into a more robust compuational model of laryngeal dynamics and acoustics.

REFERENCES

Esling, J. H. & Harris, J. G. 2005. States of the glottis: an articulatory phonetic model based on laryngoscopic observations. In *A figure of speech: a Festschrift for John Laver*. Hardcastle, W. J. and J. M. Beck (eds.). Mahwah, NJ: Erlbaum. 347–383.

Fuks, L., Hammarberg, B., & Sundberg, J. (1998). A self-sustained vocal-ventricular phonation mode: acoustical, aerodynamic and glottographic evidences, *KTH TMH-QPSR*, 3, 49-59.

Moisik, S. R., Esling, J. H., & Crevier-Buchman, L. 2010. A high-speed laryngoscopic investigation of aryepiglottic trilling. JASA 127 (3). 1548-1558.

Titze, I. R. (2008). Nonlinear source-filter coupling in phonation: Theory. JASA 125 (5), 2733-3749.

Traill, A. 1986. The laryngeal sphincter as a phonatory mechanism in !Xóõ (Bushman). In R. Singer and J. K. Lundy (Eds.), Variation, Culture and Evolution in African Populations: Papers in Honour of Dr. Hertha de Villiers. 123-131. Johannesburg: Witwatersrand University Press.

Sakakibara, K. I., Fuks, L., Imagawa, H., & Tayama, N. (2004). Growl voice in pop and ethnic styles. In Proceedings of the ISMA. Nara, Japan.

PITCH ESTIMATION FROM NOISY SPEECH BASED ON RESIDUAL-TEMPORAL INFORMATION

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

In speaker recognition, speech synthesis, coding, and articulation training for the deaf, pitch is an important speech parameter. Determining the fundamental frequency (F_0) or period (T_0) of a vocal cord vibration causing periodicity in the speech signal is the aim of pitch estimation. 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, the task of pitch estimation becomes very difficult when the speech observations are heavily corrupted by noise. Hence, many existing methods fail to provide accurate pitch estimates under noisy conditions.

In this paper, residual and temporal representations of speech are utilized for pitch estimation in a noisy environment. For a voiced speech, at the instant of glottal closure (GC), the major excitation of the vocal tract within a pitch period occurs. By careful analysis of the speech signal with the help of GC instants, it is possible to determine the pitch period. In comparison to the speech signal itself, in a residual signal (RS) of speech, some characteristics of the GC instants can be better observed. However, because of the bipolar fluctuations of RS around the GC instants, it is difficult to use the RS directly for pitch estimation. 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, we propose a circular average magnitude sum function (CAMSF) of the HE that exhibits more prominent peaks even in a heavily degraded condition. Simulation results testify that the global maximization of the temporal function, CAMSF, 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

A windowed filtered noisy speech frame is given by

$$x(n) = s(n) + v(n) \tag{1}$$

where, $s(n)$ and $v(n)$ represent the windowed and low-pass filtered version of clean speech and uncorrelated additive noise, respectively. Each windowed noisy frame of the observed noisy speech is low-pass filtered to remove very high-frequency contents. 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

Linear Prediction (LP) analysis is commonly used to derive the information about the GC instants for the extraction of pitch of speech signal. In LP analysis, in order to remove the vocal-tract information from the process of pitch estimation, an inverse-filtering operation is performed on the noise-corrupted speech $x(n)$ in a frame. The output of the inverse filter is referred as the error or residual signal (RS)

$$\Re(n) = T^{-1}\left[\hat{G}(z)X(z)\right] = x(n) + \sum_{k=1}^{p} \hat{g}_k x(n-k), \tag{2}$$

where T^{-1} represents the inverse operator of a z transform T with $T[x(n)]=X(z)$, p is the order of linear prediction, and $\hat{g}_k$ are the vocal-tract system parameters to be identified prior to inverse filtering. Here, $x(n)$ is passed through an inverse vocal-tract system filter $\hat{G}(z)$ given by

$$\hat{G}(z) = 1 + \sum_{k=1}^{p} \hat{g}_k z^{-k}. \tag{3}$$

For an N-sample frame of $x(n)$, the autocorrelation function (ACF) $\chi(m)$ of $x(n)$ can be estimated as,

$$\chi(m) = \frac{1}{N} \sum_{n=0}^{N-1-|m|} x(n)x(n+|m|), \; m = 0,\pm1,.....,\pm M, \; M < N, \tag{4}$$

where m is the discrete lag variable, and $\chi(m)$ obeys a recursive relation that relates the $\chi(m)$ values to the $\hat{g}_k$ parameters as

$$\chi(m) = -\sum_{k=1}^{p} \hat{g}_k \chi(m-k), \; 0 < m \le p. \tag{5}$$

The vocal-tract system parameters are obtained using the LP analysis based on the ACF $\chi(m)$ of $x(n)$. Since in the presence of noise, lower lags of $\chi(m)$ are generally become more corrupted than that of the higher lags, a few lower lags of $\chi(m)$ are avoided in the computation of the $\hat{g}_k$ parameters as

$$\begin{bmatrix} \chi(p) & \chi(p-1) & .... & \chi(1) \\ \chi(p+1) & \chi(p) & .... & \chi(2) \\ \vdots & \vdots & & \vdots \\ \chi(p+\lambda-1) & .... & .... & \chi(\lambda) \end{bmatrix} \begin{bmatrix} \hat{g}_1 \\ \hat{g}_2 \\ \vdots \\ \hat{g}_p \end{bmatrix} = - \begin{bmatrix} \chi(p+1) \\ \chi(p+2) \\ \vdots \\ \chi(p+\lambda) \end{bmatrix}. \tag{6}$$

A set of linear equations, which can be represented in the above matrix form is obtained utilizing the ACF coefficients $\chi(p+1),\ldots\ldots,\chi(p+\lambda)$ in (5). The number of equations to be used in (6) is governed by λ. The $\hat{g}_k$ parameters can easily be obtained from the least-squares solution of (6) to generate the RS according to (2). 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 $\hat{g}_k$ parameters. However, this ambiguity can be reduced by computing the Hilbert envelope (HE) of the RS as

$$H(n) = \sqrt{\Re^2(n) + \Re_h^2(n)}\ , \qquad (7)$$

where $\Re_h(n)$ is the Hilbert transform of the RS $\Re(n)$. The HE of $\Re(n)$ is a unipolar positive function. The correlation among the samples of the HE of $\Re(n)$ is high compared to the corresponding samples in the $\Re(n)$. However, in a severe noisy condition, by detecting the peaks at the GC instants in the HE of $\Re(n)$ 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 $\Re(n)$, we propose a circular average magnitude sum function (CAMSF) of the HE as given by

$$\eta(m) = \sum_{n=0}^{E-1} \left| H(\mathrm{mod}(n+m,E)) + H(n) \right|,\ m = 0,1,\ldots..,E-1. \qquad (8)$$

In (8), E is the number of speech samples employed to compute CAMSF for every lag m. The CAMSF of the HE is more effective in that it emphasizes the true pitch-peak even in a heavily degraded condition. By searching for the global maximum of the temporal function CAMSF, the desired pitch (F_0) is obtained as

$$\hat{F}_0 = \frac{F_s}{\hat{T}_0},\ \hat{T}_0 = \arg\max_m [\eta(m)], \qquad (9)$$

where F_s is the sampling frequency (Hz).

3. RESULTS AND DISCUSSION

By using the *Keele* reference database [2], the performance of the proposed method is evaluated. 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. In order to use the *Keele* database, we have chosen the same analysis parameters (frame rate and basic window size). 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. For windowing operation, we have used a normalized hamming window. In the estimation of $\hat{g}_k$ parameters by (6), λ is chosen as 5p.

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

Methods	Female	Male
Proposed Method	5.79	9.98
ACF Method	16.54	19.75
AMDF Method	19.40	28.75

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. For performance evaluation, we have used the voiced/ unvoiced labels included in the database as well as the true pitch value F_0.

For a speaker group, the percentage gross pitch-error is calculated considering two male (or female) speakers. We have compared the performance of the proposed pitch estimation method with the conventional autocorrelation function (ACF), and average magnitude difference function (AMDF) methods [1]. 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 is 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 temporal features is presented for pitch estimation from speech corrupted by a white noise. We computed a Hilbert envelope (HE) of the residual signal (RS) and found that the CAMSF 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 Chevengne, and Hideki Kawahara, "YIN, a fundamental frequency estimator for speech and music," *J. Acoust. Soc. Amer.* , vol. 111, pp. 1917-1930, 2002.

VOICE ACOUSTIC ANALYSIS OF TAIWANESE ADULTS WITH DYSARTHRIA FOLLOWING STROKE

Yu-Tsai Wang[1], Yuh-Mei Chung[2], Hsiu-Hsien Chen[1], Lin-Yang Chi[1], and Hsiu-Jung Lu[1]
[1]School of Dentistry, National Yang-Ming University, Taipei, Taiwan, 11221 yutsaiwang@ym.edu.tw
[2]Department of Physical Medicine and Rehabilitation, Taipei Veterans General Hospital, Taiwan, 11221

1. INTRODUCTION

Stroke is probably the most common cause of motor speech disorders, especially dysarthrias (Duffy, 2005). Although perceptual voice analysis for speakers with stroke-related dysarthria is still the most commonly used method in clinical settings, physiologic or acoustic voice analysis can provide a complementary objective evaluation of voice dysfunction in these speakers. Among the acoustic voice analysis tools, the Multi-Dimensional Voice Program (MDVP) (Kay, 1993) is commercially available and very easy-to-use. MDVP was reported to have fairly good suitability and reliability for the analysis of voice dysfunction in dysarthria (Biddle, Watson, & Hooper, 2002; Kent, Vorperian, & Duffy, 1999). However, as Kent et al. (1999) indicated 'Single trials may not be sufficient for the identification of voice dysfunction, and it appears that at least two trials are needed (p. 135)'. Moreover, the number of trials of vowel prolongation required to yield a reliable MDVP analysis for healthy speakers and speakers with dysarthria has not been determined. Investigating the discrepancy of MDVP values across trials for different populations could provide insight into this issue. Studies related to voice dysfunction in speakers with stroke-related dysarthria are still limited. It seems that objective acoustic analysis could detect voice abnormalities in dysarthrias following stroke and that group comparisons could delineate the voice profile in speakers with stroke-related dysarthria.

In order to obtain reliable acoustic measures to reach valid conclusions on the effects of stroke on voice function, it is legitimate to examine discrepancies between trials and make group comparisons based on at least two trials of vowel phonation for healthy speakers and speakers with stroke-related dysarthria. The purposes of the current study were: (1) to examine the trial-to-trial discrepancy between healthy speakers and speakers with stroke-related dysarthria; and (2) to investigate the effects of stroke on voice dysfunction in Taiwanese adults during a task of vowel phonation based on the acoustic voice parameters extracted by the MDVP.

2. METHOD

Ten Taiwanese male adults with mild to moderate stroke-related dysarthria (age 48.6 ± 13.0, ranging from 26 to 70 years) recruited from Taipei Veterans General Hospital, Taiwan. Ten neurological healthy age-and-gender matched participants were served as the controls. All participants wore an Audio-Technica PRO 8 head-mounted hypercardioid microphone connected to a digital audio tape recorder (TASCAM DA-P1). The microphone-to-mouth distance was 8 to 10 cm at an angle of about 45 degrees. They were instructed to "take a deep breath and say 'ah' for as long and as steadily as you can, until you run out of air" after the examiner's demonstration. Two trials of prolongation of the vowel /a/ for each participant were recorded on digital audio tapes (DAT R-94) at a sampling rate of 44.1 kHz and 16-bit quantization in a quiet room.

As Kent et al. (2003) suggested, the first 25 ms and the terminal phase of phonation in each trial were discarded. An interval of the most stable 3 s in the middle portion of the phonation of trial 1 and trial 2 were selected and analyzed with the Computer Speech Lab (CSL) and MDVP. The MDVP outputs of the two trials were then used to calculate the absolute mean difference (AMD-trial), standard deviation across trial (SD-trial). Finally, the means of the two outputs for each parameter were used to perform group comparisons. Based on a review of the literature that used MDVP for voice analysis, 17 out of the 33 acoustic voice parameters extracted by the MDVP were selected for this study, as shown in Table 1.

Due to the small number of subjects, limited knowledge about the nature of the data distribution, and the potential problem of non-independent relationship among acoustic measures in the current study, nonparametric statistic method (Mann-Whitney U test) was used to perform the designated group comparisons at a significance level of 0.05. Furthermore, the Bonferroni correction was applied to each group of MDVP-derived parameters to maintain the familywise error rate, i.e., 0.017 for fundamental frequency information measurements, 0.008 for short and long-term frequency perturbation measurements, 0.01 for short and long-term amplitude perturbation measurements, and 0.017 for noise related measurements.

3. RESULTS

As shown in Table 1, compared to the results of Kent et al. (1999), ten out of the 17 parameters were larger than previous reported data. The stroke group exhibited significantly larger discrepancies in fundamental frequency and its standard deviation (f_0 and STD), very long-term f_0 variation (vf_0), very short-term variability of the peak-to-peak amplitude (ShdB and Shim), and noise-to-harmonic ratio (NHR) than the control group.

Table 1. Means and standard deviations of the absolute differences between trial 1 and trial 2 (AMD-trial and SD-trial) for the 17 selected acoustic voice parameters extracted by the MDVP and statistical results of the group comparisons.

Parameter	Kent	Controls		Men with stroke		Z	p
		AMD-trial	SD-trial	AMD-trial	SD-trial		
f_0 (Hz)	8.62	2.58	2.69	**8.76**	10.54	**2.570**	0.010
STD(Hz)	3.69	0.35	0.24	**10.26**	25.17	**2.948**	0.003
PFR(semitone)	2.69	0.40	0.52	**4.40**	7.29	1.476	0.140
Jita(μs)	51.39	22.11	15.98	46.34	60.60	1.890	0.059
Jitt(%)	0.73	0.30	0.18	0.60	0.84	1.021	0.307
RAP(%)	0.41	0.18	0.11	0.34	0.50	0.492	0.623
PPQ(%)	0.48	0.18	0.12	0.38	0.70	0.076	0.940
sPPQ(%)	1.35	0.15	0.07	**1.78**	3.84	0.454	0.650
vF_0(%)	2.40	0.24	0.14	**5.80**	13.69	**3.326**	0.001
ShdB(dB)	0.18	0.06	0.05	**0.29**	0.26	**2.721**	0.007
Shim(%)	1.98	0.72	0.57	**3.05**	2.48	**2.721**	0.007
APQ(%)	1.43	0.56	0.38	**2.11**	2.05	1.739	0.082
sAPQ(%)	2.34	0.81	0.78	**2.60**	2.55	**1.965**	0.049
vAm(%)	7.28	3.82	6.79	5.97	6.02	1.739	0.082
NHR	0.04	0.01	0.01	**0.06**	0.06	**3.029**	0.002
VTI	0.03	0.01	0.01	0.02	0.03	1.102	0.270
SPI	4.30	2.80	2.10	2.23	2.41	0.907	0.364

As shown in Table 2, all mean MDVP values of the men with stroke-related dysarthria exceeded the threshold values, except for VTI, which is equal. Moreover, all mean MDVP outputs of the men with stroke exceeded those of the healthy controls. The stroke group exhibited significantly larger standard deviation and range of f_0 (STD and PFR), very long-term f_0 variation (vf_0), and long-term peak-to-peak amplitude variation (APQ and sAPQ) than the control group.

Table 2. Mean and standard deviation data for the 17 selected acoustic voice parameters extracted by the MDVP and statistical results of the group comparisons.

Parameter	Threshold	Controls		Men with stroke		Z	p
		M	SD	M	SD		
f_0 (Hz)	N/A	135	30	134	28	0.000	1.000
STD (Hz)	N/A	1.432	0.778	10.274	16.963	**2.570**	0.010
PFR (semitones)	N/A	2.100	0.459	7.300	8.782	**2.580**	0.010
Jita (us)	83.20	63.373	14.926	**122.473**	82.891	1.512	0.131
Jitt (%)	1.04	0.866	0.330	**1.670**	1.289	1.436	0.151
RAP (%)	0.68	0.515	0.202	**0.980**	0.757	1.209	0.226
PPQ (%)	0.84	0.497	0.186	**1.061**	0.916	1.209	0.226
sPPQ (%)	1.02	0.689	0.170	**2.626**	3.690	**1.965**	0.049
vf_0 (%)	1.10	1.029	0.349	**6.485**	9.842	**3.099**	0.002
ShdB (dB)	0.35	0.325	0.118	**0.675**	0.436	**2.570**	0.010
Shim (%)	3.81	3.677	1.189	**7.356**	4.233	**2.495**	0.013
APQ (%)	3.07	2.717	0.776	**5.668**	3.257	**3.024**	0.002
sAPQ (%)	4.23	4.155	1.021	**8.457**	5.107	**3.402**	0.001
vAm (%)	8.20	**13.692**	18.686	**17.897**	8.701	**2.495**	0.013
NHR	0.19	0.137	0.021	**0.195**	0.086	1.814	0.070
VTI	0.06	0.049	0.015	0.060	0.024	0.950	0.342
SPI	14.12	**14.160**	5.327	**14.424**	5.785	0.000	1.000

4. DISCUSSION

The current study indicates that, compared to the control group, (1) the stroke group exhibited significantly larger discrepancies in certain MDVP-derived acoustic parameters, such as fundamental frequency and its standard deviation (f_0 and STD), very long-term f_0 variation (vf_0), very short-term variability of the peak-to-peak amplitude (ShdB and Shim), and noise-to-harmonic ratio (NHR); and (2) stroke effects on acoustic voice parameters are mainly on the standard deviation and range of f_0 (STD and PFR), very long-term f_0 variation (vf_0), and long-term peak-to-peak amplitude variation (APQ and sAPQ).

The discrepancy across trials for the speakers with stroke in the current study is comparable to that in previous study (Kent et al., 1999), but the current study reports the degree of the discrepancy for the commonly used parameters in the literature and indicates that the values of certain parameters were more inconsistent between trials, and perhaps more than two trials for these parameters are needed to yield reliable measures to establish stable voice profile for speakers with stroke-related dysarthria. That is, less stable voice needs a greater number of trials to obtain representativeness.

Based on the average of two trials, the results of the current study roughly corresponds to previous studies (Kent et al., 1999; Kent et al., 2003), although some inconsistency exists among these studies. It is possible that the measurement discrepancy contributes to the inconsistency among studies. However, it is also possible that the differences in age, gender, lesion size and location, dysarthria severity, or other factors account for the corresponding voice stability and the resulting inconsistency among different studies.

The current study indicates changes in acoustic voice measures in stroke-related dysarthria. Discrepancies across trials of certain MDVP parameters for speakers with stroke-related dysarthria are significantly larger than those of healthy speakers. These parameters are relatively unstable and require more trials to obtain reliable measurements. Stroke effects on acoustic voice parameters are mainly on the standard deviation and range of f_0, very long-term f_0 variation, and long-term peak-to-peak amplitude variation. The results of the current study could form the basis of a larger investigation.

REFERENCES

Duffy, J. R., Motor Speech Disorders: Substrates, Differential Diagnosis, and Management: St. Louis: Mosby, 2005.

Kay Elemetrics, Multi-Dimensional Voice Program (MDVP). Pine Brook, NJ: Author, 1993.

Biddle, A., Watson, L., and Hooper, C., "Criteria for determining disability in speech-language disorders", AHRQ Publication NO. 02-E009, Evidence Report/Technology Assessment No. 52: 1-5, 2002.

Kent, R. D., Vorperian, H. K., and Duffy, J. R., "Reliability of the Multi-Dimensional Voice Program for the analysis of voice samples of subjects with dysarthria". American Jouranl of Speech-Language Pathology, 8: 129-136, 1999.

Kent, R. D., Vorperian, H. K., Kent, J. F.,and Duffy, J. R., "Voice dysfunction in dysarthria: application of the Multi-Dimensional Voice Program". Journal of Communication Disorders, 36(4): 281-306, 2003.

ACKNOWLEDGEMENTS

This work was supported in part by Research Grant number NSC 95-2314-B-010-095 from the National Science Council, Taiwan.

CONTRAST SALIENCE AND TALKER NORMALIZATION IN NONSIBILANT FRICATIVE PERCEPTION

Molly Babel[1] and Grant McGuire[2]
[1]Dept. of Linguistics, University of British Columbia, 2613 West Mall, Vancouver, BC, CANADA, V6T 1Z4
[2]Dept. of Linguistics, University of California at Santa Cruz, Stevenson Faculty Services, Santa Cruz, CA, USA 95064

1. INTRODUCTION

The contrast between /f/ and /θ/ is notoriously difficult to differentiate acoustically; Ladefoged and Maddieson (1996: 173) attribute this difficulty to individual differences in production. Such a claim would predict that when presented with auditory tokens of /f/ and /θ/, listeners would exhibit improved performance when faced with the voice of a single talker. We investigate this claim by comparing listener performance in an /f/ and /θ/ Yes-no style task where one group of participants is presented with the stimuli blocked by talker while the other is presented with a mixed-talker design.

2. METHODS

Stimuli. Five male and five female native speakers of North American English with some phonetics training provided the stimuli. Audio recordings were made in a sound attenuated room. Subjects wore a head-mounted AKG C250 microphone positioned about two inches to the side of the mouth. Productions were digitally recorded to the hard drive of a PC at a 44K sampling rate. Stimuli were displayed visually to the talker in a randomized order and consisted of the fricatives /f/ and /θ/ in CV, VCV, and VC contexts where the vowel was either /a/, /i/, or /u/ for a total of 18 stimuli. VCV tokens were consistently produced by the talkers with a H* accent on the initial vowel and L% on the second vowel.

Procedures. Both experiments used a Yes-no style task. Subjects were presented a single token per trial and responded <f> or <th> on a button-box. Subjects were encouraged to respond in less than 1000 ms. Subjects classified all tokens for the 10 talkers three times for a total of 540 trials. In the BLOCKED condition participants completed one talker before moving on to the next, creating a total of 10 blocks. In the MIXED condition participants were presented with all ten talkers in a single block; presentations were randomized across a total of 3 blocks.

Subjects. Thirty-four participants from UC Santa Cruz participated in the Blocked condition and thirteen participants from UBC completed the Mixed Condition. All participants were native speakers of North American English and had no speech, language, or hearing disorders.

3. RESULTS

Sensitivity. Accuracy scores were converted to *d'* according to MacMillan and Creelman (2005). Correct /f/ responses were assigned as 'hits' and identification of an /f/ response to a /θ/ trial was labeled as a 'false alarm'. A d' score of 0 indicates no sensitivity to the contrast and that subjects are responding randomly. Considering the unbalanced number of participants in the two conditions, the data was subjected to a linear mixed effects model for analysis. The dependent measure was d', Condition and Talker Gender were entered as fixed effects, and Subject was treated as a random effect. Main effects were found for Condition ($\beta = 0.23$, $t = 2.2$, $p < 0.05$) and Talker Gender ($\beta = 0.08$, $t = 2.2$, $p < 0.05$). These effects are shown in Figure 1. Listeners demonstrated increased sensitivity to male talkers and increased sensitivity overall in the Mixed condition.

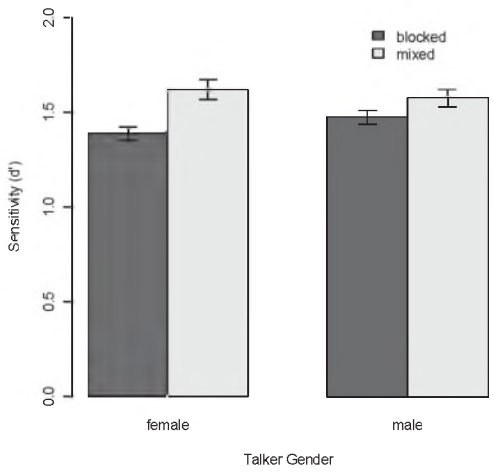


Figure 1. Main effects of Condition and Talker Gender on *d'* for the blocked and mixed conditions.

The criterion measure, or bias, was also subjected to a linear mixed effects model. Based on the arbitrary choice of calling correct /f/ identification a "hit", a positive criterion value indicates a bias to respond /f/ and a negative criterion value indicates a bias to respond /θ/. Criterion served as the dependent measure in the model, while Condition and Talker Gender served as fixed effects and Subject as a random effect. The model revealed a single main effect of Gender ($\beta = -0.34$, $t = -13.4$, $p < 0.001$). This effect is

shown in Figure 2; in both Conditions, listeners had a bias to respond /f/ to female talkers and /θ/ to male talkers.

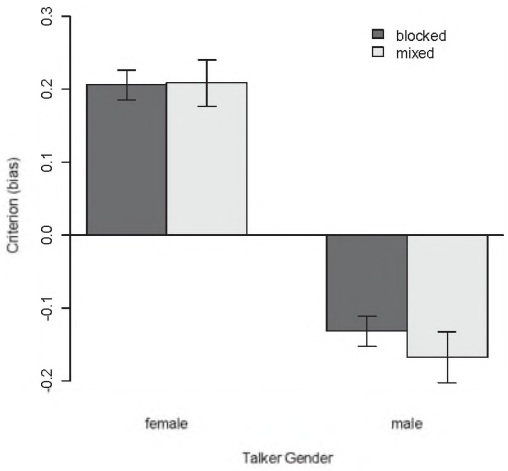


Figure 2. Main effect of Talker Gender on bias, shown for the blocked and mixed conditions.

Reaction Time. Log reaction times from correct responses were also subjected to a linear mixed effects model as the dependent measure. Prior to this, responses logged in under 200 ms were removed and outliers more than three standard deviations from the mean were removed as well. Fricative, Talker Gender, and Condition were entered as fixed effects, while Subject was a random effect. Condition (β = 0.08, t = 2.3, $p < 0.05$), Talker Gender (β=0.02, t = 4.3, $p < 0.001$), Fricative (β = 0.01, t = 2.4, $p < 0.05$) returned as main effects. There was a significant two-way Condition x Fricative interaction (β = 0.03, t = 3.3, $p < 0.001$) and a significant three-way Condition x Fricative x Talker Gender interaction (β = -0.03, t = -2.3, $p < 0.05$). Listeners responded more quickly in the blocked talker condition. Responses were also logged faster in response to female talkers and to /f/ tokens. These main effects are relatively consequential compared to the three-way interaction shown in Figure 3. Listeners responses in the mixed talker condition were fastest in response to female talkers /f/ tokens.

4. DISCUSSION

The results of this project both replicates previous work and contributes novel findings. The increased response latencies for the mixed talker condition replicates classic findings within the talker normalization literature (Martin et al., 1989; Mullennix et al., 1989), whereby increasing the number of talkers in the stimulus set prompts a delay in response latencies due to increased processing demands. However, the increase in listener sensitivity in the mixed talker condition is surprising and novel, as these same researchers who find a delay in response time in multi-talker conditions also find a decrease in accuracy. Our results suggest that for this particularly difficult acoustic contrast (Miller & Nicely 1955, Tabain 1998), even a slight increase in response latency produces more accurate (as measured by the sensitivity measure *d'*) responses. Despite an increase in response time across conditions, note that the bias effect remains constant. Regardless of whether listeners hear male and female talkers intermingled or separated, listeners have a bias to response /θ/ to male talkers and /f/ to female talkers.

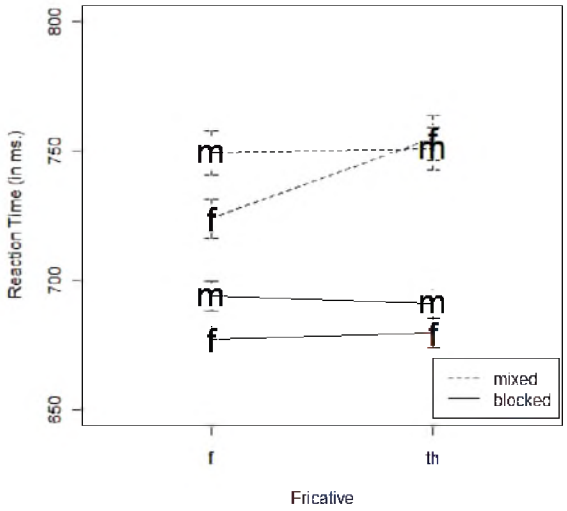


Figure 3. Reaction time by fricative for the mixed and blocked talker conditions. In this figure, "m" = male talkers and "f" = female talkers

REFERENCES

Ladefoged, P. & I. Maddieson. (1996). *The Sounds of the World's Languages*. Cambridge, MA: Blackwell Publishers.

Martin, C. S., Mullennix, J. W.. Pisoni. D. B., & Summers, W. V. (1989). Effects of Talker Varability on Recall of Spoken Word Lists. J. of Exp. Psych., 15, 676-684.

Miller, G. A. & P. E. Nicely. (1955). An analysis of perceptual confusions among some English consonants. *JASA*. 27 (2): 338-352.

Mullennix, J. W., Pisoni, D. B., & *Martin*, C. S. (1989). Some effects of talker variability on spoken word recognition. *JASA*. 85, 365-376.

Tabain, M. (1998). Non-sibilant fricatives in English: Spectral Information above 10 kHz. *Phonetica*, 55, 107-130.

ACKNOWLEDGEMENTS

Thanks to UBC and UC Santa Cruz for project funding. Thanks to Ryan Bennett Meaghan Delaney, Soraya Savji, and Travis Whitebread for assistance with this project.

AUTHOR NOTES

Author order is alphabetical; both authors contributed equally to this project.

EFFECT OF AGE ON LEXICAL DECISION SPEED WHEN SENTENCE CONTEXT IS ACOUSTICALLY DISTORTED

Marianne Pelletier, Huiwen Goy, Marco Coletta and Kathy Pichora-Fuller
Department of Psychology, University of Toronto, 3359 Mississauga Rd North, Mississauga, Ontario, Canada L5L 1C6

1. INTRODUCTION

Many factors affect a listener's ability to understand spoken language, including the availability and supportiveness of contextual information. In auditory lexical decision tasks, response times to target words are slowed when the preceding supportive context is acoustically distorted (Aydelott & Bates, 2004). In a previous study with normal-hearing younger adults, there was an effect of the amount of distortion, such that more acoustical distortion led to less facilitation by a congruent context (Pelletier, Goy, Coletta, Giroux, & Pichora-Fuller, 2010). Furthermore, distortion type affected lexical decision: when the sentence context was distorted by either time compression or low-pass filtering, congruent contexts facilitated lexical decision; however, incongruent contexts inhibited lexical decision only when the context was time-compressed.

Another factor that may affect a listener's ability to understand spoken language is age. Older listeners use congruent context in a compensatory fashion to offset age-related declines in auditory processing (Pichora-Fuller, 2008). However, when speech is time-compressed, older listeners may receive less information than younger listeners due to differences in auditory temporal processing or generalized slowing of cognitive processing speed (Pichora-Fuller & Souza, 2003).

In the current study, we tested older adults in two distortion conditions: low-pass filtering at 1750 Hz and 50% time compression. These conditions were the same as those selected in our earlier study of younger listeners because they resulted in similar reductions in word recognition for younger adults (Pelletier et al., 2010). It was predicted that, similar to younger adults, reaction times (RTs) would vary with the type of distortion, but that older adults would demonstrate more facilitation from context but a greater cost of distortion compared to younger adults.

2. METHOD

2.1 Design

Participants heard sentences in which intact target words were preceded by intact or distorted sentence contexts. The sentence contexts could be semantically congruent, neutral or incongruent with the target word. Trials with real-word items were interspersed with trials with non-word items, and participants made a lexical decision about each target word as quickly and accurately as possible by pressing one of two buttons on a button box. Sentence contexts were distorted by either low-pass filtering at 1750 Hz or 50% time compression, and 36 older adults were tested in each distortion type condition. Participants were seated in a sound-attenuating booth, and stimuli were presented binaurally at 70 dB SPL through Sennheiser HD 265 headphones.

2.2 Participants

Listeners were native English speakers with a mean age of 71.6 (SD = 4.3) and 69.7 (SD = 3.7) years in the low-pass filtering and time compression conditions respectively. They had audiometric pure-tone thresholds ≤25 dB HL from .25 to 3 kHz in the better ear. They provided informed consent and were paid $10/hr for their participation. Of the 86 participants recruited for the study, 14 participants were replaced because they failed to correctly answer at least 90% of test trials, leaving 36 in each group.

3. RESULTS

3.1. Reaction times for lexical decision by older adults

Only reaction times (RTs) for correct responses to real word targets were analyzed. The median RT for each participant in each condition was calculated. RTs were faster when the context phrase preceding the target word was intact (M=660 ms, SE=20) compared to when it was distorted (M=687 ms, SE=18). RTs were also faster when the context was congruent (M=497 ms, SE=19) than when it was neutral (M=762 ms, SE=18) or incongruent (M=762 ms, SE=23), but RTs for neutral and incongruent contexts did not differ. Furthermore, RTs were faster when the context phrase preceding the target word was low-pass filtered (M=616 ms, SE=26) than when it was time-compressed (M=731 ms, SE=26). Thus, RTs were fastest when the context preceding the word was congruent and intact congruent, with RTs for intact contexts being similar to those for low-pass filtered contexts, but RTs for time-compressed contexts being slower than those for intact contexts.

These descriptions were confirmed by an ANOVA with Distortion Type (Time Compression, Low-Pass Filtering) as a between-subjects factor, and Distortion (Intact, Distorted) and Context (Congruent, Neutral, Incongruent) as within-subject factors. There were main effects of Distortion, $F(1,70) = 5.54$, $p < .05$, Context, $F(2,140) = 310.79$, $p < .001$, and Distortion Type, $F(1,70) = 9.85$, $p < .005$. There were interactions of Context × Distortion, $F(2,140) = 17.18$, $p < .001$, and Distortion Type × Distortion, $F(1,70) = 23.67$, $p < .001$.

3.2 Age differences in reaction time

The results for older adults in the present study were compared to those of the younger adults tested under

the same distortion conditions in our earlier study (Pelletier et al., 2010). An ANOVA was conducted on median RTs to real word targets, with between-subjects factors of Distortion Type (Low-Pass Filtering, Time Compression) and Age (Younger, Older), and within-subject factors of Distortion (Intact, Distorted) and Context (Congruent, Neutral, Incongruent).

RTs were faster for younger adults (M=660 ms, SE=20) than for older adults (M=687 ms, SE=18), $F(1,140)$ = 14.23, $p < .001$. Besides an overall age-related difference in RT, there were also age-related differences specific to context and the type of distortion.

Facilitation by Context

Facilitation by congruent sentence context was calculated by subtracting the median RT for a participant's congruent trials from the median RT for the corresponding neutral trials to assess how much faster RTs were in the congruent condition compared to the neutral condition (see Fig 1).There was an Age × Context interaction, $F(2,280)$ = 5.65, $p < .005$, with older adults demonstrating more facilitation than younger adults (Fig. 1).

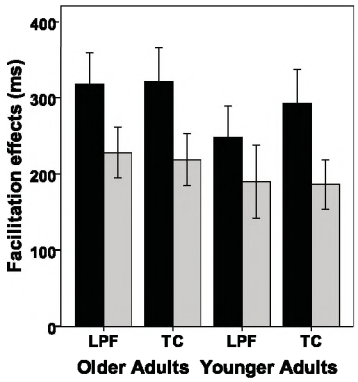


Fig. 1. Facilitation effects by age, distortion (black: intact contexts; gray: altered contexts), and distortion type (LPF: low-pass filtering; TC: time compression). Standard error bars are shown.

Distortion

As shown in Fig 2, the effect of distortion on RT was greater for older adults than for younger adults, especially in the time compression condition. We calculated the cost of distorting congruent semantic contexts as the difference RTs between each participant's distorted and intact congruent trials. There was an Age × Distortion Type interaction, $F(1,140) = 7.81$, $p < .01$.

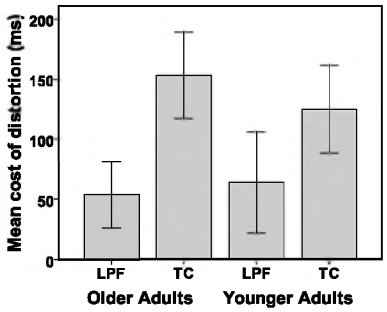


Fig. 2. Cost of distortion by age and distortion type (LPF: low-pass filtering; TC: time compression). Standard error bars are shown.

4. DISCUSSION

When tested in intact conditions, contextual information facilitated (speeded) lexical decision more for older adults than younger adults. However, for older adults, the speed of lexical decision was affected by the type of distortion, with the cost of distortion being greater for time compression than from low-pass filtering. Although the amounts of distortion used in the present experiment were matched in difficulty for younger adults, it is possible that they were not well matched in difficulty for older adults, thus inflating differences between types of distortion. Preliminary results from an ongoing experiment indicate that, compared to younger adults, older adults found the 50% time compression condition to be more difficult and low-pass filtering at 1750 Hz to be relatively easier. Thus, RT differences between distortion types seem to be minimized when word recognition difficulty is matched between conditions, but matching must be age-specific.

In conclusion, there are age-related differences in the use of contextual information and the speed of processing contextual information when it is time-compressed, and reduced uptake of the sentence context leads to consequences for later word processing.

REFERENCES

Aydelott, J. & Bates, E. (2004). Effects of acoustic distortion and semantic context on lexical access. *Lang Cog Proc,* 19, 29-56.

Pelletier, M., Goy , H., Coletta, M., Giroux, R., & Pichora-Fuller, K. (2010). *The effect of semantic context and the type and amount of acoustical distortion on lexical decision.* Manuscript submitted for publication.

Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *Int J Audiol, 47*(Suppl. 2), S72-S82.

Pichora-Fuller, M. K., & Souza, P. E. (2003). Effects of aging on auditory processing of speech. *Int J Audiol, 42*(Suppl. 2), 2S11-2S16.

NOISE REDUCTION POTENTIAL OF GREEN ROOFS

Ramani Ramakrishnan and Zarko Sopkic
[1]Department of Architectural Science, Ryerson University, Toronto, ON, Canada rramakri@ryerson.ca

1. INTRODUCTION

Green roofs on top of the commonly used roofing system design are becoming mandatory in major cities around the world[1]. One of the potential benefits of green roofs is touted to be the noise reduction of the added mass of the green roof structure. The transmission class of the green roofs has become a major research project in acoustics[2,3]. There are also skeptics who question the noise reduction potential of green roofs[4].

A model of an extensive green roof was investigated as to its noise reduction potential as part of an undergraduate thesis. Details of green roof types are presented in Lagstorm[4]. The noise reduction of different layers of the green roofs was measured in a mock-up 2 m X 2 m green roof model. The results of the noise measurement are presented in this paper.

2. THE EXPERIMENT

The green roof mock-up is shown in Figure 1. A cube of two metres is built as a box with a parapet to include green roof details on the top. An access door is also built into the side of the box to place a sound source within it. Images of the different layers that constitute green roof body are also shown in Figures (b) and (c).

The two base cases of the set-up consist of a typical roof over the plywood roof-deck. Base case 1 consists of a 2" thick extruded polystyrene insulation, water proofing membrane and ½" thick gravel. Base case 2 increases the insulation to 4" thick extruded membrane.

2.1 Green Roof Details

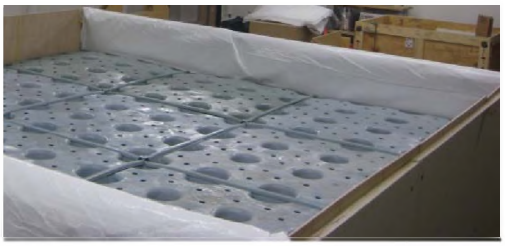
Three configurations of the extensive green roofs were tested in the current experiment. The green roof consists of: 2" retention panels (Figure 1b); bio-void, bio filter cloth (Figure 1c); 2" (or 4" or 6") layer of Bio-Mix Ecoblend soil. The green roof details are shown in Figure 2. Only two of the three extensive green roof configurations are shown in Figure 2.

2.2 Test Procedure

A boom-box DVD player was used as a sound source. A pink noise audio file was played through the DVD player. The sound levels within the box at three heights in one third octave bands were measured using Hewlett-Packard HP3569A real time analyzer. The sound levels at nine locations on top of the gravel (base case) or on top of the soil (green roof cases) were measured. The noise reduction in each third-octave band is given by the Equation (1) below:

$$NR = \{[\tfrac{1}{3}\sum_1^3 SPL_{i\ inside}] - [\tfrac{1}{9}\sum_1^9 SPL_{j\ top}]\}, dB \quad (1)$$

(a)

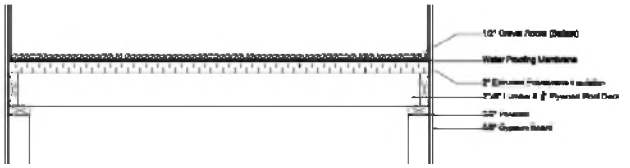


(b)

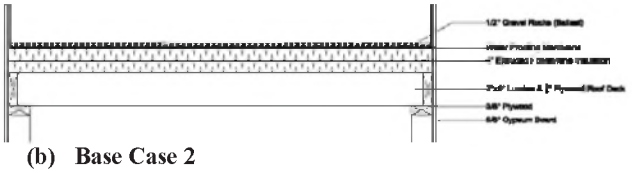
(c)

Figure 1. Details of Green Roof mock-up.

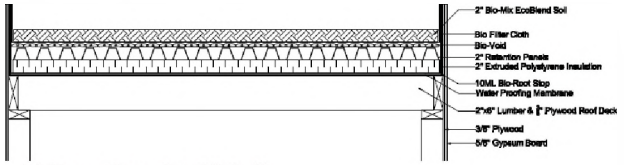
It must be pointed out that the current investigation measured only the noise reduction of the green roof configurations. No attempt was made to measure the TL (Transmission Loss) of the constructions in the same way as references 3 and 4 following a standardized procedure. The box construction was made sufficiently strong so that the noise propagation was the strongest through the roof.



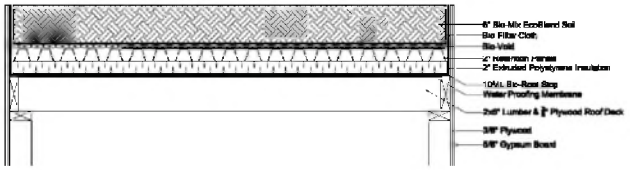
(a) Base Case 1



(b) Base Case 2



(c) Base Case 1 + 2" Soil



(d) Base Case 1 + 6" Soil

Figure 2. Details of Extensive Green Roof.

3. RESULTS AND DISCUSSION

The results of the Noise Reduction, NR, of the various configurations are presented in Figure 3.

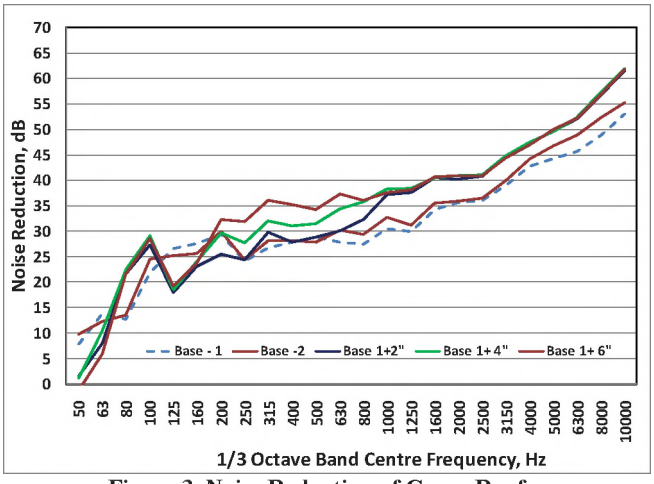


Figure 3. Noise Reduction of Green Roofs.

The first observation of the measured data is that there is hardly any difference between the two base cases. Actually, by adding another layer of insulation decreased the NR value at low frequencies for Base Case 2. This can be attributed, perhaps, to stiffness effects of the insulation materials.

The results for the three Green roof configurations were also presented in Figure 3. Two main observations were obvious from the test results: a) adding more layers of soil produce increased noise reduction in the mid-frequency range of 250 to 500 Hz; b) the thickness of the soil has no influence on the noise reduction in the high-frequency range. The reason for the low high-frequency performance is due to the leakage through the porous soil topping of the green roof construction. The performance in the high frequency range can get better due to the plant growth which was not considered during the current experiment.

The results show that the maximum additional noise reduction potential of a green roof is around 5 dB in the high frequency range and between 3 to 10 dB in the mid-frequency range. The test results of the current experiment are similar to the Transmission Loss values reported in References 2 and 3. Lagstrom[4] experiments also showed similar noise reduction of extensive green roofs.

4. CONCLUSIONS

A simple experiment was conducted as an undergraduate thesis project to evaluate the noise reduction potential of extensive green roofs. The results show that green roofs can provide around 5 dB of additional noise reduction. The results also showed that the thickness of the top soil has minimal influence on the high frequency performance of the green roofs.

REFERENCES

1. Toronto Green Roof By-Law, 31 January 2010, City of Toronto, 2010.
2. M. Connelly and M. Hodgson, "Sound Transmission Loss of Green Roofs," Conference on Greening Rooftops for Sustainable Communities, Baltimore, Maryland, April-May, 2008.
3. M. Connelly and M. Hodgson, "Sound Transmission Loss of Green Roofs-Fiewld Test Results," Proceedings of the Acoustics Week in Canada 2008, Vancouver, BC, pp 74-75, September 2008.
4. Jens Lagstrom, "Do Extensive Green Roofs Reduce Noise?" International Green Roof Institute, University of Malmo, Publication 010, Spring 2004.

MEASUREMENT OF THE SOUND ABSORPTION CHARACTERISTICS OF VEGETATIVE ROOFS

Maureen Connelly[1, 2] and Murray Hodgson[1]
[1] Acoustics & Noise Research Group, University of British Columbia, 3rd Floor, 2006 East Mall, Vancouver, BC V6T 1Z3
[2] Centre for Architectural Ecology, British Columbia Institute of Technology, 3700 Willingdon Ave., Burnaby, BC V5G 3H2

1. INTRODUCTION

Vegetative roofs have the potential to provide excellent external/internal sound isolation due to their high mass, low stiffness and their damping effect, and to provide a high level of sound absorption due to the low impedance of the vegetative substrate layer. The acoustical characteristics of vegetative roofs provide ecological contributions to the urban environment through a reduction of noise pollution from aircraft, elevated transit systems, and industrial noise. Vegetative roofs can reduce sound transmission into the interior of buildings, contributing to improved room acoustics through a reduction in noise and hence a reduction in distraction and stress (Connelly & Hodgson 2008). Through surface absorption, vegetative roofs will affect the propagation and build-up of positive and negative sounds and reduce reverberation in enclosed rooftop areas, altering the quality of the soundscape and the habitability of rooftops.

Vegetative roofs can be comprised of various material layers: root barrier, water reservoir/drainage layer, filter fabric, substrates and plants. The layer thought to have the most significant effect on the vegetative roof's acoustical characteristics is the layer of the vegetation and substrate. The vegetative substrate is complex to characterize; it varies in terms of the depth of substrate, the substrate constituents and physical properties, the plant's aerial biomass and root structure, as well as the dynamic in-situ microclimatic and conditions which vary over season and time.

2. METHOD

2.1 Laboratory Measurements of Vegetative-Roof Substrates Using an Impedance Tube

In order to understand the sound absorption characteristics of an established vegetative roof system it is first of interest to examine the absorption characteristics of the substrate independent of vegetation, before the vegetative substrate layer is investigated as a complex layer.

The most direct method to evaluate the absorption potential of vegetative-roof substrates is to measure the complex reflection coefficient in an idealized incident sound field using an impedance tube [1]. From the measured complex reflection coefficient, the normal-incidence sound-absorption coefficients and normal specific acoustic impedance ratios are calculated. A parametric experimental study was developed to determine the physical characteristics and properties of vegetative-roof substrates which contribute most significantly to the absorption of sound energy. The substrate characteristics of interest include: particle density, bulk density, total porosity, percentage organic matter, particle-size distribution. The properties which are a function of the micro-climate and site conditions of interest include volumetric water content and compaction.

Six vegetative-roof substrates and the primary three constituents – sand, compost and pumice – were measured in an impedance tube at three levels of volumetric water content: oven-dry (0%), wilting capacity and field capacity, and at two states of compaction. The permanent wilting capacity and the field capacity are percent by volume quantities which define the minimum and maximum available water content required for plant viability, and hence defined the limits of volumetric water content. The substrates were measured when either non-compacted or compacted to approximate in-situ conditions.

2.1 In-Situ Measurements of Vegetative Roofs Using the Spherical-Decoupling Method

The spherical-decoupling method has been used to measure sound-reflection properties and deduce the sound-absorption potential of locally-reactive surfaces and grounds, such as playing fields and forest floors [2,3]. This study first investigated the application of the spherical-decoupling method to vegetative roofs where the substrate surface particles and the plants of the vegetative roof do not provide a perfectly homogeneous and specular reflecting surface (i.e. no diffusion); surface properties which are inherent assumptions in the theory supporting the spherical-decoupling method. The overall objective is to understand the impact of substrate depth and plant species type and coverage on the absorption potential of vegetative roofs, so that vegetative roofs could be optimized for sound absorption.

The primary investigation was completed inside the controlled environment of an anechoic chamber with which a 1.68 m x 1.68 m test plot was constructed with material layers representational of the common vegetative-roof systems f the Pacific Northwest of North America. The less than ideal surface conditions have been accommodated in measurement methods through the determination of an appropriate geometric configuration of the sound source, the microphones and the surface plane, and through repeated measurements at multiple surface locations. The measurement method was subsequently utilized to complete in-situ measurements of 25 rooftop test plots. The test plots ranged in substrate depth in 25 mm increments from 50 mm to 200 mm; 7 test plots contained the substrate only. 18 test

plots were planted with three structurally-distinct plant species: sedums, a coastal-meadows community and grasses.

Fig.1. Set-up of instrumentation and vegetative roof test plot constructed in UBC Anechoic Chamber

3. RESULTS

3.1 Vegetative Roof Substrates

Figure 2 shows test results from the impedance tube measurements. These results suggest that within the octave bands 250 to 2000 Hz the percentage organic matter and volumetric water content are the most significant substrate characteristics relative to sound absorption. The absorption coefficient of the evaluated substrates correlated positively with percentage organic matter and negatively with moisture content and increased compaction. The state of compaction is significant in the 500 to 2000 Hz octave bands.

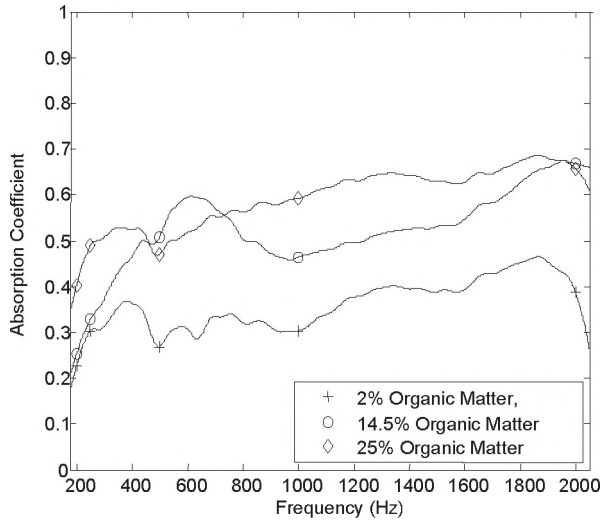


Fig.2. Absorption coefficients of Vegetative Roof substrate

3.2 Vegetative Roof Test Plots

In the anechoic chamber the absorption coefficient of the 130 mm substrate increased with frequency from 0.6 to 0.8 over the frequency range 686 Hz to 5489 Hz. Preliminary measurements of the vegetative-roof test plots – with moisture content at field capacity, plant coverage not yet fully established – indicated that the absorption coefficient of the substrate plots increased with frequency up to 3000Hz. The expectation was that the depth of substrate would increase the absorption of the vegetative roof; a positive correlation was observed - the difference between the minimum and maximum absorption coefficients was 0.15. A positive correlation was observed between absorption and total plant coverage.

4. DISCUSSION

The non-specular characteristics of the vegetative roof test plots provided additional challenges to the testing method developed in the anechoic chamber. A number of factors may have caused this disruption; this may include microclimatic conditions; wind and temperature and surface conditions..

An increase in total coverage, providing more consistency over the surface, will occur over the next several months of the spring and summer growing season. Additional investigations will focus on: validating and increasing the frequency limits; the height of the microphone above the surface and aerial biomass and the distance between the two microphones for specific plant structures.

REFERENCES

1. Allard, J.F., In situ two microphone technique for the measurement of the acoustic surface impedance of materials, Noise Control Engineering Journal, Volume 31 Number 1.
2. Attenborough, K. (1982), Acoustical characteristic of porous materials, North Holland Publishing Company – Amsterdam,
3. Kruse, R., Mellert, R. (2008), Effect and minimization of errors in in-situ ground impedance measurements, Applied Acoustics 69 884-890.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the academic input from Dr. Manfred Köhler, the support in the construction of vegetative roof infrastructure at the British Columbia Institute of Technology and funding from Western Economic Diversification and Investment Agriculture Foundation.

MEASUREMENT OF SPEECH PRIVACY OF CLOSED ROOMS USING ASTM E2638 AND SETTING CRITERIA IN TERMS OF SPEECH PRIVACY CLASS

Bradford N. Gover and John S. Bradley
National Research Council Institute for Research in Construction, Ottawa, ON K1A 0R6, Canada
brad.gover@nrc-cnrc.gc.ca

1. INTRODUCTION

The degree of speech privacy provided by an enclosed room refers to the extent to which conversations occurring inside are protected from overhearing by people outside, in the adjoining building spaces. The new ASTM standard E2638 "Standard Test Method for Objective Measurement of the Speech Privacy Provided by a Closed Room" [1] describes a test method suitable for use in enclosed rooms of nearly all sizes. It defines a measure called Speech Privacy Class (SPC) which, for a particular room and a particular listening point, is basically a fixed, physical property of the building. The measurement standard does provide some information on how to interpret and use the SPC value, but does not define criteria for "acceptable" performance.

This paper briefly reviews the new E2638 measurement method, and an approach for setting criteria using SPC.

2. ASTM E2638 MEASUREMENTS

Using a loudspeaker placed at 2 or more positions within the closed room of interest, measurements are made at locations inside the room to determine the average source room levels. (This is unlike open plan privacy measurements according to ASTM E1130 which require a calibrated loudspeaker with a specified directionality, location, and orientation.) For each source position, measurements of received level are also made, at a number of receiver positions outside the room. Unlike sound transmission loss testing (e.g., ASTM E336), the receiver positions are close to the boundaries of the room (0.25 m recommended), and at specific locations of suspected sound leaks (e.g., ducts, doors).

From the measurements, a level difference is calculated for each receiving point: the average receive level (at that point) is deducted from the average source room level to determine *LD(avg)*. Here "*(avg)*" means the one-third-octave band values are arithmetically averaged over the 16 bands from 160 to 5000 Hz. The background level $L_b(avg)$ is also determined at each receiver location. The sum of these two factors is the SPC:

$$\text{SPC} = LD(avg) + L_b(avg) \qquad (1)$$

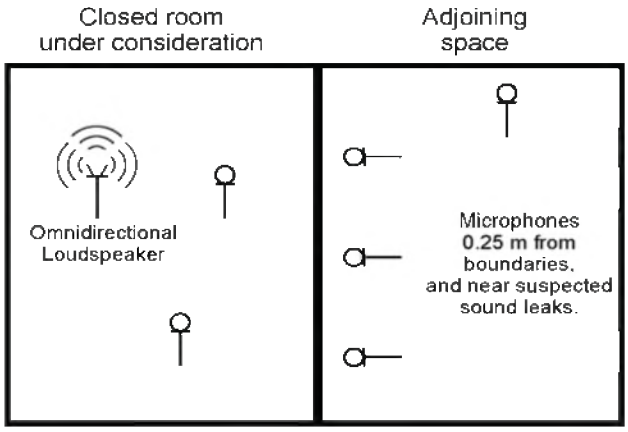


Fig. 1.Schematic indicating loudspeaker and microphone positions for ASTM E2638 measurements.

3. SPEECH PRIVACY CLASS

SPC is the sum of the measured average noise level at the position of a potential eavesdropper outside the room, and the measured level difference between a source room average and the transmitted levels at the same potential eavesdropper location. Increasing either the background noise or the sound insulation will result in a higher degree of privacy, and a correspondingly higher value of SPC. For a given speech level (inside the room), the SPC governs the "speech signal to noise ratio" at the position of a listener, which is related to intelligibility and audibility. Reference [2] defines the relevant signal to noise metric (SNR_{uni32}), and demonstrates the correlation with subjective listening test results of intelligibility and audibility.

The speech levels inside the room, however, are never fixed – they fluctuate as people speak, and can best be described statistically [3]. Viewed this way, the likelihood of transmitted speech being audible or intelligible can be related to the probability of higher-speech levels occurring in the meeting room. The SPC (for the listening position) determines the maximum speech level that is "adequately" protected. The probability of occurrence of higher speech levels is the probability of a speech privacy lapse. Higher values of SPC protect against less likely higher speech levels, and therefore relate to a greater degree of privacy.

Table 1. Descriptions of speech privacy categories, and proposed use for government designated information.

Category	SPC	Description	Proposal for suitable use
Minimal speech privacy	70	Frequently intelligible	
Standard speech privacy	75	Occasionally intelligible, and frequently audible.	Protected B information
Standard speech security	80	Very rarely intelligible, and occasionally audible.	Secret information
High speech security	85	Essentially not intelligible, and very rarely audible.	Top Secret information
Very high speech security	90	Unintelligible and essentially inaudible.	

Frequently : about 1 per 2 minutes
Occasionally: about 1 per 15 minutes
Very rarely: about 4 per 8 hours
Essentially not: about 1 per 16 hours

4. CATEGORIES

The appendix of E2638 includes a table that identifies the frequency with which speech sounds would be audible or intelligible, for various SPC values. This is summarized in the first three columns of Table 1. The descriptions of how frequently speech would be audible or intelligible are based on the statistics of speech levels measured in meeting rooms during meetings.

What the ASTM standard does not do is to indicate what likelihood of speech being audible or intelligible is suitable for end users. One possible proposal was developed recently through consultation with representatives of many federal government departments. The proposal is based on matching the speech privacy criterion to the sensitivity of the information to be discussed in the room. A draft recommendation is listed in the fourth column of Table 1. It indicates that the minimum SPC requirement for a room in which *Protected B* information is to be discussed is 75. For *Secret* information, the SPC should be at least 80, and for *Top Secret*, 85. Obviously users outside of the federal government will have different terminology and possible needs for fewer (or more) categories.

5. PRACTICALITY

In specifying criteria, it is important to be able to estimate whether it is realistic to achieve various SPC values. Measurements conducted in some new government offices indicate ranges of SPC from mid-40s to 90 or so are achievable [4].

6. DISCUSSION

Measurement of speech privacy (or speech security) of enclosed rooms is possible using the new ASTM E2638 standard. Users can use the discussion in the standard to help set criteria and to define useful categories. One proposal is to match the necessary value of SPC to the sensitivity (importance) of the information to be discussed in the room. Measurements in real rooms have demonstrated the feasibility of such an approach. What is still lacking, however, is some real-world validation that occupants' satisfaction with the rooms matches the physical measurements.

REFERENCES

[1] ASTM E2638-10, "Standard test method for objective measurement of the speech privacy provided by a closed room," ASTM International, West Conshohocken, PA (2010).

[2] B.N. Gover and J.S. Bradley, "Measures for assessing architectural speech security (privacy) of closed offices and meeting rooms," *J. Acoust. Soc. Am.* **116**, 3480–3490 (2004).

[3] J.S. Bradley and B.N. Gover, "Speech levels in meeting rooms and the probability of speech privacy problems," *J. Acoust. Soc. Am.* **127**, 815–822 (2010).

[4] C. Fortier, "Speech Privacy Class measurements in Real Office Buildings according to the new ASTM 2638 standard and related STC values," Proceedings of INTER-NOISE 2009.

ACKNOWLEDGEMENTS

This project jointly funded by Public Works and Government Services Canada (PWGSC), the Royal Canadian Mounted Police (RCMP), and the National Research Council Canada (NRC).

CASE STUDY ABOUT SPEECH PRIVACY OF INTEGRATED FURNITURE IN AN OPEN-PLAN OFFICE

Jean-Philippe Migneron[1], and Jean-Gabriel Migneron[2]
[1]Groupe de recherche en ambiances physiques, École d'architecture, Université Laval,
[2]côte de la Fabrique, Québec (Qc), Canada G1K 7P4, jean-philippe.migneron.1@ulaval.ca

1. INTRODUCTION

In the case of a retrofitting project for large open-plan offices, it was appropriate to test a prototype of integrated furniture before completing specifications for different floors. As acoustical environment was initially good in the building, it was requested to assure that the new ceiling and the new system of office furniture could offer similar performances regarding speech privacy or even better results (especially for a call center). Four modular workstations have been placed in a large room to accomplish various acoustic measurements. Speech privacy has been compared between 17 configurations of the same prototype, which included the type of acoustical tiles for the ceiling, the material and the height of screens, or by adding acoustic baffles, localized absorption, and sliding doors.

2. TESTING METHOD

The experimental procedure was based on the evaluation of speech privacy, as it stays one of the most important criteria of acoustical comfort for occupants of the building. Even if that issue is well documented, the reach of good acoustical performance in a working environment always requires some efforts. This project aimed to help owners to take decisions about renewing completely interior arrangements. Consequently, the procedure had to be relatively simple and fast to run on site, as the specification process should not be delayed by the acoustical expertise.

2.1 Studied parameters

Few parameters can quantify speech privacy considering the level of transmission from a sound source. These include the Articulation Index (AI), the Speech Intelligibility Index (SII) and the Speech Transmission Index (STI). On a 0 to 1 scale, those values must be minimized to reduce the understanding of messages, resulting in a better privacy. SII of 0.2 under is generally considered good for open-plan spaces.

2.2 Sample of office arrangement

In order to test general ergonomics and performances of proposed integrated furniture, a prototype of 4 workstations has been built. As figure 1 shows, three workstations followed the same arrangement, but unit n°3 differed because of its counter for service.

2.3 Experimental procedure

The National Research Council of Canada developed a testing software especially for diagnose problems in open-plan offices [2]. SPMSoft uses impulse responses from a sweep signal and background noise measurement to make a calculation of privacy parameters. Up to 12 propagation paths were tested using a loudspeaker and a microphone located approximately where occupants should sit (1.15 m from floor).

2.4 Studied adjustments of the prototype

The conception team established some variables before considering acoustics. For example, each workstation should size 2.14 by 2.60 m, when the height of partitions could be chosen between 1.68 and 2.06 m. The pattern of the suspended ceiling had to follow windows' position, but acoustic tiles could be changed. It was also deter-mined that integrated furniture would be chosen for it flexibility and durability. However, finishes could be adjusted, knowing that absorbing panels could replace glazed parts. In addition to those variables, other acoustical treatments have been tested. That includes the addition of sliding doors, localized absorption or acoustical baffles under the ceiling.

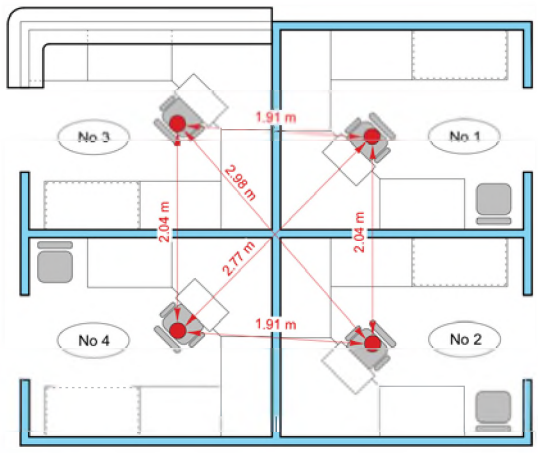


Fig. 1. Layout of integrated furniture constituting the sample of four workstations, with direct propagation paths.

Table 1. Summary of averaged results for speech privacy improvements compared to the reference case, with multiple adjustments of the same prototype.

No	Description	Ceiling tiles	Type of panels	Height of screens (m)	SII	AI	STI
1	**Reference case**	**product A[1]**	**Glazing at 8%**	**1.68**	**0.48**	**0.41**	**0.45**
2	Without computer screens	product A[1]	Glazing at 8%	1.68	-0.06	-0.08	-0.04
3	Add of localized absorbing material	product A[1]	Absorbing at 100%	1.68	-0.02	-0.02	+0.01
4	Add of absorbing baffles under ceiling	product A[1]	Absorbing at 100%	1.68	-0.06	-0.05	-0.05
5	Raise of division's height	product A[1]	Glazing at 26%	2.06	+0.04	+0.02	+0.05
6	Change between glazing and absorption	product A[1]	Absorbing at 100%	2.06	-0.06	-0.06	-0.05
7	Add of absorbing baffles under ceiling	product A[1]	Absorbing at 100%	2.06	-0.12	-0.13	-0.11
8	Add of sliding doors	product A[1]	Glazing at 26%	2.06	-0.12 [3]	-0.12 [3]	-0.11 [3]
9	Add of doors, absorption, and baffles	product A[1]	Absorbing at 100%	2.06	-0.19 [3]	-0.19 [3]	-0.20 [3]
10	Change of ceiling tiles	product B[2]	Glazing at 8%	1.68	-0.15	-0.15	-0.14
11	Change for thumbtack panels	product B[2]	Absorbing at 100%	1.68	-0.11	-0.11	-0.10
12	Change for highly absorbing panels	product B[2]	High absorbing on 32%	1.68	-0.15	-0.15	-0.14
13	Raise of division's height	product B[2]	Glazing at 26%	2.06	-0.14	-0.15	-0.14
14	Add of structure's hider	product B[2]	Glazing at 26%	2.06	-0.14	-0.14	-0.12
15	Change between glazing and absorption	product B[2]	Absorbing at 100%	2.06	-0.20	-0.20	-0.18
16	Change for thumbtack panels	product B[2]	Absorbing at 100%	2.06	-0.17	-0.17	-0.15
17	Change for highly absorbing panels	product B[2]	High absorbing on 37%	2.06	-0.21	-0.22	-0.19

Notes : 1 -High density fiberglass tiles (rated NRC-0.90) 2 - High density fiberglass tiles (rated NRC-0.85) 3 - Comparative measurement with or without the sliding door

3. RESULTS

Multiple experiments has been achieved with the prototype of office furniture. The table 1 summarizes most interesting results considering averages of each parameter for tested propagation paths. To show privacy improvement, values are compared with the first case, which is the reference proposed by project's designers. With a SII of 0.48, corrections had to be made to respect comfort specifications.

The first noticeable observation is the general effect of ceiling tiles. Even if the NRC value published by the manufacturer is lower for the second product, the perceived absorption seemed higher, as the average privacy was increased of almost 30%. Another important aspect is the height of each screen, which is normally beneficial to performances when raised. However, this assumption depends highly on the choice of material used in their construction. With the first tested case of high divisions, case n°5, the addition of modular panels at the top of partitions surprisingly reduces privacy between workstations. The choice of glazed parts, to allow natural light diffusion in the building, explains this situation, because divisions become partially sound reflectors. The substitution of those panels for standard absorbing surfaces can reduce intelligibility, but the better results are obtained with high absorption materials that are thicker.

Those results concern directly the system of integrated furniture. Other modifications represent optional elements that might be useful for particular areas with high level of privacy needed, in a call center for example. One good way to cut sound propagation between workstations is to add sliding door to screens, so each space could be closed according to occupant's wish (cases n°8 and 9). If the height of division cannot be raised because of air distribution or sprinklers' operation, then acoustic baffles become an interesting solution. Absorbing panels of only 0.20 m high suspended beneath ceiling system are enough to make a substantial difference. Finally, experiment n°2 demonstrates the significant role of computer screens in the propagation of sound as for direct reflection of user's voice.

4. DISCUSSION

As expected, results show that the choice of materials for the ceiling and divisions has the greater impact on speech privacy. The height of screens is also an important variable, but the addition of absorbing baffles can reduce that need.

The intent of this case study was to quantify acoustical performances of office furniture. Several tests made on a demonstrator of widespread commercial products has been profitable to justify owners' choices before implementation on a large scale. This practical example was constrained to comparison of most feasible modifications to interior arrangements, and few aspects were fixed in advance, as the orientation of workstations. However, it reveals the importance of collaboration between conceivers, designers and architects to assure the final efficiency of the environment procured in an open-plan office.

REFERENCES

[1] Asselineau, M. et al. (2009) *Looking at speech privacy in buildings : a few considerations*, Inter-Noise 09 Proceedings, 6p.

[2] Bradley, J.S., Gover, B.N. (2008) *Development and Evaluation of Speech Privacy Measurement Software: SPMSoft*, IRC Research Report, IRC RR-262, 45p.

[3] Bradley, J.S., Gover, B.N. (2008) *Open-Plan Office Speech Privacy Case Studies*, IRC Research Report, IRC RR-262, 35p.

REVERBERATION MEASUREMENT AND PREDICTION IN GYMNASIA WITH NON-UNIFORMLY DISTRIBUTED ABSORPTION; THE IMPORTANCE OF DIFFUSION

Kevin Packer and Clifford Faszer
FFA Consultants in Acoustics and Noise Control Ltd., 304, 605 1st Street SW, Calgary, Alberta, Canada, T2P 3S9,
info@ffaacoustics.com

1. INTRODUCTION

School gymnasia present several acoustical challenges as the rooms must support a variety of uses, mainly athletic functions, school/community gatherings and performances. Considerations such as user safety, surface durability and impact resistance, ease of maintenance and clean-ability often dictate not only a lack of acoustically absorptive finishes in the occupied portion of the room but also parallel and acoustically reflective floor and lower wall surfaces. In Alberta current government design standards [1] stipulate that reverberation times (RT) in a typical unoccupied gym not exceed RT 2.0 seconds averaged over the frequency range of 500 to 2000 Hz. Design guidelines for acceptable gymnasia finishes are also given in [1].

This paper documents the results of RT measurements conducted in several newly constructed gymnasia as part of a performance verification exercise for the builder. The gymnasia were constructed according to the design guidelines however the initial RT measurements did not meet the performance requirements.

2. ROOM DESCRIPTION

The measured rooms were rectangular in plan with two variations in room size: Type A, 27.8 m x 18.5 m, slightly sloped ceilings 9.3 m to 9.6 m above finished floor (AFF); Type B 24.0 m x 18.0 m, ceilings 9.1 to 9.5 m AFF. The finishes are painted concrete block walls to 3 m above a cushioned wood floor and painted 2-layer 16 mm gwb walls to the underside of a painted acoustic metal deck ceiling. Initially, two 1.2 m high continuous bands of exposed unpainted Tectum/mineral fibre paneling (38 mm mineral fibre behind 25 mm Tectum, edges concealed) extended around the full perimeter of the rooms on the upper walls.

3. EQUIPMENT & PROCEEDURE

A tripod-mounted Brüel & Kjaer 2270 Precision Real Time Sound Level Analyzer equipped with the latest version (3.2) of the BZ7227 Reverberation Time software was used to record archive and evaluate the RT measurements. During the initial measurements sound decays were generated either from amplified interrupted pink noise or by large diameter balloon burst impulses. For later measurements only impulses were used. The decays were measured at at least 5 locations in the rooms with the positions consistent between repeated measurement sessions.

4. INITIAL RT MEASUREMENTS

After the initial RT measurements did not meet the performance requirements a lack of adequate absorption was presumed and the gymnasia were modelled using ODEON room acoustics prediction software to determine the quantity and placement of additional absorption required to bring the RT into compliance. ODEON is based on prediction algorithms (image-source method and ray-tracing) that use scattering due to surface roughness and diffraction. A reflection-based scattering method is used that accounts for frequency-dependent scattering. Like any prediction software, ODEON's accuracy is limited by the validity of the input data, namely the room geometry, surface sound absorption and scattering coefficients. However the geometry for these rooms is not complex and with the exception of the acoustic deck and Tectum panels the absorption coefficients for the various room materials are fairly well established in literature. Both acoustic deck and Tectum panels have been used in enough projects to establish that they provide at least some absorption in the critical mid frequency bands. Scattering coefficients were chosen according to ODEON guidelines [2].

After installation of a third continuous band of panels at a height below the existing panels, the RT were re-measured and found to be significantly higher, particularly in the critical mid-frequencies, contrary to intuition and the predictions. The measurement results seemed to indicate a greater difference from the predicted RT than could be explained by invalid absorption and/or scattering coefficients. Subsequent adjustments to these model inputs confirmed this. The acoustic deck and wall panels in one of the gymanasia were inspected and no obvious problems found. Similar inspections in all of the other gymnasia by the building contractor found no obvious faults.

5. DIFFUSION AND REVERBERATION

Much past work has been done by others exploring the effects of diffusion on reverberation. The requirement for a sufficiently diffuse sound field is established for laboratory measurements in ASTM C423 - 09a [3]. This is typically achieved with fixed and/or rotating sound-reflective panels hung or distributed with random orientations about the volume of the reverberation room. ASTM C423 states that it has been found that in rectangular rooms the area (both sides) of diffusers required to achieve satisfactory diffusion is 15 to 25% of the total surface area of the room.

All of the gymnasia except for the two following cases were measured completely empty except for the measurement equipment and operator: For one Type B gym measurement session a scissor lift was located at one end of the room and a 1.2 m by 2.4 m Tectum board was leaning against a wall. During another measurement session in a different Type B gym with slightly more wall panelling a few boxes of construction materials were present on the floor. In both cases, these objects were judged at the time not to be large enough in area or volume to make a significant difference in the RT. However, the diffusion that they may have provided was not considered. In both cases lower RT were measured with the most dramatic difference in the later case: a mid-frequency average RT of 2.1 seconds, reasonably close to the ODEON prediction of 2.1 seconds.

The third band of wall panels appeared to be having some effect in the Type B gymnasia measured with the additional objects but not in the other (empty) gyms. Further investigation finally lead to the hypothesis of horizontal reflections between the untreated lower walls and a lack of sound diffusion in the lower portion of the rooms contributing to the unexpected results. It was suggested that providing some diffusive objects to break up these reflections might provide results closer to a minimally occupied condition and to the predictions. This hypothesis was tested and the RT re-measured in a Type B gym with some plywood panels and also with a few people.

Five different conditions were measured; an empty gym, addition of people, three levels of diffusion. Diffusion was varied with plywood sheets, 1.2 m x 2.4 m x 12.7 mm thick, stood on end at various locations throughout the gym. Ten of these plywood sheets were fastened together at one end to form five self-supporting A-frame units. The remaining five plywood sheets were leaned against the volleyball net and supporting end poles at the mid point of the gym. Sheets were removed and the measurements repeated. The measurements were also repeated with the room empty and again with the equipment operator plus four other adults.

6. RESULTS AND DISCUSSION

The results for the re-tested Type B gym with and without the third band of wall panels and with and without the additional sound diffusing elements are presented in Figure 1. All plotted measurements were conducted with three bands of wall panels, except as noted, and with the room empty except for the noted fittings or occupants plus the measurement equipment and operator. For the RT with plywood sheets two stepladders were also present.

The addition of as few as four people or five plywood sheets was found to significantly reduce the measured RT, closer to the modeled predictions. This decrease in the measured reverberation times is more than can be accounted for by the sound absorption provided by four additional adult bodies alone. The low frequency RT were not particularly sensitive to the addition of the plywood however the times in the 500 to 4000 Hz bands were significantly reduced.

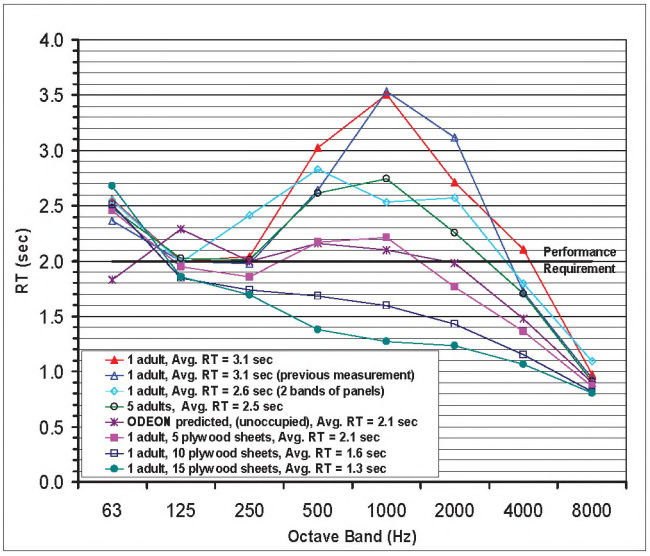


Figure. 1. Measured Reverberation Times RT60 (sec) in a Type B Gymnasium

Similar measurements were repeated by others in both Type B and Type A gymnasia built under the same contract. Their findings were similar with regards to the effect of diffusive elements on the measured RT. The importance of plywood placement was not extensively evaluated but some variations were deliberately introduced to help evaluate any effect this may have. Generally it appeared that the RT were not particularly sensitive to the location of the plywood. The physical variations between the two types of gymnasia did not result in any major differences in RT.

It has been suggested that the plywood panels change the angle of incidence of the sound waves in the lower part of the room and thus of the reflected sound incident on the acoustically absorptive wall panels and acoustic deck and thus may result in more effective absorption by the acoustic treatments. This agrees with the assertion in ASTM C423 that diffusion provides "a rapid and continuous interchange of energy between the directions of sound propagation, thereby increasing the probability that each surface area of the room is exposed to sound of the same intensity." The increase in measured RT with the addition of absorption in the empty rooms requires further study.

REFERENCES

[1] Alberta Infrastructure and Transportation, (2007). Standards and Guidelines for School Facilities. 22-24

[2] Christensen, C. L., (2009) Odeon Room Acoustics Program Version 10 Industrial, Auditorium and Combined Editions (user manual). Odeon A/S.

[3] ASTM C423 – 09a (2009). Standard Test Method for Sound Absorption and Sound Absorption Coefficients by the Reverberation Room Method. ASTM International

DESIGN AND TESTING OF AN ANTENNA ARRAY FOR SOUND-SOURCE LOCALIZATION

Murray Hodgson[1], and Vincent Valeau[1,2]
[1]Acoustics & Noise Research Group, University of British Columbia, 3rd Floor, 2006 East Mall, Vancouver, BC V6T1Z3
[2]Pprime Institute, University of Poitiers-Ensma-CNRS, UPR3346 ESIP Bat. K Campus Sud, 40 avenue du recteur Pineau, F86022 POITIERS Cedex, France

1. INTRODUCTION

This paper describes the adaptation of an existing 'eye-array' antenna [1], based on a time-delay-of-arrival algorithm, to use a beam-forming algorithm, for application to the localization of noise sources in industrial workshops and in turbulent airflows (e.g. jets), and its preliminary testing. Full details can be foound in [2].

2. ANTENNA CONFIGURATION

The antenna consists of a hemispherical array of 26 electret microphones equally spaced over a 34-cm-diameter hemisphere, with a 27th microphone at its centre, and an associated 27-channel data-acquisition system (Figure 1).

Figure 1. Antenna and DAQ system in anechoic chamber.

3. SIGNAL PROCESSING

The original antenna achieved sound-source localization based on a time-delay-of arrival (TDOA) algorithm [3]. This was tested in an anechoic chamber, and found to perform well with a single source; however, the extension of the algorithm to multiple source was considered too complicated.

Thus, a new localization algorithm was developed based on delay-and-sum beam-forming methods [3], and its performance tested.

4. EXPERIMENTATION

The antenna array, hardware and new beam-forming algorithm were tested in an anechoic chamber (see Figure 1), with one or more real sources (omnidirectional points, diameters <5 cm, radiating broadband noise) located 1-2 m from the antenna, and virtual 'image' sources created by introducing reflecting surfaces, as well as in a reverberant room (a small, empty office with dimensions 5.6 m x 4.2 m x 2.7-m high and acoustically-hard surfaces) with a single real source at various distances from the array.

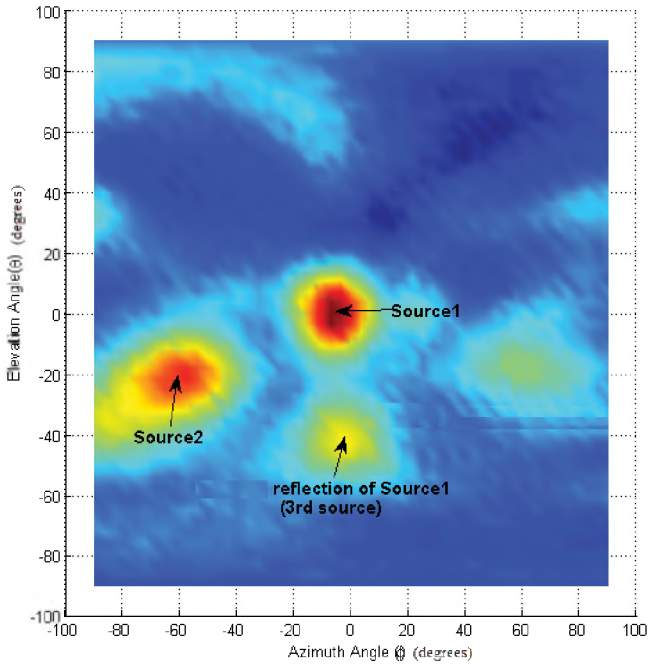


Figure 2. Directivity diagram for beamforming with two real point sources and a virtual source (reflecting surface).

5. RESULTS

Figures 2 to 4 show test results in terms of 'directivity diagrams' generated by the system. These show the distribution of sound intensity incident on the antenna over the hemisphere towards which it is pointing.

Figure 2 shows the directivity diagram for delay-and-sum beam-forming with two real sources and a virtual source (reflecting surface). The directions of the three sources are accurately localized.

Figure 3 shows the directivity diagram of the delay-and-sum beam-former with a source at a distance of 4.2 m (i.e., in the reverberant field) in the reverberant room. The directions of the real source, and of the virtual sources corresponding to first-order images in the reflecting surfaces, are well identified. Note the wrapping of the side-wall reflection

Figure 4 shows the directivity diagram of the delay-and-sum beam-former with the source in the reverberant room, and one wall covered by absorbing material. Energy arriving from the direction of the image source in the treated surface, which is now much lower, is no longer evident.

6. CONCLUSION

An existing microphone-array antenna has been adapted for sound-source localization, with new hardware, and software based on a beam-forming algorithm. The array has been tested in an anechoic chamber and in a reverberant room with one or multiple real and virtual sources. It performed well and was able to identify all sources; it displayed an angular resolution of 3-4°

Future work will improve the resolution by implementing a CLEAN or similar algorithm [3], and will adapt the system to localizing 3-D source positions. It will then be used to localize noise sources in industrial workshops and in turbulent airflows, as in jets.

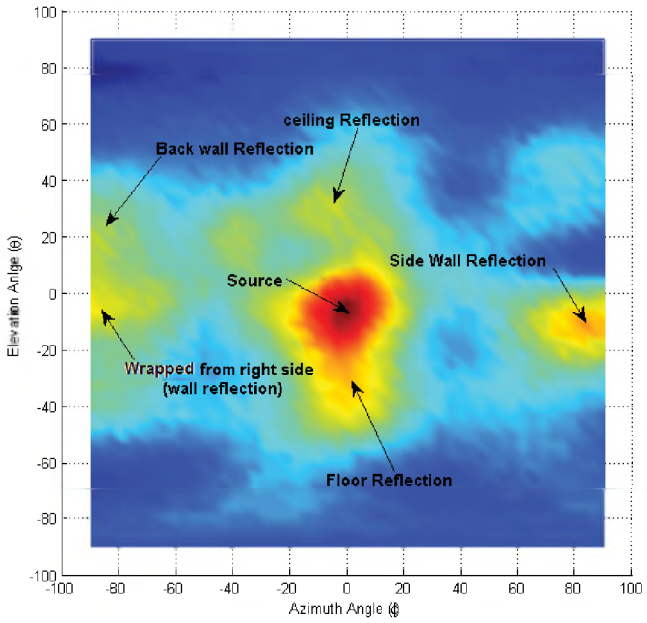


Figure 3. The directivity diagram of the beam-former with a point source in a reverberant room.

REFERENCES

1. H. Alghassi, "Eye Array Sound Source Localization", Ph.D. Dissertation, Department of Electrical and Computer Engineering, University of British Columbia (2008).
2. M. Nematifar, "Localization of Sound Sources Using a Microphone Array", Electrical and Computer Engineering co-op project report (2009).
3. D. H. Johnson and D. E. Dudgeon, "Array Signal Processing: Concepts and Techniques" (Prentice Hall, Englewood Cliffs, NJ, 1993).

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the donation of the antenna by Prof. Peter Lawrence, and the technical contributions of Chris Bibby, Glenn Jolly and Mona Nematifar to the successful completion of the work.

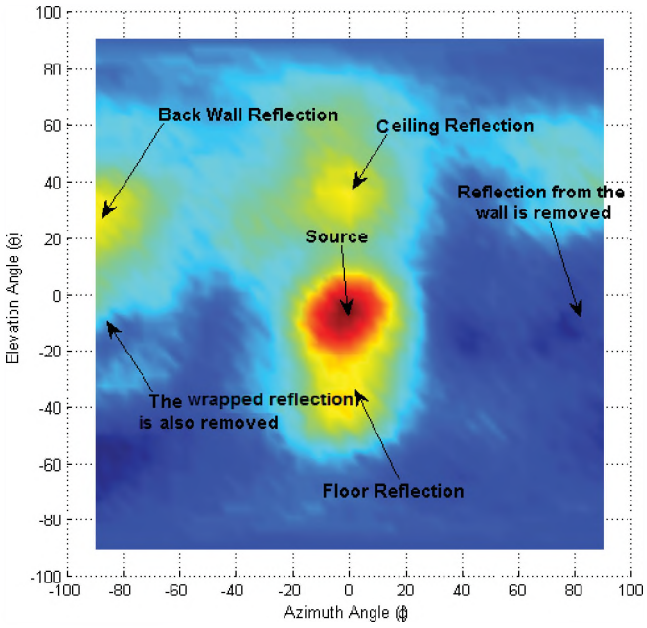


Figure 4. The directivity diagram of the beam-former with a point source in a reverberant room with one wall covered by absorbing material.

ACOUSTICAL CHALLENGES FOR A HOSPITAL CHILLER ROOM ADDITION

Zohreh Razavi
Acoustical Consultant, Stantec Consulting Ltd., zohreh.razavi@stantec.com

Providing satisfactory acoustical environments in healthcare facilities can be ensured by applying recommended minimum design requirements provided in Sound and Vibration Design Guidelines for Hospital and Healthcare Facilities. The objective of this project was to provide acoustical conditions in compliance with the Healthcare design guidelines when a giant chiller and six pumps were installed in a mechanical room, below an auditorium within a hospital. The job involved conducting measurements in a mechanical room with the same type of equipment and within an auditorium; reviews were made to the floor/ceiling assembly and all penetrations due to these additions. The result was a satisfactory acoustical condition for the occupants after overcoming all challenges.

1. INTRODUCTION

In this paper, the addition of a mechanical room and an electrical room to a hospital in the space below an auditorium discussed. The mechanical room included a Chiller and six pumps and the electrical room contained a transformer. These rooms were considered to be added to a space below an auditorium which was designed and added to the building recently. Initially, the addition of a Cooling Tower to the roof of another building close to the hospital was proposed; however, due to its exceeding noise, it was relocated to the roof of the same building as the Chiller room. The steps taken were included the environmental acoustical assessment to the post-construction evaluation measurements.

2. PROJECT STRATEGY

The goal of this project was to install the proposed mechanical equipment without exceeding the background noise level within the auditorium and in compliance with local city noise By-Laws. A review of the city noise by-law was conducted to avoid any noise violations by cooling tower additions to the roof. Background noise level measurements within the auditorium and the chiller room were conducted to evaluate the acoustical environments within the spaces before any additions, as benchmarks. All calculations and recommendations provided were based on the sound power levels of the equipment provided by the suppliers.

3. MECHANICAL EQUIPMENT

The selected Cooling Tower was Baltimore Aircoil Cooling Tower, Model 3618 C. The total noise level of the unit at 1.5 meter distance from the unit at the fan discharge, case side, and air inlet was equal to 86 dBA, 74 dBA and 80 dBA, respectively. The noise levels at the property line of the neighboring buildings were calculated to be 75 dBA and 66 dBA. The selected Chiller was a CenTraVac centrifugal CVHF 1720 with 1800 tons cooling capacity from Trane supplier. The noise levels at one meter distance from the unit case varied from 80 dBA to 82 dBA at %25, %50 and %100 loads. The two primary and two secondary chilled water pumps along with two condensing pumps were utilized in the Chiller room. The total noise level when the primary chilled pumps and the condensing pumps were running at %100 was calculated to be equal to 85 dBA at one meter distance from the unit case.

4. AUDITORIUM ACOUSTICAL PERFORMANCE

A newly renovated auditorium was located in the space above the Chiller room. The auditorium window was above the Chiller room windows, while the required air to the auditorium was through a space underneath the seats and above the ceiling in the Chiller room. The Noise Criteria – NC within the space was calculated from measured octave band frequencies sound pressure levels to be equal to 22.

5. ACOUSTICAL CHALLENGES

Installation of a Cooling Tower on top of a two storey building due to exceeding the noise levels permitted by the City Noise By-Law at the neighboring properties and all expenses for ducting has not been approved. Even by silencer additions to the air intake and fan discharge, along with quieter fans the noise level at the receivers would exceed the City Noise by-law. The enclosure wasn't an option for the unit due to the limited space on the roof and its high cost. All these limitations lead to relocation of the Cooling Tower.

6. RECOMMENDATIONS AND DISCUSSIONS

The windows in the Chiller room were replaced with an acoustically designed infill and externally louvers with the actual air intakes and exhausts. The ceiling of the chiller room were upgraded to provide required transmission loss by addition of two layer of Gypsum wall boards at 300mm distance from the ceilings to be hung with resilient hangers. The plenum was treated acoustically with acoustical spray and the chiller room was acoustically treated with addition of sound absorptive materials. Vibration isolation provided for the Cooling tower at its new location, the Chiller and pumps to avoid vibration induced noise. The noise of the transformer compared to the

chiller was not an issue. However, since the transformer was in a hospital critical situation, a higher level of vibration isolation such as seismically restrained air mounts was recommended.

7. POST CONSTRUCTION EVALUATIONS

In order to evaluate the effect of the Chiller room addition to the building, measured octave band noise levels before and after construction were compared with each other. The NC level in the auditorium before the construction was NC 22. The pumps additions did not exceed this NC level at all. The chiller additions increased the NC level by 3 and still didn't exceed the standard of NC25-30 for auditoriums.

8. CONCLUSIONS

The addition of a mechanical room and an electrical room to a hospital in the space below an auditorium resulted in retaining Stantec Consulting Ltd. for acoustical services. The applicable city of Vancouver noise by-law, along with ASHRAE standards and Sound & Vibration Design Guidelines for Health Care Facilities, were guidelines for this project. The noise attenuation measures on the walls, ceilings and all penetrations to the mechanical and electrical room were provided. The vibration isolators were recommended based on hospital critical conditions. No appreciable changes in the noise data within the auditorium were investigated through a series of post-construction noise measurements.

ACKNOWLEDGEMENTS

I would like to acknowledge all the efforts of John Maccarrone, who initiated this project and facilitated all communications.

EVALUATION OF THE NOISE-MASKING SYSTEM IN A COMMUNITY HEALTH-CARE FACILITY

Ahmed Summan, Musarrat Nahid and Murray Hodgson
School of Environmental Health, University of British Columbia, 2206 East Mall, Vancouver, BC, Canada V6T1Z3

1. INTRODUCTION

In open-plan offices speech is often the most distracting cause of noise even though a main aim of open-office design is to attenuate speech propagating between workstations so that it does not disturb the workers' concentration (Larm *et al.*, 2005). In order to achieve adequate speech privacy in open offices, appropriately designed noise-masking systems are often used (Hongisto, 2008). Noise-masking systems consist of loudspeakers located behind the suspended acoustical ceiling distributed throughout the office area; they generate a background noise over a certain area, to mask the unwanted sounds. That should result in increased speech privacy, eliminating the imposition of unwanted sound, and enhancing the general level of acoustical comfort and productivity in the space.

This pilot study was done to evaluate the advantages and the disadvantages of the noise-masking system recently installed on one of the two floors in the Vancouver Coastal Health Community Health-Care facility (CHC). In order to achieve this general goal it was important to determine the effects of this noise-masking system on background-noise levels, speech privacy, and on the workers' performance.

2. LITERATURE REVIEW

2.1 Effects of office noise

Impaired concentration or being distracted is the first major effect of speech noise in offices. Colle (1980) clarified that speech intelligibility can cause disturbance or impaired concentration. Sundstrom (1994) found that of over 2000 participants (office workers) questioned, 54% said they were often bothered by people talking and telephone ringing. Sabine and Jurgen (2009) stated that office noise can reduce verbal short-term memory and make memory retention more difficult.

2.2 Noise-masking-system effects in open offices

Installing a proper noise-masking system has been considered as a cost-effective method to achieve the above objective (Mohammad *et al.*, 2000). Lewis (2003) studied how a noise-masking system affected 136 office workers. He found that worker performance was increased and speech distraction was diminished. He found that the perceived acoustical conditions were enhanced by noise control. Hongisto (2008) investigated the effects of a noise-masking system on workers in a small department by measuring the speech privacy, and questioning the workers before and after installation of the system. The results showed that after installing the noise-masking system, speech privacy was significantly improved, speech distraction reduced, and noise-related stress was eliminated.

3. METHOD

This study was conducted at the North Shore Community Health-Care (CHC) facility of Vancouver Coastal Health on the 5th and the 6th floors of a building in North Vancouver. A noise-masking system was installed on the 5th floor only.

Acoustical measurements and calculations were carried out to estimate the background-noise levels and speech privacy for a number of workstations on both floors at the CHC. Moreover, Speech Intelligibility Index (SII) was calculated from measured noise and speech levels and reverberation times.

A questionnaire was developed especially for this study, to investigate the effects of the installed noise-masking system on the workers' performance, comfort and satisfaction with the work environment. A total of thirty-one employees from both floors, out of one-hundred and seventy (18%), returned completed questionnaires.

Different statistical analyses were used to find the relationships between the physical measurements and the questionnaire responses for the two floors. These statistical methods included Spearman correlation, T-test, Mann-Whitney test and single- and multivariable linear regression.

4. RESULTS and DISCUSSION

Total background-noise levels for both floors under occupied conditions were higher than those under unoccupied conditions by about 9 dBA. On the fifth floor, the total A-weighted background-noise levels (unoccupied condition) were 40 and 44 dBA when the noise-masking system was off and on, respectively. On the sixth floor, the total A-weighted background-noise level was 49 dBA under occupied conditions, which is higher than the A-weighted background-noise level under unoccupied conditions. It was also found that the noise-masking system alone generates a sound level of about 41.8 dBA (see Figures 1 and 2).

Speech at short distances (1 and 5 m) was more intelligible than at larger distances on the fifth floor. The noise-masking system reduced the values of SII at all distances. However, speech is always intelligible at most distances of

communication when it is occupied on the sixth floor. Statistical results showed that on the fifth floor the relation-

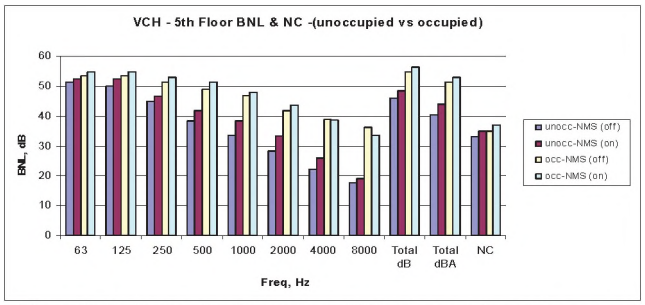


Figure 1: Background-noise levels and NC values on the 5th floor.

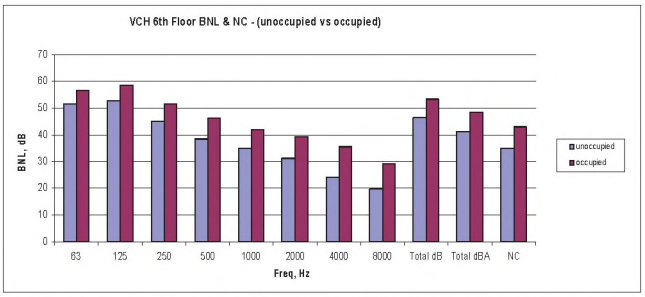


Figure 2: Background-noise levels and NC values on the 6th floor.

ship is highly significant ($p \leq 0.01$) and negative between the measured background-noise levels when the noise-masking system was on (unoccupied) and the difficulty of having confidential conversations in workstations. There is also a highly significant ($p \leq 0.01$) and negative correlation between the background-noise levels when the noise-masking system was on (occupied) and Speech Intelligibility Index values at 1, 5 and 10 m when noise-masking system was on (occupied). Speech Intelligibility Index values at 1, 5 and 10 m were statistically different between the two floors. According to the Speech Intelligibility Index results on the fifth floor, there was a significant ($p \leq 0.05$) difference between SII with the noise-masking system on and SII with noise-masking system off (calculated using casual and normal voice levels at 1, 5 and 10 m) in the unoccupied condition; however, SII at 20 m (either with normal or casual voice) didn't differ significantly. In other words, the noise-masking system seems to be effective when the workplace is unoccupied at 1, 5 and 10 m but, at much higher distances, the effect of the noise-masking system seems to decrease.

Statistical test results also showed that the employees from the two floors showed similar numbers of complaints about speech privacy in the presence or absence of the noise-masking system.

Questionnaire results showed that more than 60% of the participants on both floors were dissatisfied with the acoustical environment. Moreover, employees on the fifth floor are more bothered than those on the sixth floor by speech from other rooms. Employees are more distracted and stressed on the fifth floor than those on the sixth floor.

5. CONCLUSION

The final results showed that the background-noise levels and the noise-criterion values on the fifth floor were higher when the noise-masking system was on (occupied or unoccupied) than when it was off. Moreover, it was found that the noise-masking system provided acceptable speech privacy at short distances of 1, 5 and 10 m between workstations. The system noise levels also provided confidential speech in some workstations. However, the noise-masking system seems to be ineffective at longer distances. Due to the small sample size, statistical-analysis results (Spearman and T-test) didn't indicate significant differences between the two floors regarding the dissatisfaction with the overall environmental quality, the acoustical conditions and the speech privacy. Finally, the regression models show that it is very important to consider the acoustical conditions to achieve overall satisfaction with the workplace environmental quality in a health-care facility. In addition, intermittent noises can also affect work performance directly.

REFERENCES

Colle, H. A. (1980). Auditory encoding in visual short-term recall: effects of noise intensity and spatial location. Journal of Verbal Learning and Verbal Behavior 19, 722-735.

Hongisto, V. (2008). Effect of sound masking on workers in an open office. Proc. Acoustics'08, Paris, 537-542.

Larm, P., Keranen, J., Helenius, R., Hakala, J., & Hongisto, V. (2005). Acoustics in open-plan offices - A laboratory study. Forum Acusticum (pp. 14-18). Turku, Finland: Finish Institute of Occupational Health, Lemminkaisenkatu.

Lewis, E., Sykes, D., & Lemieux, P. (2003). Using a web-based test to measure the impact of noise on knowledge workers' productivity. 47th Annual Meeting of Human Factors and Ergonomics Society. Denver.

Mohammad, A., Hassanain, M. A., & Harkness, E. L. (2000). Noise control and speech privacy guidelines for office building design. Journal of Architectural Engineering. 6, 52-57.Sabine, J. S., & Jurgen, H. (2009). Background music as noise abatement in open-plan offices: a laboratory study on performance effects and subjective preferences. Applied Cognitive Psychology 23, 684-697.

Sundstrom, E., Town, J. P., Rice, R. W., Osborn, D. P., & Brill, M. (1994). Office noise, satisfaction and performance. Environ. Behav. 26, 195-222.

ACKNOWLEDGEMENTS

It was a real pleasure to get support and companionship from the UBC School of Environmental Health project supervisory committee, Vancouver Coastal Health, and the Acoustics & Noise Research Group.

LOCATION OF HORN SPEAKERS IN A REVERBERATION ROOM

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

1. INTRODUCTION

Considerable theoretical research has been conducted in understanding the design constraints as well as on the locations of horn speakers in reverberation rooms[1]. Many of the research methods applied simple sinusoidal source functions (tones) to evaluate the design criteria of horn speakers[1,2]. The understanding of the behaviour of horn speakers, when band limited random noise signatures such as pink noise and white noise are used as input sources, is still not clear. A hyperbolic horn with cut off frequency of 70 Hz was used in a medium sized reverberation room to study the horn behaviour. Some of the basic questions to be studied were the influence of horn location on the cut-off frequency, as well as the influence of the horn location on the diffuse sound field in the reverberation room. In addition, the influence of the input sound source on the room sound levels was also studied. Preliminary results of the experiment are presented in this paper.

2. BACKGROUND

The location of a source in a room is very much dependant on the expected sound field. Typical effects of locating the source in corners were highlighted in Bell[2] and Beranek[4]. The diffused sound in a reverberation room is supposed to reduce by the factors based on source's location in the room. Standard textbooks show that the 'Q' factors are 1, 2, 4, and 8 for the centre, single corner, double corner, and triple corner location of the source respectively. In addition, if horn (exponential, conical or hyperbolic) speakers were used in a reverberation room as the main source, the operating frequencies can be modified based on the location of the horn. It is hypothesized for example, that if the horn is located in a triple corner, the cut off frequency can be reduced or the mouth size can be reduced. The current experiment was undertaken to test the above hypothesis.

3. THE REVERBERATION CHAMBER

The reverberation chamber at Concordia University was used to conduct the experiment.

3.1 Chamber details

The reverberation chamber is located in the engineering building of Concordia University, Montreal and is used by the Building, Civil and Environmental Engineering Department (BCEE). The characteristics of the chamber are: Length, L = 6.13 m; Width = 6.96 m; Height = 3.56 m; Chamber Volume = 152.3 cu.m.

The RT_{60} varied between 0.8 sec to 3 sec. across the frequency band. The two cut-off frequencies of the room were estimated to be 188 Hz and 64 Hz[5,6,7]. The results of Reference 5 showed that for a broadband signal, the chamber had good spatial uniformity from 63 Hz (1/3 octave band) and above.

3.2 Horn Speaker details

A hyperbolic horn speaker was used for the tests. The horn in a triple corner is shown in Figure 1. The horn details are: the horn length is 92.1"; the throat area is 2.1 sq. in; the mouth area 397.5 sq. in.; band width is from 68 to 219 Hz; and the horn volume is 3.7 cu. ft.

Figure 1. The Hyperbolic horn at a triple corner (The microphone boom is in the background).

The horn was connected to an AURA NS3-193-8A speaker with frequency response from 50 Hz to 7000 Hz (±3 dB).

4. THE EXPEERIMENT

The experiment basically consisted of driving the speaker-horn combination with single sinusoids or band filleted random noise. The diffused sound field was measured by using a microphone boom at two different heights. The equivalent sound level over a 30 second traverse of the boom was calculated. The measurements were conducted for three locations of the horn – the horn speaker in triple corner (as shown in Figure 1); the horn speaker was moved diagonally by 2 feet; and the horn speaker was moved diagonally by 4 feet. The last location would represent a double corner somewhat. Different combinations of the horn speaker locations were also tested. The results for the above triple corner only are presented in this paper. As mentioned earlier, the operating frequency of the hyperbolic horn is from 68 Hz to 219 Hz. The above band width was determined by the manufacturer from the design of hyperbolic horn. The length and the mouth area were

evaluated after fixing the cut-off frequency and the upper limit frequency of the horn design.

5. RESULTS AND DISCUSSION

The room SPLs for various conditions are shown in Figures 2, 3 and 4 below.

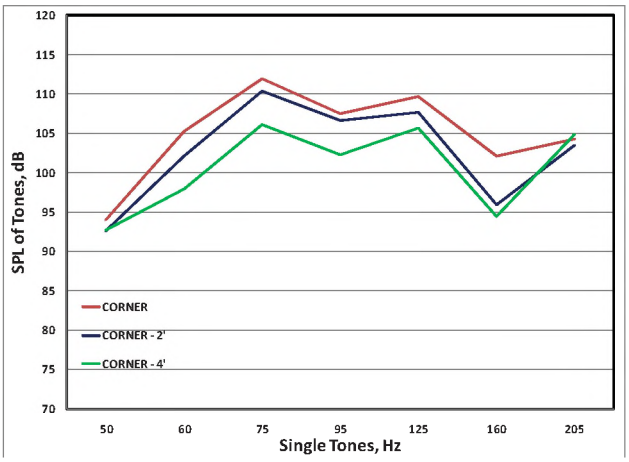


Figure 2. Room SPL variation – single tones.

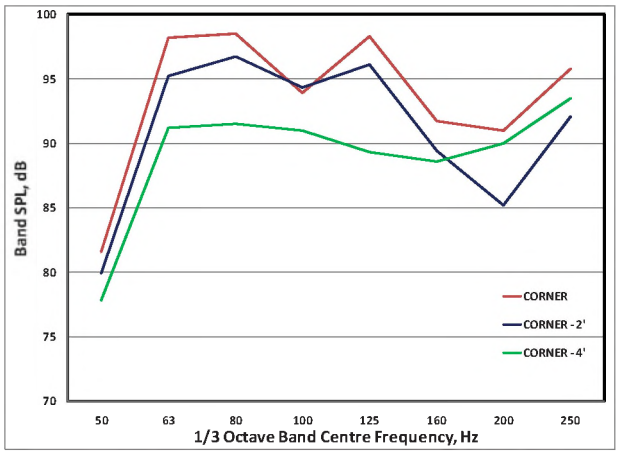


Figure 3. Room SPL variation – Band filtered noise.

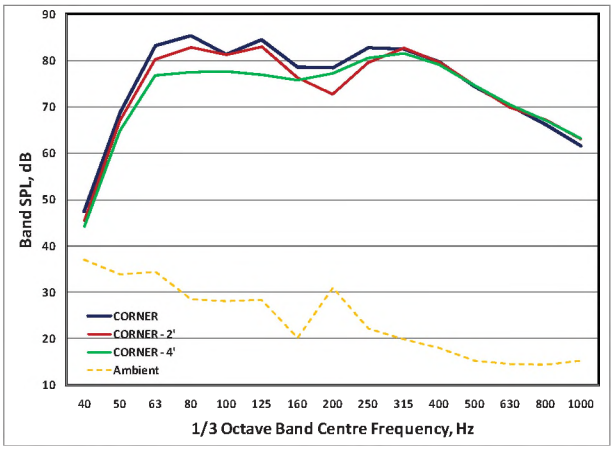


Figure 4. Room SPL variation – pink noise (40 to 10kHz).

The SPL variations in the room for sinusoidal sources are shown in Figure 2. The room SPLs between 60 Hz and 160 Hz are seen to follow the typical 'Q' factor variation of 3 to 4 dB differences. The behaviour below 60 Hz and above 200 Hz, is seen to be indifferent to the speaker location. Even though a strong signal was generated at 50 Hz, the triple corner effect is non-existent. The results for band-filtered random noise are shown in Figure 3 and much broader pink noise results are shown in Figure 4. The speaker location's effect is unpredictable for the band-filtered random noise within the operating range of the horn. The speaker location had absolutely no effect when the broader pink noise was generated. No consistent 'Q' factor effect was evident in the results of Figures 3 and 4.

6. CONCLUSIONS

The effect of the location of a horn speaker in a reverberation room was tested. The effect was evident in the sinusoidal input signals. When random noise and/or broad band signals were used as input, the preliminary results show that the speakers location had no impact on the diffused sound levels of the reverberation chamber.

ACKNOWLEDGMENTS

The reverberation room tests were conducted at Concordia University, Montreal. The kind assistance provided by the BCEE (Building, Civil and Environmental Engineering) department is duly acknowledged.

REFERENCES

1. Leo L. Beranek, *Acoustics* American Institute of Physics (1986).
2. John Murray, "The Quadratic-Throat Waveguide," Peavey Electronics Corporation (2000).
3. Lewis H. Bell, *Industrial Noise Control* Marcel Dekker Inc. New York (1982).
4. Leo L. Beranek, *Noise and Vibration Control* Institute of Noise Control Engineering. (1988).
5. R. Ramakrishnan and A. Grewal, "Reverberation Rooms and Spatial Uniformity." Proceedings of the Acoustics Week in Canada 2008, Vancouver. *Canadian Acoustics*, Vol. 36 (#3) (2008).
6. M. R. Schroeder, "Frequency-Correlation Functions of Frequency Responses in Rooms," J. Acoust. Soc. Am. Vol.34, pp. 1819, 1962. Also J. Acoust. Soc. Am. Vol.46, pp. 277-283, 1969.
7. F. Slingerland, G.M. Elfstorm and W.E. Grün, "Performance and Operational Capabilities of the Large European Acoustic Facility (LEAF)," Proceedings of the International Symposium on Environmental Testing of Space Programmes – (ESA SP-304, September 1990).

CHARACTERIZATION AND IMPROVEMENT OF SCATTERING AND ABSORPTION BY ARCHITECTURAL SURFACES WITHOUT THE USE OF SPECIALIZED FACILITIES

Chris Bibby[1], and Murray Hodgson[2]

[1] Acoustics and Noise Research Group, Department of Mechanical Engineering, University of British Columbia, Vancouver, BC, Canada [cbibby@interchange.ubc.ca]

[2] Acoustics and Noise Research Group, Department of Mechanical Engineering and School of Environmental Health, University of British Columbia, Vancouver, BC, Canada [murray.hodgson@ubc.ca]

1. INTRODUCTION

The Acoustics and Noise Research Group at the University of British Columbia is working with two companies that both manufacture a profiled wooden architectural wall panel. Currently these two surfaces are designed to be aesthetically pleasing and not intended to be used as an acoustic treatment. This work aims to first complete an acoustical characterization of the surfaces and secondly to modify them in such a way that they provide significant sound scattering and absorption, and can be marketed as an acoustical treatment. The two existing profiled wood surfaces and one flat wooden surface have been characterized in terms of their scattering and absorption coefficients. The two profiled surfaces have been modified for increased sound absorption and scattering, and re-characterized. The main challenge of this work has been to implement a method for measuring scattering on a limited budget and with limited facilities.

2. TEST SURFACES

Three original surfaces have been tested; throughout this paper they will be referred to as Surfaces 1, 2 and 3 respectively. Surfaces 2 and 3 have been modified for increased scattering and absorption. They are called 2S, 2A, 3S and 3A where S and A indicate designs altered to give increased scattering and absorption.

Surface 1 is largely flat and hard. It is composed of 5-mm-thick, 10-mm-wide softwood boards glued onto a 16-mm-thick plywood backing.

Surface 2 is made of 1- to 3-cm lengths of '2x4' (50 mm x 100 mm), '2x6' (50 mm x 150 mm) and '2x8' (50 mm x 200 mm) boards glued to a 16-mm-thick plywood backing, so that the end grain of each block is visible. Adjacent blocks are separated by a 1-mm gap. Surface 2S was made by increasing the maximum block height from 3 cm to 15 cm and defining the height of each block by the quadratic residue sequence with a period of seven. This surface was prototyped using high density polystyrene. Surface 2A was created by raising surface 2 over a 7-cm-deep, glass-fibre filled, cavity. Some of the blocks were removed and the exposed plywood cut away creating holes into the cavity to form a Helmholtz resonator type absorber.

Surface 3 is a corrugated-plywood product. The plywood is 3-mm thick, its corrugations have a 10-mm amplitude and a 40-mm period. The corrugated-plywood sheet is suspended over a 50-mm cavity, surrounded by a wooden frame. Surface 3S was created by cutting the corrugated plywood into 45-cm-square sections, curving them slightly, and inserting them into 15-cm-tall cells. Various configurations were created by varying the height and inverting the curvature of the corrugated plywood sections in the cells. Surface 3A was created by filling the cavity with fibreglass and cutting 3 x 20 mm^2 holes in the surface of the corrugated plywood to create a Helmholtz resonator type absorber. Tests were made with a variety of hole concentrations.

3. MEASUREMENTS

Three methods were considered for measuring the scattering coefficient. Two of the methods use a free-field environment [1, 2] and the third method uses a diffuse field [2, 3]. The measurement method that involves the diffuse field (ISO 17497-1) was used because the available anechoic chamber is not large enough to make measurements in the far field at full scale. In place of a reverberation chamber, a fully enclosed, concrete-walled squash court was used to create an approximately diffuse-field environment. ISO 17497-1 includes, as part of the scattering coefficient calculation, calculation of the random incidence absorption coefficient.

3.1. Measurement Facility and Apparatus

For these tests, a standard reverberation room was not available. A squash court was used as a substitute. A turntable, 3.5 m in diameter, was built to rotate at 1 rpm. All of the samples tested were 8 ft x 8 ft (2.44 m x 2.44 m) in dimensions. Test samples were recessed into the turntable to reduce scattering by the edges. Impulse-response measurements were made using an MLS signal.

4. RESULTS

It is of interest to start by considering the accuracy of the measurements. Then the scattering- and absorption-coefficient results are presented and discussed.

4.1. Measurement Accuracy

Due to time and budget constraints, the measurements were subject to non-negligible uncertainties. The uncertainties come primarily from the low quality of the diffuse sound field, and the lack of control over the ventilation system in the squash court.

The diffuseness of the sound field was limited because the wooden floors and ceiling provide low-frequency absorption, and, as the walls were brick, an uneven absorption distribution. Furthermore, the room contained no diffusers, and its width and height were very similar, resulting in high modal densities and an uneven modal distribution.

The lack of control over the ventilation system causes uncertainty as sudden changes in temperature and humidity during the measurement are possible; this method is highly sensitive to both of these factors, especially at high frequencies. Furthermore, any airflow in the room will alter the paths of the sound waves causing the later part of successive impulse responses to be incoherent. The reverberation time calculated from the average of these impulse responses would then be reduced [4].

4.2. Scattering

A selection of scattering measurement results are shown in Figure 1. The third-octave data have been averaged into octave-band data to smooth the plots. Results are shown from 100 Hz to 4 kHz as required by ISO 17497-1.

The results show that the original surfaces do not scatter significantly, except for surface 2 at high frequencies. Surface 2S showed significant scattering above 500 Hz and high scattering above 1 kHz. Surface 3S showed scattering above 250 Hz and high scattering above 500 Hz.

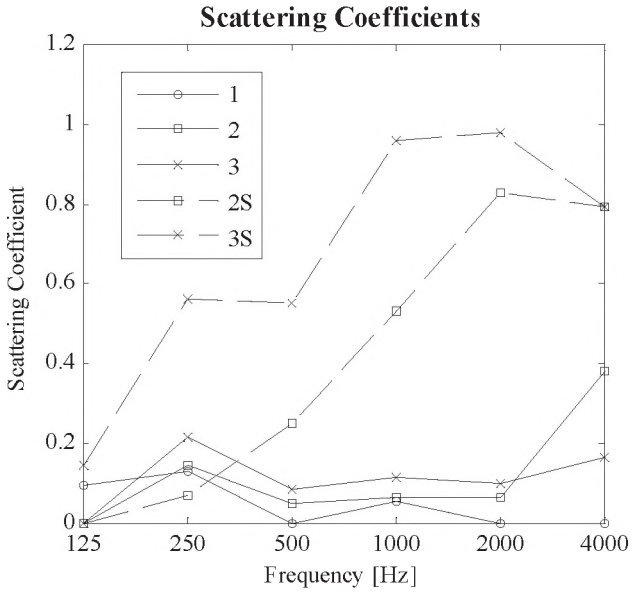


Fig. 1. Scattering results of select surfaces.

4.3. Absorption

A selection of absorption results are shown in Figure 2. As is expected for a flat, hard panel, Surface 1 has very low sound absorption at all frequencies. The absorption coefficient of Surface 2 is very low at low frequencies and slowly increases to nearly 0.4 at 4 kHz, possibly due to ¼ wave resonance in the gaps between the blocks. Surface 3 has a peak in absorption coefficient of nearly 0.4 in the 250-Hz octave band, which can be attributed to membrane absorption. Surfaces 2S and 3S both proved capable of being excellent low and mid frequency absorbers. This is typical of most Helmholtz resonator type absorbers.

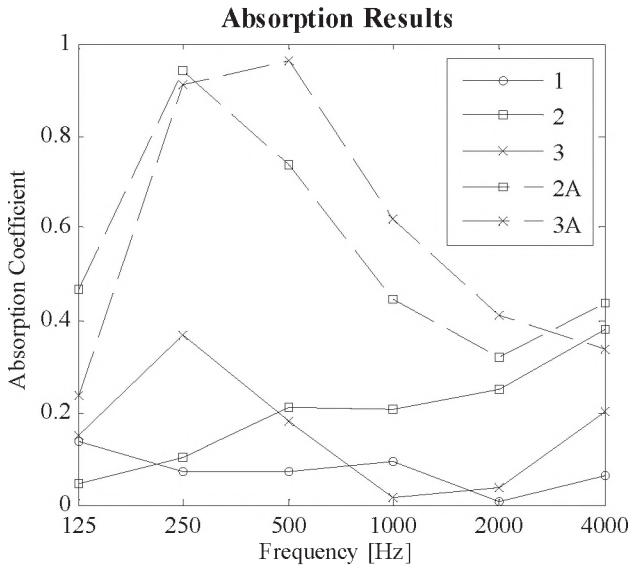


Fig. 2. Absorption results of select surfaces.

5. CONCLUSION

A method for measuring the random-incidence scattering and absorption coefficients has been implemented on a limited budget and without the use of specialized facilities. Test results from these methods, while not highly accurate, were sufficient to evaluate the general behaviour of the surfaces under study. The test method proved to be a useful tool for the development of absorptive and scattering surfaces for acoustic treatment.

REFERENCES

[1] AES-4ID-2001, Characterization and measurement of surface scattering uniformity, 2001.

[2] Vorlander, A.; Mommertz, E. Definition and measurement of random-incidence scattering coefficients, *Applied Acoustics,* Vol 60, 2000, pp.187-199

[3] ISO 17497-1:2003(E), Acoustics – Sound-scattering properties of surfaces – Part 1: Measurement of the random-incidence scattering coefficient, 2003.

[4] De Geetere, L. Analysis and improvement of the experimental techniques to assess the acoustical reflection properties of boundary surfaces (dissertation), Leuven: K.U.Leuven; 2004.

ACOUSTICAL EVALUATION OF TEMPORARY PERFORMANCE FACILITIES

Ben Gaum and Ramani Ramakrishnan
Department of Architectural Science, Ryerson University, Toronto, ON, Canada rramakri@ryerson.ca

1. INTRODUCTION

Considerable research has been conducted in understanding as well designing acoustics of performance spaces of conventional auditoria used for music, drama and speech. However, not much is known about the acoustics of spaces used for temporary performances. Typical examples of such spaces are the large group of venues used for jazz festivals. Large single arena, temporary of course, is used for the jazz festivals in Toronto and Halifax. On the other hand Montreal uses a number of venues, including one large space, for the Montreal Jazz Festival held every July. The acoustics of the three main arenas of Toronto, Montreal and Halifax was investigated. Computer simulations of the three venues were conducted. In addition, simple surveys of the stakeholders of the three venues were undertaken to compare the simulation with audience satisfaction. Complete details of the study can be found in the Master's thesis of Ben Gaum[1]. The results of the acoustics of temporary performance facilities will be presented.

2. BACKGROUND

The main phase of the research will consist of a simulation study (objective) and a subjective survey data collection and analysis. The experiment being set up will focus on looking at three different temporary structures that have been erected for the purpose of musical performances for various jazz festivals across Canada and will include: *The Atlantic Jazz Festival*, *The Montreal Jazz Festival* and the *Toronto Jazz Festival*. The three temporary structures used in this study are shown in Figure 1. The details of the three spaces are: (a) Toronto – Main tent 30 m (L) X 45 m (W) with overall height of 8.5 m; seating capacity of 1200; aluminum and steel frames with 21 ounce vinyl fabric roof and sides; (b) Halifax – 60' (L) X 130' (W) with overall height of 27'; seating capacity of 1200; aluminum and steel frames with 21 ounce vinyl fabric roof and sides; and (c) Montreal – Open air/covered stage; Stage 42' (L) X 44' (W) with overall height of 29'; seating capacity open spaced; wood and steel frames with wood panelling for the stage.

The results presented in this paper have focussed on the objective and subjective analysis techniques of Beranek[2]. The study has combined the collection of acoustical data on both the scientific engineering and subjective levels in order to find a connection between architectural form and the subjective interpretation of music. Upon gathering structural and layout information about the spaces themselves, a 3D model was constructed and analyzed using acoustical engineering software, *CATT Acoustics*[3]. As an additional element of the simulation study, an acoustics survey was also conducted to gather subjective data from the various venues themselves. In order to maintain a level of consistency for all the surveys, a single song from a single band that has been played in all three venues was used as a base for the survey analysis. The survey was given to individual musicians, technicians and engineers that were directly linked to and involved with the venues in order to obtain an accurate account of what the sound quality and acoustics were like in that specific venue when the specific song was played.

(a)

(b)

(c)

Figure 1. Temporary Performance Facilities – a) Toronto Jazz Venue; b) Atlantic Jazz Festival venue; and c) Montreal Jazz Festival venue.

3. OBJECTIVE ANALYSIS

The objective analysis involved a simulation study of the three venues using *CATT Acoustics*. The simulation model requires information about the specific venues with regards to overall dimensions, layouts, materials and surrounding site conditions. After all the important information was acquired, 3D models were built that included material

attributes for each face that would be able to be read and analyzed in CATT Acoustics. Proper locations for sources and receivers of sound had to be properly placed in the model as well as sound absorbing coefficients for each material used. A typical CATT Acoustics screen for the Halifax venue is shown in Figure 2.

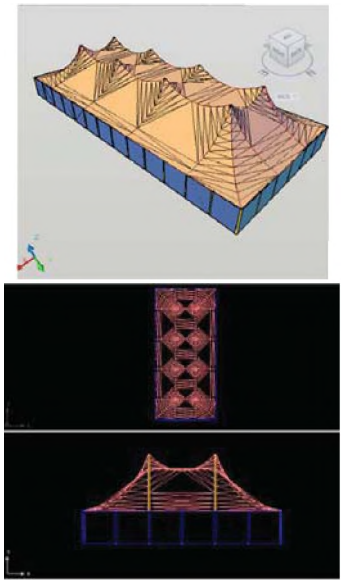


Figure 2. CATT Acoustic rendering of Halifax Venue

Upon completion of the model building for each venue, the models were then analyzed for various acoustic properties, focussing on the *quantitative* properties that would most accurately match the *qualitative* properties that would be looked at in Section 4, the survey study. The quantitative properties that were looked at in the simulation study were: a) C-80 (Clarity); b) SPL (Sound Pressure Level); c) G (Total Sound Level or Loudness); and d) RT (Reverberation time). The quantitative results for the three venues are shown in Table 1 below. It is seen that the four parameters are well within the acceptable range except the RT for Toronto venue is higher than optimum.

4. SUBJECTIVE SURVEY

The survey being used in this study is largely based on the survey questionnaire of Beranek[2] for his study. The survey focuses again on the three members of *The New Deal*, as well as other musicians who have played in the venues, technicians and engineers who have worked the venues and manufacturers of the structures themselves. The 10 survey questions are: *Clarity* (1-muddy and 10-clear); *Balance 1* – treble (1-weak and 10-loud); *Balance 2* – bass (1-weak and 10-loud); *Balance 3* – singers (1-weak and 10-loud); *Noise* (1-Intolerable and 10 inaudible); *Overall Impression* (1-poor and 10-excellent); *Reverberance* (1-poor and 10-excellent); *Envelopment* (1-poor and 10-excellent); *Intimacy* (1-poor and 10-excellent); and *Loudness* (1-poor and 10-excellent). The results of the survey were gathered and organized into an easy to read, quantitative format for analysis. In total, 30 surveys were filled out with 10 surveys per venue. The results of the survey are presented in Table 2 below. Toronto and Halifax venues are seen to be in the mid-high acceptance range where as Montreal seems to have many unhappy users due to its open air/closed stage design.

REFERENCES

1. Ben Gaum, "Sound Created Form," *M.Arch. Thesis*, Ryerson University, Toronto (2009).
2. L. L. Beranek, *Concert Halls and Opera Houses, music, acoustics and architecture*, Springer Verlag, New York (2004).
3. CATT Acoustics. V8.0g(build 2.02) CATT (1998-2007).
4. M. Mehta et.al. *Architectural Acoustics: Principles and Design*. Prentice Hall. (1998)

Table 1. *CATT Acoustics* results for the three venues .

Factor	Toronto	Halifax	Montreal
C-80 – Average of 500, 1000 and 2000 Hz bands (-1 dB to +4 dB)[4]	0 dB	5 dB (high)	0 dB
SPL: in 500, 1000 and 2000 Hz	Uniform	Uniform	Uneven
G: 0 to 5 dB acceptable in 500 and 1000 Hz bands[2,4]	9 to 10 dB	0 to 5.5 dB	0 dB (low)
RT: acceptable around 1 to 1.2 secs	3.5 secs (H)	1.2 secs	0.8 secs-low

Table 2. Survey results for the three venues.

Factor*	A	B	C	D	E	F	G	H	I	J
Toronto	7.5	5.5	6	7	7	5.5	7	7.5	5.5	7.5
Halifax	8	5.5	5.5	8	7.5	5.5	7.5	7	6.5	7
Montreal	4.5	7.5	7	4	8	5	8	6	4	6.5

***NOTE:** A-Clarity; B-Reverberance; C-Envelopment; D-Intimacy; E-Loudness; F-Balance 1; G-Balance 2; H-Balance 3; I-Background Noise; and J-Overall Impression.

EFFECT OF SOME FLOOR-CEILING CONSTRUCTION CHANGES ON FLANKING TRANSMISSION

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

1. INTRODUCTION

A large systematic study was carried out by the National Research Council-Institute for Research in Construction (NRC-IRC) to evaluate a wide range of floor-ceiling designs to control transmission of airborne and impact sound between vertically adjacent rooms. As impact sources, the tapping machine according to ASTM E492 as well as the heavy sources, the ball and the "bang-machine", according to JIS A 1418-2 were used. The first part of the study was a parametric study. The direct sound transmission of a large series of floor-ceiling assemblies was measured systematically in the NRC-IRC Floor Transmission Facility to understand the complex interaction of the various structural components and to optimize noise control measures. The second part of the study was a flanking study. A number of floor-ceiling assemblies that achieved good direct sound insulation in the parametric study were selected and their system performance was measured in the NRC-IRC Flanking Transmission Facility. The system performance is determined by the apparent sound transmission which is the sum of the direct transmission and the flanking transmission via the floor to the walls below. This paper presents some findings from this flanking study. Only impact measurements with the tapping machine are presented in this paper.

2. MEASUREMENTS AND RESULTS

The parametric study shows that adding a floor topping or decoupling the ceiling from the floor assembly are both effective ways of controlling direct transmission. Two floor-ceiling assemblies with decoupled ceilings were selected to be studied in the Flanking Facility with and without topping. One objective was to investigate the contribution of flanking transmission to the apparent sound transmission.

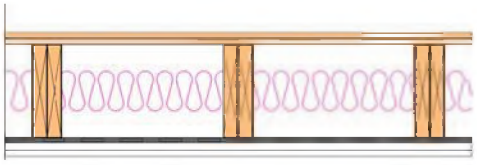


Figure 1: Floor ceiling assembly with resilient channels.

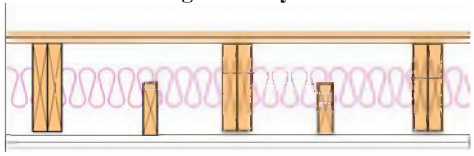


Figure 2: Floor ceiling assembly with separate joists.

The first assembly, shown in Figure 1, consisted of two layers of 16 mm plywood subfloor, scabbed 2X10 joists spaced 455 mm o/c, 100 mm glass fiber insulation, resilient channels spaced 455 mm o/c, and two layers of gypsum board (21 mm and 16 mm) for the ceiling. The second assembly, shown in Figure 2, was identical to the first except separate ceiling joists were used instead of resilient channels to decouple the ceiling. The walls in the room below the floor/ceiling assembly consisted of 2X6 wood studs spaced 455 mm o/c with a single layer of 13 mm gypsum board directly attached.

Five transmission paths contribute to the apparent sound transmission of vertically adjacent rooms:

1. One direct transmission path through the floor ceiling assembly,
2. Two floor-wall flanking paths via the floor-ceiling assembly and the load-bearing (LB) junctions, and
3. Two floor-wall flanking paths via the floor-ceiling assembly and the non-load bearing (NLB) junctions.

Figure 3 shows the direct impact sound pressure level and the flanking impact sound pressure level transmitted across the load bearing and non-load bearing junction for the two floor-ceiling assemblies. Direct transmission is reduced more effectively for the assembly with resilient channels (RC) than with separate joists (SJ) whereas flanking transmission across the load-bearing and non-load bearing junction is similar for the two assemblies. Figure 4 shows the percentage of contribution of the direct and flanking transmission to the total transmitted sound power for the two floor-ceiling assemblies. When the direct transmission is reduced effectively by decoupling the ceiling, flanking becomes more important and contributes more to the overall apparent transmission. Flanking transmission is more important for the load bearing wall than the non-load bearing wall for both assemblies. However, this relative importance could change as shown in the next part of the study.

Adding a floor treatment is a second method to improve impact sound insulation of a floor assembly. The parametric study had found a gypsum-plywood (21 mm and 15.5 gypsum and 15.5 mm plywood) floor topping to be effective in suppressing direct impact and airborne sound transmission. The gypsum-plywood topping was added to both floor-ceiling assemblies to examine its effect on flanking transmission. Figure 5 shows that the effectiveness of the topping is similar for both assemblies. However, it is different for direct and flanking transmission. Previous NRC-IRC studies[1] have shown that the floor topping

reduces power injection by the impact source but also changes how the vibration energy travels across the floor. On a bare floor, the floor joists "channel" structure-borne sound towards the load bearing (LB) junction while they "block" wave propagation towards the non-load bearing (NLB) junction. When a topping is added, both the "channeling" and "blocking" effects are diminished. The floor appears more homogeneous and isotropic and the relative importance of the load-bearing to the non-load-bearing flanking transmission changes. Figure 6 shows the percentage of contribution of the direct and flanking transmission to the total transmitted sound power for the two floor-ceiling assemblies with the topping. With the topping added, the relative importance of non-load bearing flanking path has increased. With the resilient channel ceiling and the topping added, the non-load bearing flanking path contributes more than the direct path from 100 Hz to 1 KHz and more than the load-bearing flanking path in the 200 Hz and 250 Hz bands.

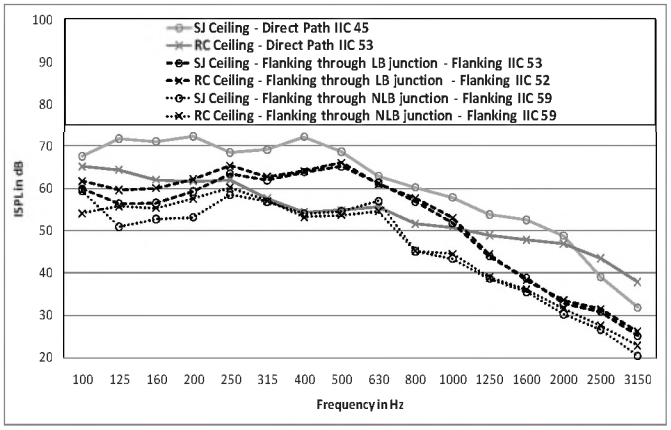


Figure 3: Direct and Flanking transmission paths across load-bearing (LB) and non-load bearing (NLB) junctions of two floor-ceiling assemblies.

3. CONCLUSIONS

The flanking study on two types of floor-ceiling assemblies with and without topping shows that both decoupling the ceiling and adding a topping are effective in reducing direct transmission. However, they affect flanking transmission differently. This changes the importance of direct and flanking transmission for the overall apparent transmission. When the ceiling is decoupled effectively and a topping is added, both the load bearing as well as the non-load bearing flanking paths are equally important and exceed the direct path in some frequency bands. Thus, it is necessary to examine the overall system performance when construction changes are made to floor-ceiling assemblies to improve sound insulation.

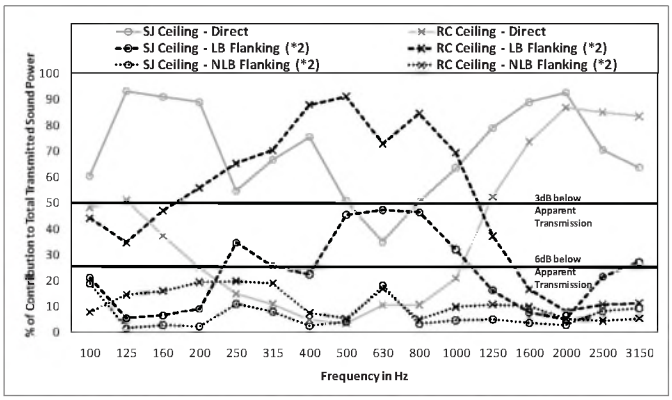


Figure 4: Relative contribution of direct and flanking transmission to total transmitted sound power for two floor-ceiling assemblies.

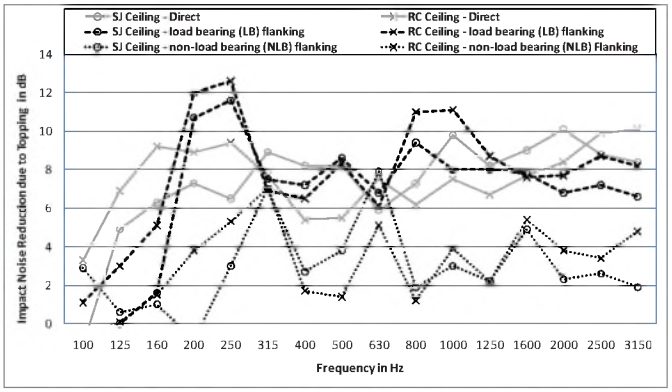


Figure 5: Impact noise reduction of two floor-ceiling assemblies with a topping.

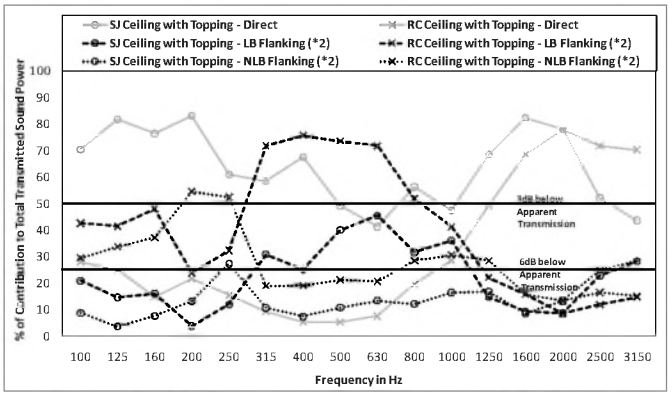


Figure 6: Relative contribution of direct and flanking transmission to total transmitted sound power for two floor-ceiling assemblies with a topping.

REFERENCES

[1]T.R.T. Nightingale "Controlling impact noise in wood-frame multi-unit buildings," Proceedings of 15th International Congress on Sound and Vibration, (Daejong, 2008).

EFFECTS OF A CONCRETE TOPPING AND MODIFIED RESILIENT INTERLAYERS ON SOUND TRANSMISSION THROUGH A CONCRETE FLOOR

Johannes Klein, Berndt Zeitler and Bradford Gover
National Research Council Canada, 1200 Montreal Rd. Ottawa, ON, Canada, K1A 0R6, johannes.klein@nrc.ca

1. INTRODUCTION

Concrete floors show good sound insulation properties. In this paper the effect of an added concrete topping on a resilient interlayer and its beneficial modification will be examined.

2. METHOD

2.2 Construction details

Assembly A shown in Figure 1 on the left, consists of a bare 150 mm concrete slab. Assembly B is the identical concrete slab with an added d=25 mm thick resilient interlayer and a 100 mm concrete topping. Assembly C on the right demonstrates the implementation of a modified interlayer of evenly distributed strips, covering 25% of the floor area.

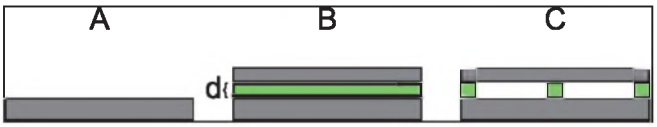


Figure 1: Sketch of the assemblies. Grey: Concrete; Green: Interlayer

2.1 Sources

The assemblies were evaluated using the standard airborne test (ASTM E90) and three standard impact sources: Tapping machine (ASTM E492, henceforth referred to as hammer), ball (JIS A 1418-2, ISO 140-11) and tire (JIS A 1418-2, KS F 2810-2). The standard impact sources can be seen in Figure 2 below and have been described in detail in an earlier paper [1].

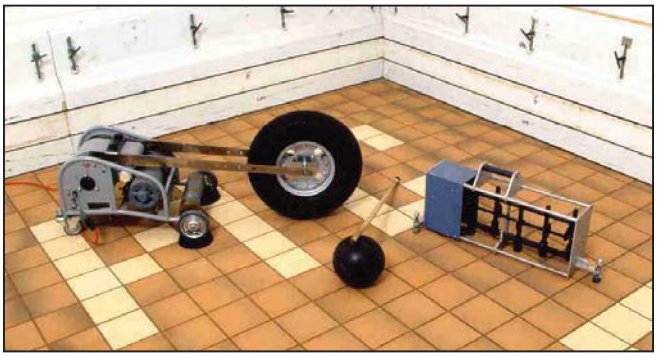


Figure 2: Standard impact sources. Tire, ball, hammer from left to right.

3. RESULTS

3.1 Bare concrete slab

Figure 3 shows the measurement results for the bare concrete slab. The tire causes the highest Impact Sound Pressure Level (ISPL), followed by the ball. Note that these two are the fast weighted maximum level of transient noise events. Compared to the other impact tests, the hammer induces more energy in the high frequency bands and therefore yields the highest signal to noise ratio in this range. The curve of the airborne test rises with increasing frequency, since it is displayed as Transmission Loss (TL).

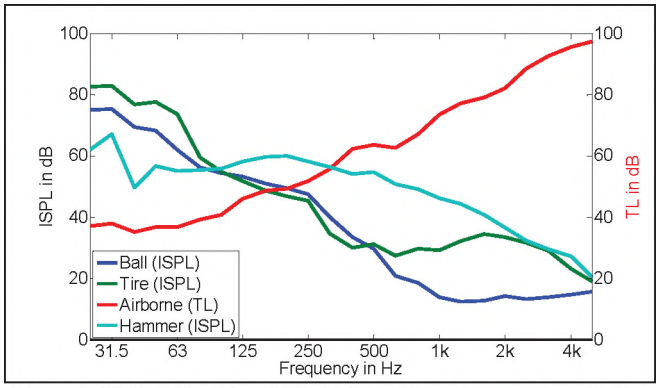


Figure 3: ISPL (left axis) and TL (right axis) measurement results for the bare 150 mm concrete slab.

Different measurements are limited in certain frequency ranges by the influence of background, application and flanking noise. The ball measurement reaches the background noise level at approximately 1 kHz. The same applies to the tire, although the level is significantly higher due to the application noise of the mechanical elements of the device (see Figure 2). The limitation due to flanking is slightly above the TL curve at high frequencies. This will be more apparent in the following measurements.

3.2 Concrete slab with concrete topping and full interlayer

Figure 4 shows the differences between the assemblies A and B. The added concrete topping on the full interlayer provides a major improvement in sound insulation.

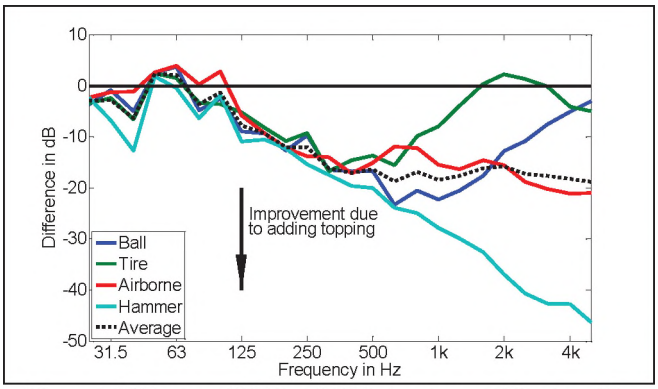


Figure 4: Difference of ISPL and TL between assemblies A and B.

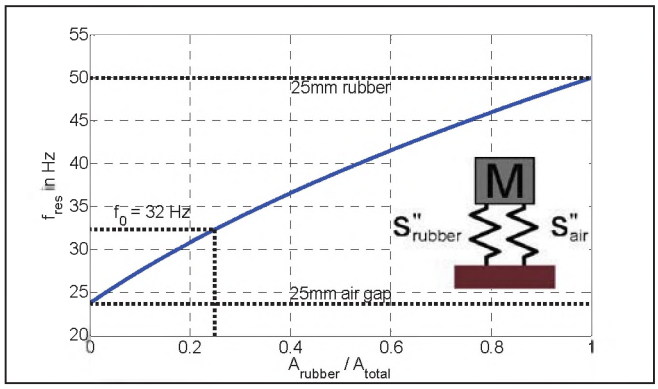


Figure 5: Nominal resonance frequency for different interlayer/air ratios.

The ball and tire measurements are limited by background noise in the high frequencies, thus the differences approach zero. The airborne measurement is restricted by flanking. Therefore there is no further improvement above 1 kHz. The noticeable worsening for the tire measurement at 2 kHz arises from different levels of application noise and can be considered as measurement uncertainty. The results of the hammer match the theory throughout the whole frequency range.

According to Cremer's parallel plate theory [2], if the upper plate is less stiff and has less mass than the lower plate, then the ISPL decreases above the fundamental resonance frequency with 6db/octave. This coincides with the theory of a simple spring mass system on a rigid foundation as investigated below. Such a simplified model is considered valid due to the big bending wave lengths of the regarded frequencies in concrete. The shape of the hammer curve confirms the validity and the fundamental resonance can be identified in the 50 Hz band in Figure 4. The sound insulation can be improved even more by shifting this resonance to lower frequencies.

3.3 Concrete slab with concrete topping on a modified interlayer

As afore mentioned, assembly B can be considered a simple spring mass system at low frequencies. This is represented in the sketch at the bottom right corner of Figure 5, with only the spring component of the resilient interlayer (s''_{rubber}). To shift the fundamental resonance frequency, the area of the rubber can be reduced, causing a rubber dependent and an air dependent spring to act in parallel. The estimate of the stiffness per area of the interlayer itself can be found by using Eq. 1 with Eq. 3 and the resonance frequency of assembly B (50 Hz).

$$\omega = \sqrt{\frac{s''_{rubber}}{m''_i}} \quad \text{Eq. 1;} \qquad \omega = \sqrt{\frac{s''_{rubber} + s''_{air}}{m''_i}} \quad \text{Eq. 2;}$$

$$m''_i = 246\,kg/m^2 \quad \text{Eq. 3} \qquad s''_{air} = \frac{\rho c^2}{d} \quad \text{Eq. 4}$$

With Eq. 2 for the resonance for the partial interlayer and Eq. 4, the resonance depending on the ratio of the interlayer and air areas can be computed as shown in Fig. 5.

For assembly C, a ratio of 0.25 was chosen, to theoretically shift the fundamental resonance to 32 Hz. The interlayer was cut into strips, filling out 25% of the area.

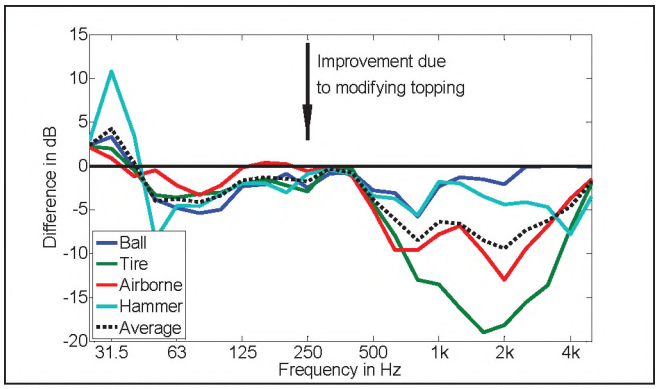


Figure 6: Difference of ISPL and TL between assemblies B and C.

Figure 6 shows the improvement of assembly C compared to assembly B. The fundamental resonance is, as calculated, shifted to the 31.5 Hz band. The new assembly has a similar improvement for all measurement methods up to 500 Hz. The deviations for the tire, ball and airborne tests are again caused by the previously mentioned measurement uncertainties and noise limitations.

4. CONCLUSION

The concrete topping and resilient interlayer improve the sound insulation properties of the concrete floor. Modifying the interlayer optimizes the results by lowering the fundamental resonance frequency of the resulting spring mass system. In this study the STC/IIC was improved from 52/28 (assembly A) to 65/58 (B) and to 66/60 (C). The interlayer strips, which introduce a directional pattern, influence the bending wave attenuation and thus the flanking transmission. Future investigations are planned.

REFERENCES

[1] Zeitler, B. et al. (2010). Parametric Study of Sound Transmission through Lightweight Floors. Proceedings of Inter-Noise 2010, Lisbon 2010

[2] Cremer, L., Heckl, M. (1996). Koerperschall. Springer

BEAM-TRACING MODEL FOR PREDICTION OF IMPULSE RESPONSES, AND EFFECTS OF SURFACE-REACTION MODELING AND EDGE DIFFRACTION IN ROOMS

Behrooz Yousefzadeh and Murray Hodgson
Acoustics and Noise Research Group, University of British Columbia,
3rd floor, 2206 East Mall, Vancouver, BC, V6T 1Z3, Canada

1. INTRODUCTION

This paper presents the development of a wave-based, triangular-beam-tracing model for predicting the transient response of rooms. The model is based on the one developed by Wareing and Hodgson [1] for studying the steady-state response of empty rooms with specularly-reflecting, extended-reaction surfaces. Room surfaces are modeled as multiple layers of elastic solid, fluid or poroelastic materials, and their acoustical properties are calculated using a transfer-matrix approach. The new model can predict the pressure impulse response between a source and a receiver, required for deriving room-acoustical parameters that correlate with the subjective perception of sound. Energy-based prediction (including diffuse surface reflection) has been implemented in the new model, and both wave- and energy-based modeling have been validated against theory for the case of sound propagation above a specularly- or diffusely-reflecting rigid plane.

The model has been further improved to include sound diffraction around wedges, based on the theory developed by Svensson *et al.* [2]. This improves the application of the model to configurations with more realistic features, such as sound propagation in fitted rooms or in long enclosures with bends, and the evaluation of screen barriers in open-plan offices. The application of the beam-tracing model is not limited to rooms; provided that the proper boundary conditions are used, it can be applied to other environments such as ventilation ducts and outdoor environments where the medium is unbounded in one or more directions.

2. TEST CONFIGURATIONS

The new beam-tracing model was used to compare the effects of different surface-reaction models on the transient response and derived room-acoustical parameters. In addition to investigating the significance of modeling room surfaces as of extended or local reaction, effects of phase change due to surface reflections have been studied by considering real and complex reflection coefficients. Moreover, wave-based energy impulse responses and room-acoustical parameters have been compared with those obtained from energy-based modeling.

Following the previous investigation of the effects of extended- and local-reaction surfaces on steady-state sound levels in rooms [3], three room geometries were studied: Room 1: a small office (3*3*3 m^3); Room 2: a corridor (10*3*3 m^3); and Room 3: a workshop (10*10*3 m^3). In each room one or more of the surfaces were modeled as multilayer walls. The configurations were as follows: a single glass panel as a large window in all three rooms (configurations G1, G2, G3); the previous configurations, but with double glass panels, (DG1, DG2, DG3); four walls modeled with double drywall panels in Rooms 1 and 2 (D1, D2); double steel panels as the walls in Room 3 (SW3); carpet on a hard backing as the floor in Rooms 1 and 2 (C1, C2); suspended acoustical ceiling modeled as a fiberglass with air backing in Rooms 1 and 2 (SAC, SAC2); double-steel panel as the ceiling in Room 3 (SC3); and fiberglass with a rigid backing as the floor for Room 3 (FG3). In all cases, other room surfaces had an angularly-invariant average diffuse-field absorption coefficient of 0.1.

3. APPLICATION OF THE MODEL

In each room configuration, predictions were made using both the energy and phase models. The energy model is the same as the phase model except that, instead of complex values of sound pressure (magnitude and phase), squared absolute values of pressure are calculated. Moreover, the effects of phase changes due to reflections have been studied by considering both complex and real reflection coefficients. The following room-acoustical parameters were calculated for each configuration, in octave bands from 63 to 4000 Hz: EDT, T20, T30, Ts, C80 and C50.

All predictions in this work have been made with 4500 beams traced for 80 reflections. The acoustical properties of the multilayer surfaces were pre-calculated in order to make runtimes reasonable. The impedance of each surface was calculated for angles of incidence from 0 to 90 degrees, with 1-degree increments. The reflection coefficients of the extended-reaction surfaces were then calculated accordingly. For local-reaction surfaces, the impedance value at normal incidence was used for all angles of incidence [3].

Pressure impulse responses were obtained in each configuration from the room transfer function via inverse Fourier transformation. Energy echograms were found by keeping track of the arrival times of beams, and their corresponding sound energies at the receiver positions. Sound-decay curves were then computed by backward integration of the squared impulse responses or the echogram, respectively, for the wave-based or energy-based models. Figure 1 shows octave-band sound-decay curves at 1000 Hz for configuration SAC2, the corridor with suspended acoustical ceiling.

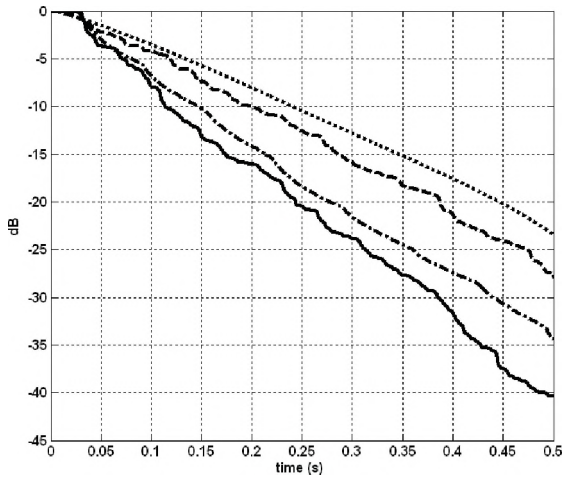


Figure 1. Octave-band sound-decay curves at 1000 Hz for configuration SAC2, the corridor with suspended acoustical ceiling. Solid line: extended reaction, wave-based; dash-dot line: extended reaction, energy-based; dashed line: local reaction, wave-based; dotted line: local reaction, energy-based.

4. SUMMARY OF THE RESULTS

Values obtained from wave-based modeling with extended-reaction surfaces are considered to be theoretically more accurate than other cases and therefore were chosen as reference values. Consequently, the standard deviation in the difference in values of the room-acoustical parameters with respect to the reference case was regarded as the comparison criteria.

The study results can be summarized as follows:

- The effects of changing between local- and extended-reaction modeling are in general smaller than the effects of changing between real and complex reflection coefficients, or changing between energy and phase models.

- Changing between local- and extended- reaction modeling is always least significant when using complex reflection coefficients with a phase model and is often most significant when using real reflection coefficients with a phase model.

- As expected, parameters obtained using a phase model with real-valued reflection coefficients are closer to reference values than those obtained using an energy model.

- When working with a phase model, extended-reaction surfaces often give results closer to reference values than local-reaction surfaces. Nevertheless, similar results were not observed for an energy model.

- Changing from complex to real reflection coefficients, and from a phase to an energy model, is, with a few exceptions, more significant in DG1, DG2 and DG3 than in G1, G2 and G3. This is due to the fact that the phase of the reflection coefficient of a double-glass panel is often larger than that of a single-glass panel.

- Changes in reverberation times (EDT, T20 and T30) follow the general trends mentioned above, except for SAC1 and SAC2 which have their most significant change in the case of local-reaction modeling with complex reflection coefficients in the phase model. Changing to real reflection coefficients and to the energy model is less significant. The suspended acoustical ceiling is the only multilayer surface in these rooms. Therefore, the large changes can be attributed to a large phase difference between the local- and extended-reaction reflection coefficients of the suspended acoustical ceiling for the incident angles of the early reflections from the ceiling.

- Similar results to the one above are observed for Ts, C80 and C50, but only for configuration SAC1. Since the dimensions of SAC1 and SAC2 are different, this can be attributed to bigger phase changes of the early reflections in SAC1.

- Changing from phase to energy modeling affects clarity indices - C80 and C50 - more than other parameters (these changes are more significant for C50). This indicates the significance of wave effects in the clarity of perceived sound. While changing from complex to real reflection coefficients makes non-negligible changes in both clarity indices, the changes are not always significant.

- Ts is the parameter which is least affected by a change of modeling conditions, with the most significant differences occurring for D1 and D2. While changing between local and extended reaction is overall not significant (*i.e.* below 10 ms), changing between phase and energy models is significant at all frequencies. Changing from complex to real reflection coefficients often makes the reverberant field stronger, and increases the center time. On the other hand, changing from phase to energy modeling often decreases the center time.

REFERENCES

[1] A. Wareing and M. Hodgson, "Beam-tracing model for prediction of sound in rooms with multilayer bounding surfaces", *J. Acoust. Soc. Am.*, **118** (4), 2321-2331 (2005).

[2] U.P. Svensson, R.I. Fred and J. Vanderkooy, "An analytic secondary source model of edge diffraction impulse responses", *J. Acoust. Soc. Am.*, **106** (5), 2331-2344 (1999).

[3] M. Hodgson and A. Wareing, "Comparison of predicted steady-state levels in rooms with extended- and local-reaction bounding surfaces", *J. Sound Vib.*, **309**, 167-177 (2008).

A NONLINEAR GEOMETRICAL ACOUSTIC MODEL FOR SONIC BOOM PROPAGATION

M. Berci[1] and L. Vigevano[2]
[1]iETSI, School of Mechanical Engineering, University of Leeds, UK, m.berci@leeds.ac.uk
[2]Dipartimento di Ingegneria Aerospaziale, Politecnico di Milano, Milano, Italy, luigi.vigevano@polimi.it

1. INTRODUCTION

Sonic boom prediction is a typical multi-scale problem, since the pressure signature generated by the shock structure at the length scale of the aircraft, L, in the "near-field", is transmitted through the atmosphere a long distance away, of the order of $100L$, to the "far-field". The propagation of the pressure perturbation is the most important aspect of the phenomenon: small amplitude nonlinear effects accumulate over long distances and distort the pressure signature significantly, giving rise to the coalescence of the pressure distribution into shocks and typically resulting in an asymptotic N-wave[1]. The goal of the present work is not to propose a new comprehensive theory of sonic boom propagation but rather to formulate an analytical model which can predict accurately, under limited conditions, the pressure signature on the ground in the vertical plane below an aircraft (where its sonic boom of maximum intensity lays, due to its minimum distance from the ground) in steady horizontal supersonic flight. To this aim, we propose to combine a revisited formulation of the nonlinear treatment of the pressure wave evolution due to Friedman and coauthors[2] with the simplified calculation of its nonlinear distortion due to George and Plotkin[3], supplemented by the "area rule" for the shock waves formation[4]; straightforward adaptations of the ray-tracing system obtained by Randall[5] and ray-tube area calculated by Pierce and Thomas[6] are then employed to complete the set of necessary equations, along with the standard atmosphere model. This combined method is simple and, although limited to a constant horizontal aircraft speed and a still atmosphere, allows a very accurate and efficient prediction of the boom propagation starting from a given pressure signature in the near-field.

2. METHOD

The sonic boom intensity at the ground $Z = 0$ is derived from an initial pressure signature generated by an aircraft in supersonic cruise flight at the altitude $Z = Z_f$ and then propagated along the acoustic ray-tubes through the real stratified atmosphere. Expanding the solution of 1D Euler equations as an isentropic perturbation of the atmospheric conditions[2], it is possible to separate the spatial dependence from the time dependence of the boom propagation problem and hence to investigate the different orders of its solution. The zeroth-order terms of the mass and momentum equations lead to:

$$P_1 = \frac{\gamma P_0}{a_0} w_1 = a_0 \rho_0 w_1, \tag{1}$$

while considering the first-order terms we obtain the following Bernoulli/Riccati ODE:

$$\frac{dw_1}{ds} + \frac{1}{2}\left(\frac{1}{A}\frac{dA}{ds} + \frac{1}{P_0}\frac{dP_0}{ds} - \frac{1}{a_0}\frac{da_0}{ds}\right) w_1 - \frac{(\gamma+1)}{2a_0^2} w_1^2 = 0. \tag{2}$$

If we neglect the second-order perturbation velocity term in the equation above, we recover the solution presented in [2] that we call here the quasi-linear model, which accounts for the variation of both the acoustic impedance $a_0\rho_0$ and the ray-tube area A through the atmosphere:

$$\Delta P_{ql} = \sqrt{\frac{a_0 \rho_0 A_h}{a_h \rho_h A}}\, \Delta P_h, \tag{3}$$

where subscript h refers to the distance r_h from the aircraft where the pressure signature is assigned. If we consider the whole ODE, we have a nonlinear theory (the best that can be done within the assumption of isentropic flow) which results in:

$$\Delta P_{nl} = \frac{\sqrt{\frac{a_0 \rho_0 A_h}{a_h \rho_h A}}}{1 + \frac{\gamma+1}{2\gamma}\sqrt{\frac{P_h}{a_h}}\, a_h^2 \left|\left(\frac{\Delta P}{P}\right)_h\right| \int_{s_h}^{s} \sqrt{\frac{a_0 A_h}{P_0 A}}\, \frac{ds}{a_0^2}}\, \Delta P_h. \tag{4}$$

In order to account for the presence of the solid ground, the ground reflection factor k_r is introduced[7], the ideal value of which is 2 (corresponding to a total reflection) and typical value 1.9 (corresponding to a standard terrain).

Randall's approach[5] is employed to derive the ray-tracing equations for the trajectory of the acoustic rays, while for the variation of the ray-tube area the expression derived from [6] and simplified as proposed in [3] is adopted.

The nonlinear distortion of the pressure signature is caused by the difference between the actual time t_r spent by each portion of the acoustic perturbation to reach the ground according to its actual propagation speed, which is the sum of the local sound speed a and the flow perturbation velocity w, and the ideal time t_i predicted by classic linear acoustics, to travel the same distance. In this boom propagation model, the pressure signature is supposed to be defined in time as $\Delta P = \Delta P(t, r)$ at any distance r from the aircraft's x axis in the vertical plane; nevertheless, it can trivially be defined also in space as $\Delta P = \Delta P(x, r)$ by considering that $x = Ut$ in steady flight, U being the flight speed. Considering the quasi-linear perturbation (3) the model by George and Plotkin[3] is recovered, whereas by substituting the ground intensity of the acoustic perturbation as calculated by (4) the nonlinear time advancement results:

$$\Delta t = -\frac{\gamma+1}{2\gamma}\left(\frac{\Delta P}{P}\right)_h \int_{s_h}^{s} \left[\frac{\sqrt{\frac{P_h a_h A_h}{P_0 a_0 A}}}{1+\frac{\gamma+1}{2\gamma}\sqrt{\frac{P_h a_h^2}{a_h s_h}}\left|\left(\frac{\Delta P}{P}\right)_h\right| \int_{s_h}^{s} \sqrt{\frac{a_0 A_h}{P_0 A}} \frac{ds}{a_0^2}} \right] \frac{ds}{a_0}, \quad (5)$$

where $ds = M_0 dr/\beta_0$, $M_0 = U/a_0$ and $\beta_0 = \sqrt{M_0^2 - 1}$. When the integral in the above formula converges to a finite number we observe the so-called "freezing effect"[8]. Typically, the ground pressure results in a non-physically multi-valued signature and the "area rule" needs to be employed in order to make it single-valued by inserting the appropriate shock waves[4], the area underneath the pressure signature being balanced on both sides of each shock.

3. RESULTS

As an example of the sonic boom propagation in the real atmosphere, we consider the pressure signature from a low-boom optimized Northrop F-5E Tiger II military aircraft[9] which is $L = 50$ft long and flies at $M = 1.4$ speed and $Z_f = 32000$ft altitude.

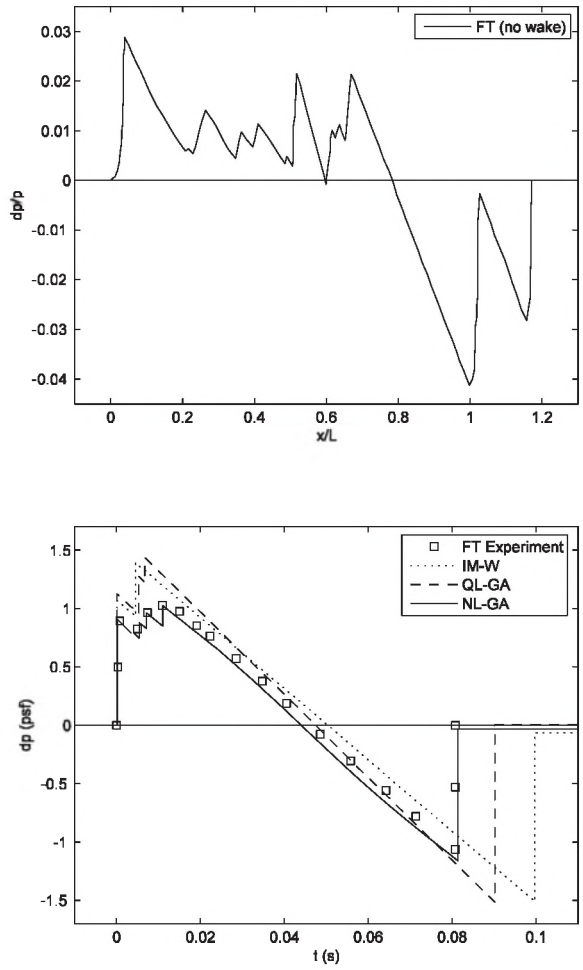


Figure 1. Initial and ground pressure signatures (wake removed)

The initial pressure signature ΔP_h is measured at a distance $r_h = 2L$ from the aircraft's trajectory during a flight test, by employing a follower aircraft. Along with the initial pressure signature, results for the ground pressure signature are shown in Figure 1 and compared with that originally obtained directly from de-turbulenced ground measurements during the flight test. The initial pressure signature has been modified by neglecting the aircraft wake contribution. For this case the ground pressure signatures calculated by the quasi-linear model does not fully agree with the measured one, in terms of amplitude of the front shock system and extension; on the contrary, the ground pressure signature calculated by the nonlinear geometrical acoustic model agrees well with the measured one.

4. DISCUSSION

Further comparisons with previously published results, obtained with well established propagation codes, show that the differences in the ground pressure signatures predicted by quasi-linear and nonlinear propagation models increase dramatically with increasing the shock intensity and the aircraft speed, due to strongly nonlinear flow in the near-field, and with decreasing the aircraft length and increasing the flight altitude, due to cumulative effect of the small nonlinearities of the flow in the far-field over a long propagation distance. Good consistency and poor sensitivity of the propagation models to the initial conditions was also found. Although simple, the proposed nonlinear method allows a very efficient and accurate prediction of the boom propagation starting from a given pressure signal in the near-field and can therefore be considered a useful tool for the aerodynamic design and multi-objective optimization of low-boom supersonic aircrafts via CFD methods.

REFERENCES

[1] R. Seebass, Sonic Boom Theory, J. Aircraft, vol. 6 (3), 177-184 (1969).

[2] M. Friedman, E. Kane, A. Sigalla, Effects of Atmosphere and Aircraft Motion on the Location and Intensity of a Sonic Boom, AIAA J., vol. 1 (6), 1327-1335 (1963).

[3] A. George, K. Plotkin, Sonic Booms Waveforms and Amplitudes in the Real Atmosphere, AIAA J., vol. 7 (10), 1978-1981 (1969).

[4] H. Carlson, W. Middleton, A Numerical Method for Calculating Near-Field Sonic Boom Pressure Signatures, NASA-TN-D-3082 (1965).

[5] D.G. Randall, Sonic Bang Intensities in a Stratified, Still Atmosphere, J. Sound Vibr., vol. 8 (2), 196-214 (1968).

[6] A. Pierce, C. Thomas, Atmospheric Correction Factor for Sonic-Boom Pressure Amplitudes, J. Acoust. Soc. of America, vol. 46 (5), 1366-1380 (1969).

[7] C. Darden, Sonic Boom Theory: Its Status in Prediction and Minimization, J. Aircraft, vol. 14 (6), 569-576 (1977).

[8] W.D. Hayes, Propagation of Sonic Boom through a Stratified Atmosphere, NASA SP-228 (1969).

[9] J. Morgenstern et al., F-5 Shaped Sonic Boom Demonstrator's Persistence of Boom Shaping Reduction through Turbulence, AIAA-2005-0012 (2005).

VALIDATION OF COMSOL MULTIPHYSICS AND ACOUSTICAL PERFORMANCE OF SPLITTER-SILENCERS

Ramani Ramakrishnan
Department of Architectural Science, Ryerson University, Toronto, ON, Canada. E-mail: rramakri@ryerson.ca

1. INTRODUCTION

One dimensional as well as pseudo two-dimensional models have been successively applied in the past to study the acoustical propagation in HVAC system ducts fitted with splitter-silencers[1, 2, 3]. Even though comparisons between experiment and theoretical finite-element models of the above formulations have been shown to be reasonably accurate within an engineering perspective, the one-D and two-D models were limited since elbows and staggered splitters were not amenable with the above modeling techniques. COMSOL multi-physics provides a powerful three-dimensional application software to solve a variety of different passive silencer designs[4]. As a first step, COMSOL was used in a two-D model mode to solve the splitter silencers whose results have been published in the literature from late 1980s and early 1990s. The results of the above exercise will be presented in this paper.

2. BACKGROUND

The schematic details of a rectangular splitter silencer are shown in Figure 1. The baffle splitters are symmetric in Figure 1. The splitter silencer details are: baffle depth is 'd'; the open air-way width is '2h'; and the silencer length including the diffuser portion is 'l'. Other combinations of rectangular silencers were presented in Ramakrishnan and Watson[1].

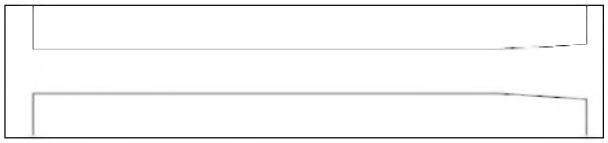


Figure 1. Details of a Rectangular Splitter Silencer

The sound propagates along the centre axis from left to right. The baffle materials are bulk reacting and hence appropriate complex wave speed and material density (complex in this case) can be obtained from References 2 and 3. The mathematical modelling details were presented in Reference 1.

2.1 Results of Ramakrishnan and Watson[1]

The mathematical model of Reference 1 assumed $e^{-\Gamma x}$ variation along the propagation axis where Γ is the complex wave number and was obtained for each of the propagating modes in the silencer configuration. The propagation assumption rendered the model into a pseudo-2-D model. Additional correction factors for reflection of sound at the silencer exit were approximated by using the expansion chamber characterization of Reference 2. Cubic finite-element discretization was used in the 'Y' (normal to the propagation direction) to solve for the complex wave number. Sample results for two silencers are presented in Table 1. The agreement between test data and the pseudo-2-D model is excellent except at high frequencies. The main reason for discrepancy is the lack of sufficient elements in the discretization process.

3. COMSOL MODEL

The splitter geometry is modelled in a 2-D configuration as shown in Figure 1. The splitter material is assumed to be isotropic and homogeneous fibrous material of given flow resistivity, 'σ'. The acoustic propagation in the splitter material uses the complex propagation constant and complex density of bulk reacting material. The inlet of the silencer has a given acoustic field and the sound is assumed to be connected to a long anechoic termination. To accommodate high frequency calculation, COMSOL suggests using a length of pipe in front of the silencer within which scattered acoustic field is calculated to provide the required acoustic field at the inlet of the silencer. The results of the acoustic propagation from COMSOL model are presented in the next section.

4. RESULTS AND DISCUSSION

Only a sample set of results are presented in this paper. The sound pressure levels within the splitter silencer are shown in Figure 2. The absorbing qualities of the fibrous material are clearly evident in Figure 2. The decay of sound levels at the silencer outlet can also be seen in the figure. One can also discern the appropriate behaviour of the anechoic termination providing a level of confidence in the COMSOL results.

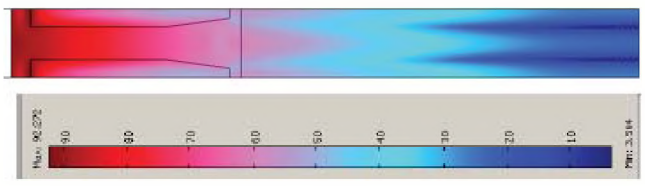


Figure 2. Sound Propagation inside Silencer #2 @ 1050 Hz.

The Insertion Loss, IL, of the silencer is given in Equation (1) below.

$$IL = \frac{W_{in}}{W_{out}}, dB \qquad (1)$$

where W_{in} is the sound power at silencer inlet and W_{out} is the sound power at the silencer outlet.

IL for silencer No. 2 is shown in Figure 3 below. Insertion Loss for a splitter silencer usually peaks at a frequency, based the following three parameters – N_1 =d/h; R = σd/z_0; and μ =2fh/c_0, where z_0 is the characteristic impedance of air and c_0 is the speed of sound in air. The IL smoothly decays around the peak frequency. Details of the three parameters are given in Reference 1.

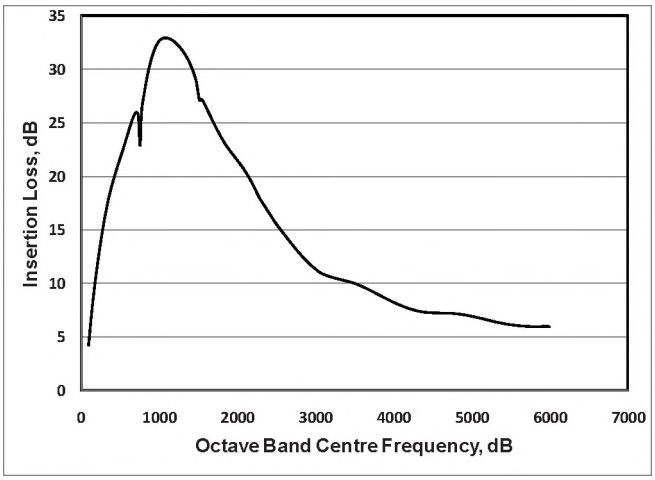


Figure 3. Insertion Loss of Silencer #2.

Similar behaviour is seen in the IL results of Silencer 2. The peak frequency for Silencer 2 is around 700 Hz. IL data increases quickly from 5 dB around 100 Hz to the peak value of 34 dB and decays slowly to 6 dB around 6000 Hz.

The insertion loss result at each octave band is calculated by summing the results within each band (the results were calculated with a constant band width of 20 Hz). The details of the summing process are given in Bies and Hansen[3]. IL results for two sample silencers are presented in Table 1. The results show that COMSOL Multiphysics model produces similar results as Reference 1 and compares well with the measured data.

5. CONCLUSIONS

Insertion loss results of rectangular splitter silencers were evaluated using two-D representation in commercial application software, COMSOL Multiphysiscs[4]. The results were compared to test data as well as predictions obtained applying a simple one-D cubic finite element model to solve the governing equations. The results show excellent agreement. COMSOL model results were seen to be closer to the test data similar to the results of Ramakrishnan and Watson[1].

6. ACKNOWLEDGEMENTS

The assistance of Arvind Chandrasekharan, a doctoral student at Concordia University, Montreal, in setting up the COMSOL modules and mesh sizes, is duly acknowledged.

REFERENCES

1. R. Ramakrishnan and W. R. Watson, "Design Curtves foir Rectangular Splitter Silencers," *Applied Acoustics*, Vol 35, No.1, pp. 1-24 (1992).
2. L.L. Beranek and I.L. Ver, *Noise and Vibration Control Engineering*, John Wiley and Sons, New York, (1992).
3. D. A. Bies and C. H. Hansen, *Engineering Noise Control- Theory and Practice*, 3rd Edition, Spon Press, London. (2003).
4. COMSOL Multi-Physics Application Software, Boston, MASS, Version 3.5a. (2008).

Table 1. Comparison of Insertion Loss Results.

Silencer Type		Insertion Loss, dB @ Octave Band Centre Frequency					
		125	250	500	1000	2000	4000
No 1. Unit Size = 305 mm, d/h = 1, and Length = 1525 mm	IL - Reference 1	4	12	26	44	36	13
	IL - Measurements	4	12	27	41	37	20
	IL – COMSOL Multiphysics	4	12	25	46	35	14
No 2. Unit Size = 408 mm, d/h = 1.13, and Length = 1525 mm	IL - Reference 1	6	14	21	29	17	8
	IL - Measurements	5	12	20	26	16	9
	IL – COMSOL Multiphysics	7	14	22	30	17	8

FIELD IMPACT INSULATION CLASS (FIIC) TESTING OF HARDWOOD FLOORING ON A VARIETY OF RESILIENT UNDERLAYMENTS IN A CONCRETE CONDOMINIUM BUILDING

Andrew Williamson
Wakefield Acoustics Ltd., 301-225 Oak Bay Ave., Victoria, British Columbia, Canada, V8R 1G5
andrew@wakefieldacoustics.com

1. INTRODUCTION

The use of hardwood flooring in new condominium buildings and the replacement of carpeting with hardwood flooring in existing buildings can result in inadequate impact noise insulation between vertically adjacent units. In Victoria, many condominium developers and strata councils are grappling with the decision of whether or not to allow hardwood floors in their buildings. This typically proves to be a difficult decision. If the developer or strata council do not permit the installation of hardwood flooring they risk turning away prospective buyers or dissatisfying strata members. If they do, they risk the creation of a serious noise problem which, in *Wakefield Acoustics Ltd.'s* experience, can lead to persistent noise complaints and, in some cases, threats of legal action.

This paper presents the results of a series of FIIC tests *Wakefield Acoustics Ltd.* recently conducted in a concrete condominium building in Victoria, B.C. The tests were conducted to assist a developer in deciding whether or not to allow the use of hardwood flooring in the building. Nine FIIC tests were conducted in which a sample of hardwood flooring was placed upon nine different resilient underlayments. A tenth test was conducted where a different sample of hardwood flooring was placed upon one of the previously tested underlayments.

2. METHODOLOGY

The ten FIIC tests were conducted in accordance with ***ASTM E 1007-04^{e1}*** with the following exceptions:

1. The hardwood flooring and underlayment samples did not cover the entire test room floor. This discrepancy may have affected somewhat the resulting impact noise spectra measured in the receiving room.

2. The volume of the receiving room was smaller (46 m^3) than required by the standard (60 m^3).

3. The tapping machine was not operated in four different positions because the hardwood flooring and underlayment samples did not provide sufficient area for multiple positions.

The tests were conducted with a CESVA MI005 Tapping Machine which conforms to the specifications of ISO 140-6, 140-7 and 717-2. All sound level measurements were performed using a Larson-Davis 2800 Real Time Analyzer - a Type 1 sound level meter. The "test floor" consisted of an approximately 1.22 m x 1.83 m sample of hardwood flooring placed over an underlayment sample on a 203mm concrete slab. The hardwood and underlayment samples were simply placed on the concrete surface and were not adhered to either the concrete or each other.

The types of underlayments tested were as follows:

1. Acoustitech Premium
2. Acoustitech Maxima
3. Durason
4. Silanzer LJ
5. Safe and Sound
6. Acoustik
7. Echo Silencer
8. Thick Cork
9. KRAUS

For these nine tests, a 14 mm thick sample of *Giorgina Engineered Hardwood* was placed over the underlayments. For the tenth test, an 8 mm thick sample of *James River Collection Prefinished Hardwood* was placed over the *Acoustitech Premium* underlayment.

3. RESULTS

Table 1 and Figure 1 present the results of the FIIC tests.

Test #	Hardwood	Underlayment	Field Impact Insulation Class (FIIC)
1	14 mm Giorgina Engineered Hardwood	Acoustitech Premium	57
2	14 mm Giorgina Engineered Hardwood	Acoustitech Maxima	57
3	14 mm Giorgina Engineered Hardwood	Durason	58
4	14 mm Giorgina Engineered Hardwood	Silanzer LJ	57
5	14 mm Giorgina Engineered Hardwood	Safe and Sound	58
6	14 mm Giorgina Engineered Hardwood	Acoustik	58
7	14 mm Giorgina Engineered Hardwood	Echo Silencer	59
8	14 mm Giorgina Engineered Hardwood	Thick Cork	58
9	14 mm Giorgina Engineered Hardwood	KRAUS	58
10	8 mm James River Collection Prefinished Hardwood	Acoustitech Premium	60

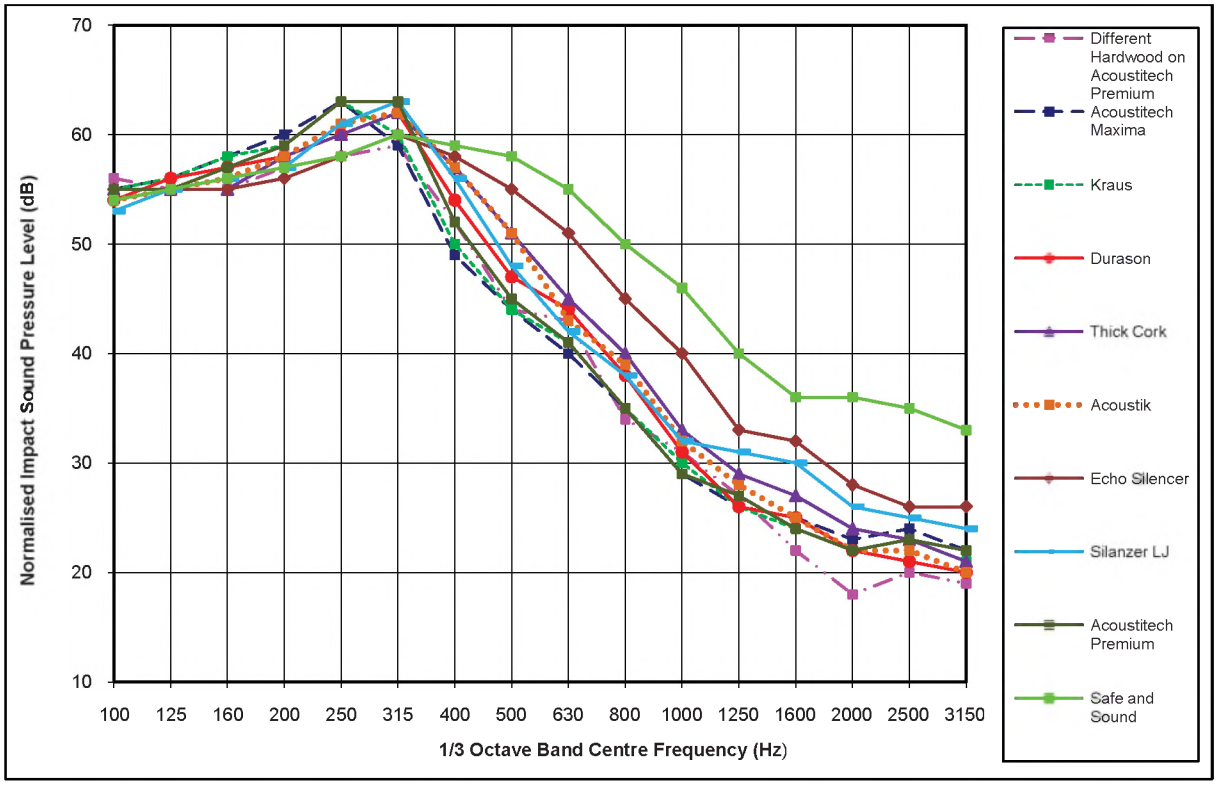


Fig. 1. Comparison of the Normalized Impact Sound Pressure Levels of the ten FIIC tests

4. DISCUSSION

Comparison of Underlayments

From Table 1 it can be seen that the FIICs achieved when the *Giorgina Hardwood* sample was placed over the various underlayments range from 57 to 59. This uniformity in FIIC ratings can be explained by referring to Figure 1 where it can be seen that there is a "spike" in the Normalized Impact Sound Pressure Levels (NISPLs) within the 250 and 315 Hz one-third octave bands. This "spike" controlled the FIIC rating of all the underlayments tested. While there is little variation in the FIIC ratings, there is some variation in the NISPLs, and particularly at frequencies above 400 Hz. So despite the underlayments achieving very similar FIIC ratings, the "character" of the impact noise experienced will vary. For example, while the *Safe and Sound* and *Thick Cork* underlayment both achieved FIIC ratings of 58, it can be seen that the NISPLs measured with *Safe and Sound* are significantly greater than those of the *Thick Cork* at frequencies above 400 Hz. As such, impact noise experienced beneath a concrete floor surfaced with hardwood flooring would have a much "brighter" character if the underlayment used were *Safe and Sound* rather than *Thick Cork*.

Comparison of Hardwoods

The tenth test conducted involved retesting the *Acoustitech Premium* underlayment with a different hardwood sample (8 mm *James River Collection Prefinished Hardwood* instead of 14 mm *Giorgina Engineered Hardwood*). It can be seen from Table 1 that changing the hardwood resulted in a 3 point improvement in the FIIC. It is interesting to note that the biggest improvements (4-5 dBA) occurred within the 250 and 315 Hz bands where the "spike" in NISPLs was observed. While it cannot be known for certain why the change in hardwood had such an effect, it may be that the 250 / 315 Hz "spike" was more a function of the resonant frequency of the hardwood samples than any properties of the underlayments. The small sample size of the hardwood flooring may have also influenced the 250 / 315 Hz spike. If the hardwood and underlayment had been actually installed (i.e. adhered to and covering the entire concrete floor), the resonant frequency of the hardwood would most likely be lower in both frequency and magnitude.

Wakefield Acoustics Ltd. typically recommends that, to avoid significant impact noise annoyance, the FIIC of a floor/ceiling system should at least exceed 60 and preferably exceed 65. Based upon these tests, however, it is unclear whether an FIIC of 60 can be consistently achieved, or whether an FIIC of 65 can be achieved at all, when hardwood flooring is used in a concrete building.

EXPERIMENTAL INVESTIGATION OF THE EFFECTS OF ABSORPTIVE SURFACES ON THE ACOUSTICAL PERFORMANCE OF A BARRIER IN AN ANECHOIC CHAMBER

Shira Daltrop[1], Murray Hodgson[1], and Clair Wakefield[2]
[1]Dept. of Mechanical Engineering, UBC, 2206 East Mall, Vancouver, BC, Canada, V6T 1Z1
[2]Wakefield Acoustics, 2250 Oak Bay Ave, Victoria, BC, Canada, V5N 1A5

1. INTRODUCTION

Roadside noise barriers are an effective way to reduce unwanted traffic noise reaching nearby areas. Barrier performance can be limited by sound reflections from the source side of the barrier, especially in the case of parallel barriers. Making barriers sound absorptive will decrease reflections from the barrier and amplification between parallel barriers. Absorbent barriers are generally more expensive than reflecting ones, and therefore it is important to know the optimum amount and placement of absorbing treatment on a barrier.

In this project, a full-scale reflective barrier was built in an anechoic chamber and absorptive material was added to create an absorptive barrier. The insertion loss (IL) was measured in third-octave bands for various configurations of the reflective and absorptive barriers; from these, total A-weighted values relevant to a typical traffic-noise spectrum were calculated.

2. METHOD

The tests were done in an anechoic chamber with dimensions 4.1 m x 4.7 m x 2.6 m to approximate outdoor conditions. A 3.66 m x 3.66 m plywood floor was built on top of the wire mesh floor, as a strong support and reflective surface for the wall to be placed on, representing a hard ground. The wall was 3.66 m long and 1.22 m high, and was made from 12-mm drywall on either side of a wooden frame, made from 2x4's. An extra layer of 12-mm drywall was screwed onto the source side of the wall to increase the transmission loss. The cavity was filled with fiberglass batt insulation to reduce sound transmission. Air gaps between the barrier and the floor were filled with putty. The wall was set up in the center of the floor in the chamber.

The sound source was placed 1 m behind the barrier, at heights of 0.25 and 1.0 m. The sound sources used were omni-directional point sources - one for high frequency (above 500 Hz) and the other for low frequency (below 500 Hz) [1].

Figure 1 shows the test configuration. The receiver was placed at distances of 0.6, 1.2 and 1.8 m on the non-source side of the barrier, at heights of 0.2 and 1.05 m. The receiver was also placed 1.8 m behind the barrier at a height of 0.75, and 1.2 m behind the barrier at a height of 0.9 m. This allowed for examination of both different angles of diffraction as well as varying distances at a constant angle.

A B&K 1/2" type 4190 microphone was used, along with a B&K type 2669 preamp. A SINUS Soundbook was used as the white-noise signal generator and the analyzer, allowing both narrow-band and limited-band analysis.

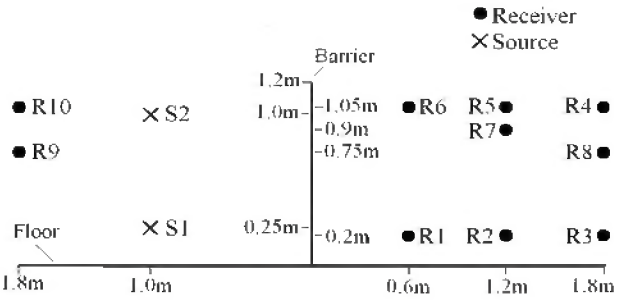


Fig. 1. The barrier configuration, including source and receiver positions.

The absorptive material used on the wall was cotton acoustic baffles. Each baffle was 1.22 m x 0.61 m and 25-mm (1") thick. Two layers of these baffles were used, giving a total thickness of 50 mm. The random-incidence absorption coefficient of these baffles was found using two methods: the spherical decoupling method [2] and the impedance-tube method [3].

3. RESULTS

The baffles were attached to the wall using insulation hangers; they were attached in five different configurations, which covered the barrier as follows: the full source side; the top half of the source side; the full receiver side; the top half of the receiver side; and the top halves of both the source and receiver sides. The ILs of these configurations and of the reflective wall were measured.

The IL in third-octave bands at receiver position R6 is shown in Figure 3 for the different baffle configurations. For the low source position, the absorptive material is more effective at increasing the IL in the frequency ranges where the baffles are highly absorptive. Between 300 and 400 Hz, where the absorption coefficient of the baffles is highest, covering the top halves of both sides increases the IL by 7-8 dB over that of the reflective wall, a 4 dB increase over any other absorptive configuration. At frequencies above 1000 Hz, where porous absorbers are expected to be highly absorptive, there is a 3-5 dB increase in IL due to having the top halves of both sides covered in baffles. For the high

source position this high-frequency increase in IL is again seen; however below 1000 Hz there is very little increase in IL due to the absorptive treatment, regardless of the configuration. This is due to the smaller diffraction angles that are present for the high source position. These trends are consistent for the other receiver positions.

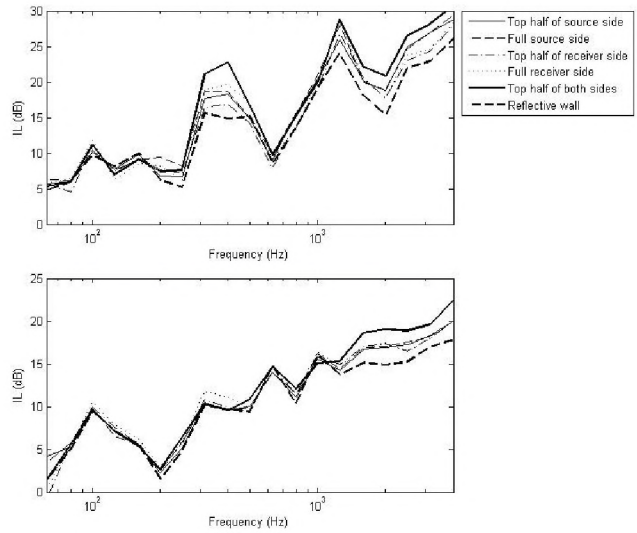


Fig. 2. Measured ILs at receiver position R6 for different baffle configurations, for low (top graph) and high source positions.

Figure 4 shows the total A-weighted IL at each receiver position for both the low and high source positions for each baffle configuration. The reflective wall provides 12-18 dB of attenuation, and the overall IL improvement due to absorption is 1-2 dB. It is seen that placing the absorptive material on the top half of both sides of the wall produces the highest IL for both source positions, while the reflective wall produces the lowest IL, as expected. Covering the full side of the barrier compared to covering half - either the source or receiver side - shows an IL increase of approximately 0.5 dB. Covering the full side of both sides of the barrier was not tested, although based on the other measurements an improvement in IL of 2-3 dB over the reflective wall could be expected.

For the higher source position, R1, R2 and R3 show the highest IL. These are the lowest receiver positions, creating the greatest diffraction angles and greatest path-length differences. Receiver R4 has the lowest IL, as it is the highest and farthest receiver position from the wall, giving it the lowest diffraction angles and the smallest path-length difference. Receivers R6, R7 and R8 are positioned such that they have the same angle of diffraction, and therefore the IL is expected to increase with path-length difference. This is what is seen in the results, increasing from R6 to R8.

For the lower source position, the IL at all receiver positions is increased compared to the higher source position, as expected due to the larger diffraction angles. However the results are not what is expected at several receiver positions. The IL decreases from R6 to R8, as opposed to what is expected. The IL at R3 is higher than at both R1 and R2, where it is expected to be lower. There may be interactions between the direct and reflected waves, either constructive or destructive interference, which affect the sound levels at particular frequencies for different receiver positions.

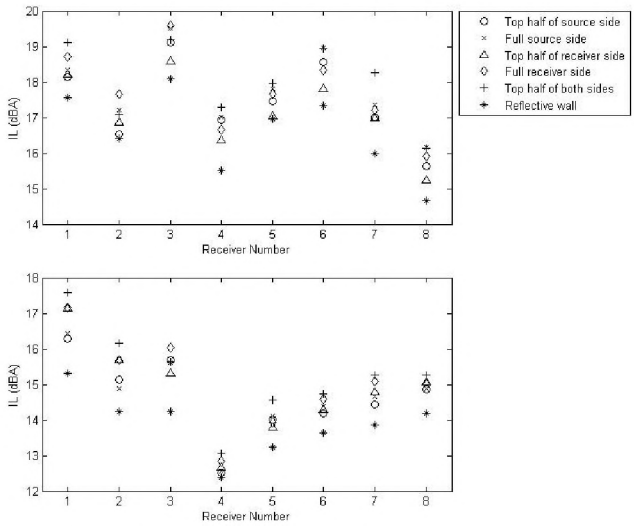


Fig. 3. Measured A-weighted ILs for the low (top graph) and high source positions at the eight receiver positions on the non-source side of the barrier for different baffle configurations.

4. CONCLUSION

Different configurations of absorbent material on a reflective stud wall in an anechoic chamber were examined. It was found that the highest IL came from covering the top half of both sides of the barrier with absorbing material. An IL of 12-18 dB was found for the reflective barrier. Absorptive material improved the IL by 1-2 dB in many of the configurations.

REFERENCES

[1] C. Bibby. Point-source design and performance. Technical Report, University of British Columbia, 2009.

[2] S. Daltrop. Measurement of the acoustical characteristics of surfaces using the spherical decoupling method. Course Project Report, University of British Columbia, 2010.

[3] ASTM C 384-04. Standard Test Method for Impedance and Absorption of Acoustical Materials by Impedance Tube Method. 2004.

FEATURES OF LOW FREQUENCY WIND TURBINE SOUND

Werner Richarz[1], and Harrison Richarz[2]
[1]Aercoustics Engineering Limited, 50 Ronson Dr., # 165, Toronto, ON, Canada werner@aercoustics.com
[2]University of Leicester, Leicester, United Kingdom, hfr1@lei.ac.uk

1. INTRODUCTION

A lot has been written about low frequency (LF) noise emitted from wind turbines. It is relatively straight-forward to document physical parameters such as the sound pressure and its spectrum. Certain individuals experience or claim to experience adverse impacts from wind turbine sound ranging from annoyance to serious health effects [1,2]. There appears to be no supporting evidence for the latter [3,4].

Although low frequency sound emitted from wind turbines continues to be the subject of much speculation, objective descriptors of the sound are lacking. Aercoustics Engineering has tried to address this by examining several aspects of sound characteristics of large wind turbines. First, amplitude modulation was identified as being attributable to source motion and source directivity [5]. The phenomenon of amplitude modulation is not a low frequency phenomenon: the principal sound energy is well above 200 Hz, which is commonly used as an upper bound for LF sound. Some dynamic features of low frequency sound and possible means of their measurement have been discussed in reference [6]. Aspects of the cyclic signal, associated with the blade passage frequency are expanded upon herein.

2. WIND TURBINE CHARACTERISTICS

The rotating lift forces generated by the motion of the wind tunnel blades extract energy from the wind, some of which is converted to electric power. Large commercial wind turbines have rotor diameters of the order of 80 to 100 m and are mounted on 80m to 120m high towers. Rated power ranges from 1.2 MW to 6MW. The majority of wind turbines installed in Canada are 1.8 - 2 MW. The nominal sound power of such machines is of the order of 104 dBA.

Typical rotation rates (N) are 9 to 15 RPM, so that the tip speed of a 90m diameter rotor rotating at 12 RPM is ~ 56m/s. Thus the 'sonic' circle of the hydrodynamic pressure field is of the order of 270m. This is the extent of the acoustic near field of the wind turbine [$R_{NF}=60c/(\pi N)$].

There is a wealth of literature on the sound emitted by propellers and (helicopter) rotors. Tip speeds and disk loading are much greater than experienced by wind turbine rotors. Nevertheless, all of have similar low frequency spectra that are dominated by the blade passage tone (nB, n=N/60) and higher harmonics. The relative strengths are governed by the pressure distribution on the blade.

For a variety of reasons, the wind turbine blade passage tone and its harmonics have received little attention. From the perspective of conventional noise impact assessment, they make no sensible contribution to the ubiquitous A weighted Leq. or L90. Most measurement and analysis systems cannot cope with frequencies of the order of 0.6 Hz.

Even when low frequency spectra are measured, they may be masked by contributions from (broad-band) wind noise. For this reason we have elected to analyze low frequency wind turbine sound in the time domain.

3. AUTOCORRELATION:CYCLIC SOUND

The auto-correlation function is ideally suited to extract cyclic signals from a time series that has both periodic and random components:

$$p(t)=n(t)+\Sigma a_m \cos(2\pi mt/T)+b_m \sin(2\pi mt/T)$$

The autocorrelation $R_{pp}(\tau)=\langle p(t)p^*(t-\tau)\rangle$ takes the form:

$$R_{pp}(\tau)=\Sigma\, c_m \cos(2\pi m(t-\tau)/T)+ R_{nn}(\tau),\ c_m=0.5(a_m^2+b_m^2)$$

The time scale of the random noise is typically short, so that $R_{nn}(\tau)$ vanish for τ large. The cyclic terms persist. The autocorrelation has a global maximum at zero time delay (t=0). It is customary to normalize autocorrelations by the $\langle p(t)p^*(t)\rangle$ so that $R_{pp}(0)=1$.

4. MEASURED AUTOCORRELATIONS

In order to capture the essence of the low frequency sound the analysis window should be capture at least 100 blade passage events. A large number of files of wind turbine sound have been analyzed. Conventional 1/2 inch ICP microphones were used. The data was recorded on a Rion DA-20 digital data recorder. The frequency response of the data acquisition system was 'flat' to 10Hz. A typical LF spectrum is shown in figure 1. The first few harmonics can be readily identified. Higher harmonics are obscured by wind noise.

A measured autocorrelation of the wind turbine sound is shown in figure 2. The maximum is at zero time delay. This peak vanished rather quickly. Distinct pulses are also evident at time delays corresponding to the reciprocal of the blade passage frequency. These pulses form the auto-correlation signature of the cyclic portion of the wind turbine sound.

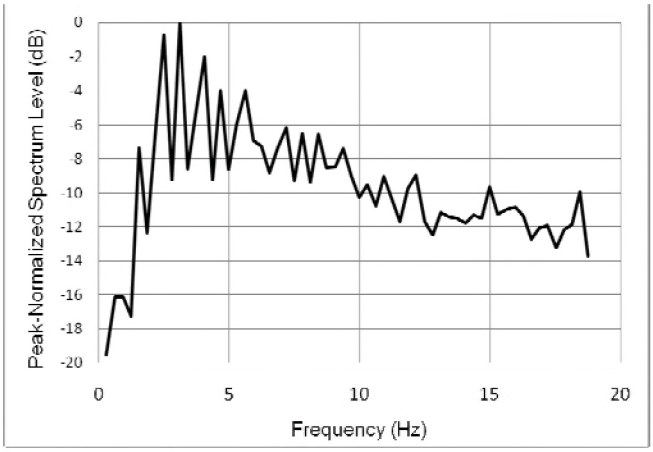


Figure 1. Low frequency wind-turbine spectrum

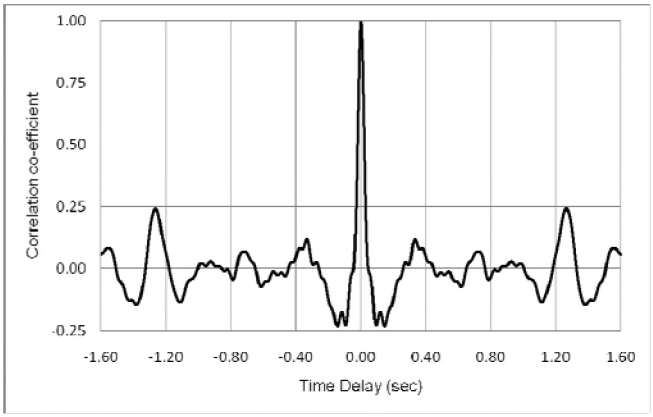


Figure 2. Measured autocorrelation of wind-turbine sound

The shape of the pulse can be image can be enhanced by time-shifting one of the records by $\Delta\tau=(nB)^{-1}$(Figure 3) This eliminates the error introduced by the so-called bow-tie correction that must be applied to FFT based autocorrelations.

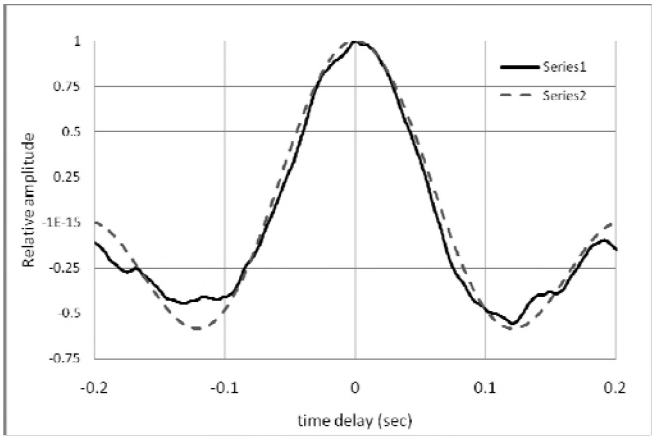


Figure 3. Cyclic pulse

The function $(1-(t/\Delta\tau)2)\cos p\ \tau/\Delta\tau)$ fits the cyclic pulse for $-\Delta\tau<\tau<\Delta\tau$. This then permits one to estimate the Fourier series coefficients (A_m) of the cyclic pulse. The series for A_m exhibits `structure` for small m, evolving into a regular pattern with slowly decaying amplitude (Figure 4). For each index m there is corresponding a frequency (f_m). For the data presented here $f_m \sim 0.8$ m Hz.

5. DISCUSSION

Since the tonal amplitudes decay at a rate of about 6dB/octave, the cyclic signal may well be audible, especially if the background noise at the point of reception is low. Although the human hearing is not very acute at low frequencies, the higher harmonics of the cyclic signal are well within the audible range. This is analogous to the well known case of large pipe organs where the fundamentals of the harmonics are below the nominal threshold of hearing.

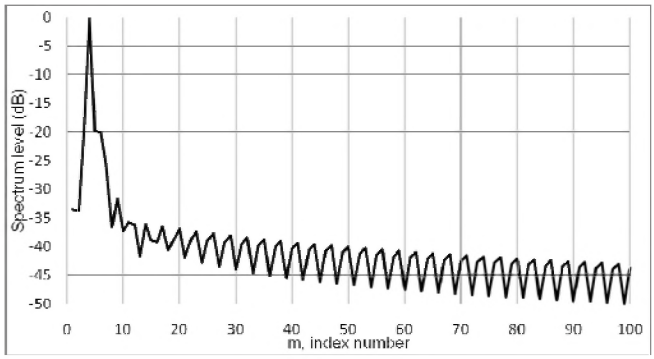


Figure 4. Spectrum level of cyclic pulse harmonics.

6. REFERENCES

[1] van den Berg, F., Pedersen, E., Bouma, J., Bakker, R., *Windfarm perception, visual and acoustic impact of wind turbines farms on residenst*, FP6-2005-Science and Society-20, Final Report, 2008.

[2] Pierpont, N., pre-publication draft, *Wind turbine syndrome:a report on a natural experiment*. http://www.windturbinesyndrome.com, 2009.

[3] Leventhall, G., *A review of published research on low frequency noise and its effects*.,Dept. For Environment, Food and Rural Affairs, 2003.

[4] Colby. W., et al. *Wind turbine sound and health effects, an expert panel review*. AWEA-CanWEA, 2009.

[5] Richarz, W., Richarz, H., *Wind turbine noise diagnostics*., Proc.3rd. Int. Mtg. on Wind Turbine Noise, Aalborg, 2009.

[6] Richarz, W., Gambino, T., *Low frequency noise monitoring of a wind turbine*, Proc. 14th Int. Mtg. Low Freq. Noise and Vib. and its Control, Aalborg, 2010.

POST CONSTRUCTION HVAC NOISE CONTROL

Werner Richarz[1], and Harrison Richarz[2]
[1]Aercoustics Engineering Limited, 50 Ronson Dr. #165, Toronto, ON, Canada, werner@aercoustics.com
[2]University of Leicester, Leicester, United Kingdom, hfr1@lei.ac.uk

1. INTRODUCTION

Newly installed HVAC systems that fail to meet acoustic design parameters share the common feature that the fault is detected just prior to occupancy. There are many configurations that give rise to the additional noise and for most there may be no effective and in-expensive mitigation. There is also some friction between the various parties: owner, architect, mechanical designer, mechanical contractor, and the acoustic consultant. The latter is expected to provide an instant solution.

2. SOURCE IDENTIFICATION

HVAC systems tend to evolve in time: differing from the lay-out set in the design phase to the as built configuration. Issues such as lack of clearance or cost constraints result 'creative' on site solutions that fit but add noise. Before recommending noise mitigation, it is necessary to identify the source of the additional noise. This is important, as a trial and error approach is really not appropriate. Experience certainly is an asset. Listening and selected sound measurements provide additional information, but may not suffice.

A two point space-time correlation is an effective tool for localizing sound sources. The concept is straight-forward in that the time of travel from the source to the listeners is measured. For low frequencies sound propagation in HVAC ducts is effectively one-dimensional. It is in this frequency range that a large percentage of the excess HVAC noise is generated. Consider a noise source at Y in a duct and transducers at X1 and X2 (Figure 1).

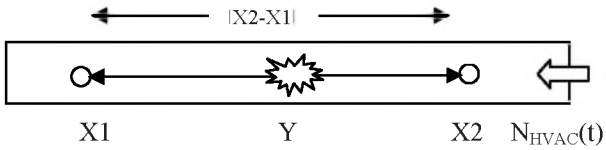


Figure 1 Schematic of HVAC system

transducer locations (X1, X2) and source location (Y)

The pressure has several components: $N_{HVAC}(t)+N(t)+S(t)$. $N_{HVAC}(t)$ is the HVAC system noise, N(t) uncorrelated noise (some of which due to the microphone-flow interaction), and S(t) the sound generated by the noise source. Referring to figure 1 it is evident that the HVAC noise detected at X2 will be measured |X2-X1|/c seconds later at X1, and $R_{HVAC}(t)$ has a maximum at t=|X2-X1|/c. Similarly, $R_{SS}(T)$ has a maximum at t=|X2-Y|/c-|X1-Y|/c.

It follows that the correlations is the sum of two distinct contributions: $R_{12}(\tau)=R_{HVAC}(\tau)+R_{SS}(\tau)$. The magnitude of the correlation is a measure of the similitude of the two signals and the coherent energy in relation to the overall signal energy. This then permits one to assess source location and the relative source strengths [1].

The measured NC rating of an occupied space exceeded the specification by almost 10 points. There was a difference of opinion as to the cause of the additional noise. One party claimed that fan noise was the issue, while others suspected a change in duct configuration near the supply air (SA) and return air (RA) openings of the occupied spaces.

Cross-correlations were measured to determine the source location and relative strength. The transducer locations were selected to detect fan noise, as well as noise generated near the SA and RA openings (Figure 2). The measured cross-correlations are shown in figures 3 and 4 with time delay replaced by an (equivalent) length. The figures support the conclusion that the sound is generated near the duct opening. An up-grade to the fittings resolved the noise problem.

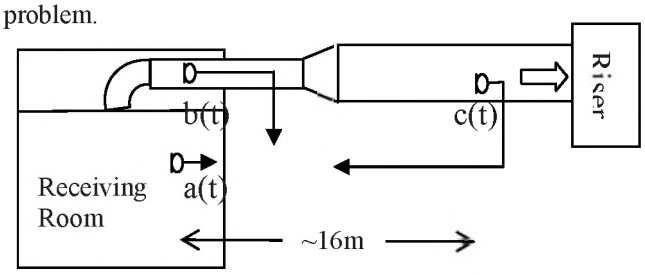


Figure 2. Cross-correlations of a(t),b(t),c(t) determine source location

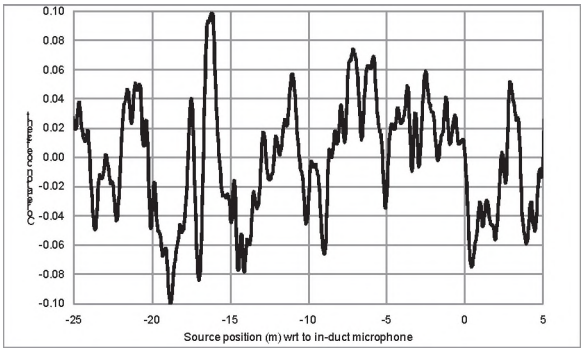


Figure 3, Cross-correlation Rac(t), low correlation indicates sound from riser is not a major source

3. LOW FREQUENCY RUMBLE

The most common acoustic deficiency in occupied spaces is an excess of low frequency noise (125Hz or 250 Hz octave band). Typically a reduction of the order of 10 dB is required. Review of narrow band data, measured at a large

number of "problem sites" suggests that attenuation over at most 1/3 octave would suffice. Conventional dissipative silencers may provide the noise reduction needed. However, cost and space constraints are always of concern. Modern active noise control systems are possible alternatives. Cost, the need to provide electrical power, and maintenance are factors that make this solution less attractive.

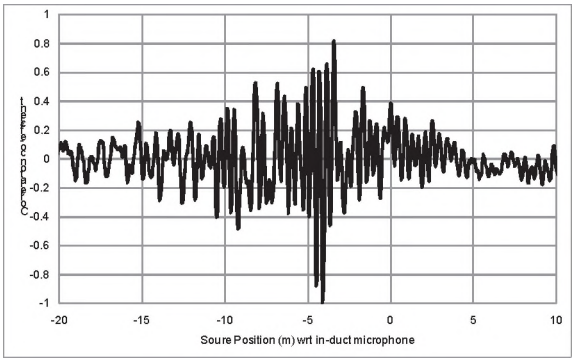


Figure 4. Cross-correlation Rab(t) shows high degree of correlation: source is located near microphone b.

4. HELMHOLTZ RESONATORS

A Helmholtz resonator may provide the relatively modest bandwidth and level reduction that is required. Helmholtz resonators require somewhat more careful design, but have been shown to provide good attenuation. Also manufacturers offer `tuned` sound absorbers that augment the dynamic insertion loss of fan system silencers. The absorber is tuned to the blade passage frequency.

It should be possible to install similar devices near SA and RA openings. A series of scale model tests was conducted by Aercoustics Engineering to obtain information on the performance of a Helmholtz resonator and an engineering design procedure.

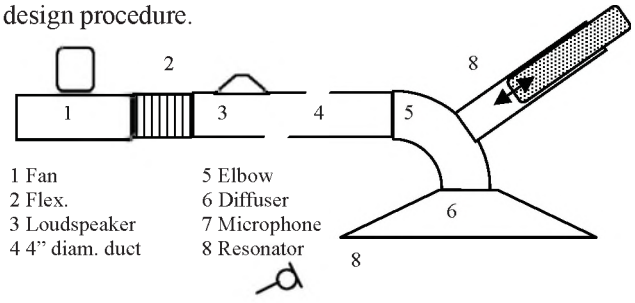


Figure 5. Schematic of measurement facility.

The experimental facility (Figure 4) was 10 cm diameter sheet metal duct. A small centrifugal fan delivered air to the duct, and a loudspeaker was the noise source. The duct was terminated with an elbow and a rectangular diffuser. Sound was measured 30 cm from the diffuser and processed by a dual channel FFT analyzer. A tunable Helmholtz resonator, 5 cm diameter ABS pipe, was installed on the elbow. A tight-fitting piston controlled the volume. The design procedure was based on the method outlined in reference [2]. The throat diameter (2.54 cm) and throat length (0.046cm) were kept constant. The change in sound pressure level afforded by the installation of resonator is interpreted as the effective transmission loss.

Measurements were conducted with and without air flow. The resonator performance was not affected by the flow, which had a nominal speed of 1000 FPM, which is representative of flows near RA and SA openings. The key parameters and performance of each configuration tested is summarized in the table below. The wavenumber ($ka=2\pi fa/c$) and diameter-wavelength ratio ($2a/\lambda$) are also listed so that the potential performance at full scale can be assessed.

Results Table

$f_{meas.}$ (Hz)	L (m)	L_{design} (m)	TL (dB)	ka	$2a/\beta$	$4L/\beta$
843	0.24	0.04	16.5	0.79	0.50	2.38
898	0.21	0.04	10.5	0.84	0.54	2.22
990	0.19	0.03	6.8	0.93	0.59	2.21
358	0.17	0.25	3.6	0.34	0.21	0.72
383	0.15	0.22	4.0	0.36	0.23	0.68
438	0.13	0.17	4.3	0.41	0.26	0.67
463	0.11	0.15	9.1	0.43	0.28	0.60
518	0.09	0.13	7.7	0.49	0.31	0.55
605	0.07	0.09	5.5	0.57	0.36	0.50
713	0.05	0.06	12.2	0.67	0.43	0.42
843	0.03	0.04	15.1	0.79	0.50	0.30
898	0.01	0.04	6.1	0.84	0.54	0.11

When the length of the resonator tube (L) is increased, the resonance frequency diminishes, until wave-effects become important ($4L/\lambda > 1$) Note that the resonator still functions, albeit at a totally different frequency. There is a discrepancy between the actual resonator volume (determined by L) and the design value. Further testing is required, with different throat diameters and throat lengths to establish a definitive design procedure.

5. SUMMARY

Cross-correlations are a useful diagnostic tools to determine source location and source strengths. Helmholtz resonators can provide substantial noise reductions. However, the standard design procedures require some empirical corrections.

6. REFERENCES

[1] Bendat, J.S., Piersol, A.G. *Engineering Applications of Correlation and Spectral Analysis*, Wiley Interscience, New York, 1980.

[2] Beranek, L. (ed.) *Noise and Vibration Control*, McGraw Hill, New York, 1971.

AIR INJECTION VACUUM BLOWER NOISE CONTROL

Tyler L.A. Mose, and Andrew C. Faszer
Noise Solutions Inc., 301, 206 – 7th Avenue S.W. Calgary, Alberta, Canada, T2P 0W7, tmose@noisesolutions.com

1. BLOWER APPLICATIONS

Air injection vacuum blowers are a widely used series of blower utilized in applications with the need of high vacuum level, over 15" Hg. These high pressure blowers can be utilized[1] in applications including: wet and/or dry waste removal (industrial, municipalities, etc.); central vacuum systems; aeration systems; and pneumatic conveyors. Air injection blowers are a lobe rotary configuration that produce high vacuum pressures through using atmospheric air injection to the blower itself to reduce overheating of the gas medium and power absorbed by the blower. A significant complaint related to the use of these blowers is the noise.

2. STUDY OBJECTIVE

Noise Solutions Inc. conducted a Noise Impact Assessment (NIA) on an air injection vacuum blower unit to identify the prevalent noise sources, as well as design custom noise abatement equipment to reduce the amount of airborne noise from the machine during operation. The reason for conducting this study was to investigate, and reduce, concern over occupational health of operators utilizing these blowers, and lessen the degree of residential/pedestrian disturbance during blower operation in urban areas.

3. IDENTIFICATION OF PROBLEMATIC NOISE

Air injection vacuum blowers often utilize diesel engines to rotate the internal 3 lobe assembly, and their respective accessories, creating many noise sources. Prevalent sources with high sound power intensity include the: driving engine casing, engine exhaust, blower exhaust, blower air injection inlet, oil cooler fan, blower casing & manifolds, as well as the blowers air filter & respective piping. Being that this style of blower is generally used in confined spaces many of the noise sources are in close proximity to one another. This proximity causes a challenge discerning the relative contribution of the many individual noise sources to the total noise signature as there is cross contamination between the individual sound fields of each source.

4. ANALYSIS AND QUANTIFICATION OF IDENTIFIED NOISE SOURCES

A Brüel & Kjær 2260 Investigator sound level meter was used to accurately quantify the noise produced by the sources deemed significant in the NIA, which were: the blower casing; the blower exhaust outlet; and the blower air injection inlet. These noise sources were found to be extremely dominant during the blowers operation, and were made the primary objectives of the investigation. The measured sound pressure levels produced by the three sources stated above are presented in Table 1. A tone centered around 142.5 Hz can be seen in Figure 1.

Table 1. Order Ranked Sound Pressure Levels No Blower Silencers

Source	Sound Level Contribution	
	dBC	dBA
Blower Air Injection Inlet	150.6	130.6
Blower Exhaust	150.5	130.5
Blower Casing	133.9	113.9
Sum	153.6	133.6

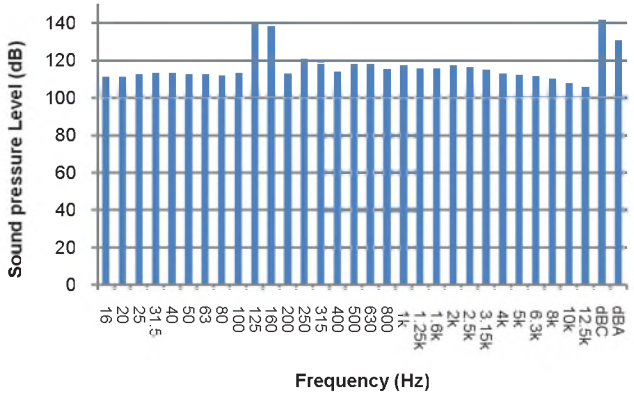


Figure 1. Air Injection Vacuum Blower Exhaust Sound Pressure Levels

5. NOISE ABATEMENT EQUIPMENT DESIGN

A priority in the design of custom noise abatement equipment for this application was to reduce the tone while staying within the design constraints.

5.1 Design Constraints

Being that this style of blower can be used in both mobile and stationary applications, many design constraints are dealing with size and weight restrictions. Each blower unit requires separate silencer bodies for the air injection inlet and exhaust, due to flow pulsations and airflow direction. In this study the blower was mounted on a mobile unit, so each silencer could be no larger than 24" in diameter, nor over 92" long. It was also stated as a design requirement that each silencer body had to weigh less than 600 lbs. Weight restrictions existed for mounting reasons, as well as mobile unit balancing.

Other design constraints included back pressure allowances, as well as sizing the silencer expansion chambers to

specifically attenuate the tone at 142.5 Hz. This blower required that the backpressure added from a silencer and attached piping would not surpass 12" H2O, so this had to be taken into account when designing the silencer internals.

5.2 Engineering

By using equations that describe the sound attenuation of a reactive expansion chamber, the tone at 142.5 Hz was specifically targeted and the sizes of expansion chambers were designed accordingly. Additional reactive and absorptive noise attenuating components were designed to reduce the entire frequency range in additiona to only attenuating the tone. Due to the fact that the sound level intensity is so high coming from the blower exhaust and air injection inlet, it was realized that not only the silencer outlet/inlet have a high level of sound emissions, but the shell of the silencer itself would radiate substantial noise. In order to eliminate this shell radiated noise, combination internal acoustically absorptive and internal acoustical lagging was incorporated into the design.

In conjunction with the acoustic design of the silencers, Computational Fluid Dynamics (CFD) analyses were conducted to optimize silencer backpressure in relation to acoustic performance. The optimum analysis estimated a maximum backpressure of 11.7" H2O under blower full load conditions.

6. POST-INSTALL EVALUATION

A set of 2 prototype silencers were constructed by Noise Solutions Manufacturing, and installed, along with a set of acoustic blankets to lag the blower casing on a blower for testing and determine equipment performance evaluation. Once the blower casing/manifolds were lagged with the acoustic blankets and the silencers were installed; noise, backpressure, and blower operation measurements were completed.

To ensure safe operation of the blower, backpressure measurements were conducted prior to any noise readings. A manometer was used to measure that under blower full load conditions the exhaust silencer had a maximum backpressure of 11.2" H2O.

Post-install noise readings were taken and are presented in Table 2.

Table 2. Order Ranked Sound Pressure Levels with Noise Solutions Blower Silencers and Blankets

Source	Sound Level Contribution	
	dBC	dBA
Blower Casing	106.3	103.0
Blower Exhaust	109.0	100.4
Blower Air Injection Inlet	105.7	98.0
Sum	112.0	105.7

It can be seen in Figure 2 that the 142.5 Hz was significantly reduced and effectively eliminating. This resulted in the level of the noise as well as the annoyance of the blower noise signature being greatly reduced.

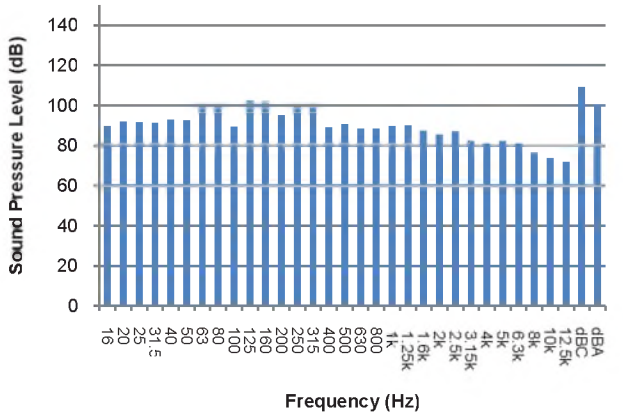


Figure 2. Air Injection Vacuum Blower Exhaust with Noise Solutions Silencer

The installation of the custom noise abatement equipment resulted in a total reduction in sound pressure levels of approximately 27.9 dBA and 41.6 dBC.

REFERENCES

[1]Common blower applications, Retrieved July 30, 2010, from Robuschi website: http://www.robuschi.com/rbdv.asp

PREDICTION OF FLOW-INDUCED NOISE IN AIRCRAFT CYLINDRICAL CABINS

Joana da Rocha[1], Afzal Suleman[1], and Fernando Lau[2]
[1]Dept. of Mechanical Engineering, University of Victoria, PO Box 3055, Stn. CSC, Victoria, BC, Canada, V8W 3P6
jdarocha@uvic.ca, suleman@uvic.ca
[2]Centro 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

Turbulent boundary layer (TBL) is a major source of aircraft cabin interior noise. In fact, jet powered aircraft cabin interior noise is mostly generated by the external flow excitation and engine noise. However, while during takeoff the engine is the dominant source of noise, in cruise flight the airflow sources are the major contribution for the interior noise [1]. As referred in [2], TBL excitation is regarded as the most important noise source for jet powered aircraft at cruise speed, particularly, as new quieter jet engines are being developed. For these reasons, reducing the turbulent flow induced noise in aircraft cabin is an important topic of research. Still, since the TBL is stochastic phenomenon and due to the complexity of the aircraft structure itself, this is an ongoing topic of investigation. In order to successfully design effective noise control systems, a clear understanding of the mechanisms involved in the aircraft cylindrical cabin TBL-induced noise, such as the sound transmission and radiation of the coupled structural-acoustic system, is crucial.

In this context, the main goal of the present paper presents an analytical framework, developed by the authors, for the prediction TBL-induced noise into a cylindrical cabin, and its validation against experimental studies. To validate the analytical framework, several studies were considered for comparison. The acoustic enclosure is of cylindrical shape, filled with air, with simply supported end caps, and with a flexible cylindrical shell. The flexible cylindrical shell is backed by random noise or by turbulent flow. Furthermore, the model was extended for the cylindrical structure divided by several curved panel. The closed-form analytical solution of the coupled response of the system is obtained, and the analytical expressions were derived. As concluded in [3-5], both the structural and acoustic resonances have a significant contribution for the interior noise in the cabin. For this reason, the dynamic response of a fully coupled vibro-acoustic system is derived in terms of the acoustic modes and structural modes. It is shown that the analytical model provides a good prediction of the reality, and that it is important to consider the coupling between structural and acoustic systems. Analytical predictions are obtained for random and TBL excitations, both for the shell vibration and sound pressure levels.

2. ANALYTICAL FRAMEWORK

2.1 TBL Model for Cylindrical Coordinates

As described in more detail in references [3, 5], the power spectral density (PSD) of the TBL wall pressure fluctuations over a flat panel, $p_{TBL}(x, y, t)$, can be defined using Corcos model by

$$S(x, \xi_x, \xi_y, \omega) = S_{ref}(x, \omega)\, e^{-\frac{\alpha_x \omega |\xi_x|}{U_c}}\, e^{-\frac{\alpha_y \omega |\xi_y|}{U_c}}\, e^{-\frac{i \omega \xi_x}{U_c}}, \qquad (1)$$

in which $\xi_x = x - x'$ and $\xi_y = y - y'$ are the spatial separations in the x- and y-directions, $S_{ref}(x, \omega)$ is the reference PSD, U_c is the TBL convective speed. Assuming that, in substitution of a flat panel one has a shallow open circular shell, then equation (1) can be written in the cylindrical coordinates systems as

$$S(x, \xi_x, \xi_\theta, \omega) = S_{ref}(x, \omega)\, e^{-\frac{\alpha_x \omega |\xi_x|}{U_c}}\, e^{-\frac{\alpha_y \omega |\xi_\theta|}{U_c}}\, e^{-\frac{i \omega \xi_x}{U_c}}, \qquad (2)$$

where $\xi_\theta = R(\theta - \theta')$ and R is the radius of the cylinder.

2.2 Cylindrical Shell Structural Model

For a given applied external pressure, $p_{ext}(x, \theta, t)$, the shell governing equation may be defined as follows

$$w + \nu_s R \frac{\partial u}{\partial x} + \frac{\partial v}{\partial \theta} + \frac{h_s^2}{12 R^2} \nabla^4 w + \frac{\rho_s (1 - \nu_s^2) R^2}{E_s} \ddot{w} + \zeta_s \dot{w} = \frac{(1 - \nu_s^2) R^2}{E_s h_s} p_{ext}(x, \theta, t), \qquad (3)$$

in which ν_s is the Poisson ratio of the shell, h_s is the shell thickness, E_s is the shell Young modulus, and ζ_s was added to account for the damping of the shell. The structural displacement may be defined as follows

$$w(x, \theta, t) = \sum_{m_x=1}^{M_x} \sum_{m_\theta=1}^{M_\theta} \alpha_{m_x}(x)\, \beta_{m_\theta}(\theta)\, q_{m_x m_\theta}(t), \qquad (4)$$

in which $\alpha_{m_x}(x)$ and $\beta_{m_\theta}(\theta)$ are the spatial functions, and $q_{m_x m_\theta}(t)$ is a function of time. For simply supported circular cylindrical shells without axial constraint at its ends, the spatial functions are defined by

$$\alpha_{m_x}(x) = \sqrt{\frac{2}{L_x}} \sin\left(\frac{m_x \pi x}{L_x}\right), \quad \beta_{m_\theta}(\theta) = \frac{1}{\sqrt{\pi}} \cos(m_\theta \theta), \qquad (5)$$

where L_x is the cabin length, while for a open curved panel with simply supported edges, the spatial functions are

$$\alpha_{m_x}(x)=\sqrt{\frac{2}{a}}\sin\left(\frac{m_x\pi(x-x_i)}{a}\right),\ \beta_{m_\theta}(\theta)=\sqrt{\frac{2}{\theta_0}}\sin\left(\frac{m_\theta\pi(\theta-\theta_i)}{\theta_0}\right), \quad (6)$$

in which a is the shell length, θ_0 is its arc, x_i and θ_i are the starting coordinates of the shell in the cabin global coordinates system.

2.3 Cylindrical Cabin Acoustic Model

The governing equation of cylindrical acoustic enclosure is the wave equation written in the cylindrical coordinates system, as

$$\nabla^2 p - \frac{1}{c_0^2}\ddot{p} - \zeta_{ac}\dot{p} = 0, \quad (7)$$

where c_0 is the speed of sound in the enclosure, and the damping term ζ_{ac} was added to account for the acoustic damping, and p is defined by

$$p(x,\theta,r,t)=\sum_{n_x=0}^{N_x}\sum_{n_\theta=0}^{N_\theta}\sum_{n_r=0}^{N_r}\psi_{n_x}(x)\,\phi_{n_r}(\theta)\,\Gamma_{n_\theta}(r)\,r_{n_x n_\theta n_r}(t), \quad (8)$$

in which $r_{n_x n_\theta n_r}(t)$ are functions of time, $\psi_{n_x}(x)$, $\phi_{n_r}(\theta)$ and $\Gamma_{n_\theta}(r)$ are the spatial functions, defined as follows

$$\psi_{n_x}(x)=\frac{A_{n_x}}{\sqrt{L_x}}\cos\left(\frac{n_x\pi x}{L_x}\right),\quad \phi_{n_r}(\theta)=\frac{1}{\sqrt{\pi}}\sin\left(n_r\theta+\gamma\frac{\pi}{2}\right),$$
$$\Gamma_{n_\theta}(r)=\frac{A_{n_\theta}}{R\,J_{n_r}(\lambda_{n_r n_\theta})}J_{n_r}\left(\lambda_{n_r n_\theta}\frac{r}{R}\right), \quad (9)$$

where J_{n_r} are Bessel functions of the 1st kind, $\lambda_{n_r n_\theta}$ factor represents the n_r^{th} root of $dJ_{n_r}(\lambda_{n_r n_\theta} r/R)/d(r/R)=0$, at r=R.

2.4 Structural-Acoustic Coupled Model

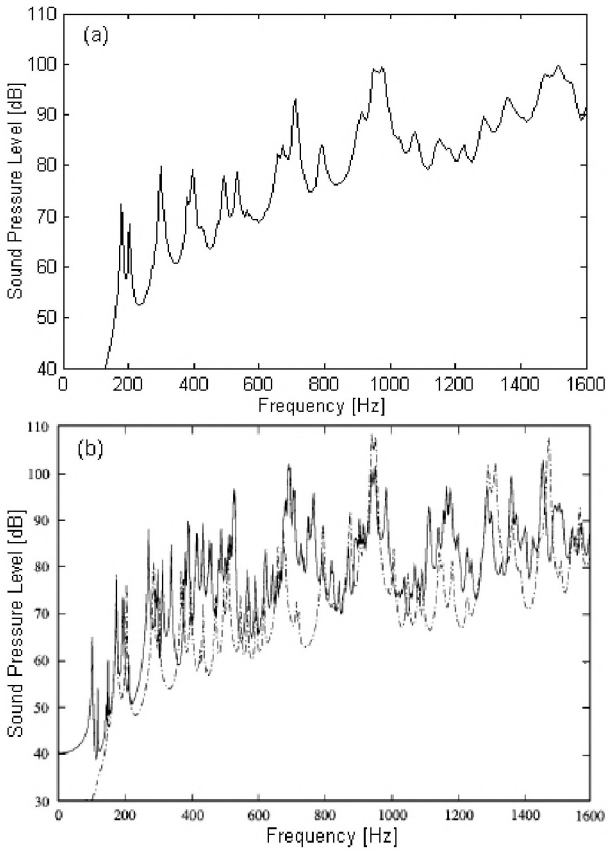
Following a similar mathematical procedure as the one described in [3, 5], the governing equations for the coupled system are obtained from the combination of the governing equations of the individual systems, and applying the boundary conditions, which in the present study are

$$\frac{\partial p(x,\theta,r,t)}{\partial r}=\begin{cases}-\rho_0\ddot{w}(x,\theta,t), & \text{at } r=R\\ 0, & \text{elsewhere.}\end{cases} \quad (10)$$

3. RESULTS AND DISCUSSION

Two independent studies were used as validation cases for the analytical framework. In this paper only one of the cases is presented. The study in [6] presents a numerical formulation based on a variational approach, in order to investigate the vibroacoustic behavior of a cylindrical shell, and the effect of an internal floor partition. The numerical results are also compared with experimental results. Figure 1 (b) displays the results from [6], showing the acoustic effect of the floor on the internal pressure filed. Figure 1 (a) shows the SPL results obtained through our analytical framework. To obtain the analytical results, a total number of M_x=15 and M_θ=15 shell modes, and N_x=10, N_θ=10 and N_r=10 acoustic modes were used to achieve convergence. Comparing the analytical predictions in part (a) with the results in part (b) (solution without attached floor), it can be observed that the developed analytical framework is in good agreement with the data from [6].

Fig. 1. SPL results: (a) obtained with our analytical framework; (b) from [6]: —— with attached floor, – · – without attached floor.



REFERENCES

[1] Bishop, D., Cruise flight noise levels in a turbojet transport airplane. *Noise Control* (1961), 7, pp. 37-42.

[2] Frampton, K., and Clark, R., Power flow in an aeroelastic plate backed by a reverberant cavity. *Journal of the Acoustical Society of America* (1997), 102(3), pp. 1620-1627.

[3] da Rocha, J., Suleman, A., and Lau, F. Prediction of flow-induced noise in transport vehicles: development and validation of a coupled structural-acoustic analytical framework, *Canadian Acoustics* (2009), 37(4), pp. 13–29.

[4] Li, D., Vipperman, S., Mathematical model for characterizing noise transmission into finite cylindrical structures. *Journal of the Acoustical Society of America* (2005), 117(2), pp. 679-689.

[5] da Rocha, J., Suleman, A., and Lau, F., Turbulent flow-induced noise in the cabin of a BWB aircraft. *Proceedings of the Canadian Society for Mechanical Engineering Forum 2010*, Victoria, BC, June 7-9.

[6] Missaoui, J., and Cheng, L., Vibroacoustic analysis of a finite cylindrical shell with internal floor partition. *Journal of Sound and Vibration* (1999), 226(1), pp. 101-123.

A LASER POSITION-SENSING SYSTEM FOR THE STUDY OF VIBRATION SHAKER TILTING

Lixue Wu
Institute for National Measurement Standards, National Research Council Canada
1200 Montreal Rd., Ottawa, Ontario, Canada K1A 0R6

1. INTRODUCTION

The calibration of accelerometers requires the use of a vibration shaker to generate sinusoidal accelerations in a main axis that is precisely defined, usually perpendicular to the mounting surface for accelerometers. Ideally, there should be no acceleration in other directions. In practice, unwanted transverse motion, tilting and tumbling cannot be avoided. Such movements occur under asymmetrical loading conditions primarily at higher frequencies. The output of the accelerometer under test increases or decreases in response to these unwanted movements. This increase or decrease in output becomes a source of measurement uncertainty in accelerometer calibration. For this reason, the international standards for calibration of accelerometers, such as ISO 16063-11[1], require that the tilting be kept sufficiently small to prevent excessive effects on the calibration results. For high accuracy calibration of accelerometers, the tilt of the shaker must be determined.

Laser-based interferometers are the most commonly used tools for accurate non-contact tilt measurement [2-6]. In these systems, a laser beam is incident on a reflective surface and the tilt of this surface is measured from the interference between the reflected beam and the reference beam. These systems, however, are not applicable to the measurement of tilt of shakers as the reference point for the reference beam is the point where the shaker is either not moving (no tilt) or at its equilibrium (only available at a specific time instance). This paper proposes a modified approach for small tilt measurement using a lateral effect position-sensitive detector (PSD). A laser position-sensing system is implemented such that the tilt measurement is displayed in real time while the shaker is operating. Loading conditions can then be adjusted in real time to reduce the tilting.

2. CONCEPT

If the relative positions of three points on the mounting surface of a shaker are described by the two-dimensional orientation of a plane joining them, then tilt is the angular amount that the orientation of the plane has changed from a previous orientation (or reference orientation). A measurement system for tilt is proposed as illustrated in Fig. 1. The measurement system is modified from a Michelson interferometer with the photo-detector replaced by the PSD and the reference beam removed.

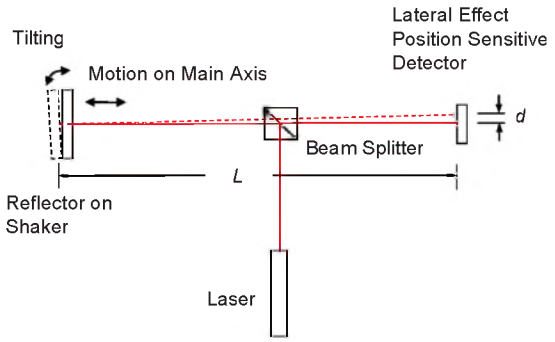


Fig. 1. Laser position sensing system for measuring angular tilt.

The angular tilt of the reflector on a shaker could be measured by measuring the displacement d, the shift of the returning beam on the PSD from its original position when the shaker is not moving, and the distance L between the reflector and the PSD. The angular tilt θ is given by $\theta = \tan^{-1}(d/L)$. However, the distance L is varying while the shaker is moving. This presents some difficulties in measuring the distance L. Furthermore, the position sensitivity of the PSD must be calibrated in measuring the displacement d. Therefore, it is proposed to directly calibrate the tilt sensitivity of the tilt measuring system by using a cantilever beam as a source of tilt.

3. IMPLEMENTATION

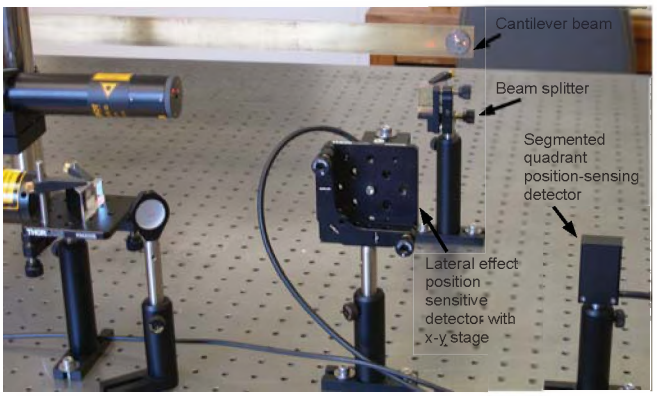


Fig.2. Calibration setup for the laser position-sensing system.

Shown in Fig. 2 is a photo of the calibration setup. The shaker is replaced by the cantilever with a mirror placed at the exact position of the shaker reflector. To ensure this positioning an additional alignment system is used as shown in Fig. 3. The position of the cantilever is adjusted such that

the returning beam hits the quadrant detector in the same place as that reflected from the shaker reflector.

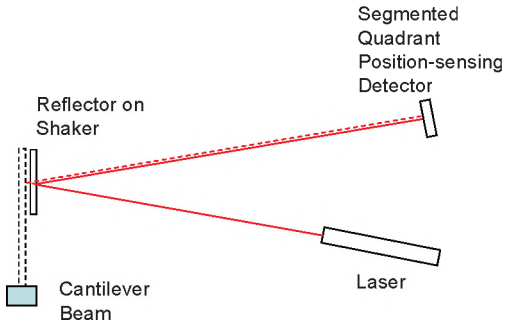


Fig. 3. System for aligning the cantilever.

Cantilevers are well-studied in literature [7]. Either the static beam equation (the Euler-Bernoulli equation) or the dynamic beam equation (the Euler-Lagrange equation) can be used to determine the deflection y of a cantilever at any position x along the cantilever. The slope of the cantilever dy/dx is the angular tilt $\theta(x)$ of the cantilever at the position x. If the static beam equation is used, the applied load can be a known force at the end of the cantilever. By changing the force applied, the desired angular tilt at the laser beam sensing point on the cantilever can be generated, which is then used to calibrate the PSD. If the dynamic beam equation is used with no transverse load on the cantilever, free harmonic vibrations can be started with a non-zero initial defection at the end of the cantilever. The acceleration at the laser beam sensing point can be measured by a calibrated accelerometer mounted at the back. Ohm et al. show that the bending (tilting) of the cantilever does not influence the main-axis (flexural) acceleration measured by the accelerometer [8]. Therefore, the tilt of the moving cantilever can be determined directly from the output of the accelerometer, the calibrated sensitivity of the accelerometer, and the physical property of the cantilever.

The laser position-sensing system is implemented using a helium-neon laser (05 STP 901, Melles Griot) operated in the intensity-stabilized mode, a 50/50 beam splitter, a lateral-effect PSD (SPOTANA-4-USB-Low, Duma), and a 24-Bit, 204.8 kS/s, 4-channel digitizer (NI PXI-1033 and NI PXI-4461, National Instruments) to measure the outputs of the PSD. The PSD has a resolution of 0.1 μm that results in a tilt resolution of 0.04 arc-second for the system if the distance L is 0.5 m. An x-y translation stage is used to align the PSD such that the mean values of the PSD outputs are close to zero. These mean values are then used to remove the bias in the output of the PSD. Fig. 4 shows the PSD outputs when the beam is freely vibrating after an initial deflection at the end of the cantilever. The x-axis output of the PSD has a typical amplitude of about 3 V which slowly decays as the vibration amplitude decays. Torsional vibration of the cantilever is observed from the y-axis output of the PSD at a level of 0.05 V. This is caused from the unbalanced loads (mirror and accelerometer) at the end of the cantilever. Better modeling for the cantilever will be considered in future [9].

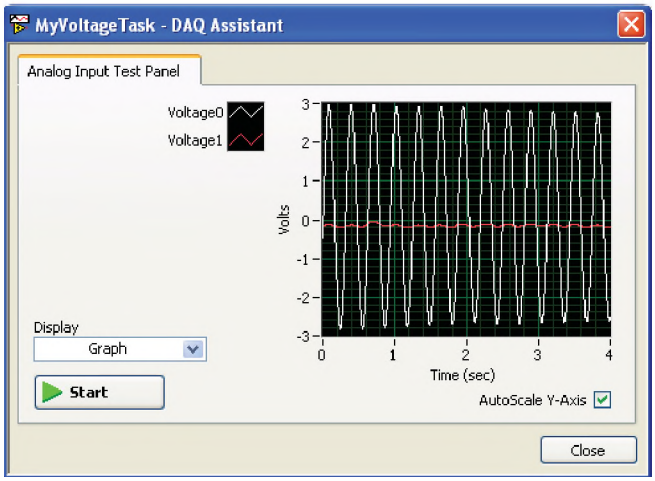


Fig. 4. PSD outputs when the cantilever is freely vibrating after an initial deflection.

REFERENCES

1. International Organization for Standardization. (1999). ISO 16063-11:1999-12, *Methods for the Calibration of Vibration and Shock Transducers – Part 11: Primary Vibration Calibration by Laser Interferometry*, Geneva, Switzerland.
2. Malacara, D. and Harris. O. (1970). "Interferometric Measurement of Angles," *Appl. Opt.* 9(7), 1630-1633.
3. Ikram, M. and Hussain, G. (1999). "Michelson interferometer for precision angle measurement," *Appl. Opt.* 38(1), 113-120.
4. Sirohi, R.S., Ganesan, A.R. and Tan, B.C. (1992). "Tilt measurement using digital speckle shear interferometry," *Opt. Laser Technol.*, 24(5), 257-261.
5. Prakash, S., Singh, S. and Rana, S. (2005). "Automated small tilt angle measurement using Lau interferometry," *Appl. Opt.*, 44(28), 5905-5909.
6. Chiu, M.H. and Su, D.C. (1997). "Angle measurement using total internal reflection heterodyne interferometry," *Opt.Eng.* 36(6), 1750-1753.
7. Gere, J.M. and Timoshenko, S.P. (1997). *Mechanics of Materials*, 4th Ed., PWS-KENT Publishing Company, Boston, MA.
8. Ohm, W.S., Wu, L., Hanes, P. and Wong, G.S.K. (2006). "Generation of low-frequency vibration using a cantilever beam for calibration of accelerometers," *Journal of Sound and Vibration*, 289, 192-209.
9. Han, S.M., Benaroya, H. and Wei T. (1999). "Dynamics of Transversely Vibrating Beams using four Engineering Theories," *Journal of Sound and Vibration*, 225(5), 935-988.

EXPERIMENTAL AND NUMERICAL COMPARISON OF VISCOELASTIC MATERIAL DAMPING TO EQUIVALENT MASS AS ACOUSTIC TREATMENT TO AIRCRAFT COMPOSITE FUSELAGE STRUCTURE UNDER VARIOUS EXCITATIONS

Esen Cintosun and Tatjana Stecenko
3M Aero Technologies - Mississauga, 7381 Pacific Circle, Mississauga, Ontario L5T 2A4

1. INTRODUCTION

Passive noise control treatments are used widespread in aircraft industry. A typical passive aircraft noise treatment (or sound package) is made up of a combination of viscoelastic damping and low density porous materials. A sound package is used to minimize the direct transmittance of sound waves (airborne) and radiation of noise (structure borne) into aircraft cabin interior.

Viscoelastic material damping is applied to fuselage skin and/or trim panel to increase transmission loss at natural frequencies and resonances and is normally applied as constraining layer damping (CLD). The viscoelastic damping characteristics are directly linked to material composition and processing, including the degree of cross-linking and the type of fillers such as carbon black [1]. The characteristics also vary with temperature and to a lesser degree with frequency.

Porous materials such as foam and fiber glass are effective noise attenuators at frequencies higher than 500 Hz depending on the thickness. For low frequency sound energy attenuation, heavy damping is used. Increased mass reduces the resonant frequencies. One issue with damping is that since the dampers are uniform materials, at resonant frequencies whole material moves which in turn may result in transfer of energy to coupled structures and cause radiated noise.

Distributed mass in porous material has been found to increase transmission loss and act as damper. Idrisi et al. [2] numerically and experimentally investigated the effect of heterogeneous blankets in reducing radiated noise in a double panel system and found that the performance of the heterogeneous blanket based on the design (location of the masses) varied with the best performer attenuating 12.5 dB and the median performer attenuating 2.6 dB in sound radiation in the 0-500 Hz frequency range. Sgard et al. [3] also investigated both numerically and experimentally the significance of heterogeneous blankets in dissipating vibro-acoustic energy. In this investigation, the major conclusions included: mass inclusions equivalent to 7% of panel weight increased energy dissipation at low frequencies up to 150 Hz. Furthermore, Kidner et al [4] investigated the effect of mass inclusions to poroelastic layer and found that the insertion loss of standard acoustic blankets can be significantly improved at low frequencies. Esteve [5] showed that optimally damped vibration absorbers and optimally damped Helmholtz resonators reduce sound transmission into payload fairing at low frequencies.

The work presented in this paper is part of a research project with a main objective of investigating the effect of equivalent masses within porous materials to the effect of viscoelastic material damping as part of sound packages as treatment to carbon composite panel. In this paper, acoustic performance of sound packages with CLD is compared numerically and experimentally to performance of sound packages with mass layer equivalent to the mass of CLD. The parameters compared were airborne insertion loss (ABIL) and radiation efficiency.

2. METHOD

Transmission loss and radiation efficiency measurements were made in order to compare the sound transfer properties of CLD to equivalent mass layer on a representative fuselage skin composite panel. TL measurements were performed according to ISO 140-3 1995 standard. 2.66 mm thick, 1.13 m by 1.8 m flat composite panel was excited using point force mechanical excitation for radiation efficiency tests and diffuse field excitation for transmission loss tests. Nova software using the FTMM numerical method (developed at the University of Sherbrooke) was used to perform the calculations. Sound package configurations tested are listed in Table 1.

Table 1: Sound package configurations

Sound package ID	Configuration (B: brick, FB: frame blanket, TC: top cover, CLD: carbon constraining layer damping)	% panel weight
cbare	Bare ribbed carbon composite panel	
CLD	Bare with 53% coverage CLD	39%
CLD+B+FB+TC	53% CLD + B(0.75" foam) + FB(1.75" foam) + TC(0.5" foam)	54%
B+FB+mass+TC	B(0.75"foam) + FB(1.75" foam) + 0.06" mass layer + TC(0.5" foam)	46%

3. RESULTS AND DISCUSSION

The comparison of CLD to equivalent mass layer in terms of measured mass corrected ABIL plots is shown in Figure 1. The sound package configuration with mass layer (cbare+B+FB+mass+TC) shows a double wall effect and is superior to the configuration with damping in the frequency range of 250 to 5000 Hz. The double wall resonance effect is seen in the third octave frequency range of 100 to 200 Hz.

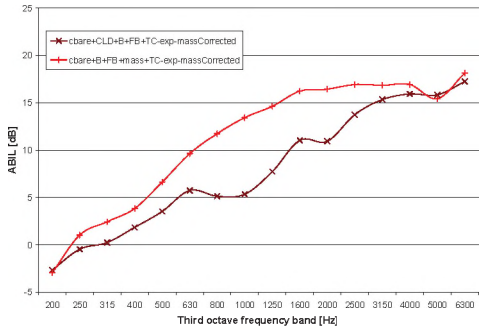


Figure 1: Comparison of experimentally obtained ABIL plots for sound packages cbare+CLD+B+FB+TC and cbare+B+FB+mass+TC

In Figure 2, radiation efficiency plots for cbare+CLD+B+FB+TC and cbare+B+FB+mass+TC are compared. Again, sound package configuration with mass layer shows superior performance with less radiation efficiency in the frequency range of 250 to 6300 Hz.

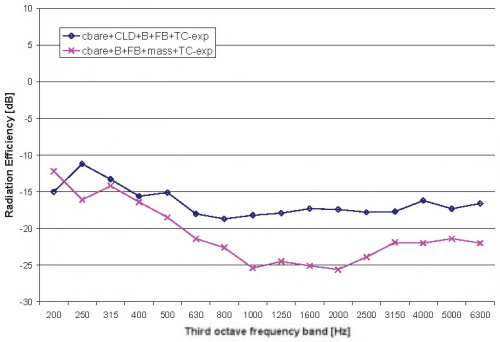


Figure 2: Comparison of experimentally obtained radiation efficiency plots for sound packages cbare+CLD+B+FB+TC and cbare+B+FB+mass+TC

ABIL plots obtained experimentally and numerically are compared in Figure 3 for cbare+B+FB+TC configuration which shows a close match. Similarly, the radiation efficiency plots for the cbare+B+FB+mass+TC sound package configuration shown in Figure 4 are closely matched.

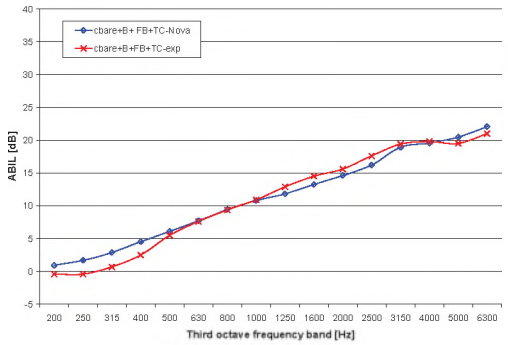


Figure 3: Experimentally and numerically obtained ABIL plots for cbare+B+FB+TC configuration

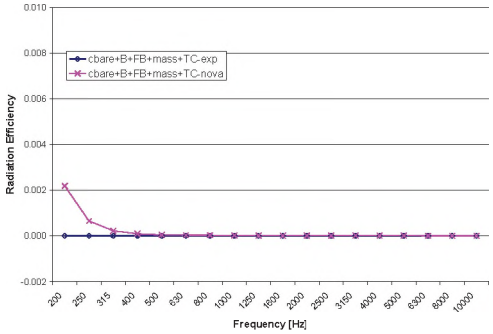


Figure 4: Experimentally and numerically obtained radiation efficiency plots for cbare+B+FB+mass+TC configuration

The sound package configuration with equivalent mass layer produced 3.6 dB measured and 5.1 dB numerical improvement over the sound package configuration with CLD in terms of the average of ABIL at octave 1000, 2000 and 4000 Hz as listed in Table 2.

Table 2: Average ABIL results for the carbon composite panel sound package configurations

Sound package	Average of measured ABIL at 1k, 2k, 4k Hz 1/1 octave	Average of numerical ABIL at 1k, 2k, 4k Hz 1/1 octave
cbare+CLD	8.6	6.5
cbare+CLD+B+FB+TC	19.9	21.7
cbare+B+FB+mass+TC	23.5	26.8

4. CONCLUSIONS

The acoustic performance of sound packages with viscoelastic material damping and equivalent mass were compared. Based on the experimental comparison, equivalent mass acoustic performance was found to be superior to damping. The experimental results were numerically validated to be used in optimization in the future. The experimental set-up was suitable for measurements at frequencies higher than 200 Hz.

REFERENCES

[1] M.D. Rao, *Journal of Sound and Vibration* 262 (2003) 457-474.

[2] K. Idrisi, M. Johnson, J.P. Carneal, *NOISE-CON 2007*, Reno, Nevada, 2007.

[3] F.C. Sgard, N. Atalla, *PACS* subject classification number: 43.50.Gf.

[4] M.R.F. Kidner, C.R. Fuller, B. Gardner, *Journal of Sound and Vibration* 294 (2006) 466-472.

[5] S.J. Estève, Control of sound transmission into payload fairings using distributed vibration absorbers and Helmholtz resonators, Ph.D. thesis, Virginia Polytechnic Institute & State University, 2004.

BRAZILS'S FIRST UNDERGRADUATE COURSE IN ACOUSTICAL ENGINEERING

Stephan Paul
Dept. of Civil Engineering, Federal University of Santa Maria, Santa Maria, Rio Grande do Sul, Brazil, 97105-900,
stephan.paul.acoustic@gmail.com

1. INTRODUCTION

It is certainly a common agreement among the participants of the CAA meeting that acoustics as a branch of science occupies an important position. In contrast to its importance it is indeed a science that is, worldwide, often only marginally incorporated into teaching curricula of universities. And while, for example, at some German universities acoustics is very well represented in their curricula, the situation in Brazil was quite different, mainly due to the lack of a comprehensive curriculum that accounts for basic education in acoustics. Contrary to this scenario, there is an enormous demand for acoustic engineers in Brazil. To account for this demand, in the end of 2008 a dedicated undergraduate course "Acoustical Engineering/ Engenharia Acústica" started at Federal University of Santa Maria, which is oriented to the German and Chilean model. As part of the article, the structure of the program will be presented and discussed.

2. BIRTH OF A NEW UNDERGRADUATE COURSE

Brazilian industry, and to a minor extent governmental agencies as well, are looking hard for engineers who have acoustics backgrounds to work in product development, environmental noise, and building acoustics. Up to now the candidates for such positions have come from the very few Master's and PhD programs dedicated to acoustics and vibration or were trained in-house. Neither solution was sufficient, because acoustics is a many faceted field of science and its knowledge can hardly be taught with a Master's or even PhD program only. Also, the Master's and PhD programs currently available in Brazil do not qualify a sufficient number of acousticians. Finally, for many smaller industries Master's and PhD degree holders are usually too ``expensive'' to be employed.

An undergraduate course in acoustical engineering was thus highly desirable. In 2008 the conditions to implement such a course, as well as others, were so far favorable as a Governmental higher education policy framework that aimed to increase the number of students in federal universities was in place.

A small acoustics working group that existed in the Civil Engineering Department of the Federal University of Santa Maria, southernmost Brazilian state of Rio Grande do Sul, was then transformed into the germ cell of a new undergraduate course in Acoustical Engineering. Beyond this, the favourable higher education policy Brazilian industry and the Brazilian Society of acoustics are also interested in such an undergraduate course.

3. CURRICULUM

The curricula was developed using experiences from Chile and Germany where similar undergraduate courses exist. Naturally, a number of rules that apply for an undergraduate course in Brazil were to be followed, e.g. a minimum 3600 hours of classes. The course in Acoustical Engineering has a total of 3750 hours, divided into 3375 h for compulsory topics (Table 1), 315 h for optional topics and 60 h for additional activities. Out of the 3375 hours of compulsory topics, 1260 h are dedicated to fundamentals, 720 h to general engineering topics, and 1395 h to acoustics topics, including an internship of 210 h in the last semester. Optional topics are either (1) technical and non-technical topics offered by the university or (2) special topics in acoustics offered by the undergraduate course or other universities, especially in foreign countries. Additional activities are activities such as participation in conferences on acoustics, research projects in acoustics, and the like. In order to provide some attractions from the very beginning of the course, dedicated lectures on topics related to acoustics are introduced in the first semester ("Music for Acoustical Engineers" starting in the first semester and "Noise, Vibration, and Humans" in the second semester). Especially the latter discipline is also intended to provide interdisciplinary points of view to the students.

4. ACADEMIC STAFF

For the topics that are related to acoustics, audio, and vibration, though currently still in the civil engineering department, the future department of acoustical engineering will have a total number of nine lecturers in the near future. Today 5 positions are filled with lecturers from Brazil, Chile, and Germany, most of them with international experience or even foreigners, and over the following two years the remaining four positions are to be completed with specialists in Structure borne sound, Audio, Digital Signal Processing, and Room- and Building acoustics.

5. FACILITIES AND EQUIPMENT

Currently the facilities that are associated to the course are acoustical laboratory facilities of the Department of Civil Engineering, composed primarily of a reverberant chamber (207 m^3) for measuring absorption and sound power, as well as another pair of reverberant chambers (60 and 67 m^3) to measure sound transmission, principally walls and other

types of room partitions. Another pair of rooms is used to measure tapping sound transmission. Recently the course acquired a large amount of new equipment, from basic tools such as microphones and accelerometers to several types of impact hammers, manikins, sound intensity probes, SPL-meters, and many others. Some software was acquired as well such as ArtemiS, Bastian, Dirac, Odeon, Pulse, SoundPlan, and VAOne. At the moment, a project is being elaborated to expand the facilities, adding a fully anechoic chamber, a recording studio, a sound booth, a new chamber for tapping sound measurements, and a library, etc. Further audio equipment and software must be still purchased.

6. RESEARCH ACTIVITIES

While the laboratory in the past worked mainly on small scale projects in building acoustics, the new staff and equipment will offer the possibility to start more dedicated research in a wider variety of fields. Research into building acoustics will be maintained and strengthened, as the reverberant rooms at the laboratory are the only ones in Brazil that allow for transmission loss measurement of whole room partitions. Other projects such as research in room acoustics, aircraft noise problems, psychoacoustics, and sound evaluation and aero-acoustics are being started.

7. NATIONAL AND INTERNATIONAL COOPERATION

National and international cooperation will be an important aspect in the development of the course and the research projects to be developed. First, cooperation started with the Federal University of Santa Catarina (Brazil) and RWTH Aachen University (Germany). Others are expected to follow soon, such as projects with McGill (Monteral), Universidad Austral (Valdivia, Chile), and the Technical University of Berlin (Germany).

8. CONCLUSIONS AND OUTLOOK

To help with education in acoustics, to provide skilled acoustical engineers to the job market, and to develop dedicated research in acoustics, Brazil's first undergraduate course in Acoustical Engineering has opened at the Federal University of Santa Maria with an attractive curriculum. It is expected that interesting research will be developed, together with other universities and private or public entities.

Table 1. Curricula (all compulsory +some optional topics)

sem	topic	hours
1	Math A	90
	Linear algebra A	90
	Introduction to geometry	45
	Introd. into Acoustical Engineering	45
	Chemistry	30
	Music for Acoustical Engineers I	30
	Physics I	75
2	Math B	90
	Linear Algebra B	90
	Sound, Vibration and Human Being	45
	Technical drawing	45
	Music for Acoustical Engineers II	30
	Physics II	75
3	Differential calculus A	60
	Numerical methods	60
	Material science	60
	Resistance of material	60
	Music for Acoustical Engineers III	30
	Physics	90
4	Differential calculus B	60
	Complex variables	60
	Fundamentals of Civil Engineering	60
	Programming	90
	Music for Acoustical Engineers IV	30
	Thermodynamics	60
	Technical English and Portuguese	45opt
5	Statistics A	60
	Fundamentals of Acoustics	75
	Electrical circuits I	60
	Fundamentals of Electronics	60
	Fluiddynamics	60
	Economics	60
6	Room acoustics	60
	Noise Control	60
	Fundamentals of Vibration	75
	Electric Circuits II	60
	App. Electronics and Instrumentation in Acoustics	60
	Urban Noise Control	60
	Cientific working	30opt
7	Vibration Control	60
	Building Acoustics	75
	Electroacoustics I	60
	Psychoacoustics	75
	Digital signal processing	60
	Measurement techniques	60
	Bachelor project I	30
	Security on the workplace	45opt
8	Numerical Methods in Acoustics and Vibration	60
	Electroacoustics II	60
	Sound Reinforcement System Theory	60
	Sound Reinforcement System Practice	60
	Digital Signal Processing II	60
	Subjective sound evaluation	60
	Auralization	60opt
9	Start-up for enginners	60
	Exp.methods in Acoustics and Vibration	60
	Loudspeakers	60
	Recording techniques	75
	Bachelor project II	30
	Project	60opt

AN ULTRASOUND INVESTIGATION OF DIDGERIDOO ARTICULATIONS

Ko Zolotas and Sonya Bird
Dept. of Linguistics, University of Victoria, PO Box 3045, Victoria BC, V8T3K5 zlolotas@uvic.ca

1. INTRODUCTION

The didgeridoo is an instrument native to Australia, traditionally made out of a eucalyptus tree naturally hollowed out by termites. These days, didgeridoos are made from a wide range of materials, including bamboo, glass, cactus, or even PVC pipe. They are simple tubes, 1-2 meters long, with bees-wax at one end serving as the mouthpiece. They have no holes or moving parts; to create different sounds players can change the amount of airflow from the lungs, manipulate their oral cavity, or even sing into the didgeridoo as they play (Schellberg & Langham 1994).

Past research on the didgeridoo has focused on vocal tract resonances (Tamopolsky et al. 2006), lip motion (Fletcher et al. 2001) and the acoustic properties of didgeridoos themselves (Amir 2004). However, no research has yet been done on how players manipulate their tongues to create different sounds. The goal of this study is to fill this gap, using ultrasound imaging (Gardner & Stone 2009). We focus on circular breathing and five typical didgeridoo vocalizations, termed here undulation, piston, alveolar trill, alveolar ejective and velar ejective. We hope that this preliminary articulatory investigation will increase our understanding of the tongue articulations used in didgeridoo playing, and will also provide a valuable visual tool for teachers and learners of the instrument.

2. METHOD

A single adult male (one of the co-authors) participated in the experiment, who has been playing the didgeridoo for approximately 3 years. Using a GE Logic E ultrasound machine with an 8C-RS convex transducer fixed under the chin, tongue articulation was recorded during circular breathing and 5 sounds of interest, described in detail below. Figure 1 illustrates the experimental set-up: with the help of a microphone stand, the player held the transducer fixed under his chin with one hand. The other hand stabilized the didgeridoo, which was supported by a box at a relatively horizontal angle - necessary in order to maintain the chin in an orientation suitable for fixing the probe. In addition to the ultrasound recording, the audio signal was recorded for acoustic analysis (not presented here) using a Senheiser ME-55 directional microphone (with a K6 capsule).

The video and audio signals were synchronized through a Mackie 1402-VLZ3 mixer and captured on an external PC computer using Sony Vegas Pro 8. From the ultrasound video, sequences of stills were extracted illustrating the tongue trajectories in each sound of interest. These are described in detail in the following section.

Fig. 1. Experimental set-up.

3. RESULTS

The following stills are taken from the ultrasound video recordings. They all represent a midsagittal (from the side) view of the tongue; the white line towards the top of the image represents the tongue contour, with the tongue tip on the right. Circles have been superimposed on the tongue contour to help view the shape of the tongue.

Figure 2 illustrates tongue movement in circular breathing, a method which allows didgeridoo players to play indefinitely, without needing to pause for air. The cheeks are initially extended, providing the extra air space that is needed for this articulation. The tongue first moves towards the back of the mouth (A-B-C), until it has completely blocked off airflow at the back of the mouth (D). At this point the cheeks begin to tense and are made smaller, aiding to push out the air that is required to maintain phonation. As soon as the oral cavity has been closed off (D), the velum is lowered, allowing the player to breathe in through their nose while simultaneously blowing out from their now closed off oral cavity. The tongue then acts as a syringe (E-F-G-H), gradually pushing air out of the mouth at a rate that is fast enough to maintain phonation, but slow enough to allow the player to breathe in a sufficient amount of air. While the tongue tip is still closing off the oral cavity (H), the velum is lowered, allowing air to flow from the lungs to the oral cavity once again.

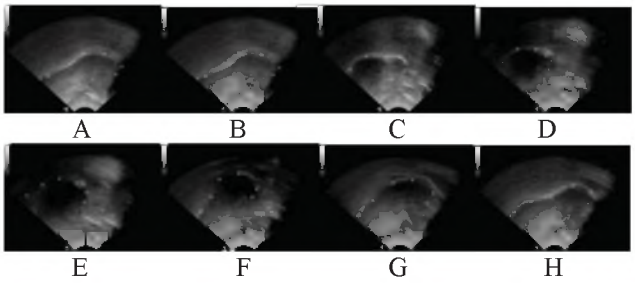


Fig. 2. Circular breathing.

In addition to circular breathing, the articulations required for five typical vocalizations were recorded. The first one is termed the "undulation" (Figure 3). The articulation begins with the tip of the tongue at the alveolar ridge (A). As the

articulation continues, the tongue body moves back through post-alveolar (B), palatal (C) and velar (D) regions. The undulation can be a cyclic articulation, so even before one cycle has been completed, the tongue tip raises in anticipation of the next undulation (E). Throughout this articulation, the tongue is narrow enough for air to pass by the sides of the tongue.

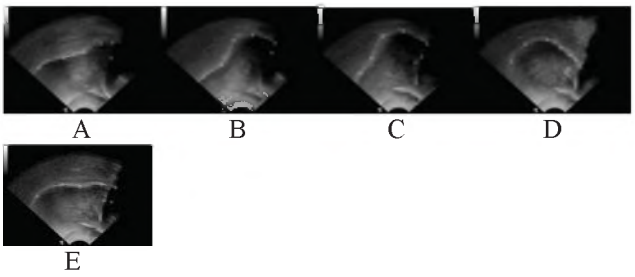


Fig. 3. The undulation.

Unlike the undulation, the "piston" (Figure 4) is a quick moving articulation. Its amplitude is mainly determined by how much volume change there is in the mouth and how quickly this occurs. The tongue begins in rest position (A), and is extremely retracted during the first part of the piston articulation (B) - a less retracted tongue would produce less amplitude. The tongue then shoots forward in the mouth (C), filling the entire oral cavity and causing a high amplitude burst - a slower movement of the tongue would produce smaller change in amplitude or no difference in amplitude at all. At this point, there is a complete stop at the front of the mouth. The tongue body begins to take on a concave shape as the air pressure builds up behind it (D), before the next articulation begins.

Fig. 4. The piston.

The "alveolar trill" (Figure 5) used in didgeridoo playing is very similar to that used in speech. From a rest position (A), the tongue is extended and flattened (B), and trilling occurs at the tongue tip (C). Acoustically, the effect of the trill can sound like a growl, and this is often what it is called by didgeridoo players.

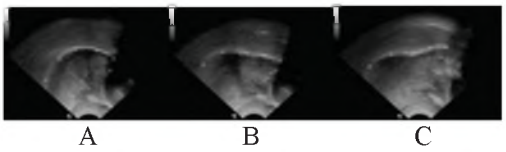


Fig. 5. The alveolar trill.

Many different stops are used in playing the didgeridoo, differing in intensity and place of articulation. The last two sounds described here are the alveolar and velar ejectives. In the alveolar ejective (Figure 6), the tongue starts in rest position (A) and moves forward to form a closure at the alveolar ridge (B), at which point a complete stop of airflow is created simultaneously in the oral cavity and at the glottis. The larynx is then raised, resulting in a sudden increase in oral pressure behind the tongue, giving it its unique shape (C). The tongue is then pulled back from the alveolar ridge (D), allowing air to rush out of the mouth. Shortly thereafter, the glottis opens and larynx lowers to provide continuous airflow from the lungs once again.

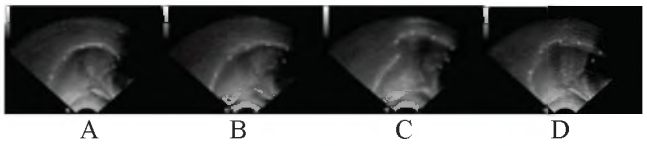


Fig. 6. The alveolar ejective.

The velar ejective (Figure 7) is similar to the alveolar ejective (Figure 6), except that the oral closure occurs in the velar region (compare Figure 6B and 7B). In Figure 7, frame A represents the initial rest position; frame C the release of the velar closure.

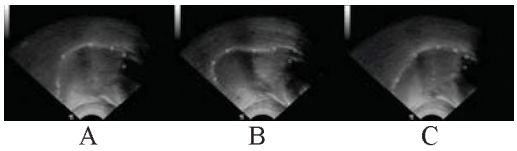


Fig. 7. The velar ejective.

4. DISCUSSION

As far as we know, this study provides the first direct articulatory view of the tongue articulations used in producing some of the typical vocalizations of the didgeridoo. We hope that this study will act as a useful guide to teachers and learners of the didgeridoo, as well as to others who are interested in understanding the articulatory basis of the didgeridoo's acoustic properties.

REFERENCES

Amir, M. (2004). Some insights into the acoustics of the didgeridoo. *Applied Acoustics* **65**: 1181-1196.

Fletcher, N., L. Hollenberg, J. Smith & J. Wolfe (2001). The didgeridoo and the vocal tract. *Proceedings of the International Symposium on Musical Acoustics*: 87-90.

Gardner & M. Stone (2010). A comparison of midsagittal tongue shapes during clarinet performance and vowel production using ultrasound. *Poster presented at Ultrafest V*. Haskins Labs, New Haven CT, Mar 20.

Shellberg, D. & T. Langham (1994). *Didgeridoo: ritual origins and playing techniques*. Diever, Holland: Binkey Kok.

Tamopolski, A., N. Fletcher, L. Hollenberg, B. Lange, J. Smith & J. Wolfe (2006). Vocal tract resonances and the sound of the Australian didgeridoo (yidaki). *Journal of the Acoustical Society of America* **119**: 1194-1204.

WHALE SONG OR WHALE MUSIC? FROM A COMPOSER'S PERSPECTIVE

Lisa Walker
The Grooved Whale Project www.groovedwhale.com e-mail: groovedwhale@gmail.com

If we were to look for music in another species what would we look for? What would we measure? While many species may eventually end up claiming musicality, for me it is the humpback whale that holds the most promise for discovery.

BACKGROUND

I was first introduced to the humpback in 1996 when I was invited to join Dr. Fred Sharpe and the Alaska Whale Foundation in southeast Alaska. It was thought that with my classical music background I might be able to identify subtle patterns in the acoustic activity of the whales that were not readably evident to the scientist.

The field work encompassed identifying individual whales, tracking their behavior and recording their feeding call - a vocalization seemingly utilized as part of their bubble net feeding strategy, where a group of whales dive down, encircle a school of herring in bubbles, blast them with a minute or two of sound, then rise up through the center of the circle with a mouthful of fish.

No feeding call we recorded was ever the same twice - sometimes the notes swooped up, sometime down, some were staccato, others long and legato (to put it in very unscientific terms). My musical senses were piqued.

After three seasons in Alaska I became curious to experience the whales' other vocalization - the winter song. In 1999 I went to Hawaii on an independent mission to record their sounds and while there fulfilled a musician's guilty pleasure of playing a 'concert' on my violin through an underwater speaker to the whales. Playing with the whales was a breeze as the patterns and rhythms in the winter song seem to follow the pattern logic of music. But were the whales interacting with me or was it I with them? When I listen back to the recording it is impossible to tell.

DIVING IN

As I began to familiarize myself with current theories of why whales sing I became curious to see how particular ideas regarding the yearly changes observed in the song – such as cultural transmission, innovation, mutation and drift - might map onto my own species musical use of sound.

To me, human music is much more than a sum of its parts – the individual notes, phrases or themes that otherwise would be simply perceived as organizations of sound. Somehow they combine to loft the resulting mix into the realm of music and away we go - our bodies and minds get wrapped up in it and we are compelled to dance and sing.

As a composer it is my job to bring new songs into being. But why do I compose? From the blank page to the finished song, why do I choose certain notes over others? Why particular patterns? Why particular changes? What is it about the human craving for music that drives my need to create?

EXPLORATIONS

To understand more about my compositional process I first had to find a way to bring it out from the murky depths of my sub-consciousness and expose it to the light of day. If music's power lies in its organization of sound, my solution was to write “de-organized” music. For instance, a bass line and a drum line, both of which have the potential to match up and create a musical unit, were purposefully misaligned in a “non-musically acceptable” way.

The technique circumvented my full immersion in the musical state, allowing me to reserve a bit of brain power to examine my decision making process along each step of the way.

In addition to revealing multiple rules and relationships between the notes, this technique and the resulting mismatched music turned my attention to the zones of “active chaos” in music where things break down and realign. These zones open up the possibility of creativity and are necessary for transitions in music. This led me to focus my attention on certain parts of humpback song, parts which lie in between the identified phrases but are not part of the phrase itself. Do they function as transitional zones for the whale as well?

My second technique of inquiry involved studying West African drumming to learn more about the basic “codes” of human rhythm - an important component of which was learning how two or more rhythms played together created “feel” or “groove”.

Learning these poly-rhythmic codes led me to look at humpback phrases, not only the “rhythms” within the phrases, how each one has its particular nuance, but how individual phrases aligned with each other when played in tandem. Are their phrases meant to relate to each other as they do in human music or does each one have its own logic?

The third technique came through my experience of joining a drum ensemble, whereby a handful of drummers meet each week to explore rhythmic idioms (often in three or more parts) which evolve and mutate as the session goes on.

Through the drum circle I have been able to participate in music creation in real time, allowing me to see how rhythmic codes shift and change over time and the logic by which these changes take place. It has also allowed me to examine the experience of creating with others, to understand more about status and hierarchy in music making and the risk and reward of deviating from the established codes to try new things.

It has also has brought my attention to particular similarities and differences between humpback and human song. For instance, in the drum circle we strive to meld our rhythms together, and in doing so strengthen the effect of the song. Humpbacks seem to be more disparate as evidenced by their singing independently of one another. Does this exclude their song as music? Or is some other musical logic at play? Are they under the same rules of competition/cooperation as us in the drum circle and do they assume the same risk and potential for reward when forging new sonic ground?

GOING FORWARD

I have developed these techniques as a way to understand humpback song through the lens of human music and understand human music through the lens of humpback song. I believe that through examining our own rules and relationships of music creation we are better equipped to investigate another species use of sound.

My hope is that by bringing to light particular aspects of music we can ground truth ideas we have about other species to see if they hold water with our own and ultimately establish a framework for a comparative analysis between human and whale song.

AUTHOR'S NOTE

The Grooved Whale Project is headed by Lisa Walker, a composer and musician fascinated by the ever-evolving structure of humpback whale song. Over the past 14 years she has dedicated herself to obtaining a scientific understanding of these animals in order to provide insight as to how the music in other species might be investigated. Her album "Grooved Whale" was released in 2001 and chronicles the beginning of her independent research. She is now preparing to release "Music For Other Ears", an album which draws upon possible permutations of musical organization to take the listener deep into the world of the whale.

Abstracts Without Summary Papers

Bioacoustics

How deep do you call? Depth localization in Southern Resident killer whales using passive acoustics
Jason Wood, Scott Veirs, Val Veirs & Dominic Tollit

The Snohomish Public Utility District of Washington State is applying for permits to install up to two in stream tidal turbines in Admiralty Inlet, Washington State, as part of a pilot project to determine the feasibility of marine tidal energy generation. To inform the permitting process, acoustic recordings from a vertical hydrophone array were made in Admiralty Inlet as Southern Resident killer whales transited through the study site to determine depth of whales in this area. The vertical array consisted of four hydrophones at ten meter offsets with the shallowest at a depth of 10 meters. A total of 682 calls and echolocation clicks were localized using hyperbolic localization techniques. The Time Of Arrival Difference (TOAD) was calculated using cross-correlation techniques for all calls and click trains. TOAD values were calculated for individual clicks by hand picking the times of arrival. Where present, surface reflections were also incorporated into the localizations of single clicks. We validated our techniques by generating signals at known depths (10 to 60 meters) and distances (100 to 500 meters). Measurement errors were calculated for each localization. Results of the validation study, error estimates, and depth measurements will be presented.

Fine-scale 3-D tracking of fish behavior in Central California using acoustic tags
Samuel Johnston, Tracey Steig & David Ouellette

Micro-acoustic tags have been used to monitor the fine-scale three-dimensional behavior and survival of fish and other aquatic life in the Pacific Northwest and Central California for years. Three-dimensional tracks are obtained at dams, lakes, open rivers, estuaries, and in marine environments. Resolution of three-dimensional positions are sub-meter with some resolutions as fine as 20 cm. In addition to juvenile and adult salmonids, other species tracked include eel, lamprey, sturgeon, shad, crab and shrimp. Salmonid smolts as small as 92 mm have been tagged and tracked with acoustic tags, which weigh as little as 0.5 g. Today's acoustic tag technology allows researchers to view fish behavior and passage in real-time. Tags operated at 307 kHz with a user-specified pulse width of 0.5-5 msec. Recent innovations include the development of smaller tags weighing 0.5 g in air, longer life tags, remote access via smart-phone, as well as various data display options. A number of advances in the analysis methods, techniques and software have been made over the past several years. Some of these improvements include the development of various fish density algorithms and advances of three-dimensional animation programs. Three-dimensional tracks of fish approaching various structures will be presented. Examples of fish tracks and fish densities will be superimposed over bathymetry, water velocities and structures. Recent Sacramento-San Joaquin River examples will be featured.

Ocean Acoustic Inversion

Bayesian model selection using evidence computations
Jan Dettmer & Stan Dosso

This paper considers approaches to computing the evidence (Z) in Bayesian inference problems for model selection in geoacoustic inversion. Bayes' theorem combines the likelihood function, model prior, and Z to form the posterior probability density (PPD). Z is difficult to compute for general problems and a common approach is to avoid its computation entirely by calculating an unnormalized estimate of the PPD which is sufficient for moment estimates. However, estimating the normalized PPD, including Z, allows for moment estimates as well as quantifying the likelihood of the model parameterization. This is commonly referred to as model selection and poses a natural way to quantifying the most appropriate model parameterization for a given data set (Bayesian razor). Several approaches for computing Z have been developed in the statistics community. Here, annealed importance sampling is applied to the geoacoustic inference problem. Annealed importance sampling follows an annealing approach and computes weighted averages along cooling trajectories. Both methods also give parameter estimates which are compared to Metropolis-Hastings results. [Work supported by the Office of Naval Research]

Transdimensional geoacoustic inversion
Jan Dettmer, Stan Dosso & Charles Holland

This paper applies a general trans-dimensional Bayesian approach to geoacoustic inversion. Trans-dimensional inverse problems are a generalization of fixed dimension inversion and include the number of model parameters as an unknown in the problem. A model is considered to be any particular choice of physical theory, its appropriate parameterization, and statistical representation for the data errors that are used to explain the physical system under examination. By including the dimension of the search space in the inversion, a joint posterior distribution is obtained that characterizes the state of knowledge about

parameters, including effects due to limited knowledge about the parametrization of the underlying environment and error processes. The inversion is implemented using a reversible-jump Markov chain Monte Carlo algorithm and the environment is parametrized with a partition modeling approach. Data errors are assumed as unknown and addressed by including a data-error model in the inversion. Jumps between dimensions are addressed with a birth-death methodology that allows the algorithm to wander dimensions by adding or removing interfaces from the seabed partition while maintaining detailed balance of the Markov chain. The approach presented can generally be applied to geoacoustic inverse problems. Here, seabed reflection-coefficient data as function of frequency and angle are considered. [Work supported by the Office of Naval Research]

Underwater Acoustics

Underwater Acoustic Levels of Southeast Alaska Cruise Ships

Blair Kipple & Chris Gabriele

Radiated acoustic levels for ten cruise ships that frequently travel Southeast Alaska waters were measured at the U.S. Navy's Southeast Alaska Acoustic Measurement Facility near Ketchikan, Alaska to quantify their underwater acoustic levels. This group of ships included diesel-electric, diesel-electric/gas turbine-electric, direct-diesel, and steam turbine propulsion plant ships ranging in size from 90 to 294 meters in length. Peak one-third octave levels for 10-knot ship speeds ranged from 158 to 172 decibels relative to 1 microPascal at 1 yard. Propulsion system type and cavitation performance were important factors in cruise ship acoustic level and spectral character. Diesel-electric ship acoustics were dominated by noise energy from diesel generators and from electric propulsion motors in combination with frequency converters and diesel generators. Propulsion diesel and reduction gear noise were important contributors in the direct diesel ships' underwater acoustic characteristics. Turbine generator, propulsion turbine, and reduction gear noise were the most significant noise items for the steam plant ship. Each ship was tested at two speeds. The sound levels of some ships were strongly speed dependent while others exhibited less speed dependence. Differences in acoustic levels between speeds were typically dependent on propulsion system and propeller cavitation noise contributions.

Acoustic propagation sensitivity to variability and uncertainty of the ocean environment: a comparison of modeled and measured data

Sean Pecknold, Cristina Tollefsen & John Osler

Sonar performance prediction models are a key tool for planning and carrying out sonar operations in littoral environments. Less well understood are the impacts that variability and uncertainty in the environment may have on propagation predictions. A sensitivity model was previously developed [Dosso et al, J. Acoust. Soc. Am. 121 (1), 2007] to determine the effects of limited sampling of, and uncertainty or variability in, oceanographic and geo-acoustic information on acoustic propagation modeling. Here, the model is applied to the conditions found during two field trials off the coast of Nova Scotia, using environmental data collected both initially from historical databases and then in-situ. The effects of the improved in-situ environmental sampling are demonstrated using the results of the sensitivity analysis. The sensitivity of the propagation and uncertainties in modeled propagation loss are then compared to the propagation loss data collected during the field trials.

Characterization of scattered acoustic intensity fields in the resonance region of a motionless rigid sphere

Robert Barton & Kevin Smith

In this study, the properties of the scattered acoustic vector fields generated by simple rigid spheroids are investigated. Analytical solutions are derived from general acoustic pressure scattering models, and analyzed for wave numbers in the resonance region. Of particular interest is the understanding of the characteristics of energy flow of the scattered acoustic vector field in the near to far-field transition region. The separable active and reactive components of the acoustic intensity are used to investigate the structural features of the scattered field components. Numerical results are presented for the near and transition region for a rigid sphere. The ability to extract scattered field features is illustrated with measurements obtained from a recent in-air experiment using an anechoic chamber and acoustic vector sensor probes to measure the scattered acoustic vector field from rigid spheroids.

Advanced Audio Applications

Transmission characteristics of two tactical communication headsets with hearing protection capabilities

Christian Giguère, Chantal Laroche & Véronique Vaillancourt

Communication headsets are increasingly used in the workplace. In some of the most challenging environments (e.g. military, law-enforcement tactical operations), the device must protect hearing against hazardous continuous and/or impulse noise while maintaining good situational awareness (e.g. sound localization, speech communication) within the immediate surrounding and

during radio communications with remote locations. These objectives must be met despite the presence of noise-induced and other types of hearing losses among users. Several analog and digital headsets are commercially available with adaptive level-dependent passive/active noise reduction, user-adjustable talk-through or surround volume, in addition to radio capabilities. Unfortunately, the technical specifications supplied by manufacturers are very limited, even for high-end products, and do not readily allow for a systematic analysis of the best devices or parameters to use in specific communication scenarios given the noise characteristics, task demands and hearing status of the user. This is in contrast to the hearing aid industry, where product testing and specification sheets are very extensive and highly standardized, and where systematic fitting procedures exit to optimize communication. In the paper, the characteristics of two high-end tactical communication devices are reported. The devices are the Peltor Powercom Plus (circumaural) and the Nacre QUIETPRO (intraaural). The test battery included measurements of (1) the passive sound attenuation, (2) the insertion gain and compression parameters at various control settings of the talk-through/surround modes, and (3) the speech intelligibility in two military noises with subjects covering a wide range of hearing profiles. [Work supported by DRDC Toronto].

DST a novel approach for noise dependent hearing protectors
Engbert Wilmink & Pieter van 't Hof

A novel concept for dynamic hearing protection has been developed based on an automatically opening and closing gate. Using miniaturized electronics this solution can be worn in the ear with a very low energy consumption meeting international standards for hearing protectors. The underlying physics of dampening sound through a varying orifice will be discussed in relation to attenuation data obtained in the lab and with test persons. Current CE certification methods of dynamic hearing protectors pose an issue to test the devices compared typical usage conditions. A lab based test method is proposed to ensure quality and performance.

Measurements of noise exposure during wildfire air firefighting operations
Hugues Nelisse, Jerome Boutin, Martine Gendron & Tony Leroux

The use of air attack on wildfire causes fixed-wing aircraft and helicopters to operate in close proximity within a designated airspace. The Bird Dog crew members are charged with the responsibility to co-ordinate the direction of air traffic over and in the immediate vicinity of wildfire alongside the ground firefighting. The crew of the Birddog consists of a pilot and a specially trained forest fire fighting person. These workers wear communication headsets during their entire mission, can receive transmissions through as much as 9 to 10 channels and can fly over long period of time ranging from 3-4 hours to 8-9 hours. There are thus great concerns regarding the noise exposure for such activities. This paper presents the comparison of two techniques for measuring the exposed levels (noise under the headset) for different workers in real flight conditions. The first technique, based on the F-MIRE technique, uses dual miniature microphones inserted into the ear muff to measure the noise outside and inside the ear muff. The second technique uses a probe-tube microphone placed close to the tympanic membrane to measure the exposed signal. A communication headset was instrumented with the necessary equipment for both techniques and simultaneous time recordings are performed during entire flights missions to allow systematic comparisons of the two approaches. Results for different workers and different flight conditions are presented and discussed.

Temporal metaphorics in auditory strategies of environmental monitoring
Joachim Gossmann

The organisation of sound for auditory monitoring is often considered from a perspective of superimposed simultaneous layers of sound: The grouping of auditory elements into streams, or the deconstruction of auditory scenes into different sound sources. A temporal perspective of sound is highlighted in Murray Schafer's investigation into soundscapes - in which the elements are regarded in their spatio-temporal relationship as part of an environmental ecology. In this paper, we would like to highlight the perspective of the human perceiver on the temporality of sound - our own contribution to the way sound occurs to us and how we can enable us as listening beings to derive more meaning from what we hear. Here, we would like to focus specifically on the temporal aspects. Marshal McLuhan highlights the nature of media as "Extensions of Man", while George Lakoff emphasises the importance of metaphorical structures to the way the world occurs to us. Temporal metaphors we could apply in this bi-directional information exchange between our approach to the world and the way the world occurs to us can come from a variety of origins (the time of music, familiar temporalities in our environment and everyday experience, temporal structures of communication and our own body, among many others). As concrete implementations scientific sonifications and audifications of seismological data are presented and evaluated: - A perceptual analysis of audified seismograms. - sonification of earthquake event catalogues under the application of different temporal metaphors This will guide us through an investigation of sound generation, how temporal structuring of sounds can target our specific perceptual abilities better, and finally, the potentials of our own openness of listening "for". This will provide us with concrete handles to make auditory applications in the field of discovery-oriented science more relevant to a human listener.

Soundscapes

The Toronto Sound Map Project

Frank Russo

Although humanistic and artistic approaches to soundscapes have flourished over the last half-century and particularly in Canada, there is very little that is known about the psychology of urban soundscapes. To this end, we have recently captured over 200 soundscapes from the metropolitan Toronto area. Each soundscape consists of a 2-minute binaural recording with accompanying sound-level measurements. The recordings will serve as stimuli for a new program of research that aims to generate knowledge about the psychology of the urban soundscape. Progress on three of the initial studies in this program will be described. Study 1 is a multidimensional scaling study in which participants will be asked to rate the similarity of soundscapes presented in pairs. The scaling solution will provide insight into the underlying cognitive representation of urban soundscapes. Study 2 combines electrophysiological and behavioral methods to examine stress response and recovery to/from commonly experienced soundscapes that have been described as aversive. Study 3 examines the influence of soundscapes on the useful field of view. In all studies, soundscapes will be presented over headphones in a double-walled IAC chamber at levels that are calibrated with the accompanying SPL measurements. We expect that physical and psychophysical dimensions that are independent of sound intensity will influence the various questions under investigation. In addition to the planned experimental work, we will report on the development of a website (torontosoundmap.com) that provides map-based navigation of the soundscapes and public dissemination of the research findings.

Speech Communication

Acoustic diagnostics of prosodic phrasing in SENĆOŦEN

Janet Leonard

Acoustic Diagnostics of Prosodic Phrasing in SENĆOŦEN (Saanich, North Straits Salish). It is widely accepted that languages organize grammatical information into prosodic units (e.g. Selkirk 1986,) and that evidence for prosodic structure should be reflected in the acoustic signal (e.g. Shattnuck-Hufnagel and Turk 1995). Building from previous work investigating the acoustic correlates associated with Salish prosodic structure, this paper seeks to examine the relationship between the acoustic signal and prosodic structure in SENĆOTEN. In particular, a set of acoustic correlates associated with the phonological phrase are determined (see Beck 1999, Koch 2008) by examining the acoustic properties of a corpus of SENĆOŦEN sentences elicited during fieldwork sessions with two fluent speakers. The results of this study illustrate that 1) pauses in pitch, along with variations in F0, coincide with the predicted boundaries between phonological phrases, 2) high F0 coincides with the predicted head of a phonological phrase and 3) longer vowel duration coincides with the end of a phonological phrase. This paper contributes to the growing body of literature investigating the nature of the relationship between phonology and phonetics in SENĆOŦEN. Beck, D. (1999). Words and Prosodic Phrasing in Lushootseed. In T. Alan Hall & Ursula Kleinhenz, (eds,). Studies on the Phonological Word (pp 23-46). Amsterdam: Benjamins. Koch, K. (2008). Intonation and Focus in Nlhe7kepmxcin (Thompson River Salish). UBC PhD Thesis Selkirk, E. (1986). On derived domains in sentence phonology. Phonology Yearbook, 3 371-405. Shattnuck-Hufnagel, S., and Turk, A. E. (1996). A prosody tutorial for investigators of sentence processing. Journal of Psycholinguistic Research, 25(2): 193-247.

Perception of stress on accented and unaccented words: A comparison between native and nonnative English speakers

Qian Wang

Native English speakers rely on F0, duration, and intensity in the perception of lexical stress. Second Language Acquisition studies have examined the use of these acoustic cues by non-native speakers in the perception of English lexical stress. A problem common to these SLA experiments was that words used for the perception tests were either in citation forms or excised from focused position in a sentence. The lexical stress on these words was, thus, confounded by phrasal accent, also cued by F0. Little is known about how well non-native speakers would discriminate lexical stress on words that bear no phrasal accent and how their performance would differ from native English speakers. In this experiment, Mandarin Chinese learners of English (CE) and native English speakers (NE) were compared in the perception of lexical stress on words excised from both ACCENTED and UNACCENTED conditions in an oddity test. Participants listened to a triad with three accented tokens or with three unaccented tokens, and decided whether the tokens in a triad all have the same stress pattern or not. An analysis of error rates showed no significant difference between CE and NE in the accented context but CE made significantly more mistakes than NE in the unaccented condition. CE¡¯s comparable performance with NE in perceiving lexical stress in accented contexts may be ascribed to their tonal background and sensitivity to F0, whereas their difficulty with unaccented words may be due to their insensitivity to duration and intensity.

The long-term retention of fine-grained phonetic details: Evidence from a second language voice identification training task
Stephen Winters

This study investigated the extent to which listeners store in memory the acoustic cues to non-native phonetic contrasts. Native English listeners were trained to identify the voices of Thai speakers, from a series of individual words, in three separate training conditions. In one training condition, the words produced by each voice consistently bore one of Thai's five distinctive lexical tones. In a second training condition, each voice was consistently associated with one of Thai's three distinctive VOT categories. After three days of training, listeners were then presented with words in a generalization test, in which the previous associations between words and phonetic properties no longer held. In a third, control condition, the voices were not consistently associated with any particular phonetic property. Evidence from both training and generalization indicated that listeners used both tone and VOT properties to learn to identify voices, including particular VOT values and tone contours that are not meaningfully contrastive in English. Talker identification accuracy in both the tone and VOT training conditions was significantly better than in the control condition; talker identification accuracy also decreased significantly in generalization testing for the tone and VOT listeners, but not for listeners in the control condition. Since listeners showed greater perceptual dependence on tone-talker associations than on VOT-talker associations, listeners may be more sensitive to longer, prosodic cues than to shorter timing cues in speech. Overall, these results indicate that listeners do store in memory low-level phonetic details, including acoustic distinctions they might normally perceive categorically.

Imagery-induced context effects
Mark Scott

This experiment examines whether context effects can be induced by auditory imagery. How a sound is perceived often depends on what was heard immediately prior – the 'context' of the sound. One well-known context effect was reported by Mann (1980) who found that a sound which is ambiguous between /da/ and /ga/ tends to be categorized as /da/ if immediately preceded by /ar/, but as /ga/ if immediately preceded by /al/. The reason for this effect is believed to be that the low F3 of /ar/ makes the F3 of the following ambiguous sound seem higher (so more /d/-like) by comparison. The high F3 of /al/ has the opposite effect, making the following F3 seem lower and so more /g/-like. The current experiment tests whether auditory imagery of the context sound can induce the same effect; if so, it would suggest that auditory imagery has detailed phonetic content (including formant structure). The experiment compares the strength of the effect across 2 conditions: normal speech and imaged speech. In both conditions subjects produce (externally or internally) one of the context sounds (/ar/ or /al/) in a rhythm, and after several repetitions, a /da/~/ga/ ambiguous target sound is played which they must categorize. The experiment is currently being run (8 participants so far) and initial results suggest that imagery does indeed induce the context effect. As not all subjects experience this context-effect, it is interesting to note that susceptibility to the effect seems to be strongly correlated across the two conditions.

A perceptual study of [liquid + stop] sequences
Terrance Nearey & Benjamin Tucker

Lotto and Kluender (1998), following Mann(1980) studied the perception of stops in the four syllables like /arga, alga, arda, alga/. These experiments and several others used a /-ga/ to /-da/ varying along an F3-transition continuum. All experiments reported finding more /d/ responses following /ar-/ than /al-/ precursor syllables. We replicated part of Lotto and Kluender's Experiment 2 in preparation for more elaborate experiments manipulating /r/ and /l/ as well as /d/ and /g/. Preliminary findings with 16 listeners (native speakers of English) show effects broadly similar to those previously reported, with /al-/ precursors leading to more /d/ responses. However, mixed-effects logistic regression analysis (Laplace approximation) suggests that more than simple boundary shifts are involved. For stimuli in the /al-/ context, the estimated /g-d/ response curve shows a steeper slope than that in the /ar-/ context. The nature of this interaction may have consequences for evaluating competing perceptual accounts (auditory contrasts vs. compensation for coarticulation) of this phenomenon. Simple logistic regression analysis of individual listeners data revealed that only eight of 16 showed significant ($p<.05$) effects of /ar-/ context in the expected direction, while three showed significant effects in the wrong direction and 5 showed non significant effects. (All but one listener showed significant main effects in the expected direction for the d to g continuum.) Further analysis of these results and results from listeners who are non-native speakers of English will be reported.

Architectural Acoustics

Providing "good", "better" or "best" acoustical plumbing system proposals to cost sensitive clients
Chip O'Neil

Providing "Good", "Better", or "Best" proposals for acoustical noise and vibration isolation of piping systems to cost sensitive

clients When it comes to the costs associated with effective acoustical isolation of a plumbing or piping system there are a variety of choices available. Though effective isolation materials and methods are available for very modest costs, there is still often the challenge of "Value Engineering" to face. As in most facets of building construction, there are a variety of quality levels available when it comes to plumbing and piping system acoustic isolation options. Become familiar with "Good", "Better" and "Best" materials and methods, in order to assist you during the planning and budgeting stages of a building project. Learn to provide valuable input during the Cost–Benefit analysis phase of a building project. Base your input to your client upon solid laboratory test data arranged by specific plumbing or piping system applications, such as through-stud isolation, riser clamp isolation, shower head attachments, etc. This paper will provide fundamental presentation points and cost analysis templates easily customized to any plumbing or piping system. These analytical tools take into account both labor and material factors, in order to generate a true "Installed Cost Analysis", while assisting you to specify a proven engineered system for your client.

Subjective ranking of low-frequency impact and footstep sounds on lightweight floor-ceiling assemblies
Bradford Gover, John Bradley, Trevor Nightingale, Berndt Zeitler & Stefan Schoenwald

To rate the acoustical performance of floor-ceiling assemblies excited by impact sources (e.g., footsteps), an objective metric must correlate well with the subjective judgments of listeners hearing the radiated sounds. For each of a series of full-scale lightweight floor-ceiling assemblies, physical measurements and sound recordings were made of the sounds generated by standard impact sources (tapping machine, rubber impact ball, bang machine), and by adult walkers (with and without shoes). The physical measurements were used to calculate standard and non-standard metrics. The recordings were played back for listening test participants, who subjectively rated the radiated sounds. The correlations among the objective and subjective ratings were calculated. Results indicate that some standardized sources and metrics are not optimal for rating subjective performance.

Noise Control

Analysis and control of bridge expansion joint "croaking" noise
Clair Wakefield

Large bridges require expansion joints to accommodate thermal expansion and contraction and seismic events. Different styles and sizes of expansion joints are employed depending on bridge type, length and other factors. Virtually all expansion joints create some additional noise over and above normal that due to normal tire-pavement interaction. Many joints create "banging" or "booming" noises due to the impact of tires on leading edges of the joint, which often feature some vertical misalignment and/or mechanical looseness. These impulsive noises can be annoying to nearby residents. It is often possible to control this noise by reducing or eliminating such misalignment and/or looseness or by installing overlapping "finger joints" which reduce tire impact forces on the joints by essentially eliminating transverse leading edges. Wakefield Acoustics Ltd. has recently been involved in assessing and controlling noise from two types of expansion joints which produce a totally different sort of noise - one that results from the excitation of resonances inside the cavities temporarily created between the rolling tires and the gaps between transverse joint elements. The two bridges employ quite difference types of expansion joint. The first utilizes a modular expansion joint consisting of series of transverse "lamella" beams (I-beams) with v-shaped rubber seals between them, while the joints on the second bridge featured corrugated, "saw-tooth" like surfaces (expansion mats) constructed from rubber-encased steel strips. In both cases a similar noise is produced which has been variously described as sounding like a "croaking frog" or a "zipper". The characteristic frequency of this croaking sound was not found to be directly related to vehicle speed as would be expected if the noise was caused simply by the sequential impact of tires on the transverse joint elements. In both situation cases, these unfamiliar intermittent noises have resulted in ongoing complaints. A series of field tests have been carried out to explore the mechanisms behind these noises and to evaluate potential control measures. The outcomes of these field trials will be discussed as will the Helmholtz resonator analogy believed to explain how these noises are created and may be controlled.

Noise survey within patient care wards at the Royal Jubilee Hospital, Victoria BC
Andrew Williamson

Hospital noise levels around the world have increased steadily over the past 50 years. High noise levels are among the top complaints of both patients and hospital staff members. High noise levels disturb patients and staff members, hinder speech intelligibility and raise the risk of medical errors. Noise levels often remain high during night and can interfere with patients sleep. Wakefield Acoustics Ltd. has recently conducted a survey of the noise environment within the West and Royal Blocks of Royal Jubilee Hospital in Victoria, B.C. The noise monitoring was conducted continuously over an eight day period at four locations within each of the patient care wards. To identify the sources of noise within the hospital, one of the four sound level meters employed within each patient ward also recorded a digital audio file and several hours of attended monitoring was conducted in each ward. The results of the noise monitoring were used to identify the most significant sources of noise and their

respective noise levels. At a later date, this noise survey will be replicated within the new Patient Care Center at the Jubilee Hospital to see if noise control measures implemented within the new facility have been effective.

Vibration and Transport Vehicle Noise

Effect of helicopter noise and vibration on healthcare facilities

Rob Jozwiak

At an increasing rate, new hospitals and renovations to existing hospitals are incorporating helicopter pads (helipad's) into the design of their facility. Hospitals contain many sound and vibration sensitive spaces, which the operation of a helicopter may affect. The vibration generated from a helicopter landing has potential to impede on the operation of sensitive diagnostic equipment in the hospital and the noise generated by the helicopter turbines and propellers can disturb patients. This paper presents sound and vibration measurements of a Sikorsky S-76 medivac helicopter landing and taking off from St. Michael's Hospital in Toronto Ontario and provides an analysis of architectural and structural design considerations to minimize the affect of helicopter operation at medical health centres.

Methods for measuring Off-Highway Vehicle (OHV) sound emissions and their correlation with "near-to-track" and "off-track" sound levels

Ian Matthew, John Emeljanow & Mark Levkoe

With the impact of noise from recreational activity being an ever increasing concern, the accurate measurement/classification of sound emissions from off-highway vehicles (OHVs) for the purposes of controlling sound levels at noise-sensitive off-track locations is becoming more and more critical. Conventionally, sound emission testing for this purpose is conducted according to SAE J1287 (or equivalent) which uses a stationary vehicle in neutral gear with the engine at approximately half-throttle. However, the poor correlation between results using this test method and measured sound levels at "near-track" and "off-track" locations is well documented. A new method has been proposed by the Fédération Internationale de Motocyclisme (FIM) which purports to provide improved correlation between the stationary vehicle sound emission level and "near-track"/"off-track" sound levels, and as such would be expected to be an improved method for qualifying or disqualifying a vehicle from operation in a given locale. A direct comparison of the measured sound emission levels as determined using the SAE method and the proposed FIM method are presented, along with comparisons with "near-track" and "off-track" measurements in order to assess the potential for improved correlation and thus, improved control of environmental noise levels.

Application of automatic data processing at the systems of condition monitoring of industrial equipment

Alexander Serov

The report is devoted to the development problems of the monitoring and diagnosis system of the industrial equipment. Modern systems of stationary type ordinary use some different autonomous subsystems of the equipment monitoring running in the synchronous mode. The structure and the set of these subsystems depend upon the functional purposes of the monitoring and diagnosis system' using. Each of these subsystems use their own set of physical methods for the detection of the equipment characteristics. It leads to arising of the development problem of multi-parametrical monitoring and diagnosis methods. Using of the simultaneous monitoring of the same technical object by the different physical methods leads to the possibility of the essential generalization of the monitoring and diagnosis problems and to the formulation of the interesting set of general scientific problems. One of the main problems of this set is the problem of the non-uniform information processing. Data acquiring by the acquisition and saving subsystem is a non-uniform one both at the problem of the detection of the current state of the monitoring object and at the problem of the prediction for the dynamics of object characteristics. The approach to the processing of data based on the methodology of Group Methods of Data Handling (GMDH) and statistical data processing is proposed at this report. It makes possible to apply the methods of self-organization of the mathematical models to the problems of the monitoring of complex technical systems. Proposed method was realized as software and applied for the problems of monitoring of turbine engines.

Musical Acoustics

The complementary roles of temporal and spectral processing in tonal perception of low-frequency tones

Frank Russo, Lola Cuddy & Alexander Galembo

We assessed tonal perception in low-frequency harmonic tones that varied the extent of interaction between partials within critical bands. Hol(e)y tones minimize the likelihood of interaction by incorporating holes (missing harmonics) in the spectrum (Mathews, 1999). Unhol(e)y tones have fewer holes, leading to more interaction. Tonal perception was assessed by the probe-tone method. A listener is presented with a key-defining context followed by a probe tone drawn from the 12-tone chromatic scale. For each probe tone, the listener rates the goodness-of-fit with the preceding context. The correlation between the set of

12 ratings, called the probe-tone profile, and a standardized profile (Krumhansl & Shepard, 1979) is referred to as the "recovery score" (Russo, Cuddy, Galembo & Thompson, 2007). Our previous work (e.g., Cuddy, Russo & Galembo, 2007) examined recovery for synthesized harmonic and nonharmonic piano tones throughout the tessitura and identified critical roles for both temporal and spectral processes. Based on this work, we made three predictions for the current experiment:
(1) recovery scores should be superior for unholy tones due to interaction between components; (2) consistent with temporal models, recovery scores for unholy tones should decrease with increasing frequency due to components falling beyond the upper limit of phase locking; and (3) consistent with spectral models, recovery scores for holy tones should increase with increasing frequency due to the increased spacing between harmonics. Results fully supported these predictions and will be discussed in the context of pitch processing models.

INSTRUCTIONS TO AUTHORS FOR THE PREPARATION OF MANUSCRIPTS

Submissions: The original manuscript and two copies should be sent to the Editor-in-Chief.

General Presentation: Papers should be submitted in camera-ready format. Paper size 8.5" x 11". If you have access to a word processor, copy as closely as possible the format of the articles in Canadian Acoustics 18(4) 1990. All text in Times-Roman 10 pt font, with single (12 pt) spacing. Main body of text in two columns separated by 0.25". One line space between paragraphs.

Margins: Top - title page: 1.25"; other pages, 0.75"; bottom, 1" minimum; sides, 0.75".

Title: Bold, 14 pt with 14 pt spacing, upper case, centered.

Authors/addresses: Names and full mailing addresses, 10 pt with single (12 pt) spacing, upper and lower case, centered. Names in bold text.

Abstracts: English and French versions. Headings, 12 pt bold, upper case, centered. Indent text 0.5" on both sides.

Headings: Headings to be in 12 pt bold, Times-Roman font. Number at the left margin and indent text 0.5". Main headings, numbered as 1, 2, 3, ... to be in upper case. Sub-headings numbered as 1.1, 1.2, 1.3, ... in upper and lower case. Sub-sub-headings not numbered, in upper and lower case, underlined.

Equations: Minimize. Place in text if short. Numbered.

Figures/Tables: Keep small. Insert in text at top or bottom of page. Name as "Figure 1, 2, ..." Caption in 9 pt with single (12 pt) spacing. Leave 0.5" between text.

Line Widths: Line widths in techincal drawings, figures and tables should be a minimum of 0.5 pt.

Photographs: Submit original glossy, black and white photograph.

Scans: Should be between 225 dpi and 300 dpi. Scan: Line art as bitmap tiffs; Black and white as grayscale tiffs and colour as CMYK tiffs;

References: Cite in text and list at end in any consistent format, 9 pt with single (12 pt) spacing.

Page numbers: In light pencil at the bottom of each page.
Reprints: Can be ordered at time of acceptance of paper.

DIRECTIVES A L'INTENTION DES AUTEURS PREPARATION DES MANUSCRITS

Soumissions: Le manuscrit original ainsi que deux copies doivent être soumis au rédacteur-en-chef.

Présentation générale: Le manuscript doit comprendre le collage. Dimensions des pages, 8.5" x 11". Si vous avez accès à un système de traitement de texte, dans la mesure du possible, suivre le format des articles dans l'Acoustique Canadienne 18(4) 1990. Tout le texte doit être en caractères Times-Roman, 10 pt et à simple (12 pt) interligne. Le texte principal doit être en deux colonnes séparées d'un espace de 0.25". Les paragraphes sont séparés d'un espace d'une ligne.

Marges: Dans le haut - page titre, 1.25"; autres pages, 0.75"; dans le bas, 1" minimum; latérales, 0.75".

Titre du manuscrit: 14 pt à 14 pt interligne, lettres majuscules, caractères gras. Centré.

Auteurs/adresses: Noms et adresses postales. Lettres majuscules et minuscules, 10 pt à simple (12 pt) interligne. Centré. Les noms doivent être en caractères gras.

Sommaire: En versions anglaise et française. Titre en 12 pt, lettres majuscules, caractères gras, centré. Paragraphe 0.5" en alinéa de la marge, des 2 cotés.

Titres des sections: Tous en caractères gras, 12 pt, Times-Roman. Premiers titres: numéroter 1, 2, 3, ..., en lettres majuscules; sous-titres: numéroter 1.1, 1.2, 1.3, ..., en lettres majuscules et minuscules; sous-sous-titres: ne pas numéroter, en lettres majuscules et minuscules et soulignés.

Equations: Les minimiser. Les insérer dans le texte si elles sont courtes. Les numéroter.

Figures/Tableaux: De petites tailles. Les insérer dans le texte dans le haut ou dans le bas de la page. Les nommer "Figure 1, 2, 3,..." Légende en 9 pt à simple (12 pt) interligne. Laisser un espace de 0.5" entre le texte.

Largeur Des Traits: La largeur des traits sur les schémas technique doivent être au minimum de 0.5 pt pour permettre une bonne reproduction.

Photographies: Soumettre la photographie originale sur papier glacé, noir et blanc.

Figures Scanées: Doivent être au minimum de 225 dpi et au maximum de 300 dpi. Les schémas doivent être scannés en bitmaps tif format. Les photos noir et blanc doivent être scannées en échelle de gris tifs et toutes les phoots couleurs doivent être scannées en CMYK tifs.

Références: Les citer dans le texte et en faire la liste à la fin du document, en format uniforme, 9 pt à simple (12 pt) interligne.

Pagination: Au crayon pâle, au bas de chaque page.
Tirés-à-part: Ils peuvent être commandés au moment de l'acceptation du manuscrit.

The Canadian Acoustical Association
l'Association Canadienne d'Acoustique

Application for Membership

CAA membership is open to all individuals who have an interest in acoustics. Annual dues total $80.00 for individual members and $35.00 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 $80.00. Many companies and institutions prefer to be a Sustaining Subscriber, paying $350.00 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
2. Engineering Acoustics / Noise Control
3. Physical Acoustics / Ultrasound
4. Musical Acoustics / Electro-acoustics
5. Psychological / Physiological Acoustic
6. Shock and Vibration
7. Hearing Sciences
8. Speech Sciences
9. Underwater Acoustics
10. Signal Processing / Numerical Methods
11. Other

For student membership, please also provide:

_____ _____ _____ _____

(University) *(Faculty Member)* *(Signature of Faculty Member)* *(Date)*

I have enclosed the indicated payment for:
[ ] CAA Membership $ 80.00
[ ] CAA Student Membership $ 35.00

Corporate Subscriptions:
[ ] $80 including mailing in Canada
[ ] $88 including mailing to USA,
[ ] $95 including International mailing

[ ] Sustaining Subscriber $350.00
includes subscription (4 issues /year)
to ***Canadian Acoustics.***

Payment by: [ ] Cheque
[ ] Money Order
[ ] Credit Card *(Indicate VISA or M/C)*

Credit card number _____

Name on card _____

Expiry date _____

_____
(Signature) (Date)

Mail application and attached payment to:

l'Association Canadienne d'Acoustique
The Canadian Acoustical Association

Formulaire d'adhésion

L'adhésion à l'ACA est ouverte à tous ceux qui s'intéressent à l'acoustique. La cotisation annuelle est de 80.00$ pour les membres individuels, et de 35.00$ 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 80.00$. Des compagnies et établissements préfèrent souvent la cotisation de membre bienfaiteur, de 350.00$ 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 electroniques, visitez le site Web: **caa-aca.ca**

Pour correspondence administative et financière:

Nom / Organisation _____

Adresse _____

Ville/Province _____ Code postal _____Pays _____

Téléphone _____ Téléc. _____ Courriel _____

Adresse postale pour la revue Acoustique Canadienne

Nom / Organisation _____

Adresse _____

Ville/Province _____ Code postal _____Pays _____

Cocher vos champs d'intérêt: (maximum 3)

1. Acoustique architecturale
2. Génie acoustique / Contrôle du bruit
3. Acoustique physique / Ultrasons
4. Acoustique musicale / Electro-acoustique
5. Physio / Psycho-acoustique
6. Chocs et vibrations
7. Audition
8. Parole
9. Acoustique sous-marine
10. Traitement des signaux /Méthodes numériques
11. Autre

Prière de remplir pour les étudiants et étudiantes:

_____ _____ _____ _____

(Université) (Nom d'un membre du corps professoral) (Signature du membre du corps professoral) (Date)

Cocher la case appropriée:

[ ] Membre individuel 80.00 $
[ ] Membre étudiant(e) 35.00 $

Abonnement institutionnel

[ ] 80 $ à l'intérieur du Canada
[ ] 88 $ vers les États-Unis
[ ] 95 $ tout autre envoi international
[ ] Abonnement de soutien 350.00 $ (comprend l'abonnement à *L'acoustique Canadienne*)

Méthode de paiement:

[ ] Chèque au nom de l'Association Canadienne d'Acoustique
[ ] Mandat postal
[ ] VISA ou M/C (Indiquez)

Numéro carte de crédit _____

Nom sur la carte _____

Date d'expiration _____

_____

(Signature) (Date)

Prière d'attacher votre paiement au formulaire d'adhésion. Envoyer à :
Secrétaire exécutif, Association Canadienne d'Acoustique, CP 74068, Ottawa, K1M 2H9, Canada

The Canadian Acoustical Association
l'Association Canadienne d'Acoustique

SUSTAINING SUBSCRIBERS / ABONNES DE SOUTIEN

The Canadian Acoustical Association gratefully acknowledges the financial assistance of the Sustaining Subscribers listed below. Their annual donations (of $350.00 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 $350.00 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 - Montréal, QC

Bruel & Kjaer North America Inc.
Mr. Andrew Khoury - (514) 695-8225
andrew.khoury@bksv.com
Pointe-Claire, QC

Dessau Inc.
Jacques Boilard - (418) 839-6034
jacques.boilard@dessau.com - Lévis, QC

H.L. Blachford Ltd.
Mr. Dalton Prince - (905) 823-3200
amsales@blachford.ca - Mississauga, ON

Hydro-Quebec TransEnergie
M. Blaise Gosselin - (514) 879-4100 ext 5309
Gosselin.Blaise@hydro.qc.ca - Montréal, QC

J.L.Richards & Assoc. Ltd.
Mr. Terry Vivyurka, P.Eng. - (613) 728-3571
mail@jlrichards.ca - Ottawa, ON

Mc SQUARED System Design Group
Mr. Wade McGregor - (604) 986-8181
info@mcsquared.com - North Vancouver, BC

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

Sound & Vibration Solutions Canada
Mr. Andy Metelka - (519) 853-4495
ametelka@cogeco.ca - Acton, ON

State of the Art Acoustik Inc.
Dr. C. Fortier - (613) 745-2003
sota@sota.ca - Ottawa, ON

True Grit Consulting Ltd.
Ina Chomyshyn - (807) 626-5640
ina@tgcl.ca - Thunder Bay, ON

Wakefield Acoustics Ltd.
Mr. Clair Wakefield - (250) 370-9302
clair@wakefieldacoustics.com - Victoria, BC

XSCALA Sound & Vibration
Jim Ulicki - (403) 274-7577
info@XSCALA.COM - Calgary, AB

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@craworld.com - Waterloo, ON

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

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

Soundtrap Inc.
Roger Foulds - (705) 357-1067
roger@soundtrap.ca - Sunderland, ON

Swallow Acoustic Consultants Ltd.
Mr. John Swallow - (905) 271-7888,
jswallow@jsal.ca - Mississauga, 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@wccl.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
johno@aercoustics.com- Toronto, ON

Dalimar Instruments Inc.
Mr. Daniel Larose - (514) 424-0033
daniel@dalimar.ca - Vaudreuil-Dorion, QC

ECORE International
Mr. Paul Downey - (416) 440-1094
pcd@ecoreintl.com - Toronto, ON

HGC Engineering Ltd.
Mr. Bill Gastmeier - (905) 826-4044
bgastmeier@hgcengineering.com
Mississauga, ON

J.E. Coulter Associates Ltd.
Mr. John Coulter - (416) 502-8598
jcoulter@on.aibn.com - Toronto, ON

JASCO Research Ltd.
(902) 405-3336; FAX: (902) 405-3337
scott@jasco.com - Halifax, NS

Owens-Corning Canada Inc.
Mr. Salvatore Ciarlo - (800) 988-5269
salvatore.ciarlo@owenscorning.com -
St.Leonard, QC

Pyrok Inc.
Mr. Howard Podolsky - (914) 777-7770
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

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

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