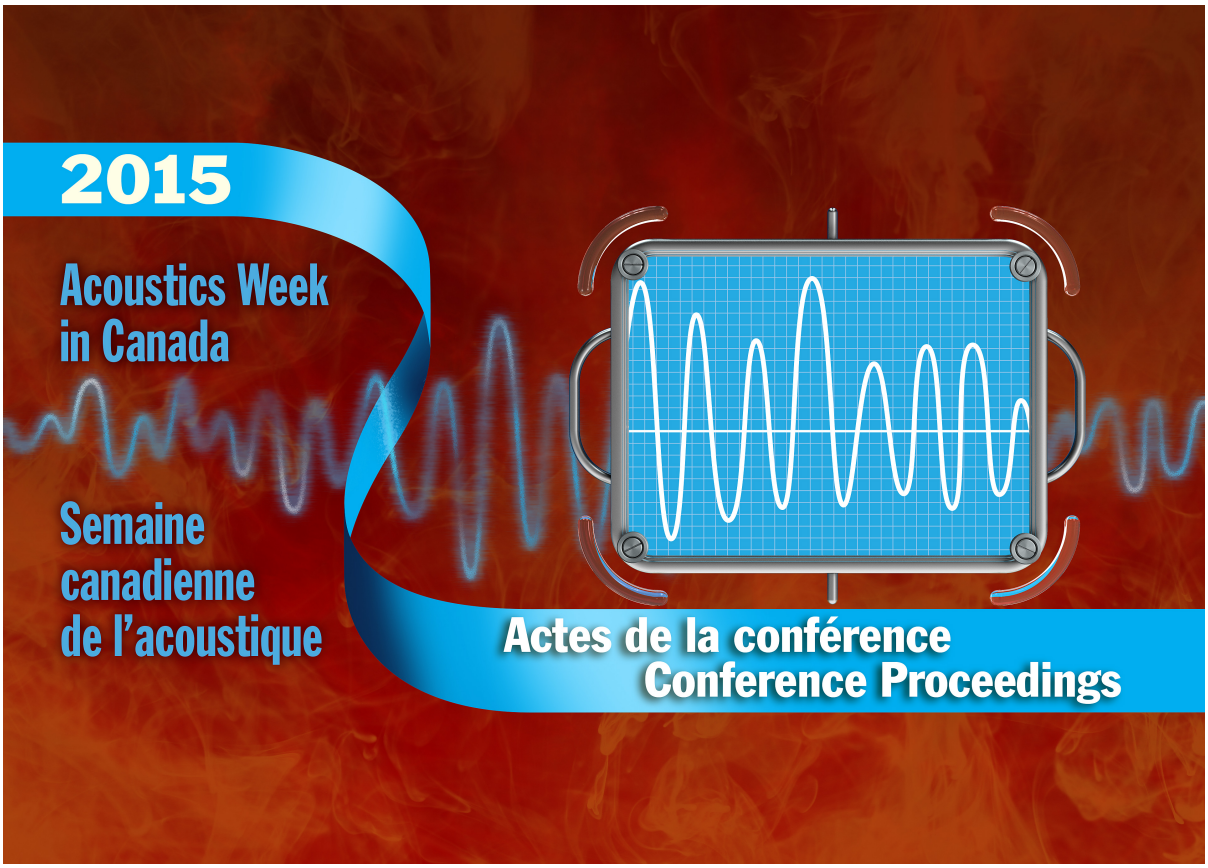
ISSN 0711-6659

canadian acoustics
acoustique canadienne

Journal of the Canadian Acoustical Association - Revue de l'Association canadienne d'acoustique

SEPTEMBER 2015
Volume 43 - - Number 3

SEPTEMBRE 2015
Volume 43 - - Numéro 3



canadian acoustics

acoustique canadienne

Canadian Acoustical Association/Association Canadienne d'Acoustique P.B. 74068 Ottawa, Ontario, K1M 2H9

Association canadienne d'acoustique B.P. 74068 Ottawa, Ontario, K1M 2H9

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 only through the journal online submission system. Complete instructions to authors concerning the required "camera-ready" manuscript are provided within the journal online submission system.

L'Acoustique Canadienne publie des articles arbitrés et des informations sur tous les aspects de l'acoustique et des vibrations. Les informations portent sur la recherche, les ouvrages sous forme de revues, les nouvelles, l'emploi, les nouveaux produits, les activités, etc. Des articles concernant des résultats inédits ou des applications ainsi que les articles de synthèse ou d'initiation, en français ou en anglais, sont les bienvenus.

Canadian Acoustics is published four times a year - in March, June, September and December. This quarterly journal is free to individual members of the Canadian Acoustical Association (CAA) and institutional subscribers. Canadian Acoustics publishes refereed articles and news items on all aspects of acoustics and vibration. It also includes information on research, reviews, news, employment, new products, activities, discussions, etc. Papers reporting new results and applications, as well as review or tutorial papers and shorter research notes are welcomed, in English or in French.The Canadian Acoustical Association selected Paypal as its preferred system for the online payment of your subscription fees. Paypal supports a wide range of payment methods (Visa, Mastercard, Amex, Bank account, etc.) and does not requires you to have already an account with them. If you still want to proceed with a manual payment of your subscription fee, please complete the application form and send it along with your cheque or money order to the secretary of the Association (see address above). - Canadian Acoustical Association/Association Canadienne d'Acoustique P.B. 74068 Ottawa, Ontario K1M 2H9 Canada - - - secretary@caa-aca.ca - Dr. Roberto Racca

Acoustique canadienne est publié quantre fois par an, en mars, juin, septembre et décembre. Cette revue trimestrielle est envoyée gratuitement aux membres individuels de l'Association canadienne d'acoustique (ACA) et aux abonnés institutionnels. L'Acoustique canadienne publie des articles arbitrés et des rubriques sur tous les aspects de l'acoustique et des vibrations. Ceci comprend la recherche, les recensions des travaux, les nouvelles, les offres d'emploi, les nouveaux produits, les activités, etc. Les articles concernant les résultats inédits ou les applications de l'acoustique ainsi que les articles de synthèse, les tutoriels et les exposées techniques, en français ou en anglais, sont les bienvenus.L'Association canadienne d'acoustique a sélectionné Paypal comme solution pratique pour le paiement en ligne de vos frais d'abonnement. Paypal prend en charge un large éventail de méthodes de paiement (Visa, Mastercard, Amex, compte bancaire, etc) et ne nécessite pas que vous ayez déjà un compte avec eux. Si vous désirez procéder à un paiement par chèque de votre abonnement, on vous invite à compléter le formulaire d'adhésion et l'envoyer avec votre chèque ou mandat au secrétaire de l'association (voir adresse ci-dessus). - Canadian Acoustical Association/Association Canadienne d'Acoustique P.B. 74068 Ottawa, Ontario K1M 2H9 Canada - - - secretary@caa-aca.ca - Dr. Roberto Racca

WELCOME TO HALIFAX!

ACOUSTICS WEEK IN CANADA 2015

October 7-9, 2015

Halifax Central Library (photo by Citobun, Wikipedia)

Halifax Public Gardens (photo by Lisa Paul, flickr.com)

Organizing Committee

Conference Chairs:	Michael Kiefte, Dalhousie University
	Sean Pecknold, DRDC Atlantic Research Centre
Technical Chair:	Steve Aiken, Dalhousie University
Exhibit Coordinator:	Lisa Cooper, JASCO Applied Sciences
Proceedings Copyediting:	Cécile Le Cocq, Université du Québec (ÉTS)
Student Prizes and Subsidies:	Hugues Nélisse, IRSST

CONFERENCE WEBSITE:
http://awc.caa-aca.ca

BIENVENUE À HALIFAX!

SEMAINE CANADIENNE DE L'ACOUSTIQUE 2015

7 au 9 octobre 2015

Halifax Citadel (photo © Parks Canada)

Comité organisateur

Présidentes de la conférence:	Michael Kiefte, Dalhousie University
	Sean Pecknold, DRDC Atlantic Research Centre
Directeur scientifique:	Steve Aiken, Dalhousie University
Coordinateur exposants:	Lisa Cooper, JASCO Applied Sciences
Correction des actes:	Cécile Le Cocq, Université du Québec (ÉTS)
Prix et subsides étudiants:	Hugues Nélisse, IRSST

SITE DE LA CONFERENCE:

http://awc.caa-aca.ca

Sponsors/Commanditaires

Exhibitors/Exposants

Éditorial invité
Guest Editor's note

Bienvenue à Halifax !

Au nom du Comité organisateur de la conférence, bienvenue à la Semaine canadienne de l'acoustique de 2015. Nous avons un programme technique palpitant couvrant un large éventail de sujets en acoustique, avec des conférences plénières de F. Stuart Foster, Dennis Phillips, et John Osler, ainsi que d'autres présentateurs distingués.

La conférence de cette année se tiendra à Halifax, en Nouvelle-Écosse, capitale du divertissement et de la culture au Canada atlantique. La conférence se tiendra à l'hôtel Westin Nova Scotian, idéalement situé près du port, et juste en face du quartier du port de mer d'Halifax et de l'historique Pier 21, qui sera l'emplacement du banquet de la conférence le jeudi soir. L'hôtel est également à quelques minutes de la bibliothèque centrale de Halifax nouvellement construite et qui a été en lice pour la construction mondiale de l'année à 2015, et sera l'emplacement de notre réception de bienvenue, le mardi soir..

Pendant que vous êtes dans la ville, nous espérons que vous pourrez profiter de quelques-uns des sites et activités qu'Halifax a à vous offrir. Promenez-vous le long du front de mer ou dans les jardins publics, visiter le lieu historique national de la Citadelle d'Halifax ou le Musée maritime de l'Atlantique, ou encore déguster les produits de l'une des nombreuses micro-brasseries de Halifax. Pour ceux qui souhaitent voir plus de la Nouvelle-Écosse en dehors de la ville, nous avons également pris des dispositions pour que le taux préférentiel accordé à la conférence puisse également s'appliquer après la conférence dans les belles hébergements de Digby Pines et Liscombe Lodge.

Nous tenons à remercier tout le comité d'organisation pour le travail acharné et le dévouement nécessaires pour planifier la conférence, en particulier Steve Aiken, notre président du comité technique, Lisa Cooper, pour la coordination des exposants et des sponsors, Jérémie Voix et Cécile Le Cocq pour les actes de la conférence, Hugues Nélisse pour les bourses et les subventions pour les étudiants, et enfin Frank Russo, président de l'ACA, pour son soutien.

Ceud Mìle Fàilte!

Michael Kiefte et Sean Pecknold
Présidents de la conférence

Welcome to Halifax!

On behalf of the Conference Organizing Committee, welcome to Acoustics Week in Canada 2015. We have an exciting technical programme covering a wide range of topics in acoustics, featuring plenary talks by F. Stuart Foster, Dennis Phillips, and John Osler, along with other distinguished presenters.

This year's conference is being held in Halifax, Nova Scotia, Atlantic Canada's entertainment and cultural capital. The conference will be held in the Westin Nova Scotian hotel, ideally located near the harbour, and just across from the Halifax Seaport District and historic Pier 21, the location of Thursday night's Conference Banquet. The hotel is also few minutes away from the newly constructed Halifax Central Library, the location of our Tuesday evening Welcome Reception, which has been shortlisted for the 2015 World Building of the Year Award.

While you are in the city, we hope you can enjoy some of the sights and activities Halifax has to offer. Stroll along the waterfront or in the Public Gardens, visit the Halifax Citadel National Historic Site or the Maritime Museum of the Atlantic, or sample the products of one of Halifax's many microbreweries. For those wishing to see more of Nova Scotia outside the city, we have also arranged for the conference rate to apply following the conference with the beautiful Digby Pines and Liscombe Lodge resorts.

We would like to thank all of the Organizing Committee for the hard work and dedication required to plan the conference, in particular Steve Aiken, our technical chair, Lisa Cooper, for coordinating the exhibits and sponsors, Jérémie Voix and Cécile Le Cocq for the conference proceedings, Hugues Nélisse for student awards and subsidies, and Frank Russo, President of the CAA, for his support.

Ceud Mìle Fàilte!

Sean Pecknold and Michael Kiefte
Conference Chairs

Informations voyageur
Travel information

Halifax Stanfield International Airport est le centre du Canada Atlantique pour le service de vols intérieurs, régionals et internationals. The Westin Nova Scotian est à environ 37 min. de l'aéroport en voiture. Plusieurs compagnies de location de voitures fonctionnent à l'aéroport (http://hiaa.ca/transportation/car-rentals/). Si vous louez une voiture et si vous conduisez de l'aéroport notez bien que le pont McDonald entre Halifax et Dartmouth est fermé du dimanche au jeudi de 19h00 à 5h30 pour des améliorations assez radicales. Durant cette période, vous pouvez utiliser le pont McKay. Un service de taxi coûte environ 63$ pour un aller simple et les taxis sont fréquemment garés juste en dehors de la zone des arrivés. Pour 22$, il y a une navette qui vous emmène à l'hôtel Westin Nova Scotian (http://maritimebus.com/halifax-airport-shuttle/) et vous pouvez acheter un billet tout près de la sortie des arrivées. Il y a aussi un autobus de l'Halifax Transit qui part toutes les 60 ou 30 minutes et qui est stationné juste en face du débarcadère-arrivé près de la zone des arrivées. Les tarifs sont de 3,50$ l'aller simple et il prend environ 1 heure 15 minutes et il nécessite un transfert à Duke Street à Halifax. Les informations de trajets d'Halifax Transit sont disponibles sur Google Maps.

Pour ceux qui arrivent par Via Rail, le Westin Nova Scotian est attaché à la gare.

Halifax a un climat maritime humide et souffre moins d'extrêmes de température que d'autres régions du Canada. La température maximum moyenne au début d'Octobre est 16°C avec minimum de 9°C la nuit. Il est difficile de prédire les précipitations, toutefois, il a tendance à pleuvoir plus fréquemment en octobre qu'en été.

Michael Kiefte et Sean Pecknold
Présidents de la conférence

Halifax Stanfield International Airport is Atlantic Canada's centre for domestic, regional and international flight service. The Westin Nova Scotian is approximately 37 min. from the airport by automobile. Several car-rental companies operate at the Airport (http://hiaa.ca/transportation/car-rentals/). If you are renting a car and driving from the airport, please note that the McDonald Bridge between Halifax and Dartmouth is closed Sunday through Thursday from 7:00pm to 5:30am for some fairly radical upgrades. At those times, you should use the McKay Bridge instead. Taxi service is roughly $63 one-way and taxis are frequently parked just outside the arrivals area. For $22, there is a shuttle from the Airport to the Westin Nova Scotian (http://maritimebus.com/halifax-airport-shuttle/) and you can buy a ticket close to the arrivals exit. There is also a public-transit bus operated by Halifax Transit that operates once every hour or half hour and is available just across the passenger pick up area close to the arrivals area. Fares are $3.50 one-way and takes approximately 1 hour 15 min and does require a transfer at Duke St. in Halifax. Routing information on Halifax public transit is available on Google Maps.

For those of you arriving by Via Rail, the Westin Nova Scotian is attached to the train terminus.

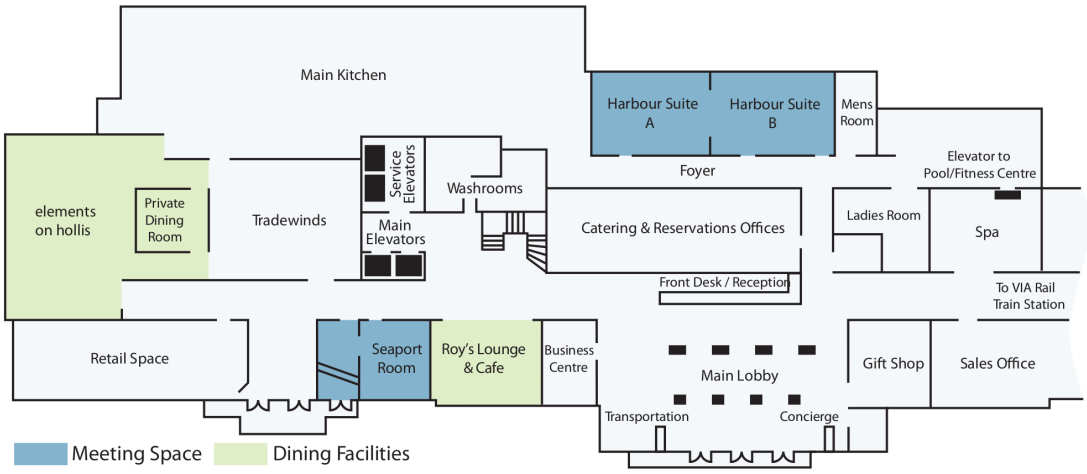
Halifax has a humid maritime climate and experiences fewer extremes in temperature than other parts of Canada. The average high temperature in early October is 16°C with the low dropping to 9°C at night. It is difficult to predict precipitation. However, it tends to rain more frequently in October than in the summer.

Sean Pecknold and Michael Kiefte
Conference Chairs

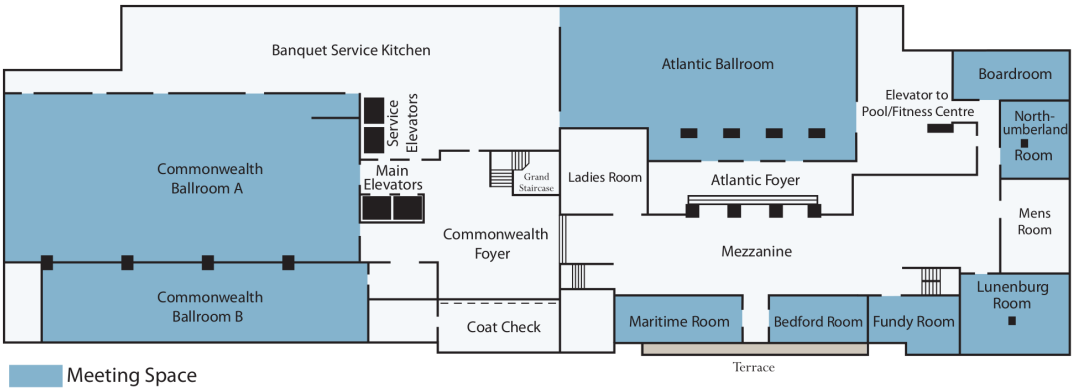
westin nova scotian

property floorplans

lobby level
ground floor



conference level
first floor



Day zero — Tuesday October 6, 2015

2:30–5:30 Board meeting (Maritime)

5:00–8:00 Acoustics Standards Committee Meeting—sponsored by Rockfon LLC (all welcome—1256 Barrington St.)

6:00–8:00 Welcome Reception—sponsored by Swallow Acoustic (Halifax Central Library, 5440 Spring Garden Rd.)

Day one — Wednesday October 7, 2015 (audio provided by Phonak)

Registration open 8:30 (Commonwealth Foyer Coat Check)

8:50–9:00 Welcome (Harbour Suite A/B)

9:00–10:00 Keynote Talk (Harbour Suite A/B) — *Prof. Dennis P. Phillips, Ph.D.*, Department of Psychology & Neuroscience, Dalhousie University — Spatial Hearing

10:00–10:30 Coffee Break—sponsored by Dalimar Inc. (Harbour Suite Foyer)

	room: Harbour Suite A		*room: Lunenburg*		*room: Harbour Suite B*	
	Speech Science		**Bioacoustics**		**Engineering Acoustics/Noise Control**	
	Chair: Bryan Gick		Chair: Hilary Moors-Murphy		Chair: Mehrzad Salkhordeh	
10:30–10:50	*Derrick, De Rybel, Fiasson*	Recording and Reproducing Speech Airflow Outside the Mouth	*Terhune*	Harbour Porpoise Presence Near Oil Tankers	*Lin, Connelly*	Traffic Noise Propagation through Four Vancouver Laneways
10:50–11:10	*Noguchi, Chiu, Wei, Yamane*	Contrastive Tongue Shapes of the Three Sibilant Fricatives in Taiwan Mandarin Read Speech	*Kowarski, Evers, Moors-Murphy, Martin*	Year-Round Monitoring of Humpback Whale (megaptera novaeangliae) Calls in the Gully MPA and Adjacent Areas	*Babic, Rodrigues*	Rumble Strip Noise
11:10–11:30	*Anderson, Gick*	Simulating Intrinsic and Extrinsic Muscles of the Soft Palate for Speech Movements	*Warner*	Localizing Bowhead Whales in the Chukchi Sea using Asynchronous Hydrophones	*Matthew*	Guidelines for New Development in Proximity to Railway Operations
11:30–11:50	*Mukai, Tucker*	Rhythm Metrics of Spontaneous Speech and Accent	*Marotte, Moors-Murphy*	Seasonal Occurrence of Blue Whale (balaenoptera musculus) Vocalizations in the Gully Marine Protected Area	*Busch, Wesolowsky, Swallow*	Theoretical Estimates of Groundborne Railway Vibration
11:50–12:10	*Cichocki*	The Timing of Accentual Phrases in Read and Spontaneous Speech: Data from Acadian French	*O'Neill*	Marine Mammal Vocalisations in the High Arctic: Resolute Bay, Nunavut	*Sun*	Investigation on Non-Point Sources Approximation in Outdoor Noise Predictions

12:10–1:20 Lunch (Commonwealth A/B)

	Speech Science		**Underwater Acoustics**		**Architectural Acoustics**	
	Chair: Bryan Gick		Chair: Nicos Pelavas		Chair: Alex Lorimer	
1:20–1:40	*Kiefte, Nearey*	Modeling Consonant-Context Effects in Dialectal Variation in a Large Database of Spontaneous Speech Recordings	*Li*	Maneuvering Vehicle Localization with an Acoustic Long Baseline System	*Mahn, Sabourin, Zeitler*	Calculating the Apparent Sound Insulation
1:40–2:00	*Milne*	Improving the Accuracy of Forced Alignment through Model Selection and Dictionary Restriction	*Rehm, Baccari, Babin*	A Baffin Bay Acoustic Navigation and Communication System	*Mahn, Zeitler, Hoeller*	Calculating the Apparent Sound Insulation in Concrete Block Buildings
2:00–2:20	*Xavier, Tkachman, Gick*	Towards Convergence of Methods for Speech and Sign Segmentation	*Zedel*	Doppler Sonar as a Tool to Observe Fish Movement: Examples from Grand Passage Nova Scotia	*Sabourin, Zeitler, Mahn*	Calculating the Apparent Sound Insulation in Cross Laminated Timber Buildings
2:20–2:40	*Yamane, Abel, Allen, Burton, Kazama, Kim, Noguchi, Tsuda, Gick*	Ultrasound-Enhanced Multimodal Approaches to Pronunciation Teaching and Learning	*Dosso*	Matched-field Source Localization with Non-synchronized Sensor Arrays	*Connelly, Rodriguez*	Field Test Evaluation of the Impact Insulation Class (IIC) of Greenroofs and Common Roof Deck Materials
2:40–3:00			*Song*	Underwater Navigation Method Based on Side-Scan Sonar Images		

3:00–3:20 Coffee Break (Harbour Suite Foyer)

	Psychological Acoustics		**Underwater Acoustics**		**Architectural Acoustics**	
	Chair: Frank Russo		Chair: Nicos Pelavas		Chair: Bill Gastmeier	
3:20–3:40	*Russo, Livingstone*	The Ryerson Audio-Visual Database of Emotional Speech and Song	*Deveau*	A Toolset for Modelling Anthropogenic Underwater Noise	*Tinianov, Lorimer*	STC Ratings of Drywall Partitions with and without Structural Sheathing
3:40–4:00	*Schutz*	Finding Unity in Impact Events: Amplitude Envelope and Multi-modal Integration	*Deveau, Zeddies, Zykov*	Modelling Exposure of Marine Mammals to Underwater Noise from Pulsed Sources in Long-Duration Surveys	*Madaras, Heuer*	Effects of Noise Flanking Paths on Ceiling Attenuation Class (CAC) Ratings of Ceiling Systems and Inter-room Speech Privacy
4:00–4:20			*Gammon*	Multistatic Sonar Operator Visualization Development Requirements	*Hasan, Hodgson*	Finite-Element Modeling of a Reverberation Room: Effect of the Room Size and Shape on Measurement Accuracy

Day two — Thursday October 8, 2015 (audio provided by Phonak)

10:00–5:00 Exhibitors: 10:00-5:00 Commonwealth B

9:00–10:00 Keynote Talk (Harbour Suite A/B) — *F. Stuart Foster, Ph. D.*, Sunnybrook Health Sciences Centre; Department of Medical Biophysics, University of Toronto — Medical Ultrasound

10:00–10:30 Coffee Break—sposored by Brüel & Kjær (Commonwealth A/B)

	room: Harbour Suite A		*room: Lunenburg*		*room: Harbour Suite B*	
	Speech Science		**Physical Acoustics/Ultrasound**		**Architectural Acoustics**	
	Chair: John Esling		Chair: Robert Adamson		Chair: Bill Gastmeier	
10:30–10:50	*Bernhard*	The Effects of Duration on Human Processing of Reduced Speech	*MacDougall, Landry, Bance, Brown, Adamson*	Optical Coherence Tomography for Clinical Otology	*Tinianov, Lorimer*	Acoustical Characteristics in Restaurants and Food Courts/Large Dining Areas with Varying Levels of Occupancy
10:50–11:10	*Nakashima, Abel, Smith*	Communication Between Native and Non-native Speakers of English in Noise	*Singh, Upmanyu*	Thermo-acoustic Investigations of Molecular Interactions in Binary Solutions of Poly Propylene Glycol- 400 and n-Alkanols	*Tinianov, Lorimer*	Considerations for Acoustical Privacy within Commercial Tenant Space Fit-Outs
11:10–11:30	*Noguchi, Hudson Kam*	Categorical Perception of Post-alveolar Sibilants by Taiwan and Beijing Mandarin Speakers	*Norman*	Ultrasonic Dry Coupling Through Tissue	*Gastmeier, Chan*	Considerations in the Acoustical Design of Black Box Theatres

11:30–11:50	*Tucker, Hawthorne* How Lexical Competition Unfolds in the Recognition of Reduced and Unreduced Word-medial Stops	*Jahns* Time-Course Characterization of High-Intensity Focused Ultrasound Exposure Using B-Mode Ultrasound Imaging	*Ramakrishnan, Dumoulin* Design of a Bar/Music Room—A Case Study
11:50–12:10	*Nearey, Tucker, Barreda* Unraveling Phonological Versus Auditory Context Effects in the Perception of Synthetic Liquid+Stop Clusters	*Latham, Zemp, Brown* Simulation of Fresnel Based Beam Focusing and Steering for a Crossed Electrode Array	*Swallow, Wesolowsky, Busch* This Ain't Your Dadd's Library—The Challenges of Modern Library Acoustics
12:10–1:20	**Lunch (Commonwealth A)**		
	Speech Science	**Hearing Sciences/Conservation**	**Bioacoustics**
	Chair: Benjamin Tucker	Chair: Frank Russo	Chair: Cristina Tollefsen
1:20–1:40	*Mitsuya, Purcell* De-adaptation of Vowel Formant Frequency with Real-time Perturbation Paradigm	*Behar, Russo* Comfort from Using Foam Earplugs—Project for a Study	*Moors-Murphy* Patterning in Northern Bottlenose Whale (hyperoodon ampullatus) Click Trains
1:40–2:00	*Mitsuya, Purcell* Adaptive Control of Vowel Formant with Insert Headphones	*Bonnet, Voix, Nélisse* The Opportunities and Challenges of In-ear Noise Dosimetry	*Fisk, Falk, Tsang, Phillmore* Do Finches Speak Russian? Songbirds can Discriminate Infant-Directed Song and Speech.
2:00–2:20	*Jones, Scheerer, Liu* Hearing Your Own Vocal Errors.	*Saremi* The Age-related Strial Degeneration Inhibits the Inner Hair Cell-Auditory Synapse Complex: A Computational Investigation.	*Patching* Traffic and Industrial Vibrations and A Snake Hibernaculum
2:20–2:40	*Keough, Ozburn, McClay, Schwan, Schellenberg, Akinbo, Gick* Acoustic and Articulatory Qualities of Smiled Speech	*Valentin, Laville* Use of Auditory Steady-state Responses in Measuring the Occlusion Effect of Hearing Protection Devices.	*Porskamp, Broome, Sanderson, Redden* Assessing the Performance of Passive Acoustic Monitoring Technologies for Porpoise Detection in a High Flow Tidal Energy Test Site
2:40–3:20	**Coffee Break—sponsored by LogiSon (Commonwealth A/B)**		
	Musical Acoustics	**Underwater Acoustics**	**Engineering Acoustics/Noise Control**
	Chair: Annabel Cohen	Chair: Garry Heard	Chair: Mehrzad Salkhordeh
3:20–3:40	*Morteza Pouraghdam, Bhat* Spectrum Analysis and Directivity Pattern of a Transducer-Driven Conch Shell	*Tollefsen, Dallas* Physical Mechanisms Underlying the Acoustic Signature of Breaking Waves	*Sharma* Acoustic Emissions from Leaky Pipelines
3:40–4:00	*Cohen, Hughes, Pan* Performance on 2 Tasks of the AIRS Test Battery of Singing Skills in Persons with Cochlear Implants	*Trevorrow, Vagle* HF Sonar Performance in Whitecaps and Wakes	*Ramakrishnan* Performance Analysis of Annular Passive Silencers—An Update
4:00–4:20	*Cohen, McKellar* Identification of Thin Slices of Music by University Students in PEI	*Barclay, Buckingham, Bevans* Ambient Noise in the Challenger Deep	
4:20–4:40	*Bonin, Smilek* The Source Dilemma: Perceptual Coherence and the Continuum of Consonance-Dissonance		
5:00–6:00	**CAA Annual General Meeting (Commonwealth A/B)**		
6:00–9:00	**AWC Banquet & Awards (Pier 21, 1055 Marginal Rd.)**		
Day three	**Friday October 9, 2015 (audio provided by Phonak)**		
9:00–10:00	**Keynote Talk (Harbour Suite A/B)**	*Harold M. Merklinger and John C. Osler* *Defence R&D Canada—Atlantic Research Centre* An Historical Perspective on a Few Canadian Contributions to Underwater Acoustics	
10:00–10:30	**Coffee Break (Harbour Suite Foyer)**		
	room: Harbour Suite A	*room: Lunenburg*	*room: Harbour Suite B*
	Speech Science	**Underwater Acoustics**	**Engineering Acoustics/Noise Control**
	Chair: Michael Kiefte	Chair: Stan Dosso	Chair: Mehrzad Salkhordeh
10:30–10:50	*Chiu, Gick* Startle-Induced Evidence for Multi-Dimensional Information in Speech Plans	*Baccari* Simulating Underwater Acoustic Propagation in an Ice-Covered Baffin Bay	*Price, Ghinet, Wickramasinghe, Chen, Grewal* Investigation of Ground and Maintenance Crew Noise Exposure for the Royal Canadian Air Force CH-147F Helicopter
10:50–11:10	*Harrigan, Tucker* Vowels Spaces and Reduction in Plains Cree	*Duan, Chapman, Yang, Ma* A Universal Expression for the Low-Frequency Attenuation Coefficient in the Surface Duct	*Jozwiak* Markham Stouffville Hospital, a Case Study for Minimizing Helicopter Noise and Vibration Impact on a Healthcare Facility
11:10–11:30	*Esling* A Parade of Illustrations of Voice Quality Settings	*Steele* Modelling Reverberation in the Northern Gulf of Mexico	*Nélisse, Boutin, Giguère, Laroche, Vaillancourt* Self-adjusting Backup Alarms in Noisy Workplaces
11:30–11:50		*McCammon, Pecknold* Sensitivity of Bellhop Intensity to Uncertainty in Sound Speed	*Young, Yu, Faszer* Measured C-Weighted Ambient Sound Levels for Use with Environmental Noise Regulations in Alberta
11:50–12:10		*Dosso* Efficiency of Trans-dimensional Bayesian Inference for Geoacoustic Inversion	
12:10–1:20	**Lunch & Student Presentation Awards—sponsored by Jade Acoustics (Commonwealth A/B)**		
1:20–5:20	**Noise and Vibration Workshop (Maritime)**		

Spatial hearing (Wednesday October 7th – 9 :00-10 :00)

Prof. Dennis P. Phillips, Ph.D. - Department of Psychology Neuroscience, Dalhousie University

Bibliographic Notice :

Dennis Phillips was born in Melbourne, Australia. He attended Monash University, and completed his B.Sc.(Hons) and Ph.D by the age of 25. His Ph.D was in the area of central auditory neuroscience as studied at the level of single-neuron stimulus selectivity in the primary auditory cortex. His thesis advisor was Dexter Irvine. Following his Ph.D, Dennis spent nearly two years as a post-doctoral fellow in the Department of Neurophysiology in the University of Wisconsin Medical School, where his preceptor was John Brugge. He then moved to Dalhousie University where he undertook a second post-doctoral fellowship and he became a faculty member there in 1985. He has been a Full Professor there since 1994. Dennis's neurophysiological studies of hearing continued into the 1990s, but his research interests then shifted towards human auditory psychophysics. Most of his human work has been in the areas of spatial and temporal processes in hearing. He has published over a hundred journal articles and book chapters. He is the winner of some nice awards including a 10-year NSERC University Research Fellowship, an American Psychological Association Distinguished Scientific Award for Early Career Contribution, a Killam Professorship, Faculty Member of the Year Award from Dalhousie's Association of Psychology Students, and the Dalhousie Alumni Association's Award for Excellence in Teaching. Among his undergraduate students, Dennis is known as "The Professor from … well, not Heaven, but the other place." His graduate students (who include Andrew Stuart, Susan Boehnke, Kerry Walker and Rachel Dingle) are going on to impressive careers. He has absolutely no sense of humour.

Medical ultrasound (Thursday October 8th – 9 :00-10 :00)

F. Stuart Foster, Ph. D. - Sunnybrook Health Center, Department of Medical Biophysics, University of Toronto

Bibliographic Notice :

F Stuart Foster is currently a senior scientist at Sunnybrook Health Sciences Centre and Professor and Canada Research Chair in Ultrasound Imaging, Department of Medical Biophysics, University of Toronto. His current research centres on the development of high frequency clinical and preclinical imaging systems, array technology, intravascular imaging, and molecular imaging. Stuart is a fellow of the American Institute of Ultrasound in Medicine. He is the founder and former Chairman of VisualSonics Inc., a company dedicated to preclinical imaging. Dr. Foster co-founded the Mouse Imaging Centre (MICe) now at the Toronto Centre for Phenogenomics. He has served on the Board of Directors of the National Cancer Institute of Canada and as Chairman of its Committee on Research (ACOR). He has previously been the recipient of the Eadie Medal for major contributions to engineering in Canada and has been the recipient of the Queen's Golden Jubilee Medal, the Manning Award of Distinction for Canadian Innovation and the Ontario Premier's Discovery Award. He serves on numerous advisory bodies and is currently on the editorial boards of Ultrasonic Imaging and Ultrasound in Medicine and Biology. Dr. Foster was the recipient of the 2010 Rayleigh Award from IEEE.

An historical perspective on a few Canadian contributions to underwater acoustics (Friday October 9th – 9 :00-10 :00)

Harold M. Merklinger – Defence RD Canada-Atlantic (retired) - John C. Osler – Defence RD Canada – Atlantic Research Centre

An historical perspective on four Canadian initiatives in underwater acoustics will be presented : Towed Variable Depth Sonar (TVDS) ; the Chapman-Harris model for reverberation ; seabed-interaction in shallow water ; and Arctic acoustics. The poor performance of hull-mounted sonars in Canadian waters during WWII prompted TVDS development in 1947. The resulting prototype (CAST/1X) was tested against a UK developmental system in 1958. Subsequently, the UK purchased the Canadian system and the US "re-packaged" their sonars according to the Canadian design. An understanding of the effects of back-scattering from the sea surface and bottom was required to model sonar system performance. Experiments began circa 1960 leading to a new model for sea surface back-scattering, and revealing the ocean volume as an additional factor. The conditions governing propagation and noise in shallow waters remained mysterious in the 1970s, until the significant role of the seabed emerged through theoretical developments (e.g., loss due to shear-waves) and experiments conducted over a variety of seabed types. Political, sovereignty and energy issues prompted Canada to investigate under-ice acoustics in the Arctic beginning circa 1958. The efforts ultimately resulted in the ability to undertake under-ice acoustic and geophysical surveys using long endurance Autonomous Underwater Vehicles (AUVs).

Editor's Note

AWC Proceedings: what you don't want to know

Many things must be considered to achieve perfectly referenced conference proceedings in today's digital world. You can trust your Editorial Board to handle it all perfectly and spare you the humdrum details, but in case you were wondering about the process behind the screen (pun intended), here are some things you do need to know.

The editorial board will produce two versions of the September issue of the Canadian Acoustics journal (JCAA) in order to include the proceedings of the Acoustics Week in Canada (AWC) conference.

- A pre-conference (PRECONF) electronic issue will be sent by email to all authors/presenters, registered attendees, sponsors and exhibitors of the AWC conference.
- A post-conference (POSTCONF) hardcopy will be mailed to all JCAA journal subscribers and journal advertisers.
- After the conference, the proceedings will be moved from the AWC's database to the JCAA's database for regular online access and proper indexing using your favorite digital scholar tools.

In order for authors to have their two-page article published in the PRECONF electronic issue, they must:

1. submit their abstract before the ABSTRACT deadline and their abstract must be accepted by the AWC Technical Chair;
2a. submit their 2-page summary article before the ARTICLE deadline and their article must be accepted by the AWC Technical Chair;
3a. have their 2-page article accepted by the JCAA copyeditor before the PRECONF deadline;
4a. register at least one of the authors for the AWC conference before the PRECONF deadline. In the case where the submitting author is not registered for the conference, the email of the registered author should be typed in exactly the way it was entered by the registered author, for proper internal cross-referencing.

Failure to comply with steps 2a., 3a. or 4a. would result in not having the article, but only the abstract, available in the PRECONF electronic issue. The author(s) will be given a second chance to have the 2-page article published in the POSTCONF proceedings issue, if they:

3b. have their 2-page article accepted by the JCAA copyeditor before the POSTCONF deadline;
4b. register at least one of the authors for the AWC conference before the POSTCONF deadline. See comment above in 4a, if the submitting author is not the registered one.

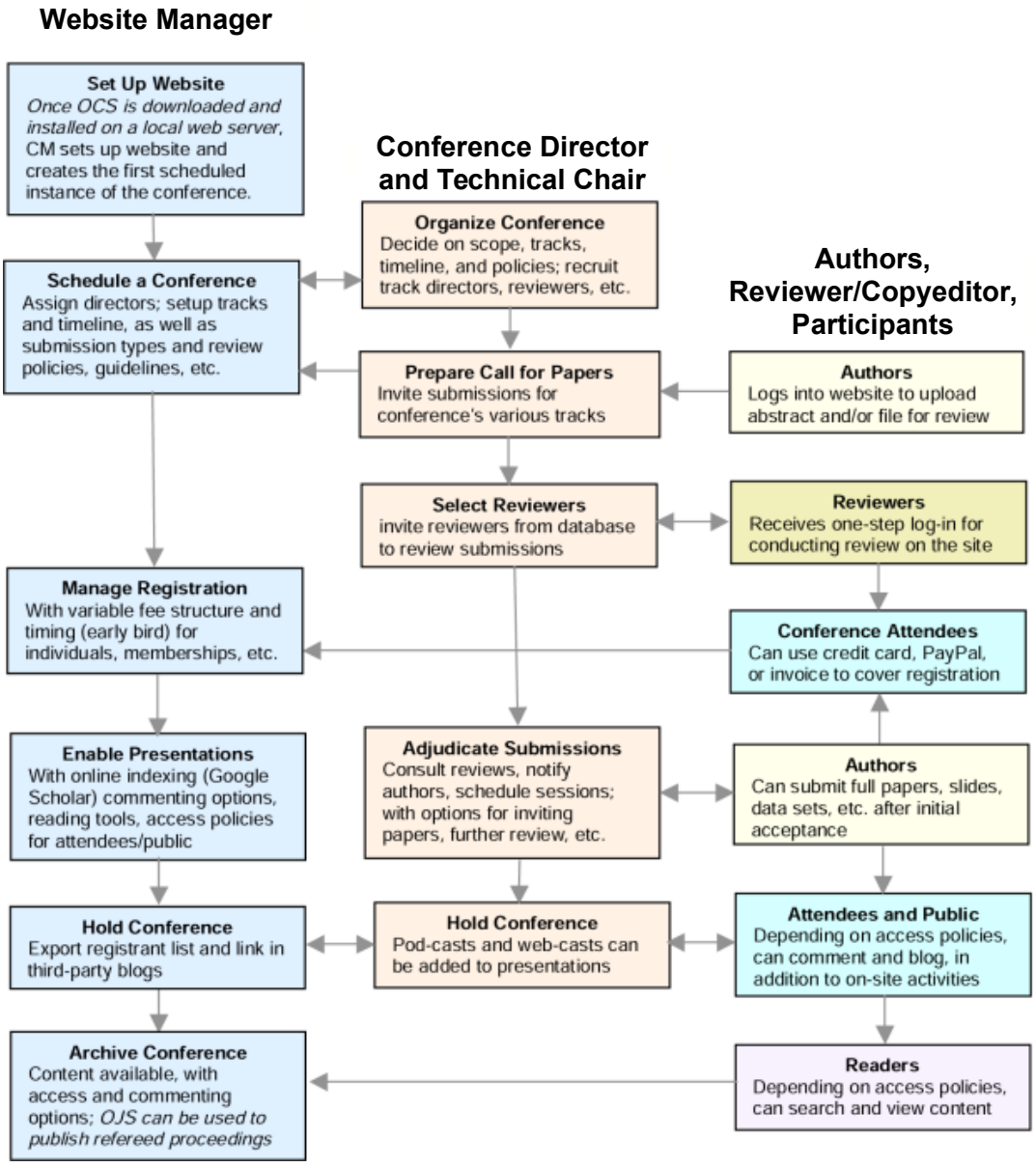
Finally, to have their two-page paper published in the POSTCONF proceedings issue as well as all subsequent publishing through the JCAA's online system, all authors must:

5. have at least one of the authors present the paper during the AWC conference.

For illustration purposes, here are the AWC 2014 planned and effective timeline (available on AWC website under "*Timeline*" at http://awc.caa-aca.ca/index.php/AWC/AWC14/schedConf/timeline)

ABSTRACT	June 15 (Postponed to June 30)
ARTICLE	August 1 (Postponed to August 31)
PRECONF	September 4 (Postponed to September 8)

And in case, you really want to learn more about how all this interact, here is some precious informations, adapted from the Open Conference System (http://ocs.sfu.ca/) that is now used by CAA.



Looking forward to meeting you all in Winnipeg!

Jérémie Voix
Editor

Why publish in Canadian Acoustics?

Because, it is...

- A respected scientific journal with a 40-year history uniquely dedicated to acoustics in Canada
- A quarterly publication in both electronic and hard-copy format, reaching a large community of experts worldwide
- An Open Access journal, with content freely available to all, 12 months from time of publication
- A better solution for fast and professional review providing authors with an efficient, fair, and constructive peer review process.

Pourquoi publier dans Acoustique canadienne ?

Parce que, c'est...

- Une revue respectée, forte de 40 années de publications uniquement dédiée à l'acoustique au Canada
- Une publication trimestrielle en format papier et électronique, rejoignant une large communauté d'experts à travers le monde
- Une publication "accès libre" dont le contenu est disponible à tous, 12 mois après publication
- Une alternative intéressante pour une évaluation par les pairs, fournissant aux auteurs des commentaires pertinents, objectifs et constructifs

ARCHITECTURAL ACOUSTICS - ACOUSTIQUE ARCHITECTURALE

CONSIDERATIONS IN THE ACOUSTICAL DESIGN OF BLACK BOX THEATRES

Mandy Chan, BASc, PEng and William J. Gastmeier, MASc, PEng
HGC Engineering, 2000 Argentia Rd., Plaza 1, Suite 203, Mississauga, Ontario, Canada L5N 1P7

1 Introduction and Background

Black Box Theatres are essential teaching/performance spaces at educational facilities for the Performing Arts. They provide flexible space for the creative training of performers, technicians and theatre staff and encourage the development of experimental performance models and new means of artistic expression. Extremely flexible by design, they can be stand-alone spaces with integral control facilities and seating or components of larger interconnected teaching facilities which include performance theaters, recital halls and multimedia production facilities.

This article deals with the fundamental acoustical design elements and collects useful information concerning the interaction of the acoustical and programme requirements in terms of multi-use functionality for Back Box Theatres. Consideration of sound isolation, acoustical treatment for interior acoustical control and background heating, ventilation, and air conditioning (HVAC) sound levels are discussed.

A discussion of several recently completed Black Box Theatres is presented along with the achieved acoustical parameters.

2 Criteria

In 2002, the American National Standards Institute Inc. published ANSI S12.60-2002 – "Acoustical Performance Criteria, Design Requirements and Guidelines for Schools", a standard developed to give guidance to designers of educational facilities to achieve good acoustical listening conditions in school classrooms and other core learning spaces. It was revised in 2010. It contains criteria for reverberation, background noise, and sound isolation between spaces as well as design guidelines for meeting those criteria. Black Box Theatres fall under the classification of special purpose classrooms which are technically exempt from the minimum requirements of the standard although those minimum requirements should be considered in the design respecting the unique acoustical requirements.

2.1 Reverberation

For teaching spaces including classrooms and theatres, reverberation must be controlled to enhance speech intelligibility, sound system performance and often audio recording capability. A lack of early reflections and the excessive use of absorptive treatments result in low levels of reverberation (a 'dead' space). While such an acoustic environment may be suitable for some audio recordings it is generally inappropriate for interactive performance activities. Excessive levels of reverberation cause the space to sound noisy or boomy, limiting speech intelligibility, sound system functionality and audio recording capability.

Reverberation is measured by the time it takes sound to die away by 60 decibels in the space, using a quantity known as the reverberation time (RT_{500}) measured in the mid-frequency (500 Hz) octave band. The ANSI standard specifies reverberation criteria based on volume, and the authors typically find that a target RT_{500} in the range of 0.7 to 1.5 s achieves a good level of acoustical control, allowing for some multiuse functionality. In many Black Box Theatres, particularly those which also function as multimedia spaces for video recording, there may be elements such as cyclorama or retractable curtains which can provide an additional level of acoustical control when deployed. The target values are for the base room without such additional controls.

2.2 Background HVAC Sound Levels

The most universally accepted criteria for background sound levels are based on Noise Criterion (NC) curves developed by the American Society of Heating Refrigeration and Air Conditioning Engineers (ASHRAE). ANSI S12.60-2010 uses A-weighted decibels (dBA).

Based on the ASHRAE Guidelines and the ANSI Standard, the following target maximum background sound levels are typically recommended. Actual background sound levels can usually exceed these targets in practice by up to 5 points due to economic considerations without serious degradation of acoustic performance for an academic environment; conversely, excesses of more than 5 points can lead to the space being considered "noisy".

Source Space	Target dBA	Target NC Level
Black Box Theatre with live recording	25	NC-20
Black Box Theatre, no live recording	30	NC-25

Table 1: Target Background Sound Levels

2.3 Sound Isolation

ANSI S12.60-2010 provides criteria for sound isolation in terms of the Sound Transmission Class (STC) rating system. Black Box Theatres are not mentioned specifically but are considered to be equivalent to music rooms, and in fact, many Black Box Theatres incorporate live or recorded music in their activities.

Source Space	Receiving Space	Minimum STC Rating
Black Box Theatre	Studios or Classrooms	STC-60
Black Box Theatre	Corridor/Lounge/ Storage	STC-55

Table 2: Recommended Minimum Room Envelope STC Ratings

3 Discussion of Practical Considerations

3.1 Reverberation

Black Box Theatres typically have high ceilings to allow for the deployment of effective lighting, seating and stage props. This results in large physical volumes and the potential for high levels of reverberation. Crucial to achieving the target range is the provision of a high performance acoustical ceiling treatment. Lower walls are reserved for other theatre related paraphernalia and acoustically are best left reflective to encourage sound propagation within and around the space, considering that the space could be used in many configurations including theatre in the round. Additional acoustical treatments are therefore generally placed on the upper walls.

3.2 Background HVAC Sound Levels

The background sound level criteria apply to all sources of mechanical and electrical noise, operating together under any normal operating conditions. For spaces that do not include dropped T-Bar ceilings these criteria apply to the space in that condition, i.e., without a ceiling. Note that this will affect the selection of HVAC terminal devices as many manufacturers' noise specifications assume a room effect consistent with a carpeted floor and a T-bar ceiling. Variable air volume (VAV) or fan powered boxes should be located outside the space as they can radiate considerable sound. The manufacturer's noise specifications typically consider that they are mounted above a mineral tile T-Bar ceiling and they would typically be exposed if mounted inside the theatre. Silencers should also be selected in the context of the interior fit-up conditions.

3.3 Sound Isolation

Black Box Theatre exterior partitions must be full height (slab to slab) for sound isolation purposes. As a result, ducted supply and return air systems should be used, or air should be returned through the full height partitions via crosstalk silencers or internally lined elbow transfer ducts designed to maintain the sound isolation integrity of the partitions.

Typically, when the sound isolation requirements are met with well-designed partitions, the doors (and windows in the case of the sound and lighting booths) become the weak link in terms of sound transmission. All penetrations (eg, ductwork, conduits, piping etc.) must be well sealed. A suitable firestop system would typically satisfy this requirement.

Acoustically rated door assemblies with integral frames and seals are required. An alternate would be the use of entry/exit vestibules. These solutions can be expensive and may interfere with entry/egress, and are often overlooked in favour of aftermarket sound sealing hardware. Such hardware can provide a level of acoustical control, but can be hard to adjust and maintain and invariably is the cause of poor sound isolation, particularly if the doors open into active areas.

4 Case Studies

The authors have been involved in the design of many Black Box Theatres over the past several years, several of which have now been built and several of which are currently under construction. The recently completed projects are discussed below.

4.1 Theatre #1

This theatre was built as a part of an active school of Applied and Performance Arts Associated with a Community College in Ontario. Reverberation and background sound measurements were conducted before final completion and the installation of acoustical treatments. Both parameters were found to be excessive with an RT_{500} approaching 3.5 s and background sound levels approaching NC-60. The HVAC systems were rebalanced and adjusted and it is our understanding that the space is now functioning as intended.

4.2 Theatre #2

This theatre was built as part of Community College facility incorporating dance, rehearsal, acting and comedy studios and two film studios. Measurements were conducted for commissioning and it was noted that the specified wall and ceiling acoustical treatments were present. The RT_{500} was found to be 1.3 s. The background HVAC sound levels were quite low, not measurable due to sound penetrating the corridor doors which were not STC rated and found to have gaps in the sealing systems. Adjustments were made.

4.3 Theatre #3

This theatre was built as part of a private school performing arts faculty upgrade. It has a unique design in that it can be acoustically isolated (large moveable firewall with integral seating) from an adjacent performance theatre, or it can remain connected to the performance theatre becoming the main theatre fly-tower/performance area. Acoustical measurements conducted for commissioning indicated an RT_{500} of 1.1 s and background HVAC Sound Levels of NC - 39. A rebalancing of the HVAC system was conducted.

5 Conclusion

Acoustical criteria and guidance are provided for the design of Black Box Theatres in academic settings. Practical considerations are discussed and actual results provided. Successful results can be achieved through a rigorous design process although budget constraints can lead to compromised results.

References

[1] American National Standard, "Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools", ANSI/ASA S12.60-2010 (Revision of ANSI/ASA S12.60-2002)

[2] American Society for Heating, Refrigeration and Air-Conditioning Engineers (ASHRAE), "HVAC Systems Applications Handbook", Atlanta, GA, 2007.

ACOUSTICAL CHARACTERISTICS IN RESTAURANTS AND FOOD COURTS / LARGE DINING AREAS WITH VARYING LEVELS OF OCCUPANCY

Alex Lorimer, MEng, PEng and Jessica Tinianov, BASc, PEng
HGC Engineering, 2000 Argentia Rd., Plaza 1, Suite 203, Mississauga, Ontario, Canada L5N 1P7

1 Introduction

Dining areas vary in type from quiet and intimate to very lively, and the acoustical character varies as well. Large spaces tend to be more lively and see more transient occupancy, but the designs often try to balance a more intimate setting for conversations, while retaining a dynamic and exciting atmosphere. The atmosphere will also vary depending on the type of space: the acoustical character will differ in food courts, smaller, more intimate restaurants, and urban restaurants with lively bar areas. In spaces without significant amplified background music, the sound level is dominated by speech, and communication quality is therefore of importance. The acoustic properties of such spaces are strongly affected by the occupancy level, due to the changes in reverberation and 'self-noise' or 'din' of the occupants themselves. Recent field measurements and experience with several dining spaces at varying levels of occupancy are reviewed herein.

2 Dining Venues

2.1 Description and Characterization

The background sound level during the mid-day rush when the space is at maximum activity was measured in three separate food courts and a sports bar/restaurant, each with various acoustical characteristics of the space, as well as demographics. All had hard-surfaced floors, drywall or glass walls and ceilings, and minimal upholstering of furniture or draperies, etc.

Food court A included a high ceiling in a suburban mall with perimeter openings. Food court B had a low panelized ceiling around the perimeter with a notably higher glass skylight in the centre. Food court C had low drywall and panelized ceilings throughout. (Panels were typically offset lattice or wooden blades, etc., offering some diffusion but not absorption). These two spaces (B and C) were in urban concourses, serving downtown office staff, and many patrons were in small groups carrying on conversations. Restaurant D included a low ceiling area, but comprised predominantly a double-height ceiling, serving the same demographic area as food courts B and C, but included a bar and wait staff. As well, D included background music with strong bass, but also required more speech communication with staff, and most patrons were in groups having lively conversations.

The reverberation time (RT60), which is the time required for sound to decay by 60 dB in a space, was also measured in each space in the speech-relevant frequency range of 500 Hz to 1 kHz. Reverberation measurements were conducted when the spaces were at 5% or less occupied to focus on the architectural acoustics, rather than the occupant acoustics.

2.2 Occupancy Sound Levels

The background sound level was measured continuously, with the 5-minute (D) or 15-minute (A, B, C) average reported as typical. As well, the range of levels (the difference between the typical maximum and minimum 1-second levels) due to variations from lulls in conversations to peaks from chair scuffles, staff activity, loud conversations or laughing, sneezing/coughing, music, etc., is also presented for each location. The data are shown for occupancies of about 5%, 40%-50%, and 80% or higher.

3 Measurement Results

3.1 Reverberation

The reverberation times for each location were measured when the spaces were essentially unoccupied to quantify the room characteristics, but also due to the difficulty of measuring with patrons present.

The measurement results for each venue are summarized below in Table 1. It is noted that the RT_{60} nominally correlates to the ceiling heights.

Venue	A	B	C	D
RT_{60}, s	2.2	1.6	1.3	1.8

Table 1: Reverberation time, RT_{60} (500 Hz) in each venue, (s).

3.2 Occupancy Sound Levels

The total sound levels in each venue were averaged over the measurement periods (5 to 15 minutes) and are summarized in Figure 1 as a function of percent occupancy. Also presented in the figure is a representation of the range of sound levels during each measurement period. For venues A, B, and C, the durations were 15 minutes and the values are the difference between the L_{10} and L_{90} slow time constant, A-weighted sound levels. For venue D, the sound levels were averaged over 5 minutes and the range is represented by the difference between the maximum and minimum slow time constant A-weighted sound levels during those periods. While these are not the same descriptors, they do both provide a similar estimate of the range of sound levels.

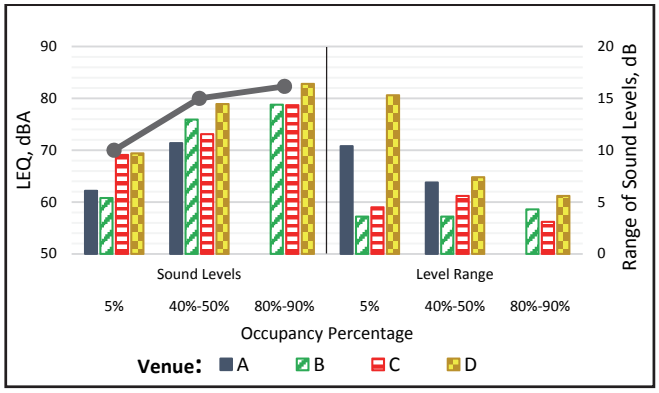


Figure 1: Representative average and range of sound levels under varying occupancy at each venue.

4 Discussion

4.1 Overall Sound Levels

The data in the figure indicates that the background sound clearly increases with occupancy as a percentage of venue capacity. It is also apparent that the sound level increase is generally greatest from low to medium occupancy, and the sound level increase is much less from medium to near-full occupancy. Note that in venue A, the occupancy did not exceed about 40%.

A reference curve is overlaid in Figure 1 which indicates a simplified model of sound level increase due to added occupancy. The model assumes that each talker has a representative sound power emission level, with the sound level at 5% occupancy arbitrarily normalized to 70 dBA. The sound level increase at 50% and 85% occupancy were calculated as a simple decibel logarithmic function of the ratio [1]. Thus, the 50% sound level is 10 dB higher than 5%, and the 85% occupancy is about 2.3 dB higher than the 50% value. This simple model appears to represent the trend quite well, especially in venue D. The two data points for venue A also appear to follow this trend.

For the two office food courts, the general trend of increase with occupancy is apparent, and is in reasonable agreement with the model from medium to high occupancy (minor change), but two different factors may account for the deviations at low occupancy. Venue B was populated by mostly single diners during low occupancy (with minimal speech), whereas, venue C included a busy coffee concession which represented a small percentage of occupancy, but a continuous line of conversing patrons.

4.2 Range of Sound Levels

Some of the above effects are also evident in the data for the range of sound levels. In venue A, the conversations are sporadic with small groups of shoppers, but as the occupancy increases, the levels are more steady. This trend is also very evident in venue D, and extends to the 85% occupancy condition, that is, decreasing range with increasing occupancy. Interestingly, these trends are not apparent in either of the urban food courts.

4.3 Reverberant Effects

It is noted that the above-noted range trends do appear to correlate to the reverberation times of the spaces; that is, the more reverberant spaces experience larger low-occupancy sound level ranges than the less reverberant food courts. Whether this is due to the reverberation or the extra volume from the high ceilings, or both, is not clear. It should be noted that the increase in absorption from the additional occupancy (that is, acoustical absorption of humans) has not been considered. Further measurements would be required to determine if the added absorption (tending to reduce sound levels) may be offset by people speaking louder as the overall sound level increases (known as the Lombard Effect [2]).

The wait staff at venue D had complained of difficulty in taking orders during high occupancy (not relevant in the food court venues), and it is noted here that the recommendation was to introduce substantial amounts of absorptive material on the ceiling and walls to reduce reverberation and hence the overall sound levels at higher at all occupancy levels. At this time, it is not known whether this has been implemented nor the resultant effect.

5 Conclusion

Sound levels were measured over the course of typical lunch time periods in four different dining venues, including a sub-urban and two urban food courts, plus a sports bar/restaurant. Measurement data confirms the expected trend of increasing sound with increasing occupancy, on an approximate logarithmic basis. The acoustical character of the spaces (reverberation and ceiling height) also appeared to be related to an increased range of sound levels during periods of lower occupancy. Further investigation would be required to determine if the simple model of level versus occupancy is accurate, or if other factors, namely occupant acoustic absorption and the Lombard effect (which may have been offsetting in these cases), should be considered.

References

[1] I.L. Ver and L.L. Beranek, Noise and Vibration Control Engineering: Principles and Applications, Second Edition, 2006.

[2] E. deRuiter, Lombard Effect, Speech Communication and the Design of Large (Public) Spaces. *European Acoustics Association, Forum Acusticum* 2011.

DESIGN OF A BAR/MUSIC ROOM – A CASE STUDY

Ramani Ramakrishnan[1] and Romain Dumoulin[2]
1 – Department of Architectural Science, Ryerson University, Toronto, M5B 2K3

1 Introduction

A vacant unit, once used by a Portuguese Deli, was converted to bar/music room in Toronto. The acoustical design aspects are critical for the music room. In addition, the acoustical separation between the two spaces is equally important. The music room/bar has been in operation since early 2015 and field measurements were conducted in June. The design results are compared to actual field measurements and are presented in this paper.

2 Background

The proposed lay out details of the music venue/bar are on Bloor Street, Toronto are shown in Figure1.

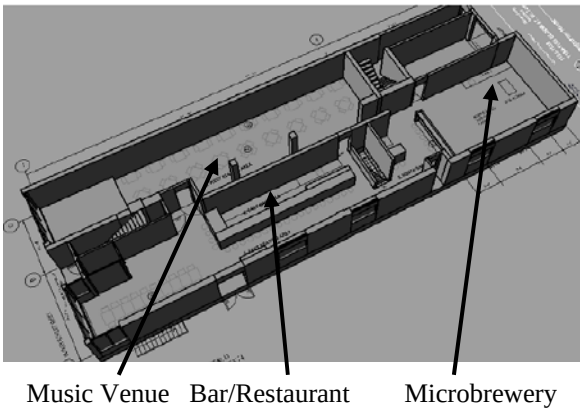


Figure 1: Music Venue/Bar layout.

The plan called for a microbrewery unit at the back of the unit. The unit was divided into two spaces along its north-south axis. The western portion was designed as a music room that would provide a performance space from a solo artist to a Jazz combo to a small rock band. The eastern part was designed as a regular bar/dining area. The bar music can be loud, while the music room can be pianissimo to forte depending on the type of performance.

2.1 Music Venue Requirements

Two different types of needs were considered: Unreinforced concerts with acoustic instruments; and music using sound reinforcement system. The main acoustical requirements are: adequate loudness in every part of the room; uniform distribution of sound energy; and optimum reverberation for music; free from acoustical defects (such as echoes); and limited background noises and vibration.

The parameters used to evaluate the quality of the venue acoustic were: reverberation time (T-30); and clarity (C50). The optimum acoustic parameters for the venue according ISO standards and well known recommendations are summarized in Table 1 below [1, 2, 3, 4].

Condition	T30, sec	C50, dB	dBA
Unreinforced music	1.0 -1.2	> 0	> 80 *
Reinforced music	0.6	- NA-	90-100

dBA is the overall sound pressure level from the activity.
* Preferred overall sound pressure levels for classical music

Table 1: Optimum acoustic parameters.

In addition, adequate acoustic separation must also be provided between the bar sound levels and the music venue.

3 Proposed Design

The microbrewery is located along the north end of the unit. The east wall faces a side street and a gap along the west wall. In addition the second level of the complex has three rental apartments. And hence it must be pointed out that the interior acoustic design is as important as the design of appropriate separation so that the noise intrusions can be kept to a minimum. Finally, the HVAC system noise levels must be adequately silenced. The details of the design are outlined in the following sections.

3.1 Acoustic Separation

The ceiling of the bar and the music venue was designed separately as a suspended double layer gypsum board construction and its details are shown in Figure 2, but without the T-Bar ceiling. The drywall ceiling is three layers for the bar and is two layers for the music venue.

A floating floor assembly was used for the music venue. ¾" thick Baltic birch board was floor mounted on rubber-in-shear mounts as shown in Figure 3.

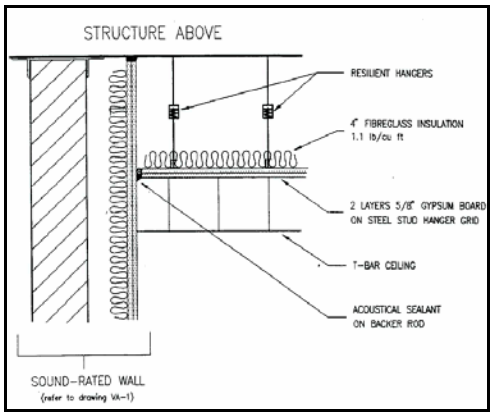


Figure 2: Membrane Ceiling Details.

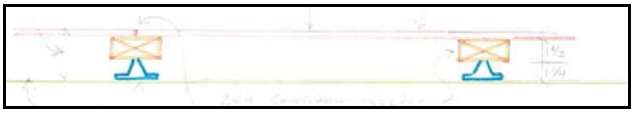


Figure 3: Floating Floor Assembly Details for the Music Venue.

Two wall assemblies will separate the music venue and the bar. Each wall assembly is two layers of drywall (5/8" thick) and is on separate metal studs. The spacing between the studs was chosen to fit the space. The air-gap between the two studs is filled with batt insulation. There are two wall assemblies for the music venue: one that separates the bar and the other one is next to the exterior wall. Both the wall assemblies, including the studs were placed on the floating floor. The studs of the wall assembly next to the exterior wall were connected thorough partition supports.

The music venue and the bar share two sets of doors, one near the stage of the music venue and the other is near the raised portion of the music venue, as shown in Figure 1. Both the doors of each set, was solid core and one of them had an STC 47 rating.

Finally, the HVAC system details were assessed and appropriate silencers were recommended so the music venue ambient sound levels were less than NC-35.

3.2 Interior Acoustics Design

The speakers for the Bar music are hung by spring hangars from the main ceiling joist and the levels are controlled to be less than 85 dBA and less than 90 dB (linear). The limiter to be used included a microphone to measure the bar noise levels.

The music venue is long and narrow with a large window on the south wall. The stage area is at the north end with the musicians on an elevated platform. It will be utilized for three different events and are: a) solo musician (with and without amplification); b) a small group such as a jazz band (with and without amplification); and c) a rock band (with amplification).

The three event scenarios were modeled in acoustic software. The results showed that the solo musician might need some amplification. Jazz Band can be loud enough without amplification. The Rock band uses amplification. Since the space is narrow and long, it needs diffusers to aid in the propagation and assist in the diffusivity of the sound. Approximately 50 square metres of diffusers were applied in the acoustic simulations and a few diffuser options were proposed. One of the main requirements is to install heavy valour drapes, operable, over the south window and behind the stage for two reasons: a) to prevent echoes from the window and b) to add acoustic absorption to the room when sound amplification is used.

The location of the speakers and sub-woofers are very critical to overcome low-frequency modal distribution in the room and appropriate locations were submitted.

4 Field Measurements

The bar/music venue was bult as designed with one exception. The north door of the music venue was built with only one door due to code constraints. The acoustic separation between the bar and music venue was tested by playing a white noise in the bar and measuring the levels in the bar and in the music venue. The results are shown in Figure 4. Between 20 and 35 dB of noise reduction can be seen. Since only a single door on the north end of the music venue was installed, high frequency sound leaked through the door. The owners will be installing a second door to increase the transmission loss. The noise levels did not intrude through the central door set, or the ceiling or the floor.

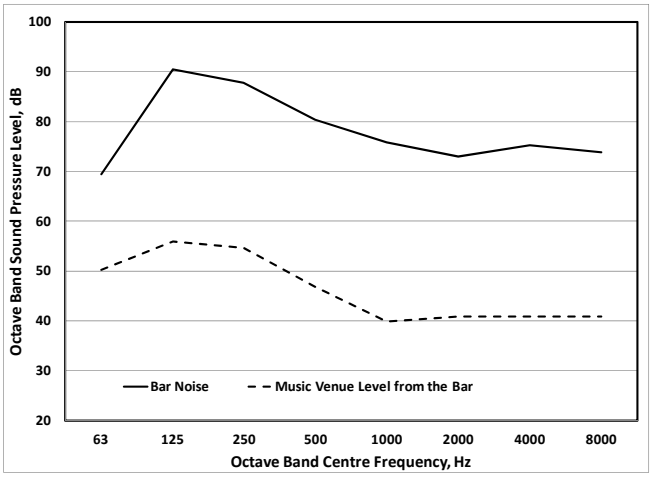


Figure 4: Noise transmission between bar and music venue.

Impulse responses were evaluated from a sine-sweep at a number of locations inside the music venue. Different acoustic parameters were evaluated (Reference 1) from the impulse responses and the main observations were: a) The average reverberation time was 0.9 sec; b) The C-80 was more than 5 dB; c) the echo potential was less than 0.5; and d) the sound level variation was ± 3dB across the eight octave bands.

5 Conclusion

The acoustical design of a proposed bar/music venue was compared to actual field measurements. The results showed the bar/music venue perform as designed.

References

[1] ISO 3382-1, "Acoustics -- Measurement of room acoustic parameters -- Part 1: Performance spaces." International Standards Organization, Geneva, 2009.
[2] ISO 3382-2, "Acoustics -- Measurement of room acoustic parameters -- Part 2: Reverberation time in ordinary rooms." International Standards Organization, Geneva, 2008.
[3] L. L. Beranek, "Concert Halls and Opera Houses – how they Sound," Chapters 2, 3 and 4, Acoustical Society of America, New York, 1996.
[4] M. Barron, "Auditorium Acoustics and Architectural Design," Chapters 2 and 3, Spon Press, London, UK, 2010.

STC RATINGS OF DRYWALL PARTITIONS WITH AND WITHOUT STRUCTURAL SHEATHING

Alex Lorimer, MEng, PEng and Jessica Tinianov, BASc, PEng
HGC Engineering, 2000 Argentia Rd., Plaza 1, Suite 203, Mississauga, Ontario, Canada L5N 1P7

1 Introduction

It is often encountered that architects or other building designers specify multiple cavity partitions, with the expectation that they provide equal or greater sound transmission loss performance than a single cavity partition of similar overall mass and dimension. Previous studies confirm that this is not generally true. In particular, the National Building Code lists wood-frame and drywall demising partitions with single large cavities without solid connections, with relatively high Sound Transmission Class (STC) ratings; however, a footnote indicates that the structural sheathing configurations in these walls can degrade the performance from slightly to unpredictably. This article presents a comparison of several field measurements demonstrating the significant acoustical effects of a commonly specified sheathing configuration.

2 Background

2.1 Wood Stud Frame Walls

In wood frame buildings, typically up to 4 storeys (and increasingly up to 6 storeys), the wood stud walls are often part of the load bearing structure, and as such, those walls require a shear panel. Commonly, this involves adding a layer of plywood or OSB on one side of the studs, followed by the required layers of drywall to achieve the fire rating. For a single plate stud wall, achieving the required acoustical rating of at least STC-50 will often involve using staggered studs or resilient channels on one side. An alternative, used for various reasons, is a double-stud wall. For a continuous floor structure, one set of studs will also receive a layer of sheathing. It is noted that builders prefer to construct the framing flat on the floor, add the sheathing layer, and then stand the wall up in place, with the stud cavities open on the suite side to add services and insulation prior to adding drywall facings. This is repeated on the other side of the demising assembly, but without the sheathing. This puts a sheathing layer within the assembly cavity, creating a three-leaf system (two cavities). In cases without a continuous floor plate (i.e., separate floor plates to limit flanking sound transmission through the floor, under the wall), a sheathing layer is required on each stud set, and builders frequently duplicate the sheathed assembly on both sides, resulting in a four-leaf system (3 cavities; two identical insulated cavities and one small un-insulated cavity).

The fully unsheathed assembly is described in the NBC supplementary table as W13a or W13b, which it rates at STC-57, with insulation between both sets of studs and 1 layer of type X drywall (13 mm or 16 mm) on each side [1]. The two-cavity assembly is not listed, however, the end notes in the supplementary table suggest a 3 point degradation in the STC rating with a single inner sheathing layer. This is consistent with past experience and discussed in reference 2. The three-cavity assembly is also noted briefly in the end notes of the supplementary table, however, it is noted that this "may drastically reduce the STC value", but no specific value is assigned [1]. These are shown schematically in Figure 1.

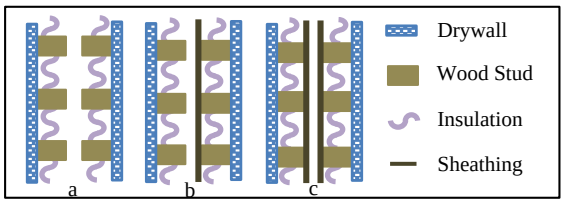


Figure 1: a) Single Cavity; b) Double Cavity; c) Triple Cavity.

HGC Engineering has recently measured the field performance (ASTC) of the unsheathed system and the three-cavity system.

2.2 Cavity Resonances

There are two main sound transmission mechanisms between the multiple layers of the assembly: solid stud connection and airborne coupling via cavity resonance (other factors like mass, damping, and stiffness are the main factors relating to transmission within each material layer).

The solid connection is absent in the unsheathed double stud assembly, hence the high rating from two single drywall layers. In the two-cavity system, the inner layer of sheathing is solidly coupled via the stud to the outer drywall layer on the same stud. This coupling is much stronger than the airborne coupling in that part of the assembly, however, this drywall-stud-sheathing assembly is still only airborne coupled to the other layer of drywall. The airborne coupling cavity depth is approximately halved, raising the resonance frequency and resulting in the 3-point degradation.

In the three-cavity system, both stud assemblies are dominated by solid coupling, and these two assemblies are airborne coupled by a small (usually 25 mm) un-insulated cavity with strong coupling at a much higher resonant frequency, which is within the STC frequency range. Approximate cavity resonance frequencies for the three assemblies are summarized in Table 1. Approximate mass-air-mass resonance frequencies are calculated using the classic empty cavity equation, and the batt-filled cavity mass-air-mass equation from NRC references. Actual resonance frequencies may vary somewhat depending on factors such as total mass, local panel stiffness, panel damping, air-cavity damping, etc.

	1-Cavity	2-Cavity	3-Cavity
1st Cavity	29	44	44
2nd Cavity	-	39	173
3rd Cavity	-	-	44

Table 1: Approximate Mass-Air-Mass Resonances in the 3 double-stud constructions [Hz].

3 Measurement Data

3.1 Laboratory Data

Published laboratory measurement data for the single and double cavity assemblies were readily available from the National Research Council [3]. Measurement results from the following constructions are summarized in Figure 2:

a) Single-cavity with a double row of 90 mm studs, separated by 25 mm, 90 mm batt insulation in both stud spaces, one layer of 16 mm type X drywall each side; no sheathing is included.

b) Double-cavity with a double row of 90 mm studs, 90 mm batt insulation in both stud spaces, one layer of 16 mm type X drywall on inside of one stud set, with a 9 mm space between the inner layer of drywall and the second stud set, one 16 mm layer of type X drywall on each face.

3.2 Field Measurements

Measurements to test the performance of the partitions were required on recently constructed low-rise condominiums to assess their compliance with the building code requirement of STC-50. Measurement results from the following constructions, as tested by HGC Engineering on recent projects, are summarized in Figure 2:

c) Single-cavity with a double row of 90 mm studs, separated by 25 mm, insulation in both stud spaces, one layer of 16 mm type X drywall each side (sheathing, where required, may be included behind the face drywall).

d) Triple-cavity with a double row of 90 mm studs, sheathing layer on inside of each stud set, separated by a 25 mm cavity, insulation in the stud spaces, one layer of 16 mm type X drywall on each face.

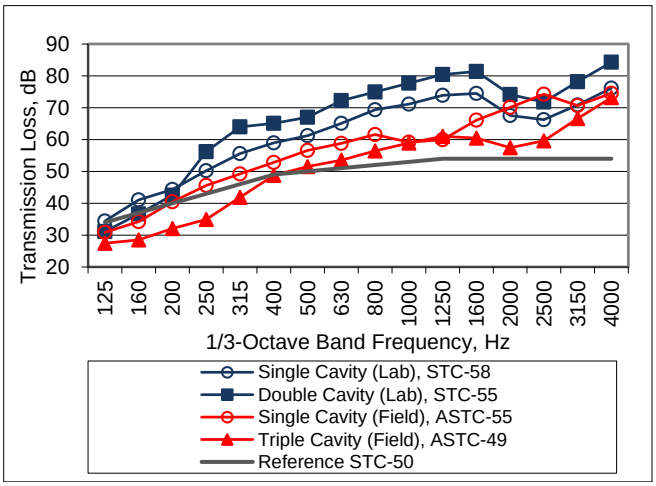


Figure 2: Transmission Loss Results

4 Discussion

4.1 Comparison of Assemblies

There is reasonable agreement between the two single cavity results (laboratory versus field); the field performance dip at 1000 and 1250 Hz is ascribed to observed leakage at back-to-back outlets (only 3 STC points below laboratory rating despite leakage). As noted in the NBC supplementary tables notes, the double cavity rating is only degraded by 3 STC-points from the single cavity, however it is clear that the low-frequency degradation (at 200 Hz and below) is responsible for the reduced STC rating, despite the boost at higher frequencies. The triple-cavity assembly measured below the target (ASTC-47 to 49) in all but one case where it managed to achieve ASTC-51. Despite adding material, it is seen that the performance is significantly degraded by the uninsulated small cavity between sheathing layers. The note in the NBC supplementary tables is accurate in discouraging these assemblies.

4.2 Possible Mitigating Assemblies

If encountered in the field (as-built), upgrading these triple cavity walls without rebuilding may include adding a second layer of drywall with a constrained damping compound, but improvement may only be marginal and should be tested. A better approach would be to remount the face layer on one or both sides of this demising wall using a resilient attachment (e.g., resilient metal channels, or rubber-based clips for snapping furring channels in). Again, this should be tested in the field to confirm whether this is required on one or both sides. In design, this issue should be flagged and alternative constructions considered (staggered studs, resilient mounts, or simply placing the sheathing under the face layers of drywall).

5 Conclusions

It has been noted in various references to avoid double cavity assemblies to not degrade the performance rating. The foregoing confirms this with laboratory data, but goes further to demonstrate the serious degradation resulting from three cavities, especially including a small uninsulated cavity, which was demonstrated to reduce the rating by up to 10 points.

References

[1] National Research Council Canada. National Building Code of Canada, Volume 2. 2010.

[2] W. Gastmeier and M. Wu, Field Sound Transmission of Demising Walls and Floor / Ceiling Assemblies. *Internoise 92 Proceedings* Volume 2. 1992.

[3] National Research Council Canada. Gypsum Board Walls: Transmission Loss Data. 1998.

[4] A.C.C. Warnock and J.D. Quirt, National Research Council Canada. Control of Sound Transmission through Gypsum Board Walls. *Construction Technology Update No. 1*. 1997.

CONSIDERATIONS FOR ACOUSTICAL PRIVACY WITHIN COMMERCIAL TENANT SPACE FIT-OUTS

Jessica Tinianov, BASc, PEng and Alex Lorimer, MEng, PEng
HGC Engineering, 2000 Argentia Rd., Plaza 1, Suite 203, Mississauga, Ontario, Canada L5N 1P7

1 Introduction and Background

Acoustical privacy is often overlooked in the design of commercial space fit-outs. This may be due to a number of factors, including designers being unaware that acoustical privacy might be needed, or of special design considerations to be considered to address such needs. For typical commercial and institutional spaces, there are generally no laws or governing bodies (in Ontario) which prescribe acoustic privacy. However, in many cases, acoustic privacy may be desired by the users of the space (i.e., between and around board rooms, executive offices, medical exam and treatment rooms, etc.). Recent projects and inquires confirm that in many cases, such acoustical privacy is not being properly addressed or implemented, or in some cases, not even considered in design. This article discusses some examples of such cases and presents some common design and field issues which lead to poor acoustic privacy, and modifications to standard designs which can be considered.

1.1 Criteria for Privacy of Personal Information

There are some laws and acts which address privacy of personal information, and under these Acts, there is an obligation for commercial and institutional spaces where personal private information may be discussed to ensure that measures are implemented to maintain privacy of the information. In Canada, the governing Act is the PIPEDA [1]. In Ontario, this is supplemented by the PHIPA [2]. In the United States, the governing document is the HIPAA [3]. The Ontario and United States Acts specifically reference health information, but the Canadian Act references personal information in general. This information may include such personal information and personal health information or financial information, as discussed in health care spaces, banking areas, or legal offices.

Further, these Acts discuss privacy of recorded information (e.g., paper or digital files), but in some cases, it is not clear whether or how these Acts apply to speech privacy. While this should be addressed where private information may be discussed and overheard, there is no specific reference in any of the documents regarding acoustic privacy, and no such limits are set in the Acts.

1.2 Standards for Design of Health Care Facilities

In Alberta, the Infrastructure department of the provincial government has a published document on Technical Design Requirements for Health Care Facilities [4], which contains an acoustics section covering various aspects of noise / vibration control, including sound isolation. This section discusses minimum STC requirements for walls between sensitive areas, and other guidelines relating to details addressing leakage and flanking paths.

There are also standards in Canada and the United States, which address design considerations for health care facilities. These include the CSA/CAN Z8000–11 and the US FGI's design and construction guideline for health care facilities. These are discussed in the paper in reference 5.

2 Common Issues Seen in Existing Facilities

HGC Engineering often receives inquiries from existing facilities, where noise transmission between sensitive spaces has become a concern. In some cases, these inquiries come from facilities which may fall under these health privacy categories due to the nature of the discussions which may occur (i.e., mental health centres, psychology offices, medical clinics, pharmaceutical call centres, etc.). In other cases, there may not be a concern directly related to health information or personal information privacy (i.e., court rooms), but there may still be other confidentiality requirements or concerns. Through various projects, HGC Engineering has had the opportunity to investigate many of these spaces, and there are a number of common issues which generally contribute to a lack of acoustical privacy.

2.1 Low Background Sound

One common issue noted is a low background sound. This can increase the audibility of even low levels of noise transmission. In one particular project, sound transmission measurements were conducted between various areas within a commercial office, and AI calculations were performed to evaluate speech privacy. Following some upgrades, the measurements were repeated and the NIC rating between a meeting room and private office across the hall increased by about 3 points. However, the AI value increased from 0.05 to 0.17 (worse speech privacy), due to a lower background sound during the second visit. Calculations with the upgraded noise reduction values were applied to the previous background sound level, indicating a confidential AI rating of 0.04 by re-introducing masking noise.

In another project, similar measurements were conducted between private offices and other adjacent private offices. In this case the background sound was in the range of 39 to 42 dBA, and AI ratings were calculated to range from 0.07 to 0.42. Calculations applying the measured noise reduction values to a typical sound masking spectrum with a level of 47 dBA reduced the AI levels to 0.02 or lower (confidential privacy), except the worst one, which was reduced to 0.07 (normal privacy). This improvement was confirmed when measurements were conducted again once a

sound masking system was implemented, demonstrating that a significant improvement in speech privacy can, in some cases, be obtained simply by increasing or optimizing the masking sound levels. Note that this approach is limited; masking levels above 48 dBA may be considered disruptive.

In some recent cases, open office designs were preferred by the tenants, including some banks, where personal financial information will be discussed. The open offices generally result in poor acoustic privacy. While these open concept systems will not likely achieve confidential acoustic privacy, sound masking systems can also be used in these cases, in conjunction with interior acoustic design considerations (optimally designed reflective and absorptive surfaces) to increase the acoustic privacy somewhat, without the use of physical barriers.

2.2 Common Ceiling Plenum

One of the most common issues identified in commercial or institutional spaces with noise transmission concerns is a common ceiling plenum. In many cases, there is a drop ceiling throughout the tenant space, and the demising assemblies between offices, board rooms, etc. only extend up to the underside of this ceiling. This creates several issues with respect to sound transmission and acoustical privacy. There can be flanking across the common ceiling, sound leakage through the gap at the top of the demising wall, and transmission through the common ceiling plenum. One example from a recent investigation is shown on the left side of Figure 1. On the right side is an example where the demising wall extends slightly above the common ceiling, addressing the first two concerns, but not the latter.

The common ceiling cavity should be avoided if possible. If this is not possible, upgraded performance can be achieved by increasing the mass of the ceilings, minimizing penetrations, and including acoustical absorption in the cavity. This is addressed in the Alberta Infrastructure document, which reiterates good practices such as including full-height wall constructions or drywall ceilings in rooms with higher STC requirements, and where this may not be possible, suggesting that the partitions should be extended slightly above the ceiling, return air openings should be well separated, and ceiling boards with high CAC ratings and NRC ratings should be used [4].

Figure 1: Demising wall shown to extend only to the underside of the ceiling (left), and slightly above the ceiling (right).

2.3 Other Sound Leakage Paths

Even when the ceilings do extend to the underside of the structure above, this junction and any penetrations through the demising wall, should still be sealed with acoustic caulking, particularly if there are no ceilings. Figure 2 shows two examples from recent projects (between adjacent commercial uses, not within tenant spaces). On the left, the penetrations and junction are well sealed, and on the right, the openings in the structure at the junction have been filled with insulation, but not sealed, allowing sound leakage.

Figure 2: Junction between demising wall and ceiling deck, and penetrations shown sealed (left), and filled but not sealed (right).

Other sound leakage paths should also be addressed. For example, air transfer ducts should be avoided without acoustically lined elbows. Additionally, in many instances, the doors are not taken into account, and there are large gaps around their perimeters, particularly at the bottom. For acoustically sensitive spaces, doors should be selected to provide high sound transmission losses, and should be well-sealed. In one recent project, ASTC testing was conducted between a privacy booth and the adjacent court room, and an improvement of 6 points was achieved by properly adjusting the bottom sweep on the door and including a threshold instead of carpet, where it could not achieve a full seal. In another recent project, a 3-point improvement in NIC was achieved by adding jamb seals to a hotel room door, and a 6 point improvement was achieved by adding jamb seals and a (poorly fitted) door sweep.

3 Discussion and Conclusion

While there are some documents and regulations regarding personal privacy, many of these do not specifically address acoustical privacy. Guidelines for health care facilities should also be considered for other commercial spaces where private health, financial, or other issues may be discussed. Some common issues have been outlined above, based on observations and measurements conducted at various office spaces where acoustical privacy has been a concern. These items among others, and their relative priorities, should be taken into consideration by the design team, as deemed appropriate.

References

[1] Government of Canada. Personal Information Protection and Electronic Documents Act. S.C. 2000, c. 5.

[2] Government of Ontario. Personal Health Information Protection Act. S.O. 2004, c. 3, Sched. A.

[3] United States Congress. Health Insurance Portability and Accountability Act of 1996, Public Law 104-191.

[4] Government of Alberta, Infrastructure. Technical Design Requirements for Health Care Facilities. “The Blue Book”, Third Edition, 2009

[5] J. Swallow and M. Wesolowsky, Acoustic Design Considerations in Modern Health Care Facility Design. *Canadian Acoustics* Vol. 42 No. 3, 2014.

EFFECTS OF NOISE FLANKING PATHS ON CEILING ATTENUATION CLASS (CAC) RATINGS OF CEILING SYSTEMS AND INTER-ROOM SPEECH PRIVACY

Gary S. Madaras, Ph.D. [†1] and Andrew E. Heuer[‡2]
[1] ROCKFON®, 4849 S. Austin Ave., Chicago, IL 60638, USA
[2] NGC Testing Services®, 1650 Military Road, Buffalo, NY 14217, USA

1 Introduction

As a potential cost savings in some buildings, interior walls are stopped at the height of a suspended, modular ceiling. They do not extend full-height up to the structural floor slab or roof above. As a result, sound potentially can transmit more easily from room to room via the open plenum above the ceiling. The sound blocking capacity of the ceiling system then becomes an important factor in the overall room-to-room sound isolation. Ceiling manufacturers test the sound blocking capacity of their ceiling panels and report the results as ceiling attenuation class (CAC) ratings. However, suspended, modular ceilings typically have recessed light fixtures, open return air grilles, supply air diffusers and other miscellaneous penetrations for sprinkler heads, loudspeakers, security/surveillance devices and Wi-Fi devices. These openings and penetrations in the ceiling system create noise flanking paths whereby noise transmits more easily from room to room. The existence of these noise flanking paths are well-known in the architectural acoustics industry. One study[1] conducted by the Institute for Research in Construction, part of the National Research Council Canada, states that some ceiling systems provide little attenuation, and even if panels with high transmission loss are used, the attenuation commonly is limited by leaks (*i.e.*, noise flanking paths) such as openings for airflow. That study found that an opening in the ceiling of only 305 millimeters (mm) by 305 mm (1 square foot) decreased room-to-room isolation by up to 10 dB in the 2 kilohertz (kHz) and 4 kHz octave bands.

2 Method

A series of five CAC tests was performed per ASTM E 1414 and E 413 on various suspended, modular ceiling systems in a dual-room chamber with a common plenum above the ceiling. The tests were performed in Buffalo, New York, USA at NGC Testing Services (National Voluntary Laboratory Accreditation Program code 200291-0). The initial test represented how ceiling manufacturers typically test their ceiling panels. The test specimen comprised just the suspension grid and ceiling panels with no additional noise flanking paths. Subsequent tests had either one or more common noise flanking paths caused by light fixtures or air distribution devices.

[†] gary.madaras@rockfon.com
[‡] aeheuer@ngctestingservices.com

2.1 Ceiling Panels & Suspension System

The ceiling panels used in this study were common, white, wet-formed, mineral fiber ceiling panels measuring 610 mm (24") (nominal) in length and width and 19 mm (3/4") thick with square, lay-in edges. Their weight was approximately 5 kilograms per square meter (kg/m^2) (1 pound per square foot). They have a marketed noise reduction coefficient per ASTM C 423 of NRC 0.70 and a marketed ceiling attenuation class of CAC 35. The measured rating was CAC 37, two points higher than the marketed value.

All of the ceiling systems used a standard 24 mm (15/16") wide, 38 mm (1-1/2") high, steel, tee-bar suspension grid. It was installed in the laboratory test chamber so that a grid member ran along the center of the demising wall in the middle of the test chamber. Ceiling panels did not span across the center wall of the chamber. The grid ran continuously over the center wall of the chamber. It was not disjoined at the center wall.

2.2 Air Distribution System

The return air grilles were aluminum, 610 mm (24") (nominal) in length and width and had a 13 mm (1/2") by 13 mm (1/2") by 13 mm (1/2") open, egg-crate grille.

The square, plaque, supply-air diffusers were 610 mm (24") (nominal) in length and width by 89 mm (3-1/2") high. They were steel with a white powder coat finish and had a 254 mm (10") round duct connection. The supply air diffusers positioned in the adjacent sides of the test chamber were connected with supply air ductwork. A rigid metal duct measuring 406 mm (16") wide by 305 mm (12") high by 3658 mm (12') long and with no internal or external lining ran through the plenum from one side to the other over the demising wall. The supply diffusers were connected to the rigid metal duct above with insulated, round, flexible ducts with a 254 mm (10") inside diameter.

2.3 Lights

The light fixtures were general purpose T8 troffers and were 610 mm (24") (nominal) in length and width. They had an eggcrate louvre with openings that were 19 mm (3/4") by 19 mm (3/4") by 13 mm (1/2") high. No bulbs were installed in the lights and they did not have electrical connections. These were judged to have no effect on the parameters being studied.

2.4 Overall Layout

Figure 1 shows the reflected ceiling plan of the last test specimen with the locations of all air distribution devices and light fixtures. Each room of the test chamber had one return air grille, one supply air diffuser and four lights.

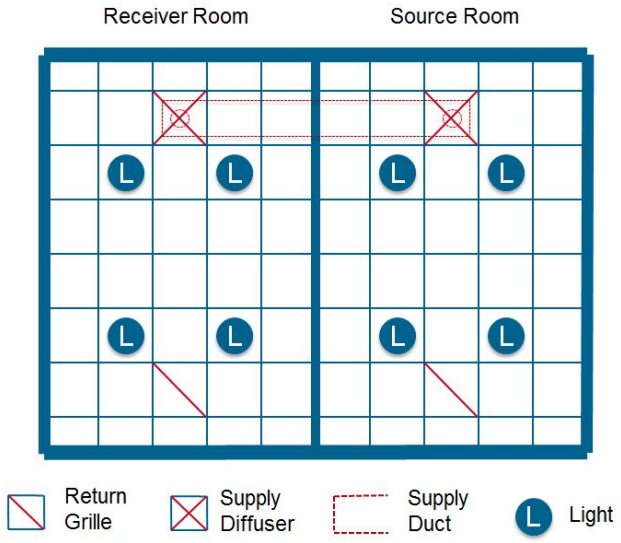


Figure 1: Reflected ceiling plan of the final ceiling system tested showing locations of the air distribution and lighting devices.

3 Results

A series of five CAC tests was performed by NGC Testing Services. The first ceiling system tested had no additional noise flanking paths; the suspension grid was filled only with ceiling panels. After this baseline test, each of three noise flanking paths (*i.e.*, lights, supply air system and return air grilles) was tested independently. Finally, all noise flanking paths were tested together. Figure 2 shows the test results in graphic form.

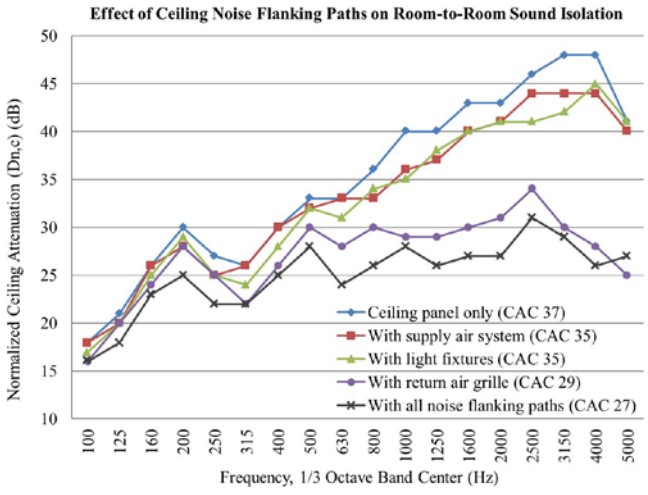


Figure 2: Normalized ceiling attenuation (transmission loss) values by 1/3 octave band for the various ceiling systems tested.

4 Discussion

The results of this study confirm that common noise flanking paths in ceiling systems created by light fixtures, supply air diffusers/ductwork and open return air grilles decrease room-to-room sound isolation compared to a ceiling system of only ceiling panels and suspension grid. The light fixtures and supply air system each degraded the room-to-room sound isolation by two CAC points. The degradation at each 1/3 octave band in the upper frequencies (800 Hz and above) was the largest and averaged 3 dB. At some frequencies, the degradation was as much as 5-6 dB.

The open return air grille resulted in a much greater degradation of room-to-room sound isolation. CAC decreased from 37 to 29 points. The degradation in the upper frequencies averaged 13 dB and at some frequencies was as much as 18-20 dB. This is subjectively equivalent to making the noise four times louder in the upper frequencies.

Combining all three flanking paths degraded the sound isolation even further. CAC decreased to 27, 10 points lower than for the ceiling panels without noise flanking paths. The degradation at the upper frequencies averaged 15 dB and at some frequencies was as much as 19-22 dB.

5 Conclusions

It can be concluded that common elements such as lights, air diffusers and grilles degrade the blocking capacity of the ceiling system and the resulting room-to-room sound isolation. The light fixtures and supply diffusers degraded the isolation to a lesser extent than did the return air grilles. However the lights and supply diffusers did contribute to further degradation when combined with the return air grilles. Overall, the noise flanking paths degraded wideband isolation 2-10 CAC points. More importantly though, they degraded high frequency isolation (1,000 Hz octave band and above), which is more relevant to whether speech is intelligible or not, by 15-22 dB. This is subjectively equivalent to making the transmitted speech four times louder.

Designers, specifiers, contractors and building owners should be aware of the common noise flanking paths that result from lights and air distribution devices in ceiling systems and the resulting degradation of sound isolation and speech privacy. They should not base their expectations of speech privacy on the CAC rating of the ceiling panel alone. Relying solely on a suspended, modular ceiling system with penetrations for sound isolation is a risky design approach.

References

[1] R.E. Halliwell and J.D. Quirt, Controlling Interoffice Sound Transmission through a Suspended Ceiling, *J. Acoust. Soc. of Am.*, **90** (3), September 1991.

COMPARATIVE FIELD TEST EVALUATION OF THE IMPACT INSULATION CLASS (IIC) OF ROOFS

Jose Rodriguez[1], Maureen Connelly[2]

[1] JRS Engineering Ltd, Burnaby, BC. jose.rodriguez@telus.net
[2] Centre for Architectural Ecology, British Columbia Institute of Technology, Burnaby, BC. Maureen_Connelly@bcit.ca

1. INTRODUCTION

As urban density is increasing, the use of rooftops to regain exterior private and/or public spaces is escalating. Rooftop use for outdoor recreation, urban agriculture and restaurant seating are examples of this increment.

Additionally, green roofs are being introduced to rooftops at an exponential rate to support occupancies and for environmental benefits. The impact noise created by users on green roofs must now be considered.

Sound tests were performed to determine how well roof/ceiling assemblies insulate against noise created by impact and airborne vibration. Where airborne sound transmission loss of green roofs has been quantified [1], it is almost unknown if green roofs perform as well as resilient deck to isolate concrete and wood decks and decrease impact noise.

Impact Insulation Class (IIC) measures the ability of the assembly to isolate impact noise like foot fall or objects being dropped or dragged. For impact sound, the nature of the surface finish plays an extremely important role in determining what is transmitted through the assembly [2]. Other factors determining impact sound isolation of an assembly include damping, stiffness and decoupling [3]. It was hypothesized that green roofs would provide a high level of impact sound isolation.

2. METHODOLOGY

The tests were performed at British Columbia Institute of Technology, Centre for Architectural Ecology's Green Roof Research Centre. The Research Centre had three independent roofs: one conventional and two green roofs. The building was 14.5m long, 6.0m wide and 7.0m high, for an approximate interior volume of 970 cubic meters. The internal temperature and relative humidity were controlled at a constant.

The dimensions of each roof were approximately 4.5x6.0m for an area of 27 square meters each. From bottom to top, the reference roof was composed of: 51x305x406mm O.C. wooden deck, vapor barrier, 38mm extruded polystyrene insulation, 6mm overlay board, torch-on modified bitumen base sheet and torch-on modified bitumen cap sheet (SBS).

The two green roofs were constructed identically as the reference roof with the addition of a drainage layer, filter cloth, growing medium substrate and plants. The two green roofs differed only by the depth of substrate. One green roof had 75mm depth of substrate and the other one had 150mm depth of substrate. The plant communities were fully established and included the following species: Sedum spurium, Sedum hybridum, Sedum forsterianum, Sedum florierum, Sedum acre, Sedum album, and Sedum divergens. The green roofs had a soft impact surface in comparison to the other roof types.

The test method is fully described in the ASTM standard E1007 "Field Measurement of tapping machine impact sound transmission through floor-ceiling assemblies and associated support structures." The standard states that the source room can be an exterior location on a roof or a deck [4]. The standard specifies that a standard tapping machine be placed at four prescribed positions on the roof to generate impact noise [4]. One-third octave band sound pressure levels were recorded at the source and at the receiver over a duration of 60 seconds. Four measurements were averaged to determine the average sound pressure level in the receiving room.

ASTM Standard E989 "Standard Classification for Determination of Impact Insulation Class (IIC)" was used to determine the IIC or FIIC. Spatially-average sound pressure levels measured in the receiving room below were adjusted for room absorption and compared to standard values to obtain an overall rating in the frequency range 40 Hz to 6300 Hz [5].

The reference roof (exposed SBS) and the following five roofs were evaluated:

- 60cm x 60cm x 5cm concrete pavers with plastic pedestal spaces directly on the SBS membrane
- 60cm x 60cm x 5cm concrete paver with rubber pads on the SBS membrane
- A wood deck with 38 x 64mm dimensional lumber on 38mm extruded polystyrene insulation providing a continuous separation from the SBS membrane
- 75mm substrate green roof with sedum cover
- 150mm substrate green roof with sedum cover.

3. RESULTS

The Impact Insulation Class (IIC) is a single figure rating scheme intended to rate the effectiveness of roof-ceiling assemblies at preventing the transmission of impact sound generated by the standard tapping machine. The higher the value of the rating, the better the roof performance.

Wooden deck and concrete pavers over conventional roof showed some improvement when using absorbent materials (polystyrene and rubber pads). The plastic pedestals used with the concrete pavers allowed for a direct transmission of the impact to the roofing structure; thus they

had almost no effect in mitigating impact sound transmission.

The wooden deck and the concrete pavers generated higher perceptible levels of air borne noise when using the tapping machine. For this reason, flanking noise would be expected inside the receiving room and lower IIC values would be obtained.

The results of the tests are summarized in Figure 1, Figure 2, and Table 1. The green roofs clearly outperformed the other roof types.

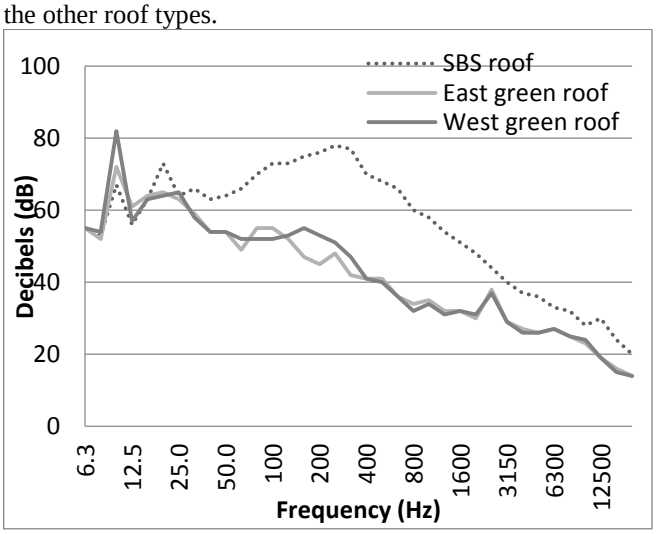


Figure 1: Comparison of SPL in receiver room under reference SBS roof and green roofs.

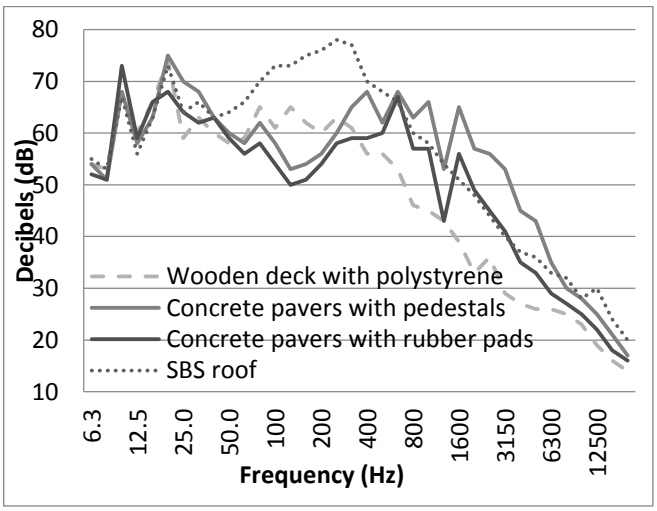


Figure 2: Comparison of SPL in receiver room under reference SBS roof and non-green roof types.

Table 1: IIC of the various tested roof assemblies

SBS roof (reference)	IIC 41
Concrete pavers with plastic pedestals	IIC 44
Concrete pavers with rubber pads	IIC 50
Wooden deck with polystyrene insulation	IIC 55
Green roof with 75mm substrate	IIC 64
Green roof with 150mm substrate	IIC 64

4. CONCLUSION

International Building Codes requires a minimum IIC of 45 for field test in new constructions for floor-ceiling assemblies. Assuming the same criteria for occupied roof top ceiling assembly, the SBS roof over wooden deck and the same roof with concrete pavers and plastic pedestal did not comply.

Some acoustical consultants define the following qualitative values based on IIC rating of floor-ceiling assemblies in new homes:

- FIIC 50 for low income or affordable (minimum quality)
- FIIC 60 for average or mid-range (medium quality)
- FIIC 65 for luxury or high end (high quality)

Under this consideration, the bare SBS roofs with concrete pavers were in the minimum quality category. Wooden deck with polystyrene was in the medium quality category. Both green roofs performed as high end or high quality roofs and outperformed the other roofs tested.

REFERENCES

1. Connelly M., Hodgson M., *Experimental investigation of the sound transmission of vegetated roofs*, Applied Acoustics 74, pp1136-1143, 2013.
2. Warnock, A., *Estimation of Sound Transmission Class and Impact Insulation Class of Steel Frame Assemblies*, NRC RP08-7 for American Iron and Steel Institute/ Steel Framing Alliance, 2008.
3. Long, M., *Architectural Acoustics*, p577, Elsevier, 2014.
4. ASTM E1007 -11e1 "*Test Method for Field Measurement of Tapping Machine Impact Sound Transmission through Floor-Ceiling Assemblies and Associated Support Structures,*" ASTM, 2011.
5. ASTM E989 "*Standard Classification for Determination of Impact Insulation Class (IIC),*" ASTM, 2012.

ACKNOWLEDGEMENTS

Thank you to **The Modal Shop, Inc**., Cincinnati OH, for the use of the tapping machine. Thank you to **Convoy Supply Ltd**. for testing materials.

THIS AIN'T YOUR DADDY'S LIBRARY – THE CHALLENGES OF MODERN LIBRARY ACOUSTICS

John C. Swallow, MASc., P.Eng., Michael J. Wesolowsky, Ph.D., P.Eng., and Todd A. Busch, MASc., P.Eng., P.E., INCE Bd. Cert.
Swallow Acoustic Consultants Ltd., 23-366 Revus Ave., Mississauga, ON, Canada, L5G 4S5

The design of modern libraries is radically different from those we knew 10 and 15 years ago. Then, libraries were quiet, often with signs reading: "Quiet", places for refuge and learning. The modern library is more of a community space, for all ages, also for learning through media other than books. The resulting acoustic challenges are enormous.

Libraries designed in the years around 1900 followed the Carnegie model, because Andrew Carnegie provided construction grants to about 1700 libraries in the United States and Canada, most of which had imposing façades, two-storeys, many stairs, and various collections of books in separate rooms. These buildings were not "accessible" by various meanings of that word and certainly not easily adaptable to changing uses. Quiet was the order of the day and it was relatively easy because heating was hot water, none had air-conditioning, the rooms tended to be small, lined with bookshelves to heights sometimes requiring ladders. Thus the rooms were inherently quiet, very sound absorptive – though they did not contain any specific sound absorbing materials – and thus conducive to quiet study. Speech intelligibility would be excellent but it wasn't needed, speech privacy of no concern because people spoke softly, noise intrusion and distraction were minimal.

The modern library is anything but quiet. Starting perhaps 25 years ago reading circles appeared, where a librarian or teacher could read to elementary school children, often as semi-circular mini-amphitheaters. But necessarily, they were pushed into a corner or separate wing in order to keep the sound of voices and laughter from the other parts of the library. Within the last decade new libraries have been designed to be open and welcoming spaces, intended as much as a community space for all ages as for quiet study. They have grown to include areas dedicated to young children, teens, collaborative study, computer terminals for different age groups, multi-purpose meeting rooms, gallery/exhibit space, performance spaces, recording studios, even a reading oasis with fireplace.

In order to be more welcoming the modern library has larger rooms, encourages line-of-sight to distant parts of the library, low partitions rather than walls and includes fewer bookstacks. Contributing to the openness, more glass is used both to the exterior and between "rooms" to provide more daylight and the use of atria provides the sense of openness and connectivity between different floors. There is much more circulation within the building, from area to area and between floors so corridors are wider and finished, for maintenance, with hard surfaces. Similarly the stairs, also hard surfaces, which were once in stairwells, are in the open. Thus, footfall noise is common throughout the building. Mechanical systems are often exposed with local heat pumps and VAV boxes needed due to the open design.

Perfect Acoustic Storm

The result is a perfect acoustic storm. In this case, it is the intersection of 1) many more noise sources: voices from children and teens, discussion groups, casual spaces; footfall noise and mechanical noise and, 2) rooms which do not attenuate sound: they are very much larger, often inter-connected, have far less sound absorption, are finished with hard floors, with glass exterior walls and low divider walls. The result is much more noise from many more sources in rooms where the sound can propagate relatively easily, where there is low transmission loss between "spaces" and little attenuation of sound all along the way.

Halifax Central Library - an "open" library. Photo: Adam Mork Courtesy of Fowler Bauld & Mitchell Architects

Recognition of this trend is quite recent; in 2002 Salter[1] in his article "Acoustics for Libraries" reviews noise control, materials and techniques for libraries, however, does not address any of these recent trends other than to say that if there are competing uses, to try to separate them. On the other hand, Madziak[2] in 2004 looked at trends in libraries including the inaccessibility of Carnegie libraries, noted the "opening up of library interiors, lower shelving units, wide aisles, lots of clear, uncluttered space…". She noted also that there was "proportionately less collection space and more people space" and in a section on "Quiet Space and Attention to Acoustics" noted that "libraries are no longer eerily quiet, almost silent places …. but have worked… in favour of being a comfortable, welcoming place for all

people, including young children." Finally, she noted "It can be argued that no other issue more challenges library planners than designing a facility that allows for both noise management and the preservation of quiet spaces". Interestingly, this was foreseen years ago, in 1931 by Dr. Ranganathan[1] in his article "The Five Laws of Library Science" of which the Fifth Law is "The library is a growing organism" which means that "A library should be a continually changing institution, never static in its outlook. Books, methods and the physical library should be updated over time".

Other Factors

Is the traditional quiet space even needed? No doubt other factors have come into play: the GenX and millennial generations are used to multimedia, eg while studying they are simultaneously listening to music and intermittently texting friends. They also multi-task regularly. In their school lives collaborative work is highly emphasized; they are much less used to quiet study than previous generations. The Internet allows huge access to information, Wikipedia, YouTube how-to videos, TED talks, available at a terminal or sitting anywhere on the floor, even outside, with a laptop.

Design Criteria

The authors' experience would indicate that a new Index or measures are needed to estimate distraction and annoyance. The acoustic design criteria for a modern Library are limited: STI (speech transmission Index) and SII (speech intelligibility Index) which measure intelligibility and privacy are relevant but do not describe levels of intrusion, distraction or annoyance. Appropriate NC or RC are not available – there is not the experience to indicate whether a lower level which is inherently calming or a higher level which minimizes intrusion and aids in privacy is more appropriate. These overlap with STC and ASTC (Apparent Sound Transmission Class) for acoustic separation between spaces, but not for spaces which are acoustically coupled. There were indices in the past to measure startle or the rate of increase in sound level to predict distraction and community reaction. An index based on impulse sound and intrusion might be more appropriate. Different Indicies and values may be needed for the different Library areas.

Noise Control Techniques, Case Studies

The acoustic goals are well known: to reduce the noise sources as much as possible, isolate the noisy and quiet areas and to reduce the propagation of noise to spaces where more quiet and less intrusion is required. However, there are not many technique options available and compromises are inevitable; these are reviewed in the case studies.

Case Study 1: Renovation Two floors of a university library, connected by an atrium with a concave curved wall opposite the floors, were renovated. Many bookstacks and all of the acoustic ceiling tiles were removed and study tables added. Users complained of the overall noise level, and distractions caused by nearby conversations particularly because of the focusing of the curved wall. Options were limited: re-install high absorption ceiling tile and provide banners at the curved wall for sound absorption. These made a marked improvement but not a total solution.

Case Study 2: Collaboration Area Four collaboration desks of 8 seats each were beside a long five-storey atrium. Openings to the atrium served for air return. Study carrels were 15 m beyond the desks. Atrium noise and noise from adjacent desks was distracting to desk patrons. Raised voices between, vs within, the desks annoyed the carrels. Acoustic material was applied to the atrium roof and floor fascia to reduce reverberant levels. A high-absorption ceiling replaced the drywall above the desks and sound masking at 45 dBA added to the desk area, the sound level tapering off toward the carrels. With masking, atrium intrusion was much reduced. It was hoped that because speech intelligibility between the desks would be much more difficult, there would be fewer raised voices trying to converse between the desks. Sound absorption was added to the walls near the carrels and an intervening wall increased in length to increase barrier effect reduces transfer to the carrels. Users report satisfaction with the results.

Case Study 3 : New; Halifax Central Library

Some noisy areas can co-exist: collaborative areas were placed adjacent to the circulation desk because each has a fairly high tolerance for intrusion. Noisy spaces were placed at one end, with less noisy spaces adjacent, then a separation of bookstacks or collaboration areas, then study carrels and finally quiet reading areas. Low bookstacks were used as partial separations. A performance space was connected to the larger rooms with an operable partition, closed when the privacy is required for performance, open when the room is used for an unamplified reading or a small gathering. However, the meeting rooms have normal walls and the recording studio has STC 60 partitions and sound lock.

Rubber floor finishes on the main traffic areas and stairs reduced footfall noise. Mechanical rooms were isolated and VAV boxes selected for low sound power. Entrances to washrooms were acoustically lined to reduce washroom noise to public spaces, eliminating doors and closing noise.

A high absorption ceiling is essential to the reduction of noise propagation through the space – a cost-effective solution was a linear ceiling with 78% open area and above it, 25mm rigid glass fibre was adhered to the concrete deck. Sound absorbing material was placed on partitions where reflection from one area to another is significant. Curtains – often needed for sunlight or blackout – assist in reducing reflections from exterior glass.

The new Library is considered a success by Library staff and the architectural design has won multiple awards.

References

[1] Acoustics for Libraries 2002, C.M. Salter, US Institute of Museum and Library Services; California State Librarian
[2] 22 Recent and Enduring Trends in Ontario Public Libraries, Anne Marie Medziak, S. Ontario Library Service.
[3] The Five Laws of Library Science, S.R. Ranganathan.

FINITE-ELEMENT MODELING OF A REVERBERATION ROOM: EFFECT OF THE ROOM SIZE AND SHAPE ON MEASUREMENT ACCURACY

Md Mehadi Hasan [†1] and Murray Hodgson [‡1]
[1]Department of Mechanical Engineering, University of British Columbia, Vancouver, BC, Canada.

1 INTRODUCTION

The reverberation-room method, which assumes a diffuse sound field, has long been used for various standardized room-acoustical measurements – i.e. absorption coefficient, source power level, transmission loss, etc. However, unsatisfactory opinions regarding the accuracy of the method, especially at low frequencies, have been reported over the years [1, 2]. This might be due to deviations from the assumed diffuse-field concept, which is very challenging to implement from an application point of view.

To investigate the problem, and find an optimal solution, a number of reverberation rooms of different sizes and shapes have been studied; their capacity to approximate a diffuse sound field is analyzed by means of descriptors like a statistically-based cut-off-frequency definition, spatial uniformity of the reverberant sound field (SPL), prediction accuracy of reverberation times (RT), etc. Results obtained with the help of a numerical finite-element-based modal approach are discussed; in particular, the effect of different room sizes and shapes on the measurement accuracy are explained. Based on these findings, recommendations are proposed regarding the sizes and shapes of reverberation rooms that will give better sound field diffuseness and, hence, better prediction accuracy.

2 THE REVERBERATION ROOMS

Four different shapes of reverberation rooms, all equipped with a number of diffusing panels hanging from the ceilings, were considered, to determine the degree of field diffuseness and, hence, the prediction accuracy. The reverberation rooms considered were: Room #1 – rectangular-shaped with the shortest vertical dimension; Room #2 – rectangular-shaped with the longest vertical dimension; Room #3 – oblique-shaped with the longest vertical dimension; and Room #4 – oblique-shaped with the shortest vertical dimension. For each of the room shapes, three volumes of 150 m^3, 125 m^3 and 82 m^3 were considered.

Predictions were performed with 2 sources and 5 receivers in seven third-octave bands ranging from 63 to 250 Hz. Positioning of the sources and receivers was performed based on the standards' prescriptions [3]. A finite impedance value is applied at boundaries to include damping.

3 RESULTS

3.1 Room #1: Rectangular-Shaped with Shortest Vertical Dimension

Predictions done in Room #1 (see Table 1) reveal that the required minimum number of modes of 20, which is the basis for the statistically-based cut-off-frequency definition [4], occurs from the 125-Hz third-octave band for all three room volumes. That means that prediction can be done from this frequency band. However, the standard deviation, which is a measure of the spatial uniformity of the reverberant sound field, yields values smaller than the ISO-prescribed limit of 1.5 dB [3] from the 160-Hz third-octave band for the 150 m^3 and 125 m^3 room volumes. In particular, for the 150 m^3 volume, the band-averaged standard deviation and the RT-prediction accuracy are relatively better than for the two other room volumes.

3.2 Room #2: Rectangular-Shaped with Longest Vertical Dimension

For Room #2, predictions are presented in Table 2. It is clear that the required number of modes 20 occurs from the 125-Hz band. The SPL deviations are also smaller than 1.5 dB from the 160-Hz band. However,

Table 1. Prediction of modal composition, SPL deviation and RT accuracy in Room #1.

Freq. band, Hz	Number of modes			SPL deviation, dB			RT- accuracy, %		
	Volume, m^3			Volume, m^3			Volume, m^3		
	150	125	82	150	125	82	150	125	82
63	5	5	5	3.2	4.2	5.1	27	22	43
80	11	8	6	2.8	1.2	3.6	26	27	48
100	18	17	11	1.7	2.1	2.4	31	32	21
125	35	28	20	2	2.5	4.1	21	15	27
160	61	53	37	1.3	0.9	0.9	12	7	14
200	116	99	69	1.3	1.1	1.5	2	9	3
250	219	188	123	0.6	1.2	1.6	1	4	1

[†] mehadi@alumni.ubc.ca
[‡] murray.hodgson@ubc.ca

Table 2. Prediction of modal composition, SPL deviation and RT accuracy in Room #2.

Freq. band, Hz	Number of modes			SPL deviation, dB			RT- accuracy, %		
	Volume, m^3			Volume, m^3			Volume, m^3		
	150	125	82	150	125	82	150	125	82
63	5	5	5	3.5	1.7	3.2	30	32	35
80	11	9	6	1.6	1.2	3	29	21	57
100	19	14	12	1.2	2.6	1.1	18	22	42
125	36	31	20	1.7	0.5	2	31	17	31
160	61	54	37	0.9	0.8	0.7	19	15	27
200	116	100	68	1	0.7	2.1	13	9	25
250	219	188	125	0.8	0.6	0.8	12	12	16

the RT-prediction accuracy is not as good as that obtained in Room #1.

3.3 Room #3: Oblique-Shaped with Longest Vertical Dimension

For Room #3, predictions are presented in Table 3. Like Room #2, the required minimum number of modes of 20 occurs from the 125-Hz band, both for the 150 m^3 and 125 m^3 room volumes. SPL deviations smaller than 1.5 dB occur from the 160-Hz band, and the RT-prediction accuracy is also better for these two volumes than for the 82 m^3 volume.

3.4 Room #4: Oblique-Shaped with Shortest Vertical Dimension

For Room #4, the results Room #3 are also true for Room #4, as presented in Table 4. However, the band-averaged SPL deviations, as well as the RT-prediction errors, are smaller in Room #4 than for Room #3.

4 DISCUSSION

Considering all the predictions discussed in section 3,

Table 3. Prediction of modal composition, SPL deviation and RT accuracy in Room #3.

Freq. band, Hz	Number of modes			SPL deviation, dB			RT- accuracy, %		
	Volume, m^3			Volume, m^3			Volume, m^3		
	150	125	82	150	125	82	150	125	82
63	6	6	4	5.6	2.5	5.9	31	29	45
80	11	9	7	3	1.7	4.8	30	28	31
100	19	15	11	1.6	2.4	1.4	21	17	21
125	34	28	19	2.5	2	2.5	17	19	20
160	63	54	37	0.9	0.5	0.7	11	18	20
200	118	99	67	0.7	1	1.9	5	4	16
250	221	187	125	0.8	0.4	1.2	9	11	17

Table 4. Prediction of modal composition, SPL deviation and RT accuracy in Room #4.

Freq. band, Hz	Number of modes			SPL deviation, dB			RT- accuracy, %		
	Volume, m^3			Volume, m^3			Volume, m^3		
	150	125	82	150	125	82	150	125	82
63	5	5	5	3.3	2.2	5.1	35	33	51
80	11	10	9	1.3	2.6	6.4	33	28	62
100	18	15	12	2.9	1.1	1	16	20	43
125	33	28	21	2.7	1.9	0.8	11	12	38
160	64	56	36	1.1	1	0.8	9	11	33
200	118	101	68	0.7	1.1	1.9	2	7	25
250	224	188	123	1.4	0.3	1.1	5	7	26

it is evident that the number of modes of 20 required to start prediction occurs from the 125-Hz third-octave band for all of the reverberation-room shapes and volumes. It is also noticeable that the 150 m^3 room gives better prediction accuracy than the two other room volumes for different room shapes. Comparing results for different reverberation-room shapes make it clear that the oblique-shaped room with the shortest vertical dimension (Room #4) gives better prediction accuracy than the other room shapes for different room volumes.

5 CONCLUSION

Based on the above findings, it can be concluded that prediction from the 160-Hz third-octave band, using the 150 m^3 oblique-shaped reverberation room with the shortest vertical dimension, will give better RT-prediction accuracy and smaller SPL spatial variation than the other reverberation rooms of different volumes and shapes.

REFERENCES

[1] C. W. Kosten, "International comparison measurement in the reverberation room", *Acustica* 10, 400–411 (1960).

[2] M. Vercammen, "Improving the accuracy of sound absorption measurement according to ISO 354", *Proceedings of the International Symposium on Room Acoustics*, Melbourne, Australia (2010).

[3] Determination of Sound Power Levels and Sound Energy Levels of Noise Sources Using Sound Pressure - Precision Methods for Reverberation Test Rooms, ISO, Geneva, 2010.

[4] R. Ramakrishnan and A. Grewal, "Reverberation rooms and spatial uniformity", *Canadian Acoustics*, vol. 36, 2008, pp. 28-31.

BIO-ACOUSTICS - BIO-ACOUSTIQUE

HARBOUR PORPOISE PRESENCE NEAR OIL TANKERS

John M. Terhune
Dept. of Biology, University of New Brunswick, Saint John, NB, E2L 4L5 Canada
terhune@unb.ca

1 Introduction

Harbour porpoises (*Phocoena phocoena*) are thought to be "shy" animals that are easily disturbed by underwater noises. There is concern that sources of anthropogenic sound, especially those associated with shipping and pile driving, will drive porpoises away from their preferred habitat [1]. A recent study, however, found that harbour porpoises were occasionally present within a few hundred m of a liquefied natural gas (LNG) tanker when it was offloading and departing from the Canaport LNG terminal at Saint John, NB, Canada [2]. The goal of the present study was to obtain more samples of porpoise presence (or absence) when noise levels were high and a tanker was in the immediate vicinity. In addition to measuring the linear (unweighted) noise levels, a weighting function to take the insensitivity of porpoise low frequency hearing into account was used to calculate the likely perceived noise levels [3]. This function is generally analogous to the dB(A) function for humans and is expressed as dB(Pp) where (Pp) represents using the hearing threshold values of *Phocoena phocoena* to calculate the weighting function.

2 Methods

An automated porpoise detector (C-POD, Chelonia Ltd.) and a sound recorder (Song Meter SM2, Wildlife Acoustics Inc., fitted with a 30 m hydrophone cable) were deployed from a trestle, <100 m from the stern of a tanker while docked (Figure 1). The hydrophone and C-POD were 4-5 m above the bottom in 12-20 m of water, depending upon the tide. The C-POD is a data logger that records the time when echolocation click trains are detected. The C-POD operates continuously and porpoise detections were expressed as one or more detections per 10 min. Porpoise presence was determined using the high and moderate certainty levels of the C-POD software. Broadband sound levels (0.1- 16 kHz) of 1 min recordings every 10 min were calculated using 1/3 octave bands measured using NoiseLab 4.0.2, (Delta Danish). The noise levels (1/3 octave band sound pressure levels; SPL) were determined using both linear and threshold-weighted functions [3]. The dB(Pp) weighting subtacted 46.4 – 35.6log(kHz) dB from each 1/3 octave noise band before the broadband levels were calculated [3].

The sound files were time-matched to the presence or absence of porpoises detected by the C-POD and the presence or absence of a tanker was also noted. The arrival and departure times of the tankers were provided by http://www.marinetraffic.com/. Descriptive statistics and a Mann-Whitney U test were used to analyze the data.

Figure 1: The underwater sound recorder and porpoise detector were suspended from a trestle at the location of the arrow.

3 Results

Three oil tankers docked, offloaded cargo and embarked during the study period of 14-25 Nov. 2014. There were high winds and waves >3-4 m for about a week in the middle of the study period. This resulted in an oil tanker having to leave the dock after only unloading 1/3 of the cargo and anchoring offshore for a while before returning to finish the offloading. The high winds also resulted in some high noise levels even when the tankers were not present.

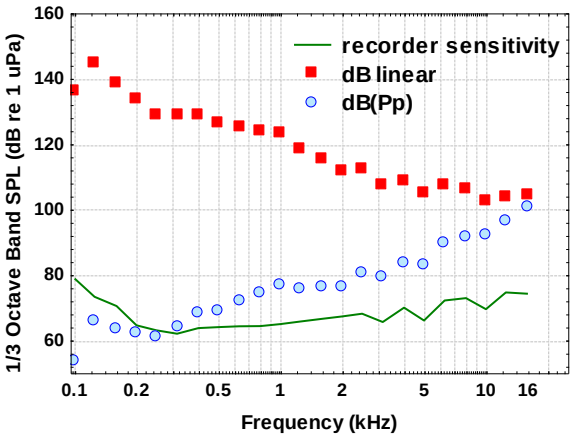


Figure 2: The 1/3 octave band sound pressure levels (SPL: linear and dB(Pp)) of the highest sound level when both a porpoise and a tanker were present.

Porpoise click detections dropped significantly from 49 ± 50 % of the 10 min periods (n = 1134) to 21 ± 41 % (n = 393) (U = 159852, z = -8.36, p<0.0001) when tankers were present. The highest noise level recorded was 119 dB(Pp) re 1 µPa (152 dB re 1 µPa) across a frequency range of 0.1-16 kHz. No porpoises were detected during this ten min period. The highest noise level when a porpoise and a tanker were both present was 107 dB(Pp) re 1 µPa (147 dB re 1 µPa). The 1/3 octave band levels and the impact of using the dB(Pp) weighting function are shown in Figure 2. The porpoise presence (using click detections) was compared to the broadband dB(Pp) noise levels allocated into 5 dB bins (Figure 3).

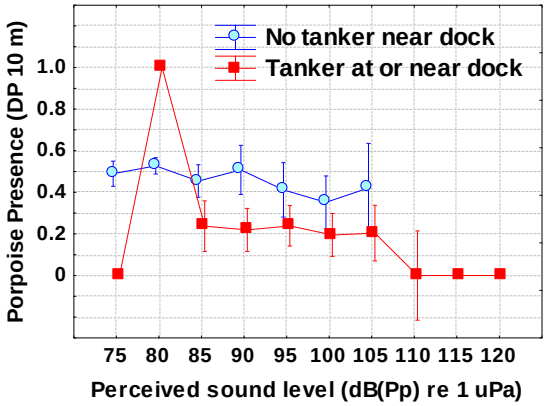


Figure 3: Harbour porpoise click presence (± 2 S.E.) at the Canaport LNG dock in the presence and absence of tankers. The error bars from sample sizes less than 5 have been removed for clarity. DP10m is detection positive 10 min periods.

4 Discussion

When no tankers were present, at least one porpoise was present near the docking site on a regular basis. A high proportion of porpoise presence had been noted at the same location previously [2]. Even when tankers were nearby or docked, porpoises remained in the vicinity for at least 20% of the time. The C-POD will only detect porpoises when they echolocate in the direction of the detector, thus the reported porpoise presence indicates the minimum amount of time that they are present in the area.

When docking or embarking, the oil tankers were accompanied by tugboats and the noise levels were highest then [2]. When offloading cargo, the engines were operating at reduced capacity and the propellor was not turning so the noise levels were lower. Once the tanker was stationary and making less noise, the physical presence of the vessel could constitute a disturbance and it would limit the locations around the porpoise detector where the porpoises could swim. The porpoise detectors have a range of a few hundred m and the ship's hull would block some porpoise detections when they were on the offshore side of the tanker. This may have reduced the number of porpoise detections somewhat.

Porpoise detections inceased once the tankers had left the area. Harbour porpoises have been observed to return to areas, often within a few minutes of the cessation of a disturbance [4].

The hearing sensitivity of the porpoises is poor at low frequencies where the highest tanker noise occurs. Thus, the possible disturbance associated with the sound levels of the tankers would be lessened due to the lower audibility of the low frequency noises. Although the sample sizes are small, the data suggest that the porpoises were displaced, or reduced echolocation activity, by noise levels above ~110 dB(Pp) re 1 µPa. The porpoises appear to tolerate low to moderate noise levels but not necessarily higher noise levels. It is not possible to determine if the porpoises simply habituated to the disturbances or if they remained there, while possibly being physiologically stressed, because it was an important feeding area.

It will not be appropriate to extrapolate dB(Pp) adjusted noise levels to assess the potential impact of noise levels on baleen whales because they have much better lower frequency sensitivity. Linear noise levels do not reflect the likely perceived noise levels by different species and thus should not be used when assessing potential noise disturbances.

5 Conclusion

The regular, although reduced, presence of porpoises when tankers are present suggests that they will tolerate moderate noise levels and related disturbances without being displaced from a location for long times.

Acknowledgements

The staff at Canaport LNG are thanked for providing access to the site and assistance with the deployment and retrieval of the recorders. The study was funded by the University of New Brunswick Retirees Research Fund. The C-POD was provided by Dr. E. Trippel of the St. Andrews Biological Station. Figure 1 was modified from a photo at www.acoa-apeca.gc.ca.

References

[1] Southall, B.L., Bowles, A.E., Ellison W.T., Finneran, J.J., Gentry R.L., Greene Jr., C.R., Kastak, D., Ketten, J.H., Miller, J.H., Nachtigall, P.E., Richardson, W.J., Thomas, J.A., and Tyack, P.L. 2007 Marine mammal noise exposure criteria: initial scientific recommendations. *Aquat. Mamm.* 33: 411-521.

[2] Babin, A. 2013 Harbour porpoise (*Phocoena phocoena*) presence in relation to underwater noise levels at four industrial areas in the Bay of Fundy, Canada. M. Sc. thesis, University of New Brunswick.

[3] Terhune, J.M. 2013 A practical weighting functions for harbor porpoise underwater sound level measurements (L). *J. Acoust. Soc. Amer.* 134: 2405-2408.

[4] Haarr, M.L., Charlton, L.D., Terhune, J.M., and Trippel. E.A. 2009 Harbour porpoise (*Phocoena phocoena*) presence at an aquaculture cage site in the Bay of Fundy, Canada. *Aquat. Mamm.* 35: 203-211.

YEAR-ROUND MONITORING OF HUMPBACK WHALE (*MEGAPTERA NOVAEANGLIAE*) CALLS IN THE GULLY MPA AND ADJACENT AREAS

Katie Kowarski[1], Clair Evers[2], Hilary Moors-Murphy[2] and Bruce Martin[1]
[1] JASCO Applied Sciences Ltd, Dartmouth, Nova Scotia, Canada
[2] Fisheries and Oceans Canada, Dartmouth, Nova Scotia, Canada

1 Introduction

In 2004, the ecologically and biologically diverse region of the Gully submarine canyon off Eastern Canada was designated a Marine Protected Area (MPA). A number of cetaceans frequent the canyon including rare and endangered species, and in order to protect the health and integrity of the Gully ecosystem the Department of Fisheries and Oceans (DFO) Canada has indicated that the year-round presence and activity of all cetaceans in the MPA should be monitored [1]. Summer surveys in the Gully have revealed that humpback whales occur in the area, particularly in late summer [2]. However, little to no research has been undertaken outside of the summer months. Here we describe the year-round acoustic occurrence of humpback whales in and around the Gully MPA.

2 Methods

Near continuous acoustic recordings were collected from the centre of the Gully (MidGul), the slope between the Gully and Shortland canyon (GulSho), and the slope between Shortland and Haldimand canyon (ShoHald) from October 2012 to September 2014 using bottom-mounted Autonomous Multichannel Acoustic Recorders (AMARs, © JASCO Applied Sciences) deployed at depths of 1400-1900 m (Figure 1). The recorders sampled at a rate of 16 kHz for 13 of every 15 min. Automated humpback whale call detectors were run during post-processing [3]. Preliminary analysis of a selection of acoustic files allowed us to define parameters to employ for the remaining visual/aural verification of humpback whale call presence. In this manner, we were able to reliably observe humpback whale calls in a large dataset, while minimizing effort and maximizing efficiency.

3 Results

3.1 Manual Analysis

The automated detector was excellent at capturing humpback whale calls, only missing ~10% of files containing calls. The wide variety of humpback whale call-types can make automated detection challenging. Preliminary analysis revealed that ~84% of files containing less than 11 detections were false and almost 100% of detections in June through September were falsely triggered by vessel noise, seismic noise, or other species. Therefore, to increase efficiency, remaining analysis was limited in

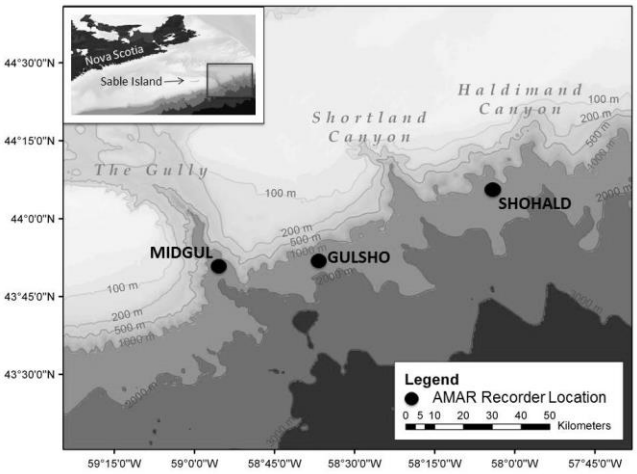


Figure 1: Bathymetric map of the MidGul, GulSho, and ShoHald stations on the Scotian Slope offshore Nova Scotia, Canada in the North Atlantic.

October through May to all files with more than 10 detections and in June through September to at least one file per detection event with more than 10 detections, where a detection event is the occurrence of 1 or more consecutive files with detections. In total, 18,037 files were manually analyzed for humpback call presence/absence (~12% of all sound files). All presence data presented here has been manually verified.

3.2 Year-round presence/absence of calls

Of the 1,835 recording days (612 at MidGul, 611 at GulSho, and 612 at ShoHald), humpback whale calls were observed on 359 days, resulting in a total of 1,855 hours with calls present.

Calls occurred in January, February, March, April, May, June, October, November and December. Call occurrence differed statistically significantly between months for all locations (Kruskall-Wallis $p<0.001$). The hours with calls present per day began to increase in early December and peaked in late December to mid-January at all stations (Figure 2). In winter 2012-2013 calls peaked on 18 Dec at MidGul, 21 Dec at GulSho, and 21 Dec at ShoHald. In winter 2013-2014 calls peaked later, on 17 Jan at MidGul, 18 Jan at GulSho, and 18 Jan at ShoHald. Calls occurred sporadically from the beginning of February to the end of May and were essentially absent during summer months across all stations (Figure 2).

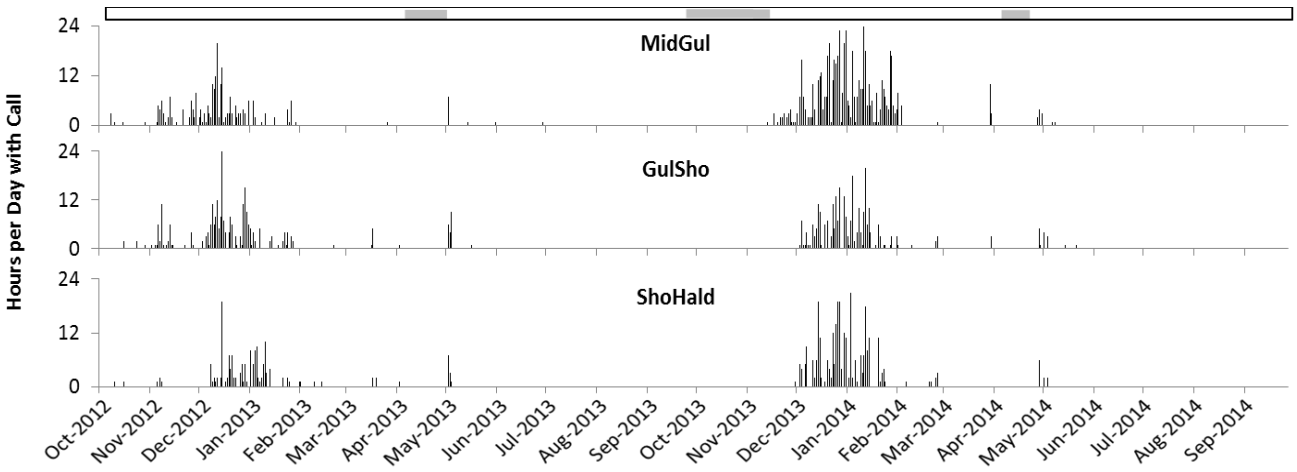


Figure 2: Plots of the number of hours per day in which humpback whale calls occurred for the MidGul, GulSho, and ShoHald stations on the Scotian Slope from October 2012 to September 2014 where the shaded area on the top bar indicates timeframes when recording did not take place.

4 Discussion

This study shows that humpback whale calls occur in and around an offshore, deep-water submarine canyon in the North Atlantic predominantly in the winter months. Following the absence of calls in the summer, call occurrence began to increase in the fall. Previously conducted visual summer surveys similarly found humpback sightings to be rare in the Gully, predominantly occurring late in the season [2]. The peak in calls in December to January observed here closely precedes a peak in calls described in the Stellwagon Bank National Marine Sanctuary (SBNMS) in October to December [4]. We propose that in December some whales move from more inshore areas such as the SBNMS to offshore productive areas such as the Gully MPA before continuing south to the winter breeding grounds. Some individuals may not migrate south, but remain in the Gully region, attributing to the sparse calls observed in February and March. The calls in April and May could either be from non-migrating whales, or whales returning from the breeding grounds, a route that has been previously documented [5].

Passive acoustic monitoring of the Scotian Slope has allowed us to collect a large amount of data at relatively low cost and effort from a highly remote MPA. However, the method has its limitations. Noise created by the mooring, vessel traffic, and seismic activities can mask whale calls; a problem amplified in the particularly noisy summer months. Whales that do not produce calls go undetected. Therefore, these results represent the minimum presence of humpback whales in and around the Gully MPA and are biased towards whales that are acoustically active during periods with low noise.

5 Conclusion

The Gully MPA and its nearby areas act as a winter migratory corridor and/or provide alternative wintering grounds for North Atlantic humpback whales, a finding that significantly expands our understanding of the importance of the MPA to this species. A more thorough analysis of call-type and structure may provide clues as to which whales are calling (male vs. female) and what behaviour is occurring (feeding vs. mating calls). Expanded acoustic monitoring efforts in the future on the Scotian Shelf are required to more fully understand how humpback whales utilize the region throughout the year.

Acknowledgments

This project was supported by DFO SPERA funds and through contributions from JASCO Applied Sciences. We would like to thank Jay Bartholette and the Program Coordination and Support Division of DFO, Andrew Cogswell, and other DFO staff, and the CCGS Hudson crew for their help with instrument deployments. We also thank all additional acoustic analysts who contributed to this work including Julien Delarue, Emma Marotte, Galaxina Renauld, and Carmen Lawrence.

References

[1] DFO. Gully marine protected area monitoring Indicators, protocols and strategies. *DFO Can. Sci. Advis. Sec. Sci. Advis. Rep.* 2010/066, 2010.

[2] S. K. Hooker, H. Whitehead, and S. Gowan. Marine protected area design and the spatial and temporal distribution of cetaceans in a submarine canyon. *Con. Biol*., 13:3, 1999. *Oceans St. John's 2014.*

[3] B. Martin, K. Kowarski, X. Mouy, and H. Moors-Murphy. Recording and Identification of Marine Mammal Vocalizations on the Scotian Shelf and Slope

[4] E. T. Vu et al. Humpback whale (*Megaptera novaeangliae*) song occurs extensively on feeding grounds in the Northwest Atlantic Ocean. *Aquat. Biol.*, 14, 2012.

[5] A. S. Kennedy et al. Local and migratory movements of humpback whales (*Megaptera novaeangliae*) satellite-tracked in the North Atlantic Ocean. *Can. J. Zool.*, 92, 2013.

TRAFFIC AND INDUSTRIAL VIBRATIONS AND A SNAKE HIBERNACULUM

Richard Patching, Leslie Noble McCleave, and Adrian Radulescu
Patching Associates Acoustical Engineering Ltd.
4825 Westwinds Drive NE, Calgary Alberta T3J 4L4

1 Presentation (English)

Anecdotal evidence suggests that hibernating snakes can be disturbed or even awoken by ground vibrations, and if they emerge in winter, this can be fatal. A concern was raised about vibrations from a pile-driving operation some 350 metres from a garter-snake hibernaculum. Preliminary measurements in winter showed that vibrations from traffic may be more of an issue than the pile-driving, and additional measurements were taken near the existing hibernaculum, which is located within 30 metres of a secondary highway. These readings may indicate levels which are acceptable to the snakes.

2 Présentation (Français)

Des preuves anecdotiques suggèrent que les serpents en hibernation peuvent être perturbés ou même réveillés par les vibrations du sol, et si elles apparaissent en hiver, cela peut être fatal. Une préoccupation a été soulevée au sujet de vibrations d'une opération de battage à environ 350 mètres d'un hibernaculum pour serpents de types Thamnophis. Des mesures préliminaires en hiver ont montré que les vibrations dues au traffic peuvent être plus un problème que le battage, et des mesures supplémentaires ont été prises près de l'hibernaculum existant, qui est situé à moins de 30 mètres d'une route secondaire. Ces relevés peuvent indiquer des niveaux qui sont acceptables pour les serpents.

3 Nature of the Issue

3.1 Snakes and Vibrations

There have been a number of instances (Haicheng China 1975, Tangshan China, 1976, and southern Taiwan 2006, and others) where snakes have reacted just prior to earthquakes. As snakes are very sensitive to ground vibrations, there is concern that industrial activities could disturb hibernation, which could lead to arousal and possible death if this occurs during winter conditions. Early investigations were conducted at a range of distances (100 to 500 metres) from an industrial operation in a forested area in winter. Traffic passing the monitor sites registered higher ground vibration levels than the industrial activities. As the hibernaculum of concern is less than 20 metres from a secondary highway, further measurements were taken in summer to assess the ground vibrations from traffic on this road, where it appears the snakes are not adversely disturbed during their hibernation.

3.2 Methodology

Accelerometers were placed at 10 and 20 metres from a secondary highway. In order to couple to the ground, the accelerometers were mounted on a small (15 cm by 15 cm) plywood plate which was designed based on a paper by Barman and Coulter (1981, ref 1). The plate was designed to accommodate a Brüel and Kjear (B&K) type 8306 accelerometer within the criterion presented in their paper. The B&K type 4370 accelerometers are about 1/10 the mass of the 8306 units, and so meet the criterion easily.

The vertical vibration levels were measured with B&K Sound Level Meters at 10 metres and 20 metres, both using the accelerometer as input and recording the levels in 1/3 octaves in 1 second intervals.

The acceleration readings were converted to vibration velocity levels using the formula:

$$L_v = L_a - 20*\log_{10}(f) + K$$

Where f is the centre frequency of each 1/3 octave band, and the constant K depends on the accelerometer sensitivity and choice of reference units used. Once the band acceleration levels are converted to velocity, the bands are added together (using deciBel math, of course).

4 Results

Vibration velocity was used as the relevant index, as it has been shown to correlate well with human perception as well as structural and architectural damage to buildings. Whether it is relevant for snakes is an outstanding question.

Figure 1 shows the location of the hibernaculum adjacent to the road. Figure 2 shows the overall vertical velocity levels over a sample hour from within the approximately 8 hour measurement period, at both the 10 metre and 20 metre distances. Figure 3 shows the levels for a large truck moving at speed, while Figure 4 shows the levels for a front-end loader moving slowly. Figure 5 shows the vertical velocity levels for a pickup truck and a passenger car, both moving at speed. It is apparent that the vibration levels depend both on the speed and size of the vehicles. The passage of the large truck (in Figure 3) presents the highest vibration level

Figure 1: Hibernaculum (Fenced) Near Road

(L_v = 111.8 dB ref 1 nm/s = 83.8 dB ref 1 µin/sec) measured over the entire measurement period.

Figure 2
Vertical Velocity Due to Traffic
One Hour Time Period

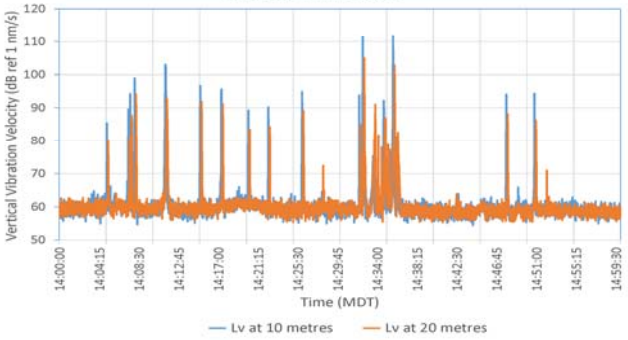


Figure 3
Vertical Velocity Due to Traffic
Large Truck in Near Lane at Speed

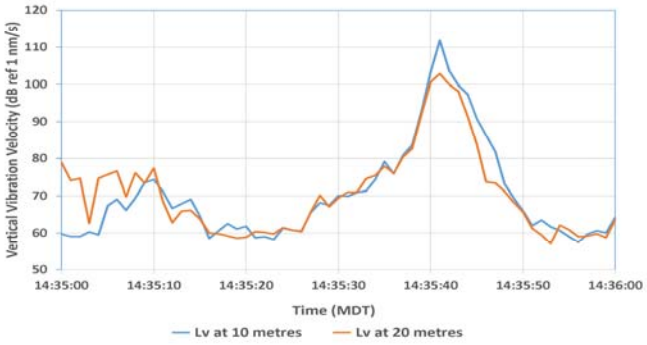


Figure 4
Vertical Velocity Due to Traffic
Loader in Far Lane (Slow)

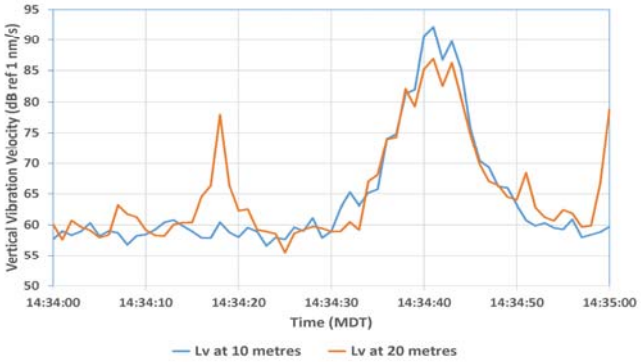
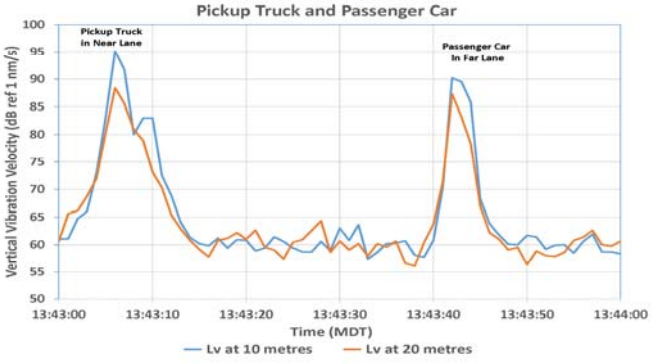


Figure 5
Vertical Velocity Due to Traffic
Pickup Truck and Passenger Car



5 Limitations of the Study

For survival, a snake hibernaculum must be below the frost line. This study measured surface vibrations only.

The surface vibrations were transmitted through undisturbed continuous ground. By its very nature, the ground in and around the hibernaculum must be disturbed and porous, so that the snakes can inhabit it. Therefore, these vibration levels are the levels which impact on the hibernaculum, but not necessarily the levels within the hibernaculum.

It is recognised that industrial vibration may involve vibratory activities taking place deeper under the ground surface, which may not be reflected in these observations.

These readings were taken in summer. During winter, the propagation of vibrational energy may be higher, but also may be restricted more to the surface (frozen) layers of the ground which can act as a wave guide. Vibrations occurring or generated before the frozen layers of ground should not be affected by this effect.

6 Conclusion

The continuing presence and use of a hibernaculum located beside a road indicates that the vibrations, generated by the passing traffic, do not disturb the hibernating snakes to a significant degree. The maximum measured vibration levels may be taken as a threshold level of vibration which is acceptable to the snakes.

Therefore, industrial vibrations well below these thresholds should not disturb the hibernating garter snakes. This finding may serve to indicate an acceptable level for the assessment of industrially generated vibrations in the vicinity of a hibernaculum.

References

1. M. Barman and J.E. Coulter. An Improved Transducer Mount for Ground Vibration Measurements. *JCAA Vol 9(1), Jan. 1981*

2. The Independent, December 2006

Acknowledgements

Thanks to Emmanuel Pradeilles for the translation of the abstract.

ASSESSING THE PERFORMANCE OF TWO PASSIVE ACOUSTIC MONITORING TECHNOLOGIES FOR PORPOISE DETECTION IN A HIGH FLOW TIDAL SITE

Peter H.J. Porskamp[*1], Jeremy E. Broome[*†1], Brian G. Sanderson[†2] and Anna M. Redden[‡1,2]
[1]Acadia University, Wolfville, Nova Scotia
[2]Acadia Centre for Estuarine Research, Wolfville, Nova Scotia

1 Introduction

Current interests in harnessing tidal energy from Minas Passage, a high flow channel in the upper Bay of Fundy, require examination of potential effects of tidal turbines on the environment, including impacts on commonly occurring harbour porpoise, *Phocoena phocoena* (Linnaeus, 1758). To collect baseline data on porpoise presence in the Fundy Ocean Research Centre for Energy (FORCE) turbine test site in Minas Passage, two Passive Acoustic Monitoring (PAM) technologies were used: the C-POD (continuous porpoise click logger, Chelonia Ltd) and the icListenHF (digital hydrophone, Ocean Sonics Ltd) (Figure 1).

Prior multi-year PAM studies in the Minas Passage involved C-POD hydrophones housed within streamlined SUB buoys suspended 3 m above the seafloor [3][4]. During high flows, these SUB buoys experienced high variability in tilt [3]. In addition, high-flow induced noise in the Minas Passage exceeded the C-POD's maximum recordable clicks per minute, resulting in "lost time", and thus under-detected porpoise click trains. To increase detection efficiency, a new mooring design and acoustic shrouds were tested.

2 Methods

2.1 Description of Study Area

The FORCE test site, located in Minas Passage, NS, is 5-6 km wide and 13 km long and features semidiurnal tides with a maximum tidal range of >13 m [2]. During spring tides, current speeds can exceed 6 m/s at the surface and be as high as 3 m/s at 3 m above the bottom [2].

2.2 Instrument Platform

A bottom moored instrument platform was deployed within the FORCE Crown Lease Area on 5 June 2014 and recovered from Minas Passage on 2 July 2014. The platform housed an acoustic release, two tilt loggers, two icListenHFs (one bare, one shrouded with 1.27 cm, 20 ppi open cell foam), two C-PODs and approximately 400 kg of anchor weight. The sensors were located about 1 m off the seafloor, within the boundary layer where current speeds are reduced (<1 m/s). The acoustic release was triggered during recovery of the unit. One C-POD, housed in a tethered SUB buoy, was located within 50 m of the platform.

[*] 099111p@acadiau.ca
[*†]broome.jeremy@gmail.com
[†]bxs@eastlink.ca
[‡] anna.redden@acadiau.ca

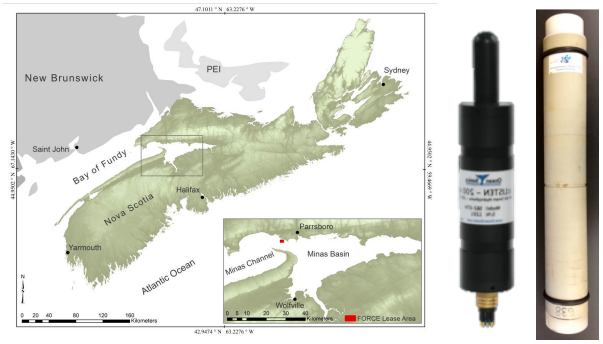


Figure 1: Left: Map of the Bay of Fundy including the location of the FORCE Crown Lease Area (red box) in Minas Passage, Nova Scotia. Right: icListenHF and C-POD hydrophones, not to scale.

3 Results

3.1 C-POD and icListenHF Performance

The platform mounted C-PODs detected greater numbers of click trains and greater detection positive minutes (DPM/day) compared to the C-POD housed in a SUB buoy (Figure 2). Platform mounted C-PODs 639 and 1615 showed similar detection peaks in early June after which C-POD 1615 became detached from the platform.

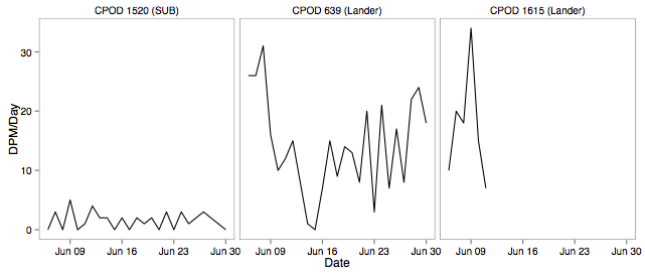


Figure 2: DPM/Day of SUB buoy mounted C-POD 1520 and platform mounted C-PODs 639 and 1615 (detached on 12 June).

Harbour porpoise detections recorded by a C-POD and icListenHF hydrophone co-located on the instrument platform showed that, over the same time period, the C-POD recorded only 19.9% of the click detections recorded by icListenHF. In marked contrast, the icListenHF recorded 99.8% of all detections recorded by the C-POD (Table 1).

Table 1: Comparison of DPMs recorded by two platform-mounted hydrophones

Hydrophone	Total DPM	icListenHF DPMs; % Recorded	C-POD DPMs; % Recorded
C-POD 639	205	19.9	-
icListenHF 1239	958	-	99.8

3.2 Time Lost Due to Noise

During flood tides, percent lost time was similar for the SUB buoy C-POD (1520) and platform mounted C-POD (639) (Figure 3). But during the less noisy ebb tide, the platform C-POD showed significantly less lost detection time (Wilcoxon signed rank test, $\alpha = 0.05$, $p<0.001$).

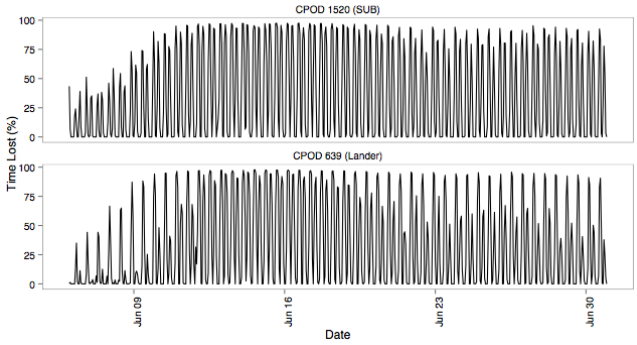


Figure 3: C-POD percent-lost time plots. Top: SUB Buoy mounted C-POD 1520; Bottom: platform mounted C-POD 639. The series of peaks represent the sequence of flood and ebb tides with flood tides showing greater % lost time compared to ebb tides. A spring neap pattern is also evident.

3.3 Shroud Test

Shrouding of one of the two icListenHF hydrophones with acoustic foam (20 ppi, 1.27 cm) reduced flow noise in tank tests but at FORCE, the shrouded hydrophone recorded a similar sound profile to the non-shrouded unit within the porpoise click frequency range of 120-140 kHz (difference was < 3dB). The difference was greatest at high current speeds on the flood tide (Figure 4).

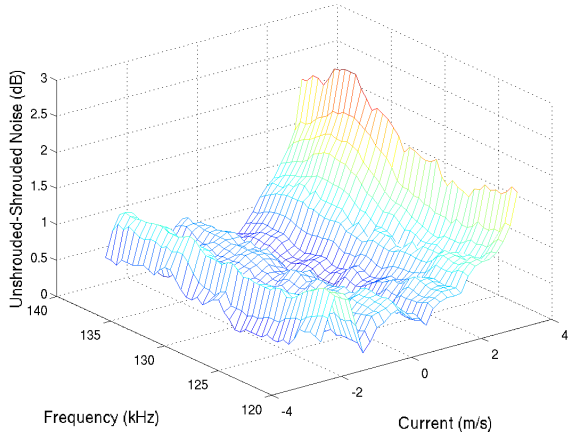


Figure 4: The difference in mean ambient noise (dB re 1µPa^2/Hz) between shrouded and non-shrouded icListenHF hydrophones. Note that the difference is low overall (mostly <1 dB) and greatest at depth averaged speeds >3 m/s on the flood tide.

4 Discussion

The two mooring types examined (tethered SUB buoy and bottom platform) differed in both mobility and distance from the seafloor, and thus experienced different current regimes. These differences are reflected in percent time lost in C-POD recordings (Figure 2) and in instrument tilt, both of which were lower for C-PODs housed in the bottom-moored platform. The tethered SUB buoy (2-3 m above bottom) experienced extreme changes in tilt, up to 60 degrees, at high flow speeds, potentially increasing the pseudonoise and time lost with C-PODs. Noise generated by shackle and chain movements, vibrations, and strumming probably also increased C-POD percent lost time. The C-PODs mounted on the platform were in a more stable, less noisy environment, and may have had greater porpoise detection range.

A comparison of the performance of both technologies, co-located on the platform, showed that the icListenHF recorded five times the number of C-POD Detection Positive Minutes, in large part due to its greater detection range (distance). The icListenHF does not have a built in setting like the C-PODs "lost time" and will only miss porpoises clicks if the ambient noise is great enough to mask the clicks. This hydrophone can also record raw audio files, which were used to validate the C-POD porpoise clicks.

Although foam shrouding has been shown to reduce noise recorded by hydrophones in the frequency range 0-30 kHz [1], this study did not find a significant reduction in the ambient sound collected at any frequency, most likely due to the extreme flows and thus very noisy conditions of the Minas Passage.

5 Conclusions

Hydrophones deployed on the instrument platform outperformed those housed in SUB buoys. Overall, the icListenHF has greater porpoise detection efficiency than the C-PODs. Shrouding the icListenHF with open cell foam did not have an effect on the ambient noise recorded, however, other shrouding options should be explored.

Acknowledgments

This project was funded by the Offshore Energy Research Association (OERA) of NS, MiTACS Accelerate in partnership with FORCE and Ocean Sonics, and an NSERC Engage grant with Ocean Sonics. DFO provided the platform and an acoustic release. M Stokesbury provided an icListenHF unit and Ocean Sonics staff assisted with shroud studies and provided a battery pack.

References

[1] C. Bassett, Ambient noise in an urbanized tidal channel. PhD thesis, University of Washington, USA, 2013.

[2] R. Karsten, J. McMillan, M. Lickley, and R. Haynes, Assessment of tidal current energy in the Minas Passage, Bay of Fundy. *J Power and Energy* 222: 493-507, 2008.

[3] J. Wood, D. Tollit, A. Redden, P. Porskamp, J. Broome, L. Fogarty, C. Booth, and R. Karsten, Passive acoustic monitoring of cetacean activity patterns and movements in Minas Passage: Pre-turbine baseline conditions (2011/2012). Final Report for FORCE and OERA, 2013.

[4] D. Tollit, J. Wood, Broome, J., and A. Redden, Detection of marine mammals and effects monitoring at the NSPI (OpenHydro) turbine site in the Minas Passage during 2010. Final Report for FORCE and OERA, 2011.

MARINE MAMMAL VOCALISATIONS IN RESOLUTE BAY, NUNAVUT

Caitlin O'Neill[12*]
[1]School of Earth and Ocean Sciences, University of Victoria, Victoria, BC, Canada
[2]JASCO Applied Sciences Ltd., 2305-4464 Markham St., Victoria, BC, Canada

1 Introduction

Resolute Bay, located in the Canadian Arctic Archipelago, on the southwest side of Cornwallis Island, faces the northern side of Parry Channel. During the short 2 to 3-month ice-free period each year, Resolute Bay hosts migratory groups of marine mammals including beluga whales, narwhals, walrus, and bearded seals. To investigate the temporal pattern of these vocalizing species, an Autonomous Marine Acoustic Recorder (AMAR) was deployed to collect broadband underwater sound. The data was then processed with automated call and click detectors.

2 Methods

An AMAR was deployed on a floating mooring, in 46 m of water, 3 m above seafloor. It recorded 24-bit broadband sound levels with duty cycles described in Table 1. A higher sample rate was chosen for the ice-free period in order to capture the high frequency sounds from beluga whales and narwhals, often present in the area during these times.

Start Date	End Date	Record (s)	Sleep (s)	Frequency Range
Aug. 5, 2013	Oct. 10, 2013	113	127	10 Hz – 48 kHz
Oct. 11, 2013	Jun. 9, 2014	340 (5.6 min)	3260 (54.4 min)	10 Hz – 8 kHz
Jun. 9, 2014	Jun. 30, 2014	113	127	10 Hz – 48 kHz

Table 1: AMAR recording schedule.

JASCO Applied Sciences' automated detector software was used to identify and classify marine mammal vocalisations within the data set. A tonal call detector was used to identify beluga and narwhal whistles, bearded seal calls, and walrus grunts. While a click detector located narwhal and beluga clicks.

2.1 Automated Tonal Call Detector

The automated tonal call detector works as follows:
1. Create normalized spectrograms by using the median value in each frequency bin for each detection window.
2. From the normalized spectrograms, locate and extract time-frequency contours of the marine mammal calls, using the tonal detector developed by Mellinger et al.[1].
3. Calculate features of extracted contours (e.g. slope, number of inflection points, max/min frequencies).
4. Employ a call sorting algorithm to determine if the contours match the definition of a specific mammal call type.

2.2 Automated Click Detector

The automated click detector works as follows. First, an 8 kHz high-pass filter is used to constrain the data that may contain marine mammal clicks. In order to remove clicks that do not originate from marine mammals, the filtered samples are then summed to create a time series with 0.5 ms root-mean-square time series. Then, a Teager-Kaiser energy detector identifies possible click events. For each possible click event, the maximum peak signal is located within 1 ms of the detected peak. The time window of the detected click is determined by searching the data for the local maxima within 12 dB of the maximum peak. Then the classification parameters are calculated for each click: the number of zero crossings, the median time separation between zero crossings, and the slope of the time/frequency separation between zero crossings. Next, the Mahalanobis distance is calculated by measuring the distance between the extracted click classification parameters and the templates of beluga and narwhal clicks. If a click exceeded the minimum specified beluga or narwhal Mahalanobis distance threshold, it is classified as that species. Otherwise, the click detection is discarded.

3 Results

3.1 Beluga and Narwhal Whistles and Clicks

Beluga and narwhal whistles are very similar in frequency and pattern, which makes it difficult to distinguish the presence of these two species using this call type. Call count results for beluga and narwhal whistles are presented in Figure 1. Whistles were found starting August 7, 2013, with the greatest number recorded on August 9, 2013. Calls were present almost daily between August 7 and November 1, 2013. After migrating out of the area for the winter, the beluga and narwhals returned June 21, 2014 and remained until the recorder stopped recording on June 30, 2014. The daily click counts from beluga and narwhal are presented in Figure 2 and Figure 3, respectively. Although fewer beluga clicks were identified than narwhal clicks, this is likely due to the recorder not having a high enough sample rate to record all the vocalizing frequencies of the beluga clicks. The greatest number of clicks occurred during freeze-up, between September 20 and 29. The click data results reflect the whistle data for the periods of higher sample rate, although with less daily beluga click counts than whistles.

*caitlin.v.oneill@gmail.com

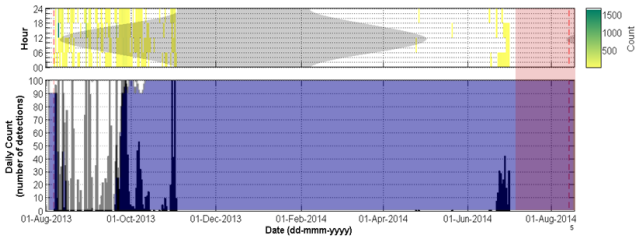


Figure 1: Combined beluga and narwhal daily whistle counts and counts per hour of day with ice coverage percentage[2] overlaid in blue and end of recording in pink.

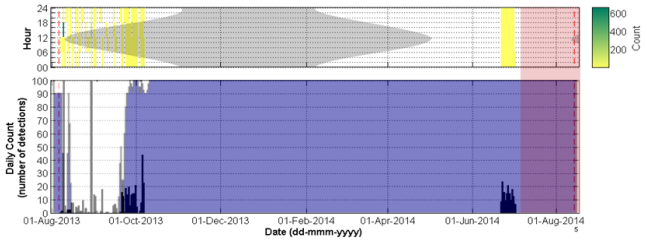


Figure 2: Beluga daily click counts and counts per hour of day with ice coverage percentage[2] overlaid in blue and end of recording in pink.

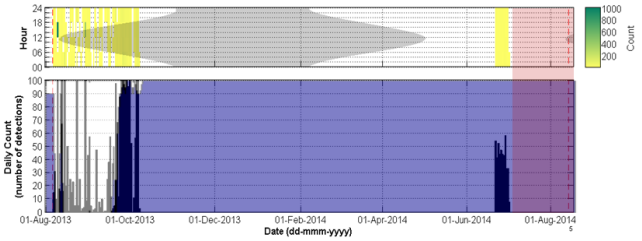


Figure 3: Narwhal daily click counts and counts per hour of day with ice coverage percentage[2] overlaid in blue and end of recording in pink.

3.2 Bearded Seal Calls

Daily call count results for bearded seals are presented in Figure 4. They follow a similar temporal pattern to the beluga and narwhal clicks and whistles, although these animals are also present in the winter, during 100% ice coverage. The greatest number of calls were recorded between June 25 and 30, 2014. Bearded seals vocalized at all times of the day.

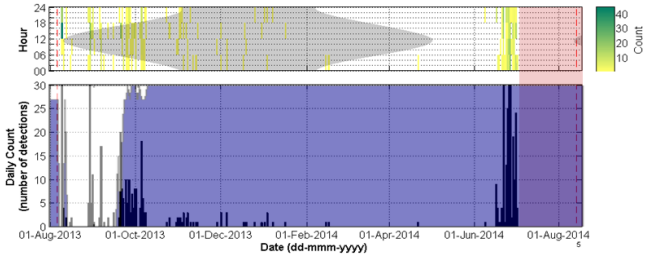


Figure 4: Bearded seal daily call counts and counts per hour of day with ice coverage percentage[2] overlaid in blue and end of recording in pink.

3.3 Walrus Grunts

Daily grunt count results for walrus are presented in Figure 5. They follow a similar temporal pattern as the beluga and narwhal clicks and whistles, also leaving the area in the winter, during 100% ice coverage. Most of the calls were recorded between October 4 and 6, 2013. Similar to beluga and narwhal, they also came back in the summer, albeit briefly, on June 25, 2014.

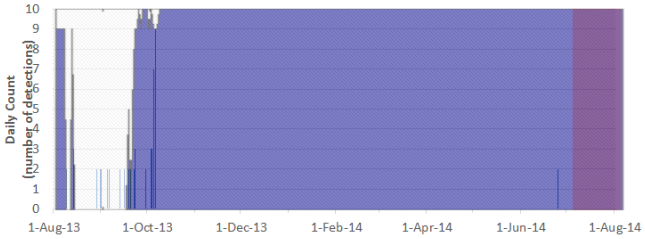


Figure 5: Walrus daily grunt counts with ice coverage percentage[2] overlaid in blue and end of recording in pink.

4 Discussion and Conclusion

It is difficult to build an automated detector that distinguishes between narwhal and beluga whistles, therefore these results were presented together. The best way to distinguish between the presences of these two species is to look at the frequencies of their clicks. Narwhals have pulsed calls between 12-40 kHz and belugas have pulse calls between 40-60 and100-120 kHz[3]. The results of the click detector presented here indicate that both belugas and narwhals were present near Resolute Bay at similar times, correlated to limited ice coverage. The whales and walrus left a couple weeks after 100% ice coverage occurred, but bearded seals remained throughout the winter. Most vocalisations for all analysed species occurred in conjunction with the freeze-up period (90-100% ice coverage). Marine mammal vocalisations did not occur at a preferred time of day throughout the year.

Acknowledgments

The author thanks the Ocean Tracking Network (OTN) for providing funding for this project. Thanks to JASCO Applied Sciences for kindly loaning the AMAR. Many thanks to Xavier Mouy (JASCO) for his assistance with the detection processing and to Ash Hill (JASCO) and Bruce Martin (JASCO) for downloading the data and running the preliminary data processing routines. Thanks to the Polar Continental Shelf Program for providing field logistical support and to Rob Currie and Dave Yurkowski for their assistance during instrument deployment and retrieval.

References

[1] Mellinger, D.K., S.W. Martin, R.P. Morrissey, L. Thomas, and J.J. Yosco. 2011. A method for detecting whistles, moans, and other frequency contour sounds. Journal of the Acoustical Society of America 129(6): 4055-4061.

[2] Canadian Ice Service archives (http://www.ec.gc.ca/glaces-ice/)

[3] Richardson, W.J., et al. Marine Mammals and Noise. Academic Press, London and San Diego, 1998.

PATTERNING IN NORTHERN BOTTLENOSE WHALE (*HYPEROODON AMPULLATUS*) CLICK TRAINS

Hilary B. Moors-Murphy[*1]
[1] Fisheries and Oceans Canada, Bedford Institute of Oceanography, Dartmouth, Nova Scotia

1 Introduction

Northern bottlenose whales (*Hyperoodon ampullatus*) are large toothed whales of the family Ziphiidae; the beaked whales. They are one of the few species of beaked whales whose vocalizations have been fairly well described and are known to produce several types of broadband impulsive vocalizations with spectral content ranging from 0.5 to > 50 kHz, including upswept frequency modulated (FM) pulses, buzz clicks, and more variable surface clicks [1-4]. [1] report a mean FM pulse duration of 0.276 msec, mean interclick interval (ICI) of 306 msec and mean centroid frequency of 47 kHz. This varies somewhat from the mean duration of 0.585 msec, mean ICI of 514 msec, and mean minimum and maximum frequencies of 10.5 and 47.5 kHz reported by [2]. These FM pulses correspond closely to the "deep-water clicks" described by [3], which had a mean duration of 0.35 msec, mean ICI of 400 msec and were heard when no whales were visible at the surface. FM pulses/deep-water clicks are likely used to echolocate prey [1,3]. Buzz clicks have shorter duration and ICI (< 14 msec), no FM structure, and are also likely used for foraging but when prey is at short range [1]. "Surface clicks" are described as clicks emitted in rapid succession with irregular timing with a mean duration of 2.02 msec and mean ICI of 70 msec, recorded when whales were visible at the surface or shortly after diving [3]. Surface clicks likely serve a different function such as echolocating on vessels or companions, or possibly social communication [3].

During northern bottlenose whale focused field studies in summer 2006 conducted in the Gully submarine canyon located offshore Nova Scotia, Canada, acoustic recordings made when northern bottlenose whales were observed at the surface revealed click trains occurring with unusual temporal patterns that have not been previously described. The objective of this study was to describe these atypical click trains in greater detail.

2 Methods

2.1 Data collection

Acoustic recordings of northern bottlenose whales were obtained from the Gully between 25 July and 3 Aug 2006. A 12-meter sailing vessel was used to approach northern bottlenose whales to within 100 m to collect data for photographic identification studies. Acoustic recordings were made on an opportunistic basis, typically when the vessel was stationary and when individuals were observed at the surface near the vessel or shortly after they dove. Length of the recording sessions varied from 1-66 minutes, and a total of 18.3 hours of recording were collected. Detailed notes on the number, sex and age class (if known) of the animals present, and their behavior while at the surface, were also collected throughout recording periods.

All recordings were made using a custom built towed hydrophone array consisting of two Benthos AQ4 hydrophones (frequency response ± 1 dB re 1 V/µPa from 5-30 kHz) spaced 1.8 m apart and housed in a fluid filled case at the end of a 100 m cable. The system contained custom built preamplifiers (which added 27 dB re 1 V/µPa to the incoming signal) and was connected to a custom built high-pass filter (set to 0.4-3.6 kHz, adjusted according to background noise levels) and a Kemo Pocketmaster 1600 lowpass filter (set to 48 kHz). The array was connected to a personal computer through a Creative Labs Audigy PCMCIA sound card with a sampling rate set to 96 kHz. The system was able to record signals up to 48 kHz.

2.2 Analysis of recordings

The recordings were visually and aurally scanned in Raven Pro 1.4 (© Cornell Lab of Ornithology) and click trains with unusual timing patterns were identified. The most frequent atypical timing pattern identified was a "doublet click train" which consisted of pairs of clicks with a typical ICI separated by longer ICIs (Figure 1). A quantitative analysis of this particular click train pattern was completed. Often multiple animals were recorded simultaneously; thus, only relatively loud clicks which were clearly part of the same click train (as determined by visually examining amplitude and frequency structure of the clicks) were included. Any echoes/reflections present were excluded from the analysis.

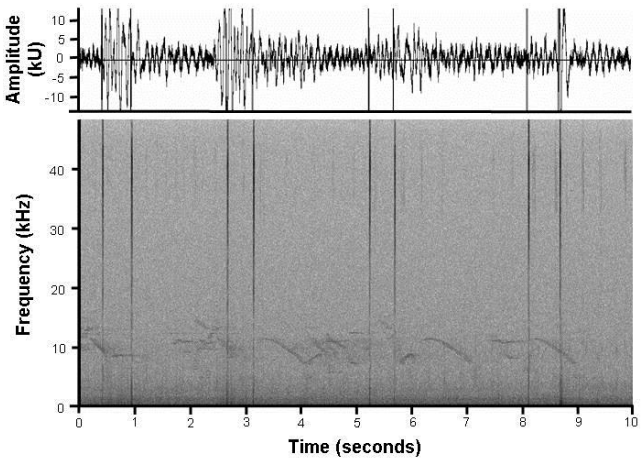


Figure 1: Waveform and spectrogram of a portion of a northern bottlenose whale doublet click train.

The start time of each click within the doublet click train was measured manually in Raven Pro. ICI durations were calculated by subtracting the start time of each click within the click train from the start time of the preceding

[*] Hilary.Moors-Murphy@dfo-mpo.gc.ca

click. The click trains consisted of both short ICI durations (between the two clicks of the click pair) and long ICI durations (between click pairs) (Figure 1). The mean, standard deviation and median ICI duration was calculated for short and long ICIs.

3 Results

Eighteen doublet click trains recorded on different days, and thus likely made by different individuals, were measured (Table 1). The mean long ICI duration was significantly longer than then the mean short ICI duration ($t_{0.05(2)} = -3.12$, df = 78, $p < 0.001$) and within a train the mean long ICI duration tended to be at least two times longer than the mean short ICI duration (Figure 2).

ICI type	Number of ICIs measured	Median (msec)	Mean (msec)	SD (msec)
Short	75	460.20	460.58	92.48
Long	57	1601.00	1486.68	617.4

Table 1: Median, mean and standard deviation of short and long ICIs within northern bottlenose whales doublet click trains.

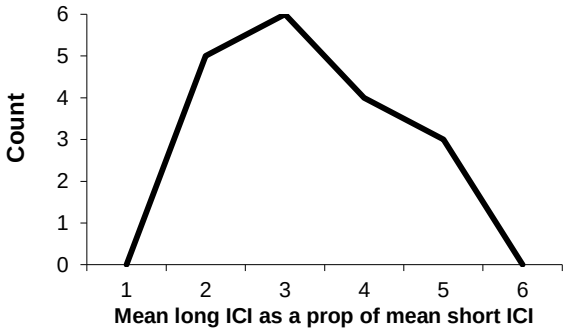


Figure 1: Histogram of the mean long ICI as a proportion of the mean short ICI for each doublet click train.

4 Discussion

The doublet click trains described here are the first click trains of this type reported for cetaceans. Click trains consisting of paired clicks have been observed in echolocating terrestrial species including swiftlets, owls and bats [5-7]. Several possible functions of this clicking pattern have been suggested including increasing signal-to-noise ratio of the clicks [8] and an anti-masking strategy used in group situations [9]. It is possible that the patterning in the doublet click trains produced by northern bottlenose whales allow individuals to distinguish their own clicks from those produced by others in the group. Another possible function is social communication. Non-click vocalizations such as the more typical social sounds produced by other odontocetes (e.g., whistles) do not appear to be a common feature of the vocal repertoire of northern bottlenose whales [1-4]. Sound production limited to clicks is known to occur in sperm whales, whose clicks serve both foraging (through echolocation) and communication functions (e.g., codas) [10]. The northern bottlenose whale doublet click trains examined in this study were recorded when groups ranging in size from 3-11 individuals were present, and often when the animals were displaying social behaviours at the surface.

5 Conclusion

Although the function of doublet click trains is not currently clear, this vocalization is extremely interesting and demands more investigation as it may be an important source of information for increasing our understanding of this species. Further research is required to determine the prevalence of these patterned click trains in the vocal repertoire of northern bottlenose whales and their function. This study highlights that there is still much to learn about the vocal behavior of northern bottlenose whales.

Acknowledgments

I thank C. Dunn for her input and support of this study. Fieldwork support was provided by the Whitehead Lab at Dalhousie University and funded by NSERC. Analysis support was provided by A. MacKenzie and E. Marotte.

References

[1] S.K. Hooker and H. Whitehead. Click characteristics of northern bottlenose whales (*Hyperoodon ampullatus*). *Mar. Mamm. Sci.*, 181: 69-80, 2002.

[2] B. Martin and H.B. Moors-Murphy. Analysis of Northern bottlenose whale pulses and associated reflections recorded from the Gully Marine Protected Area. *Proc. Meet. Acoust.*, 19: 010045, 2013.

[3] M. Wahlberg, K. Beedholm, A. Heerfordt and B. Møhl. Characteristics of bisonar signals from the northern bottlenose whale, *Hyperoodon ampullatus*. *J. Acoust. Soc. Am.*, 130: 2077-3084, 2011.

[4] H.E. Winn, P.J. Perkins, and L. Winn, Sounds and behavior of the northern bottle-nosed whale. 7th Annual Conference on Biological Sonar and Diving Mammals, Stanford Research Institute, Menlo Park, CA., 53-59, 1970.

[5] R.A. Suthers and D. H. Hector. Mechanism for the Production of Echolocating Clicks by the Grey Swiftlet, *CollocMia spodiopygia*. *J. Comp. Physiol.*, 148: 457-470, 1982.

[6] D. Crawford, J.W.H. Ferguson and A.C. Kemp. Why do grass owls (*Tyto capensis*) produce clicking calls. *J. Raptor Res.*, 33:134-142, 1999.

[7] H. Raghuram, N. Gopukumar and K. Sripathi. Presence of single as well as double clicks in the echolocation signals of a fruit bat, *Rousettus leschenaulti* (Chiroptera: Pteropodidae). *Folia Zool.*, 56:33-38, 2007.

[8] E.R. Buchler and A.R. Mitz. Similarities in design features of orientation sounds used by simpler, nonaquatic echolocators. In *Animal Sonar* Systems. Eds R.G. Busnel and J.F. Fish. Plenum Press, NY: 871-874, 1980.

[9] J.H. Fullard, R.M.R. Barclay. Echolocation in free-flying Atiu Swiftlets (*Aerodramus sawtelli*). *Biotropica.*, 25:334-339, 1993.

[10] W.A. Watkins and W.E. Schevill. Sperm whale codas. *J. Acoust. Soc. Am.*, 62: 1485-1490, 1977.

SEASONAL OCCURRENCE OF BLUE WHALE (*BALAENOPTERA MUSCULUS*) VOCALIZATIONS IN THE GULLY MARINE PROTECTED AREA

Emmaline Marotte[*1, 2], **Hilary Moors-Murphy**[†1]

[1]Fisheries and Oceans Canada, Bedford Institute of Oceanography, 1 Challenger Drive, Dartmouth, Nova Scotia, Canada, B2Y 4A2

[2]Dalhousie University School of Resource and Environmental Studies, 6100 University Avenue, Halifax, Nova Scotia, Canada, B3H 4R2

1 Introduction

The North Atlantic blue whale (*Balaenoptera musculus*) is known to frequent waters off eastern Canada during summer months. Many sightings have occurred in the Gully, a large submarine canyon located approximately 200 km off the coast of Nova Scotia, designated a Marine Protected Area (MPA) in 2004 because of its ecological significance [1]. During summer field studies, sightings of blue whales mostly occurred in the Gully in August while few were seen during limited survey efforts in other nearby areas [2]. Their winter occurrence in the area remains poorly understood. Blue whales have been listed as endangered since 2002 and information on their distribution throughout the year is key to establishing effective conservation strategies for the species. The purpose of this study was to examine acoustic recordings obtained from the Gully MPA for the presence of blue whale calls during both summer and winter months.

2 Method

2.1 Acoustic dataset

Acoustic data for this project were collected from bottom-moored hydrophones called Marine Autonomous Recording Units (MARU;© Cornell Lab of Ornithology) deployed by Dr. Hal Whitehead's lab at Dalhousie University as part of a larger cetacean monitoring project along the Scotian Slope [3]. For the present analysis, recordings from the head (GULH) and the mouth (GULM) of the canyon (Figure 1) collected during the summer (June-Sept 2006 at both locations) and winter (Dec 2006-Jan 2007 at GULM and Dec 2007-Feb 2008 at GULH) were examined.

Data were collected at a sampling rate of 50 kHz, thus frequencies up to 25 kHz were recoverable. Recordings were duty-cycled and a single seven or ten minute recording was collected every hour.

2.2 Analysis of call presence/absence

The dataset consisted of > 853 hours of recording and was too large to fully analyze manually, so an automated blue whale call detector developed by JASCO Applied Sciences was used to detect possible blue whale calls on the recordings. Each recording that contained at least one blue whale detection was aurally and visually inspected using spectrographic analysis software Raven Pro 1.4 (© Cornell Lab of Ornithology). When a detection was verified to be a blue whale call, it was deemed a true positive. The minimum proportion of recordings with blue whale calls present on them (P_{min}) was calculated as the number of true positives divided by the number of recordings in the dataset.

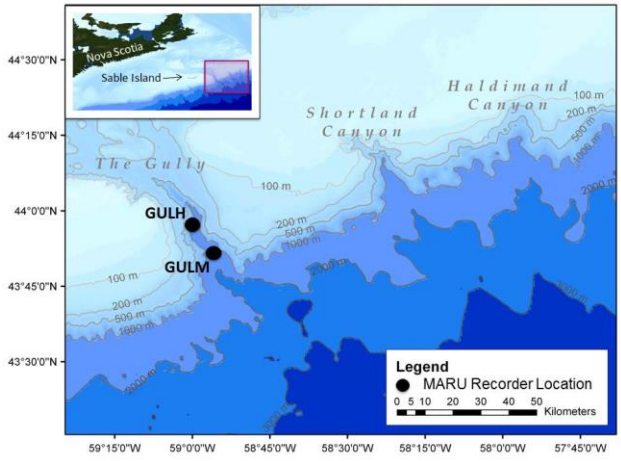


Figure 1: Location of the MARUs within the Gully.

When using automated detectors, two types of detection error need to be considered: the proportion of false positives (no blue whale call is heard or visually observed on the recording despite there being a detection, i.e., a false alarm) and the proportion of false negatives (no detection despite there being a verified blue whale call on the recording, i.e., a missed call). To determine the proportion of false negatives, 100 recordings with no detections on them were randomly selected from each deployment and inspected for the presence of blue whale calls.

3 Results

3.1 Call presence/absence

A total of 5695 recordings from the two locations were obtained. Of those, 787 had detections on them. Manual analysis revealed that 432 (55%) of those recordings contained blue whale calls (Table 1) resulting in an overall false positive rate of 45%.

Of the 400 analyzed recordings with no detections, 41 (10%) contained blue whale calls (Table 2). The proportion of false negatives was higher in summer than in winter; however, the false negative rate was relatively low overall for all four locations.

* emma.marotte@dal.ca
† hilary.moors-murphy@dfo-mpo.gc.ca

Table 1: Blue whale call presence at all four locations. S = summer and W = winter. P_{min} is the minimum proportion of files within a dataset with blue whale calls present on them.

Deployment	Number of recordings with detections	Number of true positives	Proportion of false positives	P_{min}
GULM S	267	214	0.199	0.169
GULM W	162	44	0.728	0.035
GULH S	215	94	0.563	0.074
GULH W	143	80	0.441	0.042
TOTAL	787	432	0.451	0.549

Table 2: Blue whale call absence analysis at all four locations. S = summer and W = winter.

Deployment	Number of recordings with no detections	Recordings with no detections examined	Number of true negatives	Proportion of false negatives
GULM S	1000	100	77	0.230
GULM W	1086	100	98	0.020
GULH S	1051	100	90	0.100
GULH W	1771	100	94	0.060
TOTAL	4908	400	359	0.103

3.2 Temporal and spatial variation

P_{min} was higher overall in summer than in winter (Table 1) but varied between months (Figure 2). The highest proportion of calls occurred in September, followed by August and only a small number of calls occurred in July. Call presence diminished but continued into the winter months, with the highest proportion occurring in January.

P_{min} was higher at GULM than at GULH in late summer and early winter. However, throughout January and February, P_{min} was higher at GULH (Figure 2). No recordings were obtained from GULM in February.

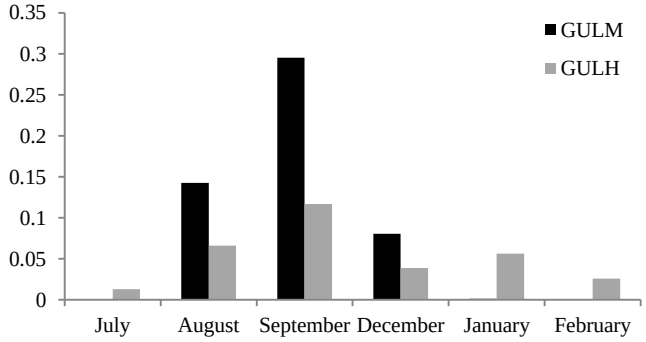


Figure 2: The mean minimum proportion of recordings with blue whale calls present on them (P_{min}) for each month in which blue whale calls were identified.

4 Discussion

Because the detector missed some calls, the results presented here are likely to be underestimates of blue whale presence, especially for the summer months. The percentage of calls missed by the detector is generally low but does vary with deployment. This is likely a reflection of varying levels of system self-noise present on the recordings and differences in background ambient noise levels between seasons.

Many studies suggest that North Atlantic blue whales migrate to southern breeding areas during the winter [4]. Due to limited survey effort, blue whale occurrence in the Gully in winter has not been previously assessed. Results from this study present the first evidence that at least some blue whales remain in the Gully during winter months. Increased calling rates in summer may indicate a higher number of individuals in the area, or that individuals are simply more vocal during that time of year. Reasons for why blue whales call remain debated but it is believed that calling may be related to food availability or to mate selection [5].

5 Conclusions

That blue whales continue to use the Gully in the winter months suggests that the area may be an important winter habitat for individuals. Knowledge of blue whale seasonal occurrence is important for understanding their distribution, which in turn has implications for effective management and protection of this endangered species. Further studies are needed to better describe the how blue whales use the area throughout the year relative to other nearby areas.

Acknowledgments

We would like to thank JASCO Applied Sciences, the Whitehead Lab of Dalhousie University and the Bioacoustics Research Program of Cornell University for their contributions to this study. Funding was provided by the Natural Sciences and Engineering Research Council of Canada, Fisheries and Oceans Canada, The Okeanos Foundation, World Wildlife Fund Canada/Environment Canada Endangered Species Recovery Fund, the Whale and Dolphin Conservation Society.

References

[1] Department of Fisheries and Oceans Canada. The Gully Marine Protected Area Management Plan. *Oceans and Coastal Management Division of the Department of Fisheries and Oceans Canada Maritimes Region*, 74 pp, 2008.

[2] H. Whitehead. Trends in cetacean abundance in the Gully submarine canyon, 1988-2011, highlight a 21% per year increase in Sowerby's beaked whales (*Mesoplodon bidens*). *Can. J. Zoolog.* , 91:3, 2013.

[3] H. B. Moors. Acoustic Monitoring of Scotian Shelf Northern Bottlenose Whales (*Hyperoodon ampullatus*). Dalhousie University, 2012.

[4] J. Beauchamp, H. Bouchard, P. de Margerie, N. Otis and J-Y. Savaria. Recovery Strategy for the blue whale (*Balaenoptera musculus*), Northwest Atlantic population, in Canada. *Species at Risk Act Recovery Strategy Series*, Fisheries and Oceans, Canada, 62 pp, 2009.

[5] E. M. Oleson, J. Calambokidis, W. C. Burgess, M. A. McDonald, C. A. Leduc and J. A. Hildebrand. Behavioural context of call production by eastern North Pacific blue whales. *Mar. Eco. Prog. Ser.*, 330:2, 2007.

LOCALIZING BOWHEAD WHALES IN THE CHUKCHI SEA USING ASYNCHRONOUS HYDROPHONES

Graham A. Warner *[1,2], **Stan E. Dosso**[2], **David E. Hannay**[1], **and Jan Dettmer**[2]
[1]JASCO Applied Sciences, 2305-4464 Markham St, Victoria BC V8Z 7X8
[2]School of Earth and Ocean Sciences, University of Victoria, 3800 Finnerty Rd, Victoria BC V8P 5C2

1 Introduction

Marine mammal localization is required for assessing behavioural effects from underwater anthropogenic noise. Passive acoustic monitoring using widely-distributed autonomous recorders is often used for detecting marine mammals over large areas ; however, localization is more difficult because non-linear recorder clock drift desynchronizes underwater recorders and precludes correlation-based localization methods that use time difference of arrival (TDOA) data. In this paper, we localize bowhead whales using Bayesian inversion of calls detected on asynchronous hydrophones in the Chukchi Sea north of Alaska. The shallow water environment acts as a dispersive waveguide and supports a limited number of propagating modes. The mode arrival times depend on the whale location, water sound-speed profile (SSP), seabed geoacoustic properties, and relative recorder clock drift, all of which are treated as unknown parameters estimated in the inversion. The Bayesian formulation also provides rigorous uncertainty estimates for the model parameters.

2 Methods

2.1 Data Processing

Bowhead whales make low-frequency calls that propagate tens of kilometres as normal modes in the shallow waters of the Chukchi Sea. The modes propagate with different group speeds (i.e., they disperse) but it can be difficult to distinguish mode arrivals in a spectrogram if, for example, the whale-hydrophone range is small or the instantaneous frequency (IF) of the call is slowly-varying with time. Recently, Bonnel *et al.* [1] developed a method to apply mode-warping to frequency-modulated (FM) bowhead calls for filtering modes from each other. We use the same approach here to filter the modes and define arrival times as the time of maximum amplitude for each frequency and mode using a spectrogram of the mode-filtered signal.

2.2 Bayesian Inversion

The frequency-dependent mode arrival times depend on the whale location, source IF, environment, and relative recorder clock drifts as given by

$$t_{wam}(f) = \frac{\sqrt{(x_w - x_a)^2 + (y_w - y_a)^2}}{v_m(f)} + \tau_w(f) + \Delta_a, \quad (1)$$

where x_w and y_w are the easting and northing coordinates of whale w, x_a and y_a are the coordinates of recorder a (considered known in this paper), $v_m(f)$ is the group speed for mode m, $\tau_w(f)$ is the source IF, and Δ_a is the recorder clock drift relative to a reference recorder. We use a trans-dimensional (trans-D) Bayesian formulation for the unknown range-independent environmental properties [2] which determine the mode group speeds. The environmental model for the water column has unknown depth and SSP that is defined by an unknown number of depth/sound-speed nodes. The seabed model consists of an unknown number of homogeneous layers overlying a halfspace. The seabed properties (layer thickness, sound speed, and density) are considered unknown. The normal-mode code ORCA [3] is used to calculate the mode group speeds for a given environment and are converted to predicted modal arrival times using the whale location, source IF, and relative recorder clock drift [Eq. (1)], all of which are unknown parameters in the inversion. In a Bayesian formulation the solution consists of properties of the posterior probability density (PPD) of the model parameters given the measured data and prior information. The PPD is approximated numerically for non-linear problems using a reversible-jump Markov-chain Monte Carlo [4] algorithm where model transitions are accepted with a probability that depends on the prior, proposal, and likelihood ratios. See Warner *et al.* [2] for details on the implementation of this algorithm in a related inverse problem that uses the same trans-D environmental model.

2.3 Likelihood

In this paper, we assume residual errors between measured and predicted data are uncorrelated and Gaussian-distributed for each whale call (with standard deviation σ_w). If $\mathbf{d}_{wam}$ is a vector of modal arrival times at N_{wam} frequencies, the likelihood function is the product

$$L(\mathbf{m}) = \prod_{w=1}^{W} \prod_{a=1}^{A} \prod_{m=1}^{M_{wa}} \frac{1}{(2\pi\sigma_w^2)^{N_{wam}/2}} \times \exp\left[-\frac{|\mathbf{d}_{wam} - \mathbf{d}_{wam}(\mathbf{m})|^2}{2\sigma_w^2} \right], \quad (2)$$

where W is the total number of whale calls, A is the total number of recorders that detected a call, and M_{wa} is the number of modes for call w on recorder a. A maximum-likelihood estimate for the inter-recorder clock drift Δ_a can be derived by substituting Eq. (1) into this equation and setting $\partial L / \partial \Delta_a = 0$, giving :

$$\hat{\Delta}_a(\mathbf{m}) = \left[\sum_{w}^{W} \sum_{m}^{M_{wa}} N_{wam}/\sigma_w^2 \right]^{-1} \times \sum_{w}^{W} \sum_{m}^{M_{wa}} \sum_{f}^{N_{wam}} \left[d_{wam}(f) - \tau_w(f) - \frac{|\mathbf{r}_{wa}|}{v_m(f)} \right] / \sigma_w^2, \quad (3)$$

where $\mathbf{r}_{wa}$ is $(x_a - x_w, y_a - y_w)$. Equation (3) provides an estimation of unknown Δ_a without having to explicitly sample

*. graham.warner@jasco.com

this parameter. This formulation assumes the clock drifts do not change significantly in the time interval between the inverted whale calls (3.25 mins).

3 Bowhead Whale Call Data

JASCO Applied Sciences recorded thousands of bowhead whale calls in the Chukchi Sea on several Autonomous Multichannel Acoustic Recorders (AMARs) during Aug.–Oct. 2013 [5]. In this paper, we invert mode arrival times for nine whale calls recorded on up to seven AMARs (herein named A–G) over a 3.25 min period on 11 Oct. Figure 1 shows spectrograms of whale call 1 on three selected AMARs with the mode arrival time picks ; the relative modal time separation indicates the whale was closest to AMAR D since that recording showed the least dispersion.

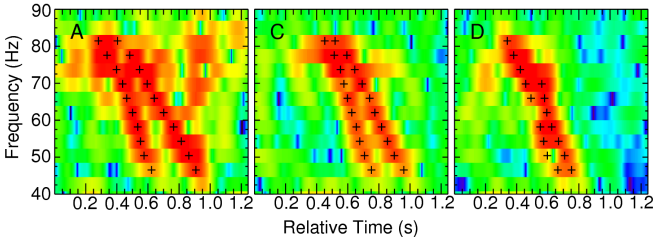


Figure 1 : Example spectrograms of whale call 1 recorded on AMARs A, C, and D. Data picks are shown with + symbols.

4 Inversion Results

Nine whale calls were inverted independently (scenarios S1–S9) and then jointly for one multi-call inversion (S10). Figure 2 shows the two-dimensional marginal probability distributions for whale locations for S1–S3, S9, and S10 (results for S4–S8 are similar to those of S3 and are omitted here for brevity), with the recorders that detected calls shown with × symbols. Scenario 10 clearly shows the localization improvements from joint inversion of multiple whale calls. Whale locations in S10 are estimated to accuracies of 30–160 m.

The clock drift uncertainty estimates varied between 8 and 728 ms for the single-call inversions. Uncertainties were correlated with distance from the reference recorder and strongly reflected localization uncertainty. The corresponding drift uncertainty estimates for the multi-call scenario varied between 3 and 26 ms (a reduction of up to 96% compared to the single-call scenarios).

5 Discussion and Conclusions

Each recorded call provides a whale range estimate through the amount of modal dispersion, and having more estimates from a distributed cluster of recorders better constrains the location of the whale, which in turn improves estimated clock drifts. Localization and clock drift results significantly improved when jointly inverting multiple calls. Multi-call inversions require all data to be fit simultaneously, which effectively requires an averaging of the clock drifts. The benefit is greatest when the whales are spread out among the hydrophone cluster. The Bayesian inversion was applied to bowhead whale call modal dispersion data recorded on asynchronous hydrophones and accounted for unknown whale location, source IF, water SSP, subbottom geoacoustic properties, and relative recorder clock drifts. The synchronization provided by the inversion may allow other localization methods to be applied to other types of marine mammal calls.

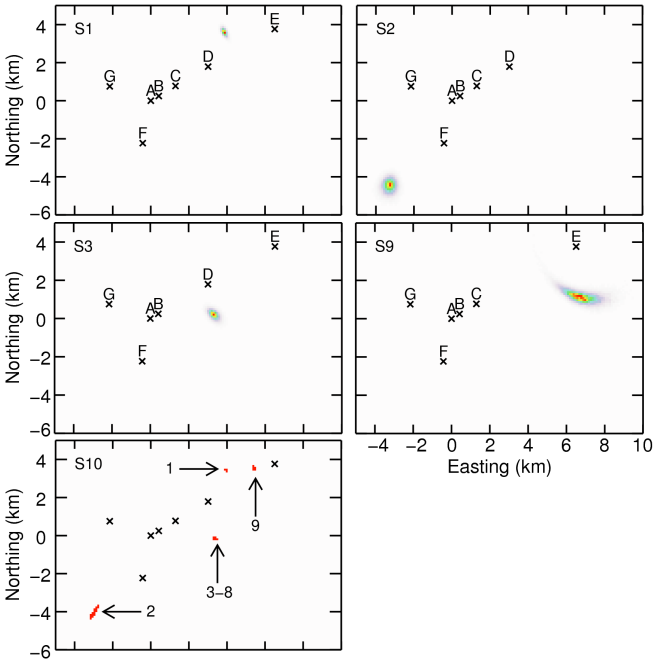


Figure 2 : Marginal probability densities for bowhead whale location(s) for selected scenarios. Probability distributions for S10 are quite compact and are shown as a binary image for clarity.

Acknowledgments

The authors thank Shell Exploration and Production Company for permission to publish this analysis of their data.

References

[1] J. Bonnel, A. M. Thode, S. B. Blackwell, K. Kim, and A. M. Macrander. Range estimation of bowhead whale (*Balaena mysticetus*) calls in the Arctic using a single hydrophone. *J. Acoust. Soc. Am.*, 136 :145–155, 2014.

[2] G. A. Warner, S. E. Dosso, J. Dettmer, and D. E. Hannay. Bayesian environmental inversion of airgun modal dispersion using a single hydrophone in the Chukchi Sea. *J. Acoust. Soc. Am.*, 137 :3009–3023, 2015.

[3] E. K. Westwood, C. T. Tindle, and N. R. Chapman. A normal mode model for acousto-elastic ocean environments. *J. Acoust. Soc. Am.*, 100 :3631–3645, 1996.

[4] P. J. Green. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82 :711–732, 1995.

[5] J. Delarue, J. MacDonnell, B. Martin, X. Mouy, and D. Hannay. Northeastern Chukchi Sea joint acoustic monitoring program 2012–2013. Technical Report 00808, JASCO Applied Sciences, 2014.

ENGINEERING ACOUSTICS / NOISE CONTROL - INGÉNIERIE / CONTRÔLE DU BRUIT

MEASURED C-WEIGHTED AMBIENT SOUND LEVELS FOR USE WITH ENVIRONMENTAL NOISE REGULATIONS IN ALBERTA AND BRITISH COLUMBIA

Victor W. Young [*1], Zhaohui Yu [†1] and Andrew C. Faszer [‡1]
[1]Golder Associates Ltd. 102, 2535 3 Avenue SE, Calgary, Alberta, T2A 7W5

1 Introduction

Noise Impact Assessments (NIAs) for industrial facilities in Alberta are typically conducted in accordance with Alberta Energy Regulator (AER) Directive 038 [1] or Alberta Utilities Commission (AUC) Rule 012 [2]. Both regulations assess noise cumulatively and require that the contribution of natural and non-industrial sources be included when testing noise compliance for industrial facilities. Noise regulations in British Columbia (BC) are very similar [3].

The Alberta regulations and the comparable BC regulation require that A-weighted and C-weighted cumulative noise levels be compared. A difference between C-weighted and A-weighted noise levels greater than or equal to 20 is considered indicative of a potential Low Frequency Noise (LFN) issue.

Both Alberta regulations and the comparable BC regulation provide desktop techniques for estimating the A-weighted Ambient Sound Level (ASL) associated with natural and non-industrial noise sources. The problem is that these regulations do not provide a methodology for estimating C-weighted ASL values. As such, it is difficult to apply the LFN test using cumulative noise levels and instead NIAs often consider a facility in isolation when testing for LFN issues. This can lead to potential LFN issues being identified even when facility noise levels are very low – so low that they would be completely obscured by the background ASL if it could be included in the LFN test.

This paper establishes a link between A-weighted and C-weighted ASL values for three classes of noise receptors: remote areas far from human activity; isolated rural dwellings; and dwellings adjacent to noisy roads. C-weighted ASL values that can be used in conjunction with the A-weighted ASL values when applying regulatory LFN tests are presented. Links between A-weighted and C-weighted ASL values are established by examining measured noise spectra from 15 receptor locations.

2 Regulatory Context

AER Directive 038, AUC Rule 012, and the comparable BC regulation are all very similar. All three require a cumulative assessment of noise levels at receptor locations corresponding to occupied dwellings. In all three regulations, A-weighted ASL values at receptors are estimated via the same desktop technique, which accounts for time of day, population density, and proximity to transportation infrastructure. ASL values calculated using this technique can range from 35 A-weighted decibels (dBA) to 61 dBA. The minimum ASL value corresponds to the nighttime period (i.e., 10 pm to 7 am) at a receptor located in an area with population density less than nine dwellings per quarter section and more than 500 metres from heavily travelled roads or rail lines. The maximum ASL value corresponds to the daytime period (i.e., 7 am to 10 pm) at a receptor located in an area with population density greater than 160 dwellings per quarter section and less than 30 metres from heavily travelled roads or rail lines.

The exact process by which the regulatory bodies established the desktop technique for calculating ASL values is rather obscure, but seems to be based on field measurements conducted in the 1970s on behalf of the Alberta Department of the Environment [4]. Wherever the original ASL data came from, the professional experience of this paper's authors is that the desktop technique does a reasonably good job of establishing ASL values for most situations. As such, the purpose of this paper is not to challenge the A-weighted ASL values obtained using the desktop technique. Instead, the purpose of this paper is to establish complementary C-weighted ASL values that can be used when applying the LFN test. In particular, this paper seeks C-weighted ASL values to complement the A-weighted ASL values for the receptor classes presented in Table 1.

Receptor Class	Regulatory Daytime ASL	Regulatory Nighttime ASL
A – remote area far from any human activity	45 dBA	35 dBA
B – occupied dwelling more than 500 metres from busy roads or rail lines in an area with population density less than nine dwellings per quarter section	45 dBA	35 dBA
C – occupied dwelling located between 30 metres and 500 metres of a busy road in an area with population density less than nine dwellings per quarter section	50 dBA	40 dBA

Table 1: A-weighted ASL values for three receptor classes

3 Data and Analysis

3.1 Methodology

Golder Associates Ltd. (Golder) has been conducting noise monitoring in Western Canada for more than a decade and during that time has built up an extensive measurement database, which was sampled for this paper. Monitoring

[*] victor_young@golder.com
[†] zhaohui_yu@golder.com
[‡] andrew_faszer@golder.com

data for five class A receptors, five class B receptors, and five class C receptors were identified as being representative of the range of environmental conditions observed at these receptor classes, and were processed to establish a link between A-weighted and C-weighted ASL values.

The raw monitoring data for each receptor consisted of a series of one-third octave-band noise spectra, with each spectra representing a one-minute energy equivalent noise level ($L_{eq,1min}$). To capture circadian variations in noise levels and noise sources, daytime and nighttime data were considered separately for each receptor.

All monitoring data considered in this paper were collected in accordance with AER Directive 038, AUC Rule 012, or the comparable BC regulation using an appropriately-calibrated Type I integrating sound level meter. Likewise, all monitoring data were filtered in accordance with regulatory guidance to remove invalid noise sources, such as rain on the microphone. As a result of this filtering process, there were periods for some class B receptors where data were not available.

The individual one-minute spectra were weighted and the spectral values were summed to obtain a series of $L_{eq1,min}$ values in dBA and C-weighted decibels (dBC). All available $L_{eq,1min}$ values for a given receptor were plotted on a dBC vs. dBA scatter plot and a statistical t-test was used to check the significance of the correlation between dBA and dBC values. For all receptors, the correlation was found to be both strong and significant.

Linear regression was separately used to fit the dBC vs. dBA data for each receptor. The regression coefficients were then used to predict complementary C-weighted ASL values for use with the A-weighted ASL values presented in Table 1. A statistical t-test was used to establish 95% confidence bounds on the predicted C-weighted values for each receptor. Individual predictions for each receptor class were then averaged, using the 95% confidence bounds as uncertainty weights, to obtain C-weighted ASL values representative of the class as a whole.

3.2 Results

Table 2 presents results for receptor class A – remote area far from any human activity. Table 3 presents results for receptor class B – isolated dwelling. Table 4 presents results for receptor class C – occupied dwelling adjacent to a noisy road.

Receptor	C-Weighted ASL for use with 45 dBA	C-Weighted ASL for use with 35 dBA
A1	47 ± 7 dBC	38 ± 4 dBC
A2	57 ± 10 dBC	34 ± 3 dBC
A3	51 ± 6 dBC	47 ± 4 dBC
A4	51 ± 7 dBC	44 ± 3 dBC
A5	53 ± 11 dBC	37 ± 2 dBC
Weighted Average	*51 dBC*	*39 dBC*

Table 2: Results for Receptor Class A

Receptor	C-Weighted ASL for use with 45 dBA	C-Weighted ASL for use with 35 dBA
B1	49 ± 6 dBC	43 ± 7 dBC
B2	52 ± 10 dBC	no valid nighttime data
B3	no valid daytime data	43 ± 3 dBC
B4	53 ± 8 dBC	40 ± 6 dBC
B5	50 ± 9 dBC	41 ± 5 dBC
Weighted Average	*51 dBC*	*42 dBC*

Table 3: Results for Receptor Class B

Receptor	C-Weighted ASL for use with 50 dBA	C-Weighted ASL for use with 40 dBA
C1	53 ± 3 dBC	50 ± 5 dBC
C2	58 ± 6 dBC	53 ± 6 dBC
C3	59 ± 7 dBC	53 ± 6 dBC
C4	64 ± 8 dBC	61 ± 7 dBC
C5	58 ± 7 dBC	49 ± 8 dBC
Weighted Average	*56 dBC*	*53 dBC*

Table 4: Results for Receptor Class C

4 Discussion

By analyzing $L_{eq,1min}$ spectra measured at 15 different receptor locations, this paper has established C-weighted ASL values that can be used in conjunction with A-weighted ASL values when performing a regulatory LFN test.

The authors do not claim that the C-weighted ASL values presented in this paper are true or accurate representations of C-weighted noise levels that would be observed for a particular receptor or receptor class. Indeed, the variability inherent in the environment at most receptors makes characterization of noise levels by a single number effectively impossible. Instead, the authors claim that the C-weighted ASL values presented in this paper are an appropriate match for A-weighted ASL values calculated using the desktop technique described in AER Directive 038, AUC Rule 012, and the comparable BC regulation.

It is hoped that use of the C-weighted ASL values presented in this paper when applying the LFN test will reduce the number of false-positives that often occur when facility noise levels are predicted to be low, and thereby allow for a more realistic assessment of LFN issues in future NIAs conducted in Alberta and BC.

References

[1] Energy and Utilities Board. 2007. Directive 038: Noise Control. Approved February 16, 2007.

[2] Alberta Utilities Commission. 2013. Rule 012: Noise Control. Effective April 1, 2013.

[3] BC Oil and Gas Commission. 2009. British Columbia Noise Control Best Practices Guideline. Issued March 2009.

[4] Bolstad Engineering Associates Ltd. 1977. Compressor Station and Gas Plant Noise Emission. Prepared for Alberta Department of the Environment.

PERFORMANCE ANALYSIS OF ANNULAR PASSIVE SILENCERS – AN UPDATE

Ramani Ramakrishnan
Department of Architectural Science, Ryerson University, Toronto, M5B 2K3

1 Introduction

One dimensional analysis of the performance of annular silencers containing porous absorbers was presented in early 1990s by Ramakrishnan and Watson [1]. The impact of the perforated covering of the porous materials was not included in the earlier results. Further, the influence of duct end reflections were also not included. The multi-physics software, COMSOL, can handle three dimensional propagation as well as the effect of perforated sheets. Six simple annular silencers, tested using ASTM standards, were used in a two-D and three-D models developed in the multi-physics simulation software. The impact of the perforated sheet was also included in the simulations. The predicted results were compared to the test data. The main aim of the current investigation is create updated versions of design curves for performance evaluations of passive silencers. The comparison results are presented in this paper.

2 Background

Details of typical annular silencers are shown in Figure 1. The main focus is to evaluate the transmission loss of the sound as the flows traverses the area with porous materials.

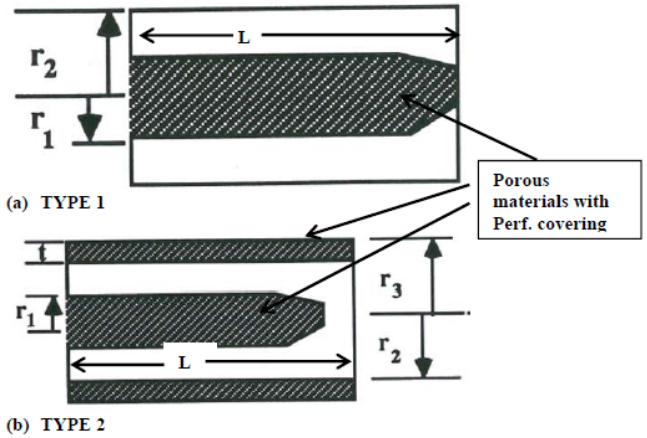


Figure 1: Schematic details of annular silencers.

The acoustic performance is evaluated by solving the governing wave equations between the inlet and outlet of the silencers. The porous materials are considered bulk materials and solution domain includes propagation within the bulk materials [1, 2, 3]. The insertion loss, IL, is given by,

$$IL = 10\log\left(\frac{L_{w,in}}{L_{w,out}}\right) \qquad [1]$$

where, $L_{w,in}$ is the sound power at the inlet plane and $L_{w,out}$ is the sound power at the exit plane. FEM (Finite Element Analysis) methods were applied to evaluate the IL. The powerful software COMSOL Multiphysiscs was used as the FEM solver [4]. The only major disadvantage of COMSL is the computing time and storage capacity of the machine used to solve the fundamental wave equations governing both the free air and porous material regions.

The propagation through the material requires the complex wave speed and acoustic density and they can be obtained from either of the two References 2 and 3. The main input parameter for the determination of the wave speed and density is the flow resistivity of the porous material.

To test the performance of the silencers, the experimental results of Reference 1 will be used for comparison with COMSOL results. The details of two different types of annular silencers are presented in Table 1.

Number	Type	L, cm	r_2, cm	r_1/r_2	t, cm
SIL 1	1	105	61	2	10
SIL 2	2	122	61	1.5	10
SIL 3	2	122	61	2	10
SIL 4	2	122	61	3.43	10
SIL 5	2	61	31	1.71	10
SIL 6	2	61	31	2.4	10

Table 1: Six Annular Silencer Types.

The six silencers were simulated in COMSOL as Two-D and Three-D models and the results are discussed below.

3 Results and discussion

The results of the comparison between the Two-D and Three-D models are shown first in Figure 2, 3 and 4 for silencers SIL 1, SIL 2 and SIL 5.

Figure 2: Comparison results for SIL 1.

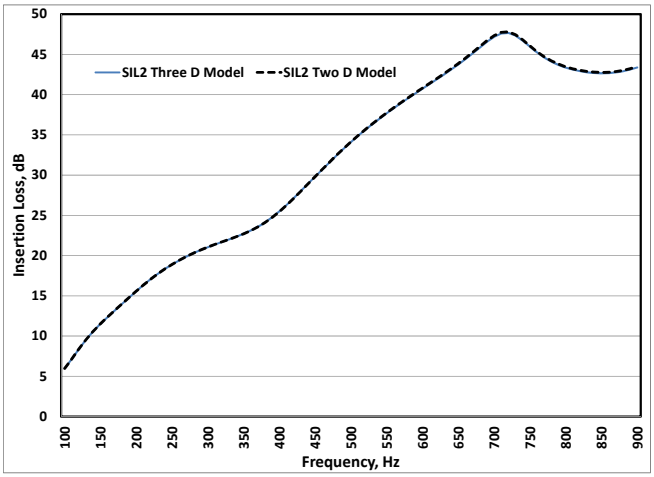


Figure 3: Comparison results for SIL 2.

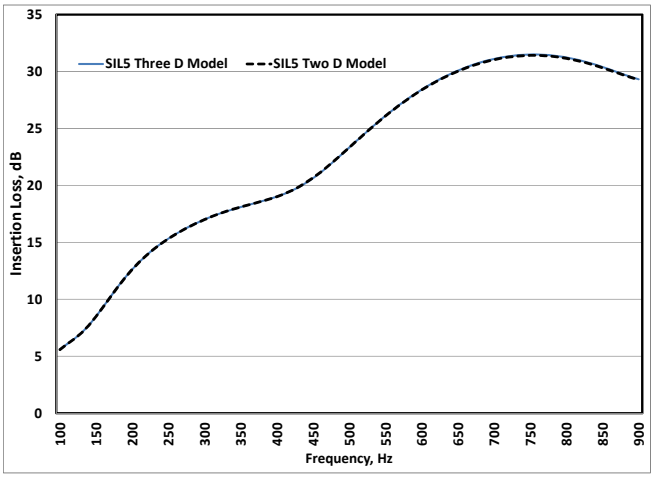


Figure 4: Comparison results for SIL 5.

The simulation model requires a minimum of 5 elements per wavelength. The required number, even though the six silencers are not large, for Octave band numbers 6 (2000 Hz band) and 7 (4000 Hz band) would be in the millions. The 'run time' would be around 36 hours. Hence comparison results between the Two-D model and the Three-D model for three silencers were presented from 100 to 900 Hz in Figures 2, 3, and 4. The results clearly indicate that Two-D modelling is more than adequate.

The IL results for the six silencers were evaluated using Two-D simulation, with the flow resistivity set at 20,000 MKS rayls. The results were compared to the test data and One-D model of Reference 1. The results are shown in Table 2. It is seen that the Two-D model results compare well with test data except at a few high frequency cases. It must be pointed out the flow resistivity of the porous material used in the experiments (Reference 1) was assumed to be 20,000 MKS rayls. However, the test results of Logawa and Hodgson showed that the flow resistivity is a function of frequency and can vary by about 10% to 15% across the frequency bands. Three dimensional modelling is also critical to provide results with engineering accuracy. Even the One-D model of Reference 1 was seen to compare well test data. It must be pointed out the COMSOL simulation is more realistic as it includes the effect of the perforate cover and the reflection effects at the silencer exit.

SIL #	Condition	Octave Band Number					
		2	**3**	**4**	**5**	**6**	**7**
1	TEST Data	**2**	**4**	**7**	**8**	**7**	**6**
	Reference 1	2	4	7	10	5	4
	Two -D Model	**2**	**3**	**8**	**11**	**6**	**3**
2	TEST Data	**10**	**19**	**29**	**38**	**41**	**24**
	Reference 1	8	19	27	37	38	15
	Two -D Model	**9**	**18**	**30**	**46**	**43**	**21**
3	TEST Data	**8**	**15**	**23**	**29**	**29**	**17**
	Reference 1	6	16	27	36	30	13
	Two -D Model	**7**	**15**	**25**	**33**	**25**	**11**
4	TEST Data	**8**	**12**	**17**	**23**	**20**	**14**
	Reference 1	5	13	23	27	16	9
	Two -D Model	**5**	**13**	**20**	**22**	**13**	**5**
5	TEST Data	**7**	**14**	**18**	**26**	**34**	**25**
	Reference 1	5	12	21	31	37	29
	Two -D Model	**7**	**15**	**23**	**32**	**35**	**23**
6	TEST Data	**6**	**13**	**15**	**22**	**27**	**18**
	Reference 1	5	12	18	25	28	19
	Two-D Model	**7**	**14**	**20**	**27**	**26**	**15**

Table 1: Insertion Loss of Annular Silencers.

4 Conclusion

The insertion loss results of annular silencers were determined from experiment and numerical analysis. The results showed the importance of flow resistivity of absorbers. The Two-D model was seen to provide good comparison with test data.

Acknowledgments

We would like to thank Prof. Hodgson and his research group at the University of British Columbia for providing performance results of porous materials.

References

[1] R. Ramakrishnan and W. Watson. Design Curves for Circular and Annular Silencers. *Noise Control Engineering Journal,* vol. 36 (3), pp. 107-120, 1991.

[2] M. Delany and E. Bazley, Acoustical Properties of Fibrous Absorbent Materials. *Applied Acoustics,* vol. 3, pp. 105-116, 1970.

[3] Y. Miki. Acoustical Properties of porous materials - Modifications of Delany-Bazley models. Journal of the Acoustical Society of Japan. vol. 11, no. 1, 1990.

[4] COMSOL Multiphysics, COMSOL Inc. Burlington, MASS, USA.

[5] B. Logawa and M. Hodgson. Duct-Liner Acoustical Characterization. Technical Report, Acoustics and Noise Research Group, University of British Columbia, July 2013.

INVESTIGATION ON NON-POINT SOURCES APPROXIMATION IN OUTDOOR NOISE PREDICTIONS

Sheying Sun, Neil Morozumi, Justin Caskey, Richard Patching
Patching Associates Acoustical Engineering Ltd.
4825 Westwinds Drive NE, Calgary Alberta T3J 4L4

1 Présentation (Français)

La norme ISO 9613 spécifie une méthode de calcul de l'atténuation de la propagation des bruits de l'extérieur, qui a été reconnu et accepté par les différentes autorités réglementaires. Comme cette méthode a été mise en œuvre dans la plupart des logiciels de pointe, tels que Cadna/A et SoundPLAN, cela rend possible la prédiction et la résolution de problèmes de bruit dans un environnement complexe ou l'évaluation du bruit d'un grand établissement. Cependant, une importante décision est à prendre sur la façon de modéliser chaque source de bruit, en particulier, en décidant quelle situation de la répartition des sources de bruit d'éléments individuels associés à une installation complexe aura une incidence sur le motif du champ de bruit qui en émane. La norme ISO 9613 indique que les sources linéiques et surfaciques (« line sources » et « area sources ») peuvent être divisées en sections, chacune représentée par une source ponctuelle en son centre. Toutefois, cela ne peut garantir la précision de calcul pour certaines distances de l'unique point source équivalent. Par conséquent, une grande segmentation des sources est nécessaire pour calculer avec précision le niveau de bruit sur le terrain à proximité de sources linéique et surfacique, ce qui peut augmenter la charge de calcul et peut être nécessaire seulement en champ proche. Ce document met l'accent sur la comparaison des résultats entre les niveaux de bruit prévus et mesurés, en général et dans bandes d'octaves qui prouvent la validité de la méthode de calcul proposée. Ensuite, le modèle validé sera utilisé pour étudier l'applicabilité et la précision dans la prédiction de bruit dans l'environnement. Les critères de distance pour la méthode de source ponctuelle peuvent être appliqués efficacement pour prédire et résoudre les problèmes de bruit dans un environnement complexe pour la prédiction de bruit extérieur et aussi pour les mesures sur le terrain.

2 Presentation (English)

ISO 9613 specifies a method for calculating the attenuation of outdoor sound propagation, which has been recognized and accepted by various regulatory authorities. As this method has been implemented in most advanced software packages, such as Cadna/A and SoundPLAN, this makes it feasible for predicting and resolving noise problems in a complex environment or the noise assessment of a large facility. However, an important decision to be taken is on how to model each noise source, in particular, deciding which situation the distribution of the individual component noise sources associated with a complex facility will affect the pattern of the noise field that emanates from it. ISO 9613 guides that line and area sources may be divided into sections, each represented by a point source at its center. However, this may only ensure the calculation accuracy at some distance from the single equivalent point source. Hence, a large segmentation of the sources is required to precisely calculate the noise level at the near field from line and area sources, which may increase calculation burden and may be only required in the near field. This paper will focus on comparing results between predicted and measured noise levels, overall and in octave-bands, which prove the validity of the proposed calculation method. Then, the validated model will be used to investigate the applicability and accuracy in the environmental noise prediction. The distance criteria for the point source method can be effectively applied for predicting and resolving noise problems in a complex environment for outdoor noise prediction and also for field measurement.

3 Introduction

A major challenge in acoustics is to accurately predict what the sound level will be at some location, far or near. The sound level at some location is the combination of all the noise sources.

In order to predict or model noise from equipment we need to understand, or better, define the sound fields and the predicted sound level associated with those fields. ISO 9613-2, *Acoustics – Attenuation of sound during propagation outdoors*, is the standard used for modeling outdoor sound propagation and predicting far field sound levels. The propagation of sound is generally described using the following expression,

$$L_p = L_w + 10\log\left(\frac{Q}{4\pi r^2}\right) - \sum A_i$$

In this method, the sound pressure level at some location is modeled by a set of geometric terms ($Q/4\pi r2$) and losses (ΣAi). Q defines the reflective surfaces (boundary conditions) that are around the source of noise having a sound power, L_W. ΣAi is the term used to account for all the elements that can affect the sound level (directivity, atmospheric loss, barriers, ground effects, trees, etc.).

The far field starts where the sound field becomes more stable and propagation is fairly uniform. The location is frequency (wavelength) dependent and is usually two to four major source dimensions (width and height of the source) away from the noise source.

In modeling, all sources can be modeled as a point source if the concerned receiver location is far enough based on the frequency consideration. This simplification will get the modeling of complex plant more efficient with reasonable accuracy.

4 Comparison on Model Scenarios

4.1 Comparison settings

The noise modeling was conducted using CadnaA by Datakustik. CadnaA is an advanced noise propagation model that considers geometric spreading, atmospheric sound absorption, ground impedance effects, site topography and geometry, vegetation and environmental conditions. The calculations performed in CadnaA were conducted in accordance with ISO 9613.

In order to investigate the point source simplification of extended noise sources such as line sources and area sources, the following scenarios were simulated using CadnaA software:

1) Single noise source (point source vs line, area and vertical area source;

2) Single gas driven compressor unit, sound power levels of the noise sources validated with field data;

3) Multiple compressor units (10x) to simulate a plant with dozens of point sources mixed with extended (line or area) noise sources.

4.2 Data analysis

In these three scenarios, the predicted overall and octave band sound pressure levels at various receiver locations using point sources were compared with the model results built from extended noise sources. The accuracy of the noise source simplification was checked at various receiver locations.

5 Results and Discussion

Figures 1-3 show the predicted overall sound pressure levels at various receiver distances in four directions for the three scenarios above. The ground absorption was modelled as 0.6 for a typical grass field. The results show that simplification of point sources could get reasonable prediction alignment compared with the extended source modelling, usually from 50m away and greater.

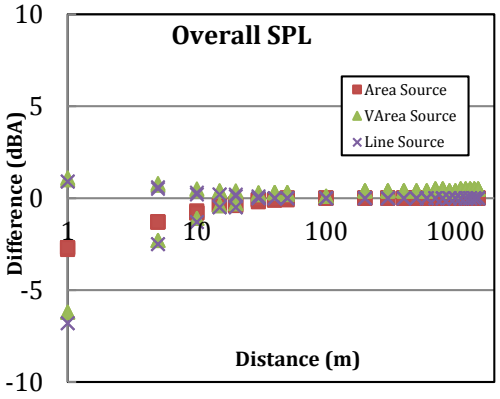


Figure 1: Predicted SPL Difference of Single Source - Point Source Compared with Extended Source (20m Length)

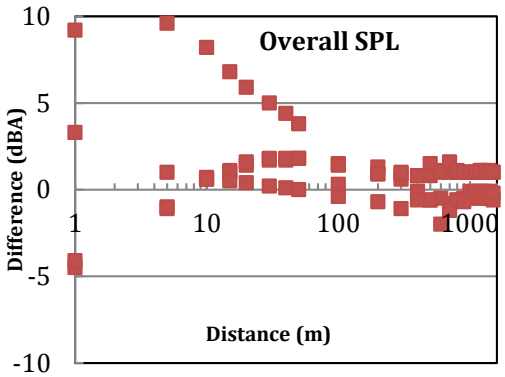


Figure 2: Predicted SPL Difference of Single Compressor Unit - Point Source Model Compared with Extended Source Model

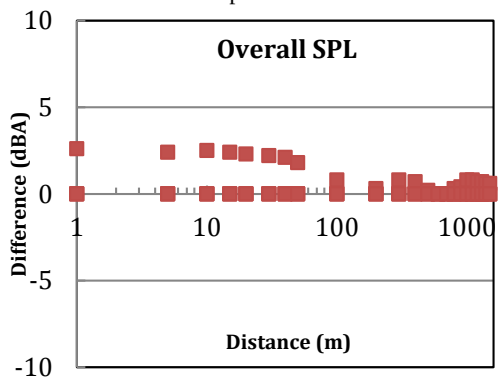


Figure 3: Predicted SPL Difference of Simulated Plant - Point Source Model Compared with Extended Source Model

Big differences in the near field is because the sound energy is fluctuating and propagation is far from uniform, especially within the wavelength of the lower frequencies. The differences are even bigger for the single compressor unit, which are also caused by the building shadow effects in the near field. In the plant scenario, less fluctuation results because the receiver locations are based on the plant boundary, which is approximated to the far field.

6 Conclusion

To create simplified point sources, the locations of the acoustical center of extended sources such as line and area sources and accurate sound power input are very important for the simplification of an accurate prediction.

Even though with the enhanced computer power of today's systems, the extended noise source can be divided automatically based on the distance and frequency consideration, which may result in a more accurate prediction in the near field. However, the simplification of point sources will make the model building more efficient and result in less calculation time. This is especially useful when the details of the noise sources are not available and near field prediction is not necessary.

References

[1] ISO 9613-2:1996, *Acoustics – Attenuation of sound during propagation outdoors.* International Organization for Standardization

[2] DataKustik GmbH. (2013). Cadna Manual

RUMBLE STRIP DESIGN CHARACTERISITICS ON WAYSIDE NOISE

Frank Babic*, Alfredo Rodrigues †
Amec Foster Wheeler, 160 Traders Blvd East, Suite 110 Mississauga Ontario CANADA L4Z 3K7

1 Introduction

Rumble strips are a traffic safety measure designed to help drivers that are either distracted, sleepy or not paying attention, and as a result unintentionally stray off their lane. Rumble strips do this by generating noise inside the vehicle cabin. Noise is generated when the vehicles pass over irregularities created in the rolling surface of the road. The noise generated inside the cabin of the vehicle is intentional; in the sense that the higher noise levels generated from the rumble strips are desirable to be effective in alerting drivers. However, the use of rumble strips also generates increased noise level in the outside environment (wayside) when compared to pass-by of vehicles on a section of road without the rumble strips.

This paper reviews milled rumble strip design characteristics, with respect to exterior noise levels for automobiles. Overall noise levels, frequency spectra and tonal characteristics are considered with respect to rumble strip design.

2 Strip Design Characteristics

Although the optimum design dimensions for milled rumble strips depend on operating conditions, two key dimensions that have been noted in the literature to create the greatest effect on the vehicle alerting sound and vibration inside the vehicle: depth and width, where longitudinal to the road. Although several documents can be found about the design of rumble strips, not many review and take into consideration the wayside noise generated by traffic passing over the rumble strips.

Of the five milled rumble strips designs that were used for testing of exterior noise levels, there are three design characteristics that are the focus of this paper, with respect to their impact on overall noise level, frequency spectra and tonal characteristics:

a) Width of strip
b) Distance between strips
c) Straight versus angled strip

Within construction tolerances, we expect that the rumble strips tested had the same relative depth.

3 Methodology

Milled rumble strips were installed in the section of Highway 58A in Welland, Ontario. All of the rumble strips designs are located at the edge of the shoulder on the north side of the highway.

Four light vehicles were used in the test, and included a Subaru Forrester, Ford C-Max, Chevrolet Volt and Toyota Camry. Vehicles were running on the north shoulder of the highway, in order that measurement locations could be taken in the center of the road. This meant that the testing was conducted without other traffic on the highway to ensure that the results would not be affected by extraneous traffic.

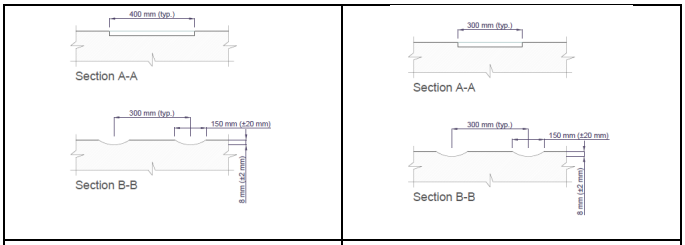
Vehicle test runs were conducted to achieve 80 km/h over the rumble strips. Exterior noise measurements were done following the guidelines of AASHTO standard TP 98-13 [2], which defines that measurements are to be taken:

- 7.5 m away from the center of the travel lane; and
- 1.5 m above the roadway plane.

We note that the number of vehicle test runs do not comply with the minimum test run requirements in the standard, and as such the testing is not in full compliance with the AASHTO standard for data analysis.

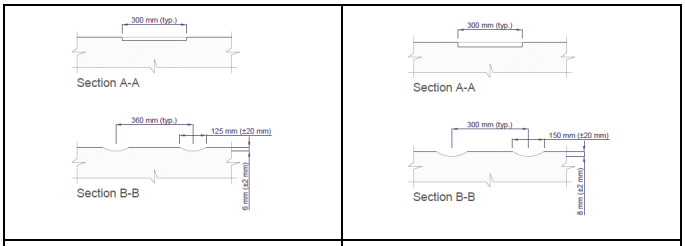
4 Results

The following rumble strip groups are defined for the purpose of assessing rumble strip design characteristics. One rumble strip design was used as a control, and others with varying characteristics were compared to its results.



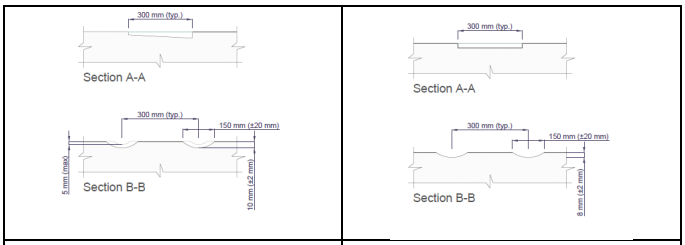
Rumble Strip #2	Rumble Strip #3 (Control)

Figure 1 - Width of Strip



Rumble Strip #1	Rumble Strip #3 (Control)

Figure 2 - Distance between Strips



Rumble Strip #4	Rumble Strip #3 (Control)

Figure 3 - Straight versus Angled Strip

The overall noise level change of the strip types compared to the control strip are provided in Table 1.

Table 1 – Rumble Strip Change in Overall Sound Level

Group	Control Strip dBA Level	Strip Type dBA Level	Difference[Δ in dB]
Width of Strip	86.4	86.9	0.5
Distance Between Strips	86.4	91.9	5.5
Straight versus Angled Strip	86.4	84.0	-2.4

Figure 4 shows the frequency spectra of the control strip compared to the strips with varying characteristics.

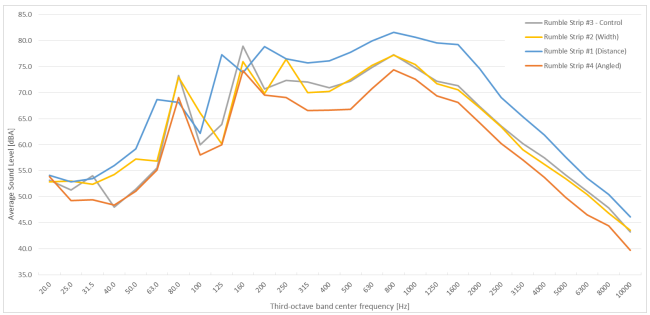


Figure 4 - Rumble Strip Change in Frequency Spectra

Table 2 compiles the tonal characteristics[1] of the control strip, and the changes in tonality with respect to the rumble strip design.

Table 2 – Rumble Strip Change in Tonal Quality

Group	Control Tonal Components [Hz]	Strip Type Tonal Components [Hz]
Width of Strip	80, 160	80, 160, 250
Distance Between Strips	80, 160	none
Straight versus Angled Strip	80, 160	80

5 Conclusion

Tests on milled rumble strips were conducted on Highway 58A in Welland, Ontario. The tests were conducted following as basis the AASHTO standard TP 98-13. The results presented in this paper were grouped in order to assess the change in wayside noise levels due to changes in rumble strip design.

The results shows that in terms of overall sound pressure level the following characteristics seem to affect the resultant wayside sound level in order of decreasing importance:

- An increase of the distance between strips seems to have a direct effect on the sound level, with an increase of 5.5 dB when distance between rumble strips is changed from 300 mm to 360 mm;
- An increase of the width of the rumble strip has marginal to no effect on the noise level, with an increase of only 0.5 dB; and
- An angled[2] design seems to have a reduction of the sound level. It cannot be determined if it is the angled design that makes it quieter, or if this is a consequence of having the smaller longitudinal width of the rumble strips facing the measurement location.

For the tonal quality of noise, in terms of the presence of tonal components, it seems that an increase of the distance between rumble strips reduces tonal components, while an increase of the width will have the opposite effect.

The angled strip[2] design has a small mitigating effect on the presence of tonal components. As stated for the sound levels assessment above, it cannot be determined from this investigation whether this is a consequence of the angled design, or the fact the longitudinal width facing the measurement location is smaller.

References

[1] Amec Foster Wheeler, “Rumble Strips Noise Study”, October 2014

[2] AASHTO standard TP 98-13 - Standard Method of Test for Determining the Influence of Road Surfaces on Vehicle Noise Using the Statistical Isolated Pass-By (SIP)

[3] Federal Highway Administration, "Technical Advisory - Center Line Rumble Strips, T 5040.40, Revion 1," FHWA, November 7, 2011.

[4] NCHRP 641 – National Cooperative highway Research Program Report #641, “Guidance for the Design and Application of Shoulder and Centerline Rumble Strips”.

[1] A tonal components was defined to exist when the sound pressure level for a third-octave band exceeded simultaneously the level of the adjacent bands by more than 5 dB

[2] It should be noted that in the case of the angled design the sound levels where only evaluated on one side – the side with a smaller longitudinal width.

TRAFFIC NOISE PROPAGATION THROUGH FOUR VANCOUVER LANEWAYS

Rosa Lin* and Maureen Connelly
Centre for Architectural Ecology, British Columbia Institute of Technology, Burnaby, BC, Canada. *rlin35@bcit.ca

1 INTRODUCTION

The changing built environment of Vancouver's residential laneways towards increasing density, such as by infill with laneway housing, has acoustical consequences for laneway home residents. The small dimensions of laneways and increase of reflective surfaces within them contribute to the urban canyon effect—limited geometrical spreading and reduced sound decay inside a laneway relative to open terrain [1, 2]. A study showed that Vancouver laneways near noisy roads exceeded outdoor residential noise limits by 10dB [2]. For laneways exposed to excessive noise, the canyon effect exacerbates noise conditions by keeping noise levels high further into the laneway.

The physical form and environmental characteristics of residential laneways were investigated in terms of reflectivity and absorptivity, which affect sound attenuation. Factors taken into consideration included degree of vegetation, surface textures, ratio of reflective building surfaces, laneway dimensions, and ground conditions. It was hypothesized that the more reflective a laneway is—typically achieved by laneway pavement and larger buildings along the lane—the more it would behave like an urban canyon. An empirical test method and two modelling methods were used to test the hypothesis.

2 METHOD

Four case study Vancouver laneways were investigated for physical characteristics that affected sound propagation (see Figure 1). These were quantified in terms of overall reflectivity. A was a new development laneway estimated to have 90% reflective material surfaces. B was a high-density mixed commercial laneway estimated to have 80% reflective material surfaces. C was an older regular laneway estimated to have 60% reflective material surfaces, and D was a highly-vegetated, low density and unpaved laneway estimated to have 35% reflective material surfaces.

Figure 1: Four case study laneways in Vancouver, B.C.

Sound propagation was measured in SPL (dBA) at increments of 10m along each case study laneway, in the direction away from a stationary point source emitting pink noise from one end of the laneway. Attenuation describes the decrease in measured SPL from the source level. Each laneway's attenuation trend described its unique capacity to absorb sound. The laneway sound decay trends were compared to the standard traffic noise model for theoretical hemispherical sound decay over an open, unobstructed area, with adjustments for terrain type [2, 3, 4]. Figure 1 shows the theoretical curves as reference trends for point source sound propagation over open area.

Odeon software and CMHC were used to model sound propagation through each case study laneway. Point source propagation data modelled in Odeon was compared to empirical data. CMHC predicted traffic noise levels based on line source sound propagation theory [4]. In this application, the laneways were modelled using one perpendicular road for line source. The modelling tools were assessed for laneway application.

3 RESULTS

3.1 Empirical results of point source decay

Figure 1 shows that SPL attenuation through at least the first 95m of all laneways was lower than that in an open area, differing by as much as 15±1dBA for reflective A and B. However, beyond 95m, absorptive D exhibited higher attenuation than open area by up to 6±1dBA. A and B were 13.8±1dBA louder than D at 75m (about six houses down the block), and B was 19.3±1dBA louder than D at 130m. For reference on perceived loudness, 3dB is a noticeable difference, 6dB is a substantial difference, 10dB is double the perceived loudness and 20dB is quadruple the perceived loudness [5].

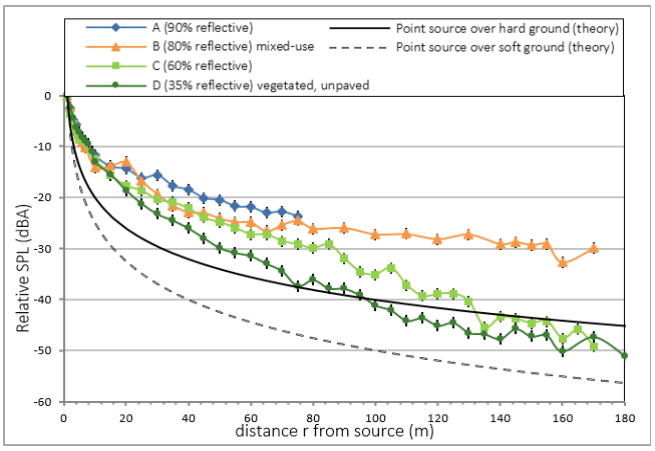


Figure 2: Point source sound attenuation through laneways, empirical data

3.2 Odeon model for point source decay

Point source Odeon model data was calibrated for the near field effect (based on empirical data) and compared to

empirical data. Figure 2 shows that at r > 80m, empirical data better demonstrated the effects of reflections (e.g. laneway B) and absorption (e.g. laneway D) than Odeon data, with a discrepancy reaching 9±1dBA. At 75m, Odeon predicted B to be 7.5±1dBA louder than D where empirical data was 13.8±1dBA louder.

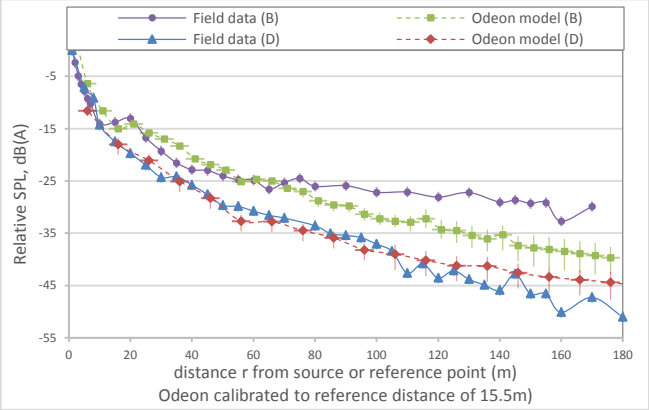


Figure 3: B and D point source attenuation; empirical data vs. Odeon model data

3.3 CMHC model for line source decay

Figure 3 shows CMHC line source attenuation model results. The more reflective cases A, B and C attenuated similarly to one another: roughly 2±1dBA more than CMHC's reference line source decay trend over open hard ground and up to 4±1dBA less than point source decay over open hard ground. Significant discrepancy was predicted between the reflective cases and the absorptive case D. At 75m, A was 9±1dBA louder than D and B and C were 7±1dBA louder. At 93m, A, B and C were 9±1dBA louder than D. D even attenuated more than the reference line source decay trend over soft absorptive ground, by up to 2±1dBA.

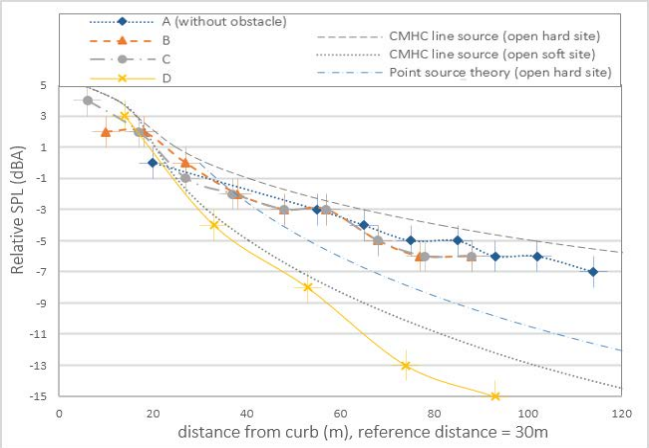


Figure 4: CMHC single line source attenuation results

4 Discussion

Empirical point source propagation results confirmed the hypothesis that the four case study laneways exhibited urban canyon effects. Even the least reflective, most absorptive laneway D maintained notably higher SPL than open field theory for the first 95m. The four case studies demonstrated an inverse correlation between surface reflectivity and sound attenuation through a laneway. All methods showed that high absorptivity could reverse the urban canyon effect, as D began to decay more than open field beyond 95m in empirical data, and beyond 50m in the CMHC model. The most reflective lane was measured to be as much as four times louder than the most absorptive lane.

Odeon adequately modelled the more absorptive cases C and D, but did not satisfactorily predict urban canyon effects in reflective cases A and B. CMHC conservatively estimated the three more reflective laneways to be minimally attenuating and that D was the most effective at noise-attenuation. Both models underestimated the magnitude of SPL discrepancy between B and D at 75m by half of the measured data (7dB vs 14dB).

5 Conclusion

This study confirmed that Vancouver laneways exhibited urban canyon effects, with a positive correlation between laneway surface reflectivity and SPL resulting from sound propagating over distance through a laneway. This is a concern particularly at locations exposed to excessive noise and with increasing laneway construction that increases reflectivity. Vancouver's current laneway housing design guidelines do not specify acoustical requirements [2]. This study shows the need for acoustically designed laneways and laneway housing to provide adequate acoustical protection for residents.

Some acoustical design strategies can be learned from the highly absorptive laneway D, where the canyon effect diminished significantly and eventually reversed with distance. Since the high absorptivity of D was attributed to unpaved dirt road; open yard spaces; shorter, smaller and rougher-surfaced buildings along the laneway; and dense vegetation, these factors can be incorporated as "noise-regulating" design solutions for a quieter laneway living environment. Absorptive and scattering materials such as living walls and green fences, and porous ground cover such as rubber, gravel, dirt or grass, will help attenuate noise through laneways while being space-efficient and beautiful.

Based on the results of this study, both modelling tools required improvement for better prediction of laneway noise levels.

References

[1] P. S. Warren et. al., "Urban bioacoustics: it's not just noise," *Animal Behaviour (71)*, pp. 491-502, 2006.
[2] Lin, R., "Sound Living in Vancouver's Laneway Housing" (Master of Applied Science thesis), BCIT, 2014.
[3] Long, M., "Architectural Acoustics," p91, Elsevier, 2006.
[4] CMHC, "Road and Rail Noise: Effects on Housing," Canadian Mortgage and Housing Corporation, 1981.
[5] Sengpiel, E., "Sound level change and loudness ratio," www.sengpielaudio.com/calculator-levelchange.htm, 2011.

GUIDELINES FOR NEW DEVELOPMENT IN PROXIMITY TO RAILWAY OPERATIONS

Ian Matthew, M.Sc., P.Eng., A.D. Lightstone, Ph.D., P.Eng.
Valcoustics Canada Ltd., 30 Wertheim Court, Unit 25, Richmond Hill, Ontario, L4B 1B9

1 Background

In 2003, the Federation of Canadian Municipalities (FCM) and the Railway Association of Canada (RAC) identified a common desire to reduce proximity issues which arise from the construction of new development in proximity to railway operations. To this end, the FCM and the RAC signed a memorandum of understanding outlining the common goals of the partnership and established a Community-Rail Proximity Initiative with a mandate to develop and implement a strategy to reduce misunderstanding and avoid unnecessary conflicts arising from railway-community proximity.

In 2004, the FCM/RAC Proximity Initiative published a report identifying best practices and guidelines for new development in proximity to railway operations. In 2013, the Proximity Initiative updated that report, publishing the "Guidelines for New Development in Proximity to Railway Operations" ([1], referred to herein as "the Guideline"), the subject of this article.

2 Legislative Framework

The legislative framework regarding the planning and approval of new development in proximity to railway operations has been identified as lacking a consistent approach across the country. In the 2007 review of the Railway Act of Canada, *Recommendation 34* stated that "the Railway Safety act should be amended to require the developer and municipalities to engage in a process of consultation with railway companies prior to any decision respecting land use that may affect railway safety" [2]. However, regulation of the two sides of the equation noted here fall to two different jurisdictions:

- In general, railways and railway operations are regulated at the federal level under the Railway Safety Act [3] and the Canada Transportation Act [4].
- In general, decisions and legislation relating to land use planning fall to provincial jurisdiction, and in most cases, further on to municipal governments as the land use planning authority.

The jurisdictional ambiguity means that a national approach to addressing proximity issues is unlikely, and as such efforts by the Proximity Initiative are aimed at provincial and municipal governments.

2.1 Canada Transportation Act

The Canada Transportation Act [4] regulates transportation carriers under federal jurisdiction. In the context of this paper, the Act contains two specific sections which address Noise and Vibration specifically. Section 95.1 requires that "when constructing or operating a railway, a railway company shall cause only such noise and vibration as is reasonable, taking into account its [level of service] obligations, its operational requirements, and the area in which the construction or operation takes place" [4]. Section 95.2 requires that the Canada Transportation Agency (under the authority provided in the Act) publish guidelines with respect to noise and vibration requirements as well as the dispute resolution process for complaints arising from railway operations.

2.2 Province of Ontario, Land Use Planning Regulations

The Province of Ontario is the only provincial authority across the country which specifically requires that a railway operation be notified of land use planning applications (Official Plan Amendments, Plans of Subdivision, Zoning By-laws, etc.) affecting lands in proximity to their railway operations.

2.3 Municipal Government Regulations

In the absence of provincial legislation across much of the country, municipal governments have begun to create legislation of their own in an attempt to require better communication between prospective land use applicants and the nearby railway operators – all in an attempt to reduce proximity issues. Specific municipalities with legislation include Montreal and Moncton with cities like Edmonton researching possible implementations as well.

3 Guideline Document

3.1 Noise and Vibration Proximity Issues

Various issues arise when new development is located in close proximity to railway operations. These can include safety issues, trespass issues, and storm water management issues, amongst others. However, perhaps the most notable of the issues related to the close relationship of rail and new development are noise and vibration.

Noise and vibration are, at least to some extent, inherent to rail operations. Vibration is generated as wheels roll over steel rail, and rail noise is produced wheel/rail interaction, locomotives, and whistles.

3.2 Standard Mitigation

The Guideline proposes a standard set of mitigation measures which are intended to deal with many noise and vibration scenarios arising from a railway line. See Figure 1. It should be noted that although these mitigation measures are referred to as "standard", the Guideline still recommends that noise and vibration studies be prepared for a proposed development

site (the Guideline advocates the use of the "Railway Noise Measurement and Reporting Methodology" ([5], published by CTA) for the preparation of Noise and Vibration reports. The studies are to confirm that the standard mitigation will be sufficient to achieve the desired noise and vibration guideline levels for the new development while also providing a concrete communication tool for the approval authority (typically the municipality) as well as the railway operator itself. Alternatively, the noise/vibration studies need to determine what mitigation is needed if the standard mitigation is not appropriate.

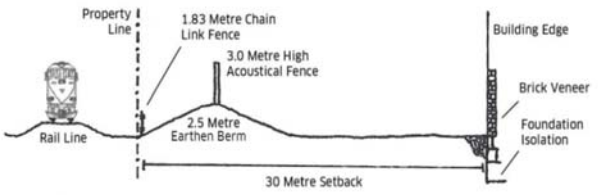


Figure 1 - Standard Mitigation Measures

3.3 Measurement and Prediction

The approaches to quantifying noise and vibration are typically treated differently. In the case of noise, sound exposure levels (L_{eq}) are typically determined by prediction rather than measurement. There are several reasons for this approach:

- Sound exposures are based on future train volumes and train traffic levels that may not be realized on any given day (thus, precluding measurement); and
- Models allow for the prediction of sound exposures at many locations throughout a proposed development that does not yet exist, allowing sound reflections and screening by future buildings as well as for specific determination of mitigation requirements for specific receptor locations.

As per the "Railway Noise Measurement and Reporting Methodology" document, the most commonly used model for the assessment of rail pass-by noise is "Sound from Trains Environmental Analysis Method" (STEAM, [6]). This model is prescribed by the Ontario Ministry of the Environment and Climate Change for use in Ontario, but can be used in any jurisdiction in Canada.

Unlike noise, vibration mitigation requirements are typically determined by measurement, largely because vibration prediction models are not as reliable as those for noise. Since vibration requirements are based on a "worst-case" vibration magnitude (rather than time-averaged as in noise), train traffic on any busy day usually yields sufficient data for determination of mitigation requirements. Measurements of the vibration magnitude should be done at several setback distances (relative to the rail right-of-way) as well as several lateral positions along the length of the rail right-of-way in order to determine:

- The site-specific decay rate with distance; and
- The variation in vibration decay along the interface with the rail right-of-way.

3.4 Noise Guidelines

Appendix AC1.4 of the Guideline provides noise criteria for indoor spaces as well as outdoor living areas as shown in Table 1.

Table 1 – Noise Criteria

Type of Space	Time Period	Sound Level Limit, 1hr L_{eq}, dBA	Outdoor Sound Level Limit, 1hr L_{eq}, dBA
Bedrooms	2300 to 0700 hours	35	50
Living Rooms	0700 to 2300 hours	40	55
Outdoor Living Areas	0700 to 2300 hours	55	N/A

The indoor sound levels limits are used to determine the architectural components required for the structure(s). The outside (façade) sound level limits are used to determine the air conditioning requirements.

It should be noted that the numerical criteria provided above are consistent with the previous Ontario noise guideline, LU-131. The new Ontario noise Guideline, NPC-300, contains the same criteria noted above but also provides daytime/evening (0700 to 2300 hours) criteria for bedrooms and nighttime (2300 to 0700 hours) criteria for living rooms.

3.5 Vibration Guidelines

Appendix AC2.5 of the Guideline provides a vibration limit of 0.14 mm/s RMS (1-second average, between 4 and 200 Hz) at a receptor location. If the rail vibration magnitude is above this limit, the Guideline states that appropriate vibration isolation measures are recommended at the receptor location (dwelling) such that the vibration magnitude inside the dwelling at the first floor (and above) is below the guideline limit.

4 Conclusions

If implemented across the country, the "Guidelines for New Development in Proximity to Railway Operations" would be expected to provide a unified approach for addressing issues related to new development in proximity to railway operations.

References

[1] Railway Association of Canada, The, & Federation of Canadian Municipalities, The. Guidelines for New Development in Proximity to Railway Operations, 2013.

[2] Railway Safety Act Review Secretariat. Stronger ties: A shared commitment to railway safety, 2007.

[3] Railway Safety Act, R.S.C. 1985, c. 32 (4th Supp).

[4] Canada Transportation Act, S.C. 1996, c. 10.

[5] Canadian Transportation Agency. Railway noise measurement and reporting methodology, August 2011.

[6] Ontario Ministry of the Environment (MOE). Sound from Trains Environmental Analysis Method (STEAM), 1990.

SELF-ADJUSTING BACKUP ALARMS IN NOISY WORKPLACES

Hugues NÉLISSE[1], Jérôme BOUTIN[1], Christian GIGUÈRE[2], Chantal LAROCHE[2], Véronique VAILLANCOURT[2]
[1] Institut de Recherche Robert Sauvé en Santé et Sécurité du Travail, Montréal, Canada
[2] Audiology & SLP Program, University of Ottawa, Ottawa, Ontario K1H 8M5, Canada

1 Introduction

Self-adjusting backup alarms are often used to reduce annoyance in residents living near industrial and construction sites with significant traffic from heavy vehicles. These devices automatically adjust, within a certain operational range, the level of the emitted warning signal to exceed that of the surrounding background noise, which is sampled by an integrated sensor. However, apart from the work of McDaniel *et al.*[1] and the test requirements set out in ISO 9533[2], little information is available in the literature and in the form of technical manufacturer's data to explain their functioning and to document their effectiveness in real workplaces. Questions therefore arise as many prevention specialists, local authorities and residents regularly claim that self-adjusting alarms fail to minimize noise annoyance. This paper presents research aimed at describing the inner-workings of self-adjusting backup alarms, more specifically how they adjust their output level based on the surrounding background noise in typical noisy workplace environments. Using microphones, a recorder and a current clamp, the backup alarm level can be estimated, even amidst background noise, when the device is installed on a heavy vehicle operating in typical noisy workplace conditions. Laboratory and field data are presented for two backup alarms, a tonal alarm and a broadband alarm.

2 Methodology

2.1 Laboratory measurements

To understand how a self-adjusting alarm operates, the setup shown in Figure 1(a) was used during laboratory measurements carried out in an anechoic room. Two microphones (one near the alarm and the other a meter away) were used to measure alarm levels, while a loudspeaker placed two meters from the alarm device generated various levels of a white noise and a current clamp fixed to the device's wiring measured its input current. To verify how the alarm levels are set based on noise levels in effect prior to the activation of the device, two time sequences of pulsed background noise were played as depicted in Figure 1(b). In the "Out-of-sync" sequence, the background noise pulses are progressively increased in level but appear only during the "off" (quiet) portions of the alarm, thus allowing the microphones to measure exclusively the sound pressure level of the alarm during the "on" portions since both signals (alarm and background noise) appear at different times. Results obtained with this sequence provide an insight into how alarm levels increase with increasing background noise levels. They also allow defining the relationship between the current flowing through the device and the alarm levels measured acoustically at 1 m, knowledge which will prove useful to obtain equivalent free-field estimates of alarm levels in field trials where continuous background noise sources representing real workplace conditions are encountered. In the "In-sync" sequence, the background noise pulses appear only during the "on" portions of the alarm. Alarm levels should then remain stable (at the minimal value) in this case since the preceding "off" portion was free of noise.

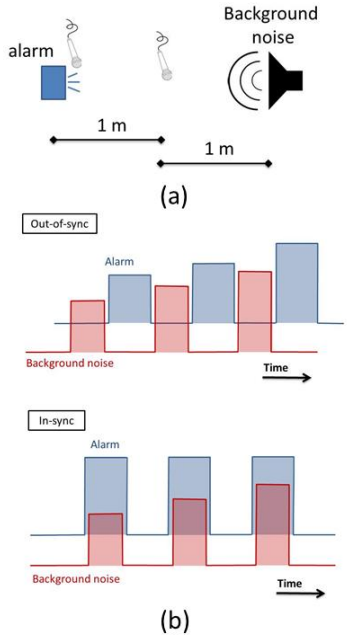


Figure 1: (a) Alarm device, microphones and loudspeaker setup; (b) Time sequences for presenting the background noise amidst the "on" and "off" portions of the alarm.

2.2 Field measurements

During field trials, a microphone placed on top of the alarm device, which was mounted on a heavy vehicle, measured the surrounding sound field. Results at this measurement point served to characterize the background noise levels used by the device to set alarm levels. A second microphone placed directly at the rear of the vehicle (exterior side and centered horizontally) was used to provide more representative readings of sound pressure levels within the danger zone behind the vehicle, levels which could differ significantly from those obtained at the other microphone position. Measurements of current were also performed. Both microphones and the current clip were connected to a recorder and continuous time recordings were performed using a 24 kHz sampling rate. Recordings were thereafter analyzed using an in-house Matlab™ code to extract and plot measured sound pressure levels (SPL) as a function of time, and to obtain estimates of equivalent free-field alarm levels based on readings from the current clip.

3 Results and discussion

3.1 Laboratory data

For both alarms, a linear relationship was found between sound pressure levels (SPL) measured at 1 meter during the "on" portions and measurements of current, results which support the use of current measurements to make free-field estimates of alarm SPLs. Such estimates are shown for the two alarms in Figures 2 and 3, along with the background noise level measured in the "Out-of-sync" sequence. The vertical lines represent alarm SPL estimates during an "on" portion of approximately 400 ms (in red) and background noise levels measured 300 ms prior to each "on" portion (in blue), while the two horizontal lines (in black) define the labeled operating range of the alarm device.

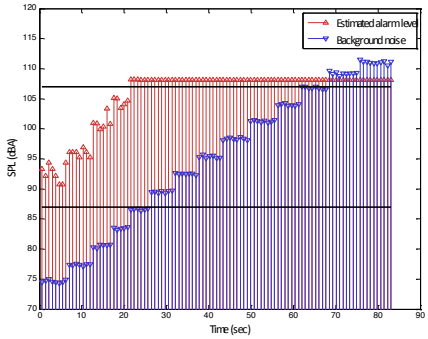


Figure 2: Alarm SPL estimates and background noise SPLs as a function of time for the broadband alarm ("Out-of-sync" sequence).

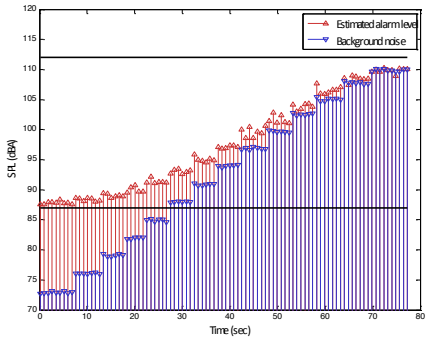


Figure 3: Alarm SPL estimates and background noise SPLs as a function of time for the tonal alarm ("Out-of-sync" sequence).

As can be seen, the levels of the broadband and tonal alarms exceed those of the background noise; however, the resulting signal-to-noise ratio or SNR (alarm level minus background noise level) appears to depend on background noise levels and alarm type. Compared to the tonal alarm, the broadband alarm produces higher SNR values and reaches its maximum output level more rapidly with increasing background noise levels.

For both alarms, current measurements (and associated alarm SPL estimates) remained constant at the lower level of the alarm operating range during the entire "In-sync" sequence, thereby confirming that the two self-adjusting alarms are only sensitive to background noise in the "off" portions of the alarms.

3.2 Field data

SPL estimates and background noise SPLs over a short time period are shown in Figure 4 for a self-adjusting broadband alarm device mounted on a heavy truck. While in this example alarms level are clearly well above the background noise, it should be noted that the resulting SNR fluctuates significantly over a few seconds, even when the background noise appears to be relatively stable. Similar results were obtained with a tonal alarm.

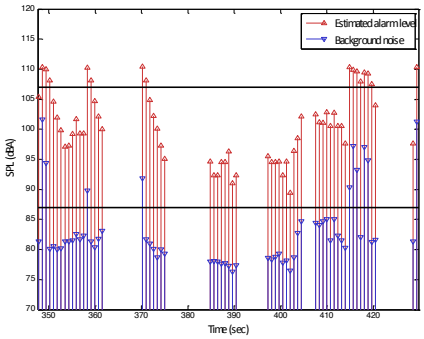


Figure 4: Example of alarm SPL estimates and background noise SPLs as a function of time for a broadband alarm device mounted on a heavy truck under normal operating conditions.

Overall, results of the field trials showed significant SNR variations and alarm levels far exceeding those of the background noise. This suggests that, in the presence of low to moderate background noise, the alarm produces levels well above the minimal SNR specified in [2]. Further data analysis is currently underway to get a better understanding of these sharp fluctuations and high SNR values since both factors may impact worker safety and noise annoyance arising from the alarms, and therefore need to be more thoroughly investigated and addressed.

4 Conclusion

A simple methodology, developed and tested in a laboratory, was used to estimate the levels of self-adjusting backup alarms in the field when mounted on heavy vehicles operating in typical noisy workplace conditions. Preliminary results show that alarm levels can fluctuate significantly even when the background noise is relatively low and stable. This impact of these fluctuations on the safety of workers and pedestrians, as well as the noise annoyance reported by nearby residents, must be further analyzed.

Acknowledgments

The authors would like to thank the Institut de Recherche Robert-Sauvé en Santé et Sécurité du Travail (IRSST) for its financial support.

References

[1] M.S. McDaniel, D.C. Copley & D.E. Zimmermann, "Application of a self-adjusting audible warning device as a backup alarm for mobile earthmoving equipment", *Proc.eedings of the 2005 SAE Commercial Vehicle Engineering Conference*, Rosement, IL, USA: (2005).

[2] ISO 9533, "Earth-moving machinery--Machine-mounted forward and reverse audible warning alarm--Sound test method", *International Standards Organization*: (2010).

INVESTIGATION OF GROUND AND MAINTENANCE CREW NOISE EXPOSURE FOR THE ROYAL CANADIAN AIR FORCE CH-147F HELICOPTER

Andrew Price, Sebastian Ghinet, Viresh Wickramasinghe, Yong Chen, Anant Grewal
NRC Aerospace, Flight Research Laboratory, Ottawa, Canada

1 Introduction

Helicopter ground and maintenance personnel, in addition to helicopter aircrew, are exposed to high noise levels while working inside and around aircraft. The noise environment may include tonal rotor noise, high frequency engine, transmission and hydraulic systems noise as well as broadband noise. Prolonged exposure may lead to hearing loss if the designated hearing protection is inadequate or improperly fitted. A thorough interior and exterior noise level measurement campaign was conducted on the Royal Canadian Air Force's CH-147F helicopter. A number of representative operational states for maintenance and tarmac conditions were investigated and referenced to the Canadian Aviation Occupational Health and Safety Regulations. The results were analyzed to determine the acoustic performance of ground maintenance crew hearing protection equipment.

2 Flight Test Procedures and Equipment

The objective of the flight test was to measure interior cockpit and cabin noise levels, exterior tarmac noise levels and aircrew whole body vibration exposure during typical aircraft manoeuvres. The CH-147F Chinook helicopter was outfitted with nine interior microphones, 18 interior accelerometers and eight exterior microphones were held on the tarmac. The aircraft was operated in open door and closed door configurations. The ground measurements took place in the open door configuration.

ISO 5129 [1] and MIL-STD-1294A [2] were the primary standards followed for the measurement procedure. The two standard procedures outlined requirements for microphone locations, measurement conditions and testing procedures. Details of the flight test plan and microphone configuration may be reviewed in Ref [3].

2.1 Instrumentation

The microphones were mounted at a standing height as per ISO 5129 on fixed hard-mount stands and fitted with windscreens. Each stand and cable combination was managed by a ground crew member whilst in the vicinity of the running aircraft. The approximate locations of the sensors are shown in Figure 1 and Figure 2. The microphone locations are labelled in Figure 1. The eight locations on the starboard side of the helicopter are the exterior ground measurement locations. The remaining locations are interior measurements which will not be discussed in this paper.

The helicopter exterior noise data was collected using an LMS SCADAS front end, eight ICP Piezotronics 378B02 microphones and a Dell precision M6500 laptop workstation. Post processing was completed in LMS Test.Lab.

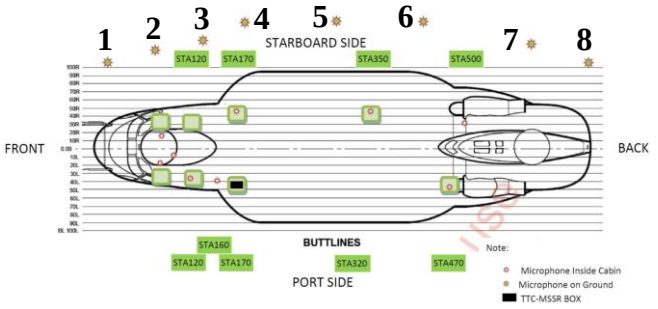


Figure 1: CH-147F Measurement Locations [3]

Figure 2: Exterior Noise Measurement

3 Results

A number of measurements were completed during the aircraft's standard start up procedure. It was found that the microphones situated near the stern of the aircraft experienced higher sound pressure levels (SPL).

3.1 Measurement Conditions

The measurement conditions of interest are shown in Table 1.

Table 1: Measurement Conditions

ID	Description
1	Exterior Background Noise
2	Auxiliary Power Unit (APU) ON, Avionics ON
3	APU ON, Engine Setting: Idle
4	APU ON, Engine Setting: Flight
5	APU OFF, Engine Setting: Flight

3.2 Aircraft Exterior Noise Levels

Condition 4 (APU ON, Engine Setting: Flight) exhibited the highest noise levels. The 3rd octave band results for Condition 4 have been shown in Figure 3.

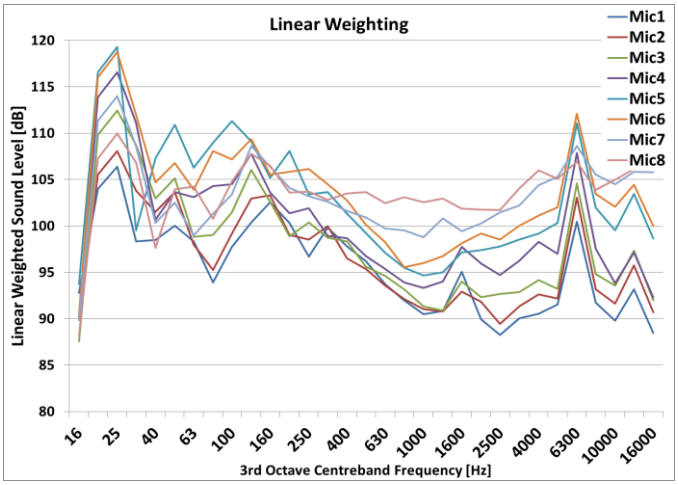


Figure 3: Condition 4, Linear Weighting

The Mic5 and Mic6 positions were dominantly louder in the lower frequencies of the measurement. These locations were underneath the tips of both the front and rear rotors whose nominal rotation speed was 225 RPM. The Mic7 and Mic8 experienced greater broadband high frequency noise as they were nearest the engine outlet.

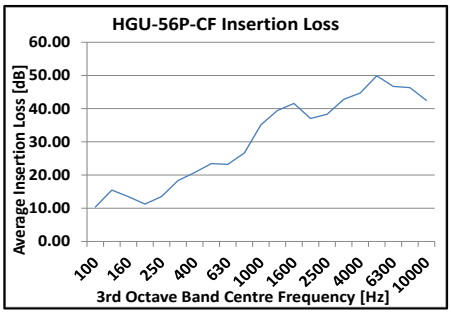


Figure 4: HGU-56P-CF Insertion Loss [5]

The performance of the HGU-56P-CF helmet was tested in the NRC acoustic chamber facilities and the mean insertion loss has been shown in Figure 1. The helmet exhibited better attenuation at higher frequencies.

The results from Figure 3, with A-weighting, helmet attenuation and windscreen correction factors applied, are shown in Figure 5. The helmet insertion loss attenuates the majority of the high frequency engine noise. The majority of the remaining noise is associated with the low frequency main rotor harmonics.

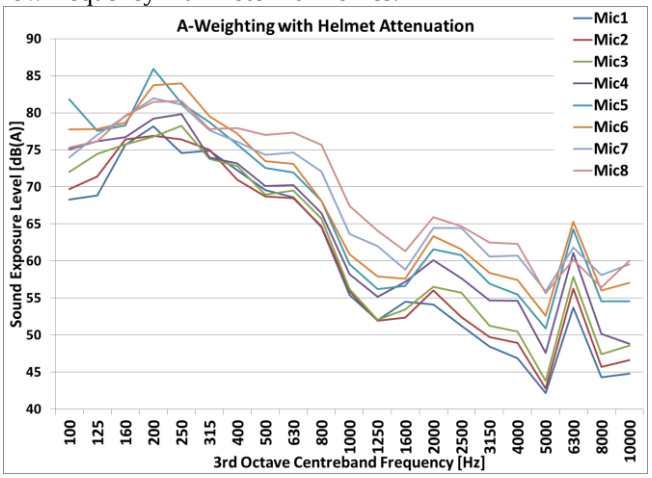


Figure 5: Condition 4, A-Weighting and Helmet Attenuation

The overall SPL for Condition 4 with linear weighting, A-weighting and the HGU-56P-CF helmet attenuation at the eight microphone positions surrounding the helicopter have been shown in Table 2.

Table 2: Condition 4 Overall SPL

Mic #	L-Weighted [dB]	A-Weighted [dB(A)]	Helmet [dB(A)]
1	112.97	105.73	83.48
2	114.82	106.96	84.08
3	117.59	108.18	84.61
4	120.72	111.29	86.33
5	123.37	113.50	89.81
6	123.14	115.14	88.86
7	120.32	115.07	87.75
8	119.65	115.60	88.63

The highest overall SPL (after applying the attenuation of the helmet) was 89.81 dB(A) at the Mic5 position during Condition 4. In accordance with the Canadian Aviation Occupational Health and Safety Regulations (within the Canada Labour Code) a worker may be exposed to 90 dB(A) for a maximum of four hours [4].

Without helmet attenuation the highest overall SPL was 115.60 dB(A). At 116 dB(A) a worker reaches their maximum exposure limit in 36 seconds [4]. Therefore, a correctly fitted helmet is vital under these circumstances.

4 Conclusions

The HGU-56P-CF helmet satisfies the Canadian Labour Code for work around the Canadian Forces CH-147F aircraft with a few caveats:

1. A more detailed assessment of a standard work day's activities (and their associated noise) may be required to ensure safe working limitations.
2. The HGU-56P-CF is currently providing attenuation on the order of 20 dB. Improper fittings or adjustment and removal due to discomfort will greatly impact a worker's exposure level and safety.

Considering the high level noise environment, active noise cancellation may be added as a simple solution to improve the HGU-56P-CF helmet attenuation in the low frequency range as the main rotor harmonic noise is static and tonal within the low frequency regime.

Acknowledgments

The authors would like to acknowledge the support of the DTAES Human Factors Engineering and 450 THS squadron staff without whom this testing could not have been accomplished.

References

[1] ISO, "Acoustics - Measurement of Sound Pressure Levels in the Interior of Aircraft During Flight," Organization for Standardization, Geneva, Switzerland, 2001.

[2] "Acoustic Noise Limits in Helicopters," Department of Defence, Washington, 1985.

[3] Flight Research Laboratory, "Flight Measurement Procedure to Evaluate Cabin Noise and Vibration Environment of Royal Canadian Air Force CH-147F Helicopter," National Research Council, Ottawa, Ontario, 2014.

[4] "Aviation Occupational Health and Safety Regulations," Government of Canada, Canada Labour Code, 2014.

[5] S. Ghinet, "Acoustic Test Evaluation of DND Pilot Helmets," Flight Research Laboratories, Aeroacoustics Lab NRC, Ottawa, 2014.

THEORETICAL ESTIMATES OF GROUNDBORNE RAILWAY VIBRATION

Todd Busch, M.A.Sc., P.Eng., P.E., INCE Bd. Cert. [1], Michael Wesolowsky, Ph.D., P.Eng. [2] and John Swallow, M.A.Sc., P.Eng., LEED AP [3]
Swallow Acoustic Consultants Ltd, 366 Revus Avenue, Unit 23, Mississauga, Ontario, Canada L5G4S5
[1] tbusch@swallowacoustic.ca, [2] mwesolowsky@swallowacoustic.ca, [3] jswallow@swallowacoustic.ca

1 Introduction

The United States (U.S.) Federal Rail and Transit Administrations (FTA/FRA) publish guidelines [1, 2] for railway noise and vibration purposes showing generic curves for estimating vibration from various transportation modes with distance. The data that were applied by the U.S. FTA/FRA towards creating these generic curves are not readily available. This paper compares the generic FTA curves to theoretical estimates based upon independent contributions from three parametric components of groundborne vibration: surface waves, longitudinal/pressure waves, and shear waves. Utilizing a theoretical equation that is comprised of these three components, it has been found that it is not possible to recreate the generic curves that are promulgated by the U.S. FTA; a prior review suggests that the U.S. FTA curves greatly overpredict groundborne vibration at distances beyond about 30 m [3]. As such, this paper presents explorations of a new theoretical equation that is comprised of surface, spherical, and shear waves, which is thought to provide more accurate estimates of groundborne vibration from railways as a function of distance.

2 Literature Review

The U.S. FTA guideline [1, pp.10-3 – 10-4] has the following to say about the generic curves, "The curves in [Figure 1] were developed from many measurements of ground-borne vibration. Experience with ground-borne vibration data is that, for any specific type of transit mode, a significant variation in vibration levels under apparently similar conditions is not uncommon. The curves in [Figure 1] represent the upper range of the measurement data from well-maintained systems. Although actual levels fluctuate widely, it is rare that ground-borne vibration will exceed the curves in [Figure 1] by more than one or two decibels unless there are extenuating circumstances, such as wheel- or running-surface defects." Further adjustments can be made for speed, wheel and rail type and condition, type of track support system, type of building foundation, and number of floors [within a building] above basement level. The interpretation of these generic curves for the purposes of "General Assessment" is such that projected vibrations are compared to a vibration-velocity-impact criterion; for example, 65 VdB re: 1 microinch/sec is the criterion for vibration-sensitive instrumentation, such as electron microscopes. When projected vibrations are: (1) below the criterion, vibration impact is unlikely; (2) 0 to 5 VdB above the criterion, site-specific "Detailed Analysis" is recommended; and (3) 5 VdB or more above the criterion, "Detailed Analysis" is required for environmental study that complies with U.S. FTA regulation. The generic curves within U.S. FRA regulation [2] are specifically for trains traveling at speeds of 150 mph and are not subject to analysis within the scope of this paper.

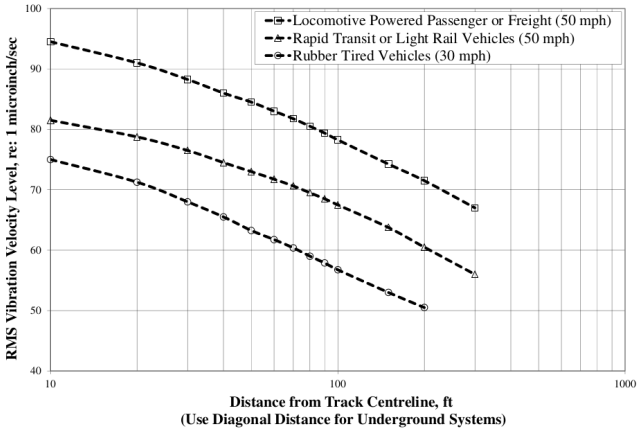


Figure 1: U.S. FTA Generic Vibration Curves .

Richart, Hall and Woods [4, pp. 88–92] note that Rayleigh/surface waves, spherical/pressure waves, and shear waves display different behaviors at the horizontal surface of an ideal, elastic half-space than within the elastic media itself. For a vertical force loading due to a small circular footing vibrating as a sinusoid at the elastic half-space boundary, Rayleigh/surface waves decay along the surface as $r^{-0.5}$; with longitudinal/pressure and shear waves as r^{-2} with distance. Below the horizontal surface, Rayleigh/surface waves diminish in amplitude exponentially; longitudinal/spherical pressure and shear waves as r^{-1} with distance. Wave velocities are such the longitudinal/pressure wave arrives first, followed by the shear wave, and Rayleigh/surface wave. Miller and Pursey [5] reported the percentage of energy that is distributed amongst the three wave types for displacement waves from a circular footing on a homogeneous, isotropic, elastic half-space, as follows: Rayleigh/surface 67%; shear 26%; and longitudinal/pressure 7%. Klein and Rainer [6] explore wave propagation within elastic, homogeneous, isotropic media, and provides an Attenuation Formula that varies based upon: (1) geometry of the vibration source (point versus line); (2) type of excitation (stationary versus impulsive); and (3) predominant type of wave (Rayleigh on the surface, longitudinal/pressure waves at depth (i.e., body waves)). Exponents are described that are specific to each of the possible variations as summarized in Table 1.

Source	Surface Wave		Body Wave	
Geometry	Stationary	Impulsive	Stationary	Impulsive
Point	0.5	1.0	1.0	1.5
Line	0.0	0.5	0.5	1.0

Table 1: Exponents for use within Attenuation Formula

3 Results

3.1 Stationary Line Source

Figure 2 shows the results of curve fitting using an exponent of 0.0 for a stationary line source and 0.5 for a stationary body source. Stationary within the context of these results refers to the stability of vibration emissions over time. The variable is the distance-dependent propagation loss due to material properties. As shown, it is possible to match the closest and furthest points of the calculated curves relative to the U.S. FTA curve, but at the expense of greater overprediction at intermediate distances. As such, this does suggest that the U.S. FTA curves, based on measurements, are indicative of vibration propagation that is neither a line geometry nor is stationary in time.

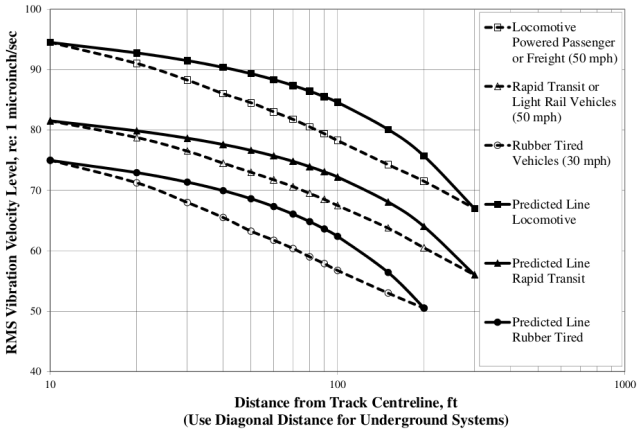


Figure 2: U.S. FTA Generic Vibration Curves versus Stationary Line Source Predictions .

3.2 Impulsive Line Source / Stationary Point Source

Figure 3 shows the results of curve fitting using an exponent of 0.5 for the Rayleigh/surface wave type and 1.0 for the body wave type. The only variable in question is the distance-dependent propagation losses due to material properties. As can be seen, it is possible to match of the calculated curves relative to the U.S. FTA curve at both the closest point at a distance of 100 ft (30 m), but at the expense of some error at other distances of interest. The improved match does suggest that the transportation modes of interest are generating vibrations in more complicated ways in terms of behaving as an impulsive line source and/or stationary point source.

4 Discussion

The U.S. FTA generic vibration curves are to be applied for situations involving both surface transportation and underground sources. As such, the combination of these two data sets to develop a single set of generic curves may be a factor limiting the prospects for matching measurement data to the parametric equations developed for this paper; the differences from within Figure 3 being 3.8 VdB or less. Comparing new theoretical predictions to the 38 data points used to construct the generic curves, the agreement is within 1.0 VdB or less for 30 of those data points, with the largest discrepancies at the greatest distances.

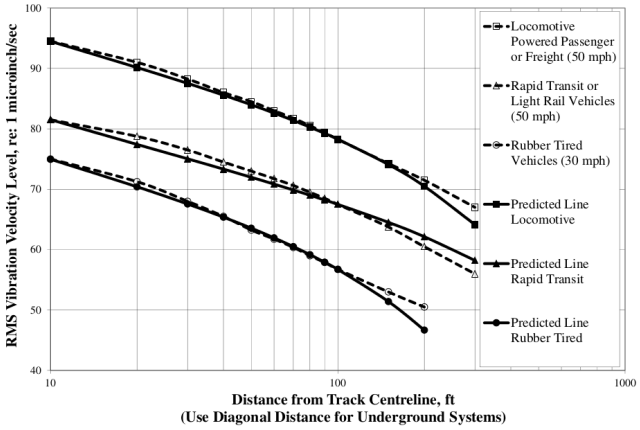


Figure 3: U.S. FTA Generic Vibration Curves versus Impulsive Line Source / Stationary Point Source Predictions .

5 Conclusion

This paper presented explorations of a parametric theoretical equation comprised of surface, spherical, and shear wave types. The results indicate an improved match between U.S. FTA generic curves for estimating vibration velocity level as a function of distance and underlying parameters for each wave type within a new theoretical equation.

Acknowledgments

The authors wish to acknowledge the support of professional colleagues and family.

References

[1] Transit Noise and Vibration Impact Assessment, U.S Department of Transportation, Federal Transit Administration, Office of Planning and Environment, Washington, D.C.,.May 2006.

[2] High-Speed Ground Transportation Noise and Vibration Impact Assessment, U.S Department of Transportation, Federal Rail Administration, Washington, D.C.,.September 2012.

[3] Todd Busch, Michael Gendreau and Hal Amick, "*Vibration and Noise Criteria used to Evaluate Environmental Impacts of Transportation Projects on Sensitive Facilities*", Proc. SPIE 5933, Buildings for Nanoscale Research and Beyond, 593306, San Diego, August 29, 2005.

[4] F.E. Richart Jr., R.D. Woods, and J.R. Hall Jr., Vibrations of Soils and Foundations, Prentice-Hall of Canada Ltd, Toronto, 1970.

[5] G.F. Miller and H. Pursey, " *On the Partition of Energy between Elastic Waves in a Semi-Infinite Solid* ", Proc. Royal Society, v.233, pp. 55–69, 1955.

[6] Gunter K. Klein and Johan H. Rainer, Chapter 6: " *Wave Propagation*", Vibration Problems in Structures: Practical Guidelines, Ed. Hugo Bachman, Birkhauser Verlag, Boston, 1995.

ACOUSTIC EMISSIONS DUE TO LEAKAGE IN PIPELINES

Vishash K. Sharma , R. Arief Budiman
Department of Mechanical and Manufacturing Engineering
University of Calgary, 2500 University Dr. NW, Calgary, AB, T2N 1N4

1 Introduction

Using acoustic waves in non-destructive testing of pipelines is a growing industry. The precision required necessitates the consideration of radiation loading. This paper outlines a mathematical analysis to obtain the acoustic emissions emanating from leaky pipeline by simulating the propagation of a sound wave in a fluid-filled, steel pipe. This work allows for three-dimensions distribution of the pressure and sound pressure level along the axial and circumferential directions which can facilitate a novel and very precise approach in detecting the pipeline leak using acoustic detection technology.

2 Methodology

2.1 Propagation of sound wave in a pipeline

Sound propagates by transmission of time-dependent pressure fluctuations due to leaky pipeline about the ambient pressure. The transfer of pressure fluctuation from fluid into an elastic medium such as steel pipeline is known as radiation loading. A portion of the acoustic waves will be absorbed and a portion will be reflected. This absorption and reflection must be included in a complete analysis, as must the attenuation of acoustic wave as it propagates through both the fluid and elastic structure. The equation of sound wave in a fluid is derived using continuity equation and Naviers stokes' equation for fluid flow:

$$\frac{\partial \rho}{\partial t} + \rho div \upsilon + \upsilon.grad\rho = 0$$

$$\rho\left[\frac{\partial \upsilon}{\partial t} + (\upsilon.grad)\upsilon\right] = -grad\ p + \eta\Delta\upsilon + \left(\varsigma + \frac{1}{3}\eta\right) grad\ div\ \upsilon$$

where $p,\rho,\rho,\upsilon,\varsigma$ *are pressure, density, velocity and coefficients of viscosity*

The polar co-ordinates (r,ϕ,z) are used considering the cylinderical geometry of pipeline. The sound wave equation for the perturbation of pressure p is given as

$$\nabla^2 p = \kappa\rho\frac{\partial^2 p}{\partial t^2} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}$$

where κ and c are bulk modulus and sound velocity in the fluid in pipeline.

The leakage in pipeline creates perturbation in pressure and hence the acoustic source is developed . To determine the acoustic pressure at any location r,ϕ,z in pipeline due to a simple acoustic source from a leakage in a pipeline at any arbitrary location at r_0,ϕ_0,z_0 in the pipeline wall(z being the pipeline axis) the following Green's function is used, which also satisfies the boundary condition at the pipeline wall given by the relation between the gradient of the acoustic pressure normal to the pipeline wall surface and the specific acoustic admittance β of the pipeline material[2]

$$\frac{\partial p}{\partial n} = ik\beta p \quad where \quad k = \frac{\omega}{c}$$

The Green's function takes the form

$$g_\omega(r,\phi,z \mid r_0,\phi_0,z_0) = \frac{i}{\pi b^2}\sum_{m,n}\frac{\Psi_{m,n}(r,\phi)\Psi_{m,n}(r_0,\phi_0)}{\Lambda_{m,n}k_{m,n,z}}\exp(ik|z-z_0|)$$

where $\Psi_{m,n}(r,\phi)\Psi_{m,n}(r_0,\phi_0)$ *are orthogonal functions*

$$\Lambda_{m,n} = \frac{1}{\varepsilon_m}\left[1 - \frac{m^2 + (\beta kb)^2}{(\pi q_{m,n,r})^2}\right] J_m^2(\pi q_{m,n,r})$$

$\varepsilon_m = 1$ *when* $m = 0$ *and* $\varepsilon_m = 2$ *when* $m > 0$.

$q_{m,n,r}$ *is eigen value for radial wave number for* $(m,n)th$ mode *shape*

$k_{m,n,z}$ *is axial wave number for* $(m,n)th$ mode *shape*

$J_m(\pi q_{m,n,r})$ *Bessel function* m_{th} *order*

b = *inside diameter of the pipeline*

Using the Green's function, the acoustic pressure distribution p can be determined by simulating the acoustic source due to leakage as quadrupole acoustic source[3].

2.2 Absorption and attenuation of acoustic energy

When a sound wave is incident on the pipeline wall, some of the energy is reflected and some is absorbed by the surface. These losses of the sound energy are in addition to the internal energy losses of the sound energy due to viscosity and heat conduction.The coefficient of viscosity is modified as per Kirchoff's equation to account for effects of viscosity and heat conduction which in turn affects the sound velocity in the fluid(specifically for gases).

$$\mu_c = \mu\left\{1 + \left(\gamma^{1/2} - \frac{1}{\gamma^{1/2}}\right)\left(\frac{K}{\mu C_p}\right)^{1/2}\right\}$$

additionally the Korteweg-Lamb correction[1] is applied to calculate the modified sound velocity in the pipeline fluid.

$$c_c / c = \left[1 + \left(\frac{2b\rho c^2}{h\rho_p c_p}\right)\right]^{-1/2}$$

To account for losses due to reflection of acoustic waves from the pipeline wall the absorption coefficient is calculated as if the surface impedance is the function of frequency[2].

$$\alpha = \frac{1}{2\rho cI}\int_0^{2\pi} d\phi \int_0^{\pi} \alpha(\theta)|p|^2 \cos\theta \sin\theta d\theta$$

θ = *Angle of incident*; ϕ = *Polar angle for pipeline geometry*

$$\alpha(\theta) = 1 - |C_r|^2; C_r = \frac{z(\omega)\cos\theta - \rho c}{z(\omega)\cos\theta + \rho c}; z(\omega) = Impedance$$

I = *Incident Power per unit area*

$$= \frac{1}{2\rho c}\int_0^{2\pi} d\phi \int_0^{\pi} |p|^2 \cos\theta \sin\theta d\theta$$

ANSYS was used to calculate the natural frequencies for the continuous simply-supported pipineline which was then incorporated to calculate impedance for the pipeline.

2.3 Acoustic wave transmission thru the pipeline wall

The radiation of sound from an infinite elastic cylindrical shell excited by an internal sound source are anlysed by coupled equations of motion. The Donnell's formulation is used [1] which is based on the assumption that the expression for the change in curvature and the twist of the cylindrical pipeline are the same as those of the flat plate and that the effect of the transverse shearing-stress resultant

on the equilibrium of the forces in the circumferential direction is negligible. The radiated acoustic pressure from the surface of the pipeline is given by [5][1]:

$$p_{radiated} = (p/a)(\rho_f^0/\rho_s)\Omega^2(h/r)^{-1}\sum_{m=0}^{\infty}\varepsilon_m \ldots$$

$$\int_{-\infty}^{\infty}[J_m(k_{m.n.ri}r)/J_m'(k_{m,n,ri}r)k_{m,n,ri}r][H_m(k_{m.n.r0}r)/H_m'(k_{m,n,r0}r)k_{m,n,r0}r]I_{33}..$$

$$\times\exp[i(k_{m.n.z}z)(z/r)]d(k_{m.n.z}r)$$

$$\text{where}: J_m(k_{m.n.ri}r)/J_m'(k_{m,n,ri}r)k_{m,n,ri}r = \frac{z}{\rho ckb} \quad \text{on inside of pipeline}$$

$$H_m(k_{m.n.r0}r)/H_m'(k_{m,n,r0}r)k_{m,n,r0}r = \frac{z}{\rho_f^0 cka} \text{ on outside of pipeline}$$

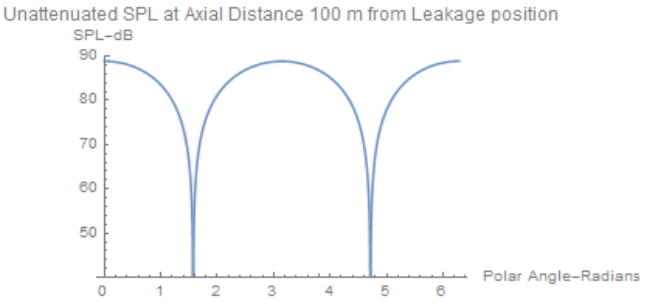
(a is mean radius of the pipeline)
h = pipeline wall thickness;
p = Calculated acoustic pressure after acounting for losses
I_{33} *= coeffiecient calculated from equations of motion from Donnell's formulation for radial displacements*

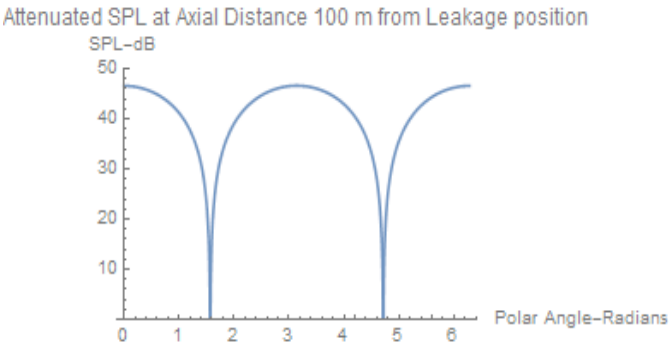
Using the above described methodology, the complete profile of acoustic pressure inside the pipeline and outside the pipeline due to leakage, can be plotted as shown in the following section.

3 Results

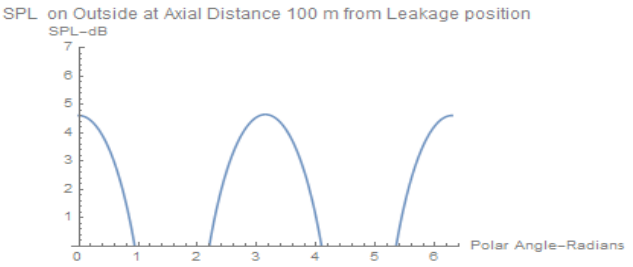
3.1 Acoustic pressure distribution inside the pipeline wall(Mode shape m=0,n=1)

Data used for the acoustic pressure analysis is for 16", carbon steel pipeline. Pipeline wall thickness is 9.525 mm. Unattenuated and attenuated acoustic pressure distributions on the inside of the pipeline at a distance of 100 m from one end are as shown below. The driving frequency is 5000 Hz.The size of leakage hole is 2 mm dia and rate of leakage is 10 litres per minute:





3.2 Acoustic pressure distribution outside the pipeline wall (Mode shape m=0,n=1)



Discussions

On the inside radius of the pipe, a significant pressure level attenuation – approximately a 48% reduction in sound pressure level(SPL) from the unattenuated case – shows the internal energy losses due to Korteweg-Lamb correction at the boundary and the internal reflection from the pipeline inside wall surface. At the outside radius of the pipe, the significant pressure losses – 90% reduction in SPL in comparison to the attenuated inside radius – stem from the radiation of the pressure wave from the surroundings. At this surface, the attenuation of the pressure wave is much more rapid than at the inside radius, which is reflected in the lower SPL seen at the outside radius.

4 Conclusion

The significance of the results shown above may not be readily apparent, but they show that the method developed in this paper is powerful and versatile. Using this method, given an initial source frequency, the SPL can be accurately calculated at any distance from this source given a fluid and an elastic pipe. This includes a distance outside the pipe, away from its surface. The effects of the fluid type, fluid velocity and profile, pipe material and thickness, and surrounding medium can all be tested and compared efficiently using this method. The different sensitivities of stiffness-controlled systems may also be easily studied using this method.

Acknowledgments: This is a self funded research paper.

References

[1] C.J. Miguel, and D.Feit. Sound, structures and their interaction, 1986,The MIT press, pp. 378-387,216-222,37-43.

[2] Philip M. Morse, and K. Uno Ingard. Theoretical Acoustics 1968, McGraw-Hill, pp.306-318,501-514,129-139.

[3] XuQingquing, Zhang Laibin, and Liang Wei. Acoustic detection technology for gas pipeline leakage. Process safety and environmental protection 91(2013)253-261.

[4] Michael J. Buckingham. Causality, Stokes' wave equation and acoustic pulse propagation in a viscous fluid Phys. Rev., E 72,026610,2005.

[5] C.R.Fuller, and F.J. Fahy. Characteristics of wave propagation and energy distribution in cylindrical elastic shells filled with fluid , Journal of sound and vibration 1982,81(4), 501-518.

Markham Stouffville Hospital, A Case Study For Minimizing Helicopter Noise And Vibration Impact On A Healthcare Facility

Rob Jozwiak

Markham Stouffville Hospital was renovating their facilities and this included the redesign of the helipad, moving it from a nearby location to be located on the roof of the building. The incorporation of the heliport introduced the potential to significnatly impact the neo-natal Intensive Care Unit (NICU) located directly below the helipad, as well as impede the use of sensitive diagnostic imaging equipment an additional floor below.This study discusses the work carried out to evaluate the impact from the helicopter landings, designing the architectural elements of the building to meet stringent noise & vibration criteria, providing auralization feedback for the stakeholders and post-construction follow-up testing of the performance of all elements to confirm the helicopter impact.

HEARING SCIENCES/CONSERVATION - SCIENCES DE L'AUDITION

THE AGE-RELATED STRIAL DEGENERATION INHIBITS THE INNER HAIR CELL-AUDITORY NERVE COMPLEX: A COMPUTATIONAL INVESTIGATION.

Amin Saremi *[1],

[1] Computational Neuroscience and Cluster of Excellence "Hearing4all", Department for Neuroscience, University of Oldenburg, Germany.

1 Introduction

The stria vascularis is a vascularized tissue at the lateral wall of the mammalian cochlea which injects electrically-charged potassium ions (k+) into the scala media [1]. This is directly linked to the 89-mV endocochlear potential (EP) measured inside the healthy human cochlear duct [1]. Because of the age-related degeneration of the stria vascularis, the k+ recycling is impaired and is not capable of maintaining the optimal EP, leading to a common type of hearing loss known as the metabolic presbyacusis [2].

The effects of the age-related deficits of the EP on the outer hair cell (OHC) motile forces and the associated decline of the 'cochlear amplifier' have been investigated both experimentally [3] and theoretically [4]. Nevertheless, much less is known about the additional effects of the age-related EP reduction on the sensory inner hair cells (IHCs) function. Here we quantitatively analyze the effects of the EP reduction on the IHC membrane potential and, thereby, the auditory nerve (AN) neurotransmitter release by means of two physiologically-based computational models of the IHC [5] and the AN synapse [6].

2 Methods

2.1 Modeling Method

Figure 1 shows the two publicly-available models which have been integrated and used for this computational analysis. The IHC model used here [5] is a lumped-element biophysical model that consists of cell-level parameters such as the capacitance and conductance of the cell body, the mechanical sensitivity of the potassium channels, and the EP. The input to this model is the IHC stereocilia displacement and its output is the IHC membrane potential.

The auditory synapse model used here [6] to simulate the neuronal activity of the AN is developed according to the parameters based on the time-varying three-store diffusion synapse model. The model is excited by the IHC membrane potential and gives the neuronal spikes on the high spontaneous rate (HSR) AN fibers in its output. Based on the spikes, the post-stimulus time histograms (PSTH) can be estimated for a bin width of 0.4 ms.

2.2 Definition of the Hearing Threshold

It is believed that the hearing threshold occurs in the human auditory system at a 2-nm displacement of the IHC stereocilia [1]. This amount of displacement is regarded as the reference intensity (0 dB) in this analysis. Accordingly, the IHC model is excited by a sinusoid at 0 dB and the output of the AN model is assessed. Due to the probabilistic nature of the neuronal activity, the simulation is repeated five times for any given level. The level of the input is increased with steps of 1 dB until a spike is formed at the output of the AN model, in all five repetitions (PSTH=1). The corresponding level that caused this is regarded as the hearing threshold. In other words, the threshold is defined as the minimum stimulus intensity that yields a firing probability of 100% on the HSR AN fibers.

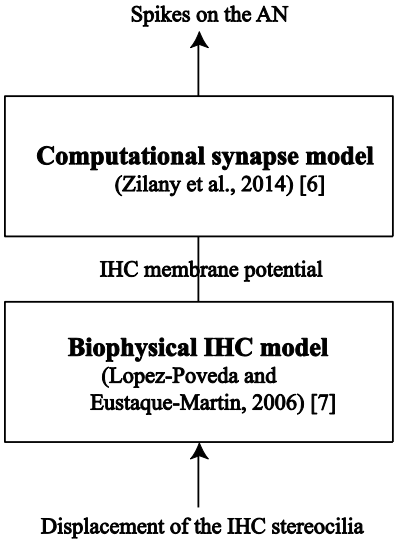


Figure 1: Our modeling approach consists of a biophysical model of the IHC [5] and a computational model of the AN synapse [6]. These two models were integrated to simulate the AN neuronal activity in response to the displacement of the IHC stereocilia.

2.3 Simulation Method

The EP parameter in the IHC model is set to -89 mV and the model is excited by a sinusoid at 0 dB. The amplitude of the sinusoid is increased until the hearing threshold is reached. This is repeated for characteristic frequencies (CFs) from 0.1 to 6 kHz to estimate the hearing thresholds for the condition in which the EP is optimal (healthy condition).

The EP parameter in the IHC model is then decreased to 50% of its optimal value and the hearing thresholds are assessed for CFs between 0.1 and 6 kHz, according to the procedure explained above. These hearing thresholds are compared with the thresholds estimated in the healthy condition, to yield the corresponding threshold elevations (hearing loss) caused by the EP reduction.

3 Results

Figure 2(A) shows that the hearing thresholds vary between 0 and 4 dB for the healthy cochlea (Mean=0.44 dB, SD= 0.84 dB). This average hearing threshold (0.44 dB) corresponds to 2.1 nm displacement of the IHC stereocilia. Figure 2(A) also illustrates the hearing thresholds for the presbyacusis condition (EP=-44.5 mv) which show a frequency-dependent elevation. Figure 2(B) shows the hearing loss associated with the hearing threshold elevations seen in Fig. 2(A). The hearing loss was calculated by subtracting the hearing thresholds of the presbyacusis cochlea from those of the healthy cochlea.

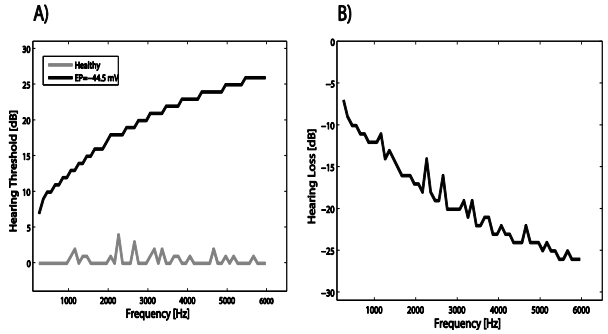


Figure 2. A). The hearing thresholds for two conditions : (i) the healthy condition (EP=-89 mV), and (ii) the EP reduced to half of its optimal value (EP=-44.5 mV). **B)** The hearing loss corresponding to the threshold elevations seen in panel A.

4 Discussion

Our simulations show that the IHC-related hearing threshold elevations caused by a 50% decrease of the EP are frequency dependent as the threshold elevations increase from 6.5 dB at 100 Hz to 27 dB at 6 kHz. This is in general agreement with the observed high-frequency profile of presbyacusis [2] [3] [4].

Due to the lack of explicit experimental data on the AN activity in response to the IHC stereocilia displacement, our simulations have not been compared to any experimental references, and therefore, remain solely theoretical in nature. However, the models used here have been validated by successfully reproducing various experimental data in other studies [5] [6]. This makes the model predictions presented here noteworthy.

The presented predictions suggest, at least theoretically, that the EP decline caused by the age-related degeneration of the cochlear lateral wall affects the hearing thresholds by inhibiting the IHC-AN functions while the OHCs and IHCs were totally intact and healthy. This is opposed to the minimalist view that reduces the origins of all cochlear hearing losses to merely structural lesions of the OHCs and IHCs, undermining the potential roles of the pathologies in the stria vascularis.

5 Conclusion and Future Works

Several investigations have quantified the effects of the age-related EP reduction on the decline of the motile active forces generated by the OHCs, leading to the deficit of the cochlear amplification, and thereby hearing threshold elevations [3], [4]. They showed that a 50% decrease of the EP leads to a sloping high-frequency hearing loss that reaches approximately 40 dB at 4 kHz.

Here, we studied an additional contribution of the EP reduction to the hearing loss, using a computational model of the IHC-AN complex (Fig. 1). The hearing thresholds, defined as the minimum displacement of the IHC that results in a 100% probability of the neurotransmitter release on the AN, were assessed for the healthy condition (EP=-89 mV) and the presbyacusis condition (EP=-44.5 mV). These two hearing thresholds were subtracted from each other to give an estimate of the hearing loss associated with the EP decrease.

The presented results demonstrate that a 50% decrease of the EP suppresses the IHC-AN functions and yields threshold elevations that increase from 6 dB at 100 Hz to 27 dB at 6 kHz. This additional IHC-related hearing loss is independent of the OHC-related hearing loss, caused by the decline of the cochlear amplification due to the EP reduction. In other words, our results suggest that age-related EP reduction not only restrains the OHC motile forces [4], but also causes additional hearing loss by directly inhibiting the membrane potential of the IHC. The analysis presented in this paper can be developed to pave the way toward differentially understanding the underlying mechanisms involved in the age-related hearing loss.

Acknowledgments

This work was funded by the DFG Cluster of Excellence EXC 1077/1 "Hearing4all", University of Oldenburg, Germany.

References

[1] N. B. Slepecky (**1996**). "Structure of the mamalian cochlea," in The Cochlea, edited by Dallos, P., Popper, A., N., Fay, R., R. (Springer-Verlag, USA), pp. 44-129.

[2] H., F. Schuknecht (**1971**), “Presbyacusis”, In: Pathology of the Ear. Cambridge, Harvard university press, MA, USA. pp 1-178.

[3] R. A. Schmiedt, H. Lang, H. O Okamura., and B. A. Schulte (**2002**). "Effects of furosemide applied chronically to the round window: a model of metabolic presbyacusis," J. Neurosci. **22**, 9643-9650.

[4] Saremi, A., and S. Stenfelt (**2013**). "Effects of metabolic presbyacusis on cochlear responses: A simulation approach using a physiologically-based model," J. Acoust. Soc. Am. **134** (4), 2833-2852.

[5] E. A. Lopez-Poveda, and A. Eustaquio-Martin (**2006**). "A biophysical model of the inner hair cell: the contribution of potassium currents to peripheral auditory compression," JARO. **7**, 218-235.

[6] M. S. Zilany, I. C. Bruce, and L. H. Carney (**2014**). "Updated parameters and expanded simulation options for a model of the auditory periphery," J. Acoust. Soc. Am. 135, 283-286.

THE OPPORTUNITIES AND CHALLENGES OF IN-EAR NOISE DOSIMETRY

Fabien Bonnet [1], Jérémie Voix [*1] and Hugues Nélisse [2]
[1]École de Technologie Supérieure, Université du Québec, 1100 rue Notre-Dame Ouest, Montréal, Québec H3C 1K3, Canada
[2]Institut de Recherche Robert-Sauvé en Santé et Sécurité au Travail (IRSST), 505, Boulevard de Maisonneuve Ouest, Montréal, Québec H3C 1K3, Canada

1 Introduction

Most industry employers are required to follow certain regulations when it comes to protecting their employees from Noise Induced Hearing Loss (NIHL). Although ideally, noise reduction measures should be directly applied to the damaging noise sources by reducing their acoustic emission below acceptable limits, these mitigations are often unrealizable for practical or economical reasons. Hearing Protection Devices (HPDs) then become the only option left to prevent the effects of noise overexposure. Yet despite a huge choice of HPDs, their performance is not guaranteed as it mostly depends on how well they are adapted to a particular worker's situation. A compliant HPD model should offer sufficient attenuation while leaving unaltered speech and communication signals, when possible. Overprotection should be avoided, as it may lead the worker to remove his HPD to better communicate and, thus, decrease its performance drastically [1]. To help in the selection process, it is therefore essential that each worker's individual noise exposure be precisely known when wearing a particular HPD model, which is rarely the case. In fact, noise exposure is generally computed from combined measurements of the unprotected noise levels and the attenuation of the HPDs, although the latter is subject to significant intra- and inter-subject variability [2]. After a short review of the existing methods for HPD attenuation measurement and individual noise exposure assessment, this paper will focus on the new emerging in-ear dosimetry approach. The benefits and opportunities of such method will be described together with its challenges with regard to hearing-loss prevention.

2 Review of existing methods

2.1 Measurement of HPD attenuation

HPDs are typically characterized by attenuation ratings that rely on measurements performed under well-controlled laboratory conditions. Those measurements can use subjective methods such as the real-attenuation-at-threshold (REAT) procedure, or objective techniques like the microphone-in-real-ear (MIRE) method. Unfortunately, these are known to fail at predicting the HPD performance in the real-world work environment [2], which encouraged the emergence of so-called Field Attenuation Estimation Systems (FAES) [3]. Considering various quality factors (speed, accuracy, repeatability and practicality), it is a modified MIRE method, termed field MIRE (F-MIRE) [4], that appears like the best current compromise to evaluate the in-field performance of HPDs.

[*] jeremie.voix@etsmtl.ca

Yet despite the progress achieved with field measurement technologies, some uncertainties remain as all of the above methods rest on timely measurements that generally fail to account for poor re-insertion, HPD removal or even potential loss of acoustic seal over time [3]. The real-world HPD attenuation is hence expected to deviate from these well-controlled measures, which was confirmed by findings looking at the effective daily protection as a function of time during work shifts [2]. Future research should also address the need for fully individualized attenuation data, which could be computed on the basis of an identification function for each particular subject's ear dimensions [4].

2.2 Measurement of Individual Noise Exposure

When the effective protected noise levels (those actually received at the subject's eardrum) cannot be measured directly, the tracking of individual noise exposure requires the knowledge of both the attenuation and the unprotected levels. The unprotected noise exposure varies based on environmental factors and can be measured effectively using personal dosimeters (body-worn derivatives of integrating sound level meters), although errors due to placement effects can be quite significant in a directional field [5]. But the main obstacle to individual noise exposure assessment arises from the difficulties in measuring the HPD attenuation. It is to circumvent this problem that later works have been focusing in measuring the effective (protected) noise exposure rather than the unprotected noise levels.

An in-ear dosimetry system, the QuietDose, was commercialized by Sperian Protection following their 2008 acquisition of doseBusters™ USA. This consists of a generic eartip adapter with an integrated miniature microphone that inserts into compatible eartips and connected to a dosimeter. When the HPD is being worn, the dosimeter measures the protected level and when removed, the microphone continues to measure the level of exposure (unprotected). Such device appears to take into account the performance of the protector as well as proper fit, but does not provide any insight as to why a particular worker is over his dose as it gives little information regarding the exposure level when the HPD is worn. Furthermore, the convenience of this system is hampered by the necessity of downloading the exposure data at the end of the day.

This lack of real-time data was then overcome by some of the authors of this paper with the development of a smart earplug with integrated in-ear dosimetry [6], which also benefits from a microphone doublet; one microphone measures the sound reaching the outer part of the HPD while the other is attached to a sound bore that travels through the earplug and monitors the protected noise levels.

3 Aspects relating to In-Ear Noise Dosimetry

3.1 A promising avenue

Those recent technologies, along with the miniaturization of electronics during the last decade, are a clear step towards more intuitive approaches for assessing individual noise exposure. While previous methods aimed in the prediction of personal exposure based on misrepresentative HPDs ratings, in-ear noise dosimetry can ensure compliance with safety regulations through an 'upstream' control of the effective noise dose received by the employee.

Besides, the monitoring of the unprotected noise levels that comes with the latest instrumentation available [6], added to the protected noise levels, should permit to gather indications as to why a particular worker is overexposed (unusually noisy equipment, sound field spectral balance, HPD removal or poor fit, insufficient HPD attenuation, etc.) and help improving the effectiveness of HPDs.

3.2 Challenges with regard to NIHL prevention

NIHL has been studied as far as the 1700's but remains challenging because of the extreme complexity of the human auditory system, and it is beyond the scope of the proposed research project to go through the biological factors that may drive phenomena like individual susceptibility. Nevertheless, we have identified three main elements of on-going research that are to be considered when using in-ear noise dosimetry to prevent NIHL by limiting the received noise dose of a given individual.

The first research element concerns the acoustical corrections due to the Transfer Function of the Open Ear (TFOE), as well as the Occluded Ear Canal Resonance and Probe tube Effect (OER). The TFOE represents the amplification of the sound pressure caused by the resonance in the open ear canal, which varies with the geometry of the human head, torso, pinna, and shape of ear canal as well as eardrum impedance. The OER is mostly dependent on the length of the probe tube through the HPD and the length of the residual part of ear canal between earplug and eardrum. These correction factors, which are both dependent on the user's morphology, are essential for comparison with noise regulations since most noise criteria are expressed as free-field values. Estimates such as those measured on a head and torso simulator have already been used [6], but future instrumentation would benefit from individualized factors that can take the high variability of the mentioned ear characteristics into account. An analytic model was developed as part of this study that uses the geometry given by Stinson [7] to calculate individual correction factors based on ear dimensions. This showed encouraging results, but the search for a practical method to identify the main parameters used in this model remains a major challenge.

The second research element deals with the noise induced by the user himself, which was clearly identified by several authors as a dominant sound source in medium-level noise environment [2] [6]. This can be caused either by low-frequency noise generated by movements from the wearer or by his own voice that directly contributes to the recorded noise levels. An automatic detection method is currently under investigation to discriminate such time events This method could use some of the algorithms used as part of the development of a Voice Activity Detector that can operate in low signal-to-noise situations [8].

Finally, some researchers have raised questions about the potential influence of ear occlusion on noise susceptibility. According to Theis et al. [9], "human subject data is extremely important in developing and validating calibration factors for any type of noise dosimeter but particularly so for in-ear dosimetry". This statement comes along with data supporting the idea that in-ear dosimetry overestimates the noise dose. To verify this finding, further studies will be undertaken involving loudness-balance tests performed on a group of human subjects.

4 Conclusions

Despite the great advantages that in-ear dosimetry systems can offer for hearing conservation, progress is to be made with regard to the acoustical corrections needed to relate noise measured in the occluded ear to risks of hearing loss.

Acknowledgments

The support of the IRSST (Quebec Occupational Health and Safety Research Institute) is gratefully acknowledged.

References

[1] R. Neitzel and N. Seixas. The Effectiveness of Hearing Protection Among Construction Workers, *J. Occup. Environ. Hyg.*, vol. 2, 227:238, 2005.

[2] H. Nélisse *et al*. Measurement of Hearing Protection Devices Performance in the Workplace during Full-Shift Working Operations, *Ann. Occup. Hyg.*, vol. 56, no. 2, 221:232, 2012.

[3] J. Voix *et al*. Field Fit-Testing and Attenuation Measurement Procedures, in *Noise Manual. 6th ed. Fairfax, VA: American Industrial Hygiene Association Press*, 2015 (in press)

[4] J. Voix and F. Laville. The Objective Measurement of Earplug Field Performance, *J. Acoust. Soc. Am.*, vol. Vol. 125, no. Issue 6, 3722:3732, 2009.

[5] D. Byrne and E. Reeves. Analysis of Nonstandard Noise Dosimeter Microphone, *J. Occup. Env. Hyg.*, vol. 5, no. 3, 197:209, 2008.

[6] K. Mazur and J. Voix. Development of an Individual Dosimetric Hearing Protection Device, in *Inter-Noise 2012 : The 41th International Congress and Exposition on Noise Control Engineering*, 2012.

[7] M. R. Stinson and B. W. Lawton. Specification of the geometry of the human ear canal for the prediction of sound-pressure level distribution, *J. Acoust. Soc. Am.*, vol. 85, no. 6, 2492:2503, 1989.

[8] N. Lezzoum *et al*. A Low-Complexity Voice Activity Detector for Smart Hearing Protection of Hyperacusic Persons, in *Interspeech 2013*, 4:8, 2013.

[9] M. A. Theis *et al*. Hearing protection with integratedin-ear dosimetry: a noise dose study, in *Proceedings of the Internoise 2012/ASME NCAD meeting August 19-22, NY, New York*, 2012.

USE OF AUDITORY STEADY-STATE RESPONSES IN MEASURING THE OCCLUSION EFFECT OF HEARING PROTECTION DEVICES

Olivier Valentin* **and Frédéric Laville†**
Dept. de génie mécanique, École de technologie supérieure (ETS),
1100, rue Notre-Dame Ouest, Montréal (QC), Canada, H3C 1K3.

1 Introduction and objectives

Worldwide hearing loss estimates increased from 120 million people in 1995 [1] to 250 million in 2004 [2]. A common solution to protect workers from noise exposure consists of using hearing protection devices (HPD). An important parameter about HPD is the wearing time, since it can decrease the effective protection provided by HPD [3]. However, the recommended wearing time for limiting exposure to noise is not always respected. The occlusion effect (OE) is one of the reasons often given to justify the non-use of HPD: the occlusion of the ear canal induced a modification of the wearer's voice perception, which creates a discomfort that sometimes brings people to remove their HPD. Present methods of OE measurement have limitations. Objective measurements using microphone do not asses bone conducted sounds directly transmitted to the cochlea. Psychophysical measurements at threshold are biased due to the low frequency masking effects from test subjects' physiological noise and contain variability of measurement due to subjective responses. The present study reports an attempt to overcome the limitations of these methods through the recording of auditory steady-state responses (ASSR), which has been adapted from the methodology used in previous work for the measurement of HPD attenuation [4]. Due to the time consuming nature of ASSR recording, the study was conducted using only two stimuli having 250 and 500 Hz carriers, chosen in the low frequency range (below 1000 Hz) where the greatest positive OE is expected. ASSR results were compared to the results obtained with a subjective psychophysical method.

2 Method

2.1 Participants

Eight men with ages from 22 to 26 and hearing thresholds below 20 dB SPL (from 125 Hz to 8 kHz) were assessed. A typical experimental procedure included two steps:

2.2 Step 1: Psychophysical measurements

The psychophysical OE for each subject was measured on three trials during a single visit to the laboratory. Each trial consists of a paired open and occluded threshold, the order being counterbalanced across subjects. The hearing protectors, a pair of 3M™ foam earplugs with a 10mm insertion depth, were refitted by the experimenter for each trial. The test signals consisted of bone-conducted pure tones at 250 and 500 Hz presented to a Radioear B-81 [5] bone vibrator which was coupled to the forehead by an elastic headband with approximately 400 to 450 g of force. For both stimuli, the level of the first presentation was 30 dB HL. The level of each succeeding presentation was determined by the preceding response. After each failure to respond to a signal, the level was increased in 1-dB steps until the first response occurred. After the response, the intensity was decreased 10 dB and another ascending series was started. The minimum number of responses needed to determine each BC threshold was three responses out of five presentations at a single level.

2.3 Step 2: Physiological measurements (ASSR)

ASSRs are electrophysiological responses, recorded from the human scalp, and often evoked by one or more carrier frequencies (Fc) that are amplitude-modulated at a specific frequency (Fm). In practice, when a subject is exposed to such a stimulus, spectral power of the EEG frequency spectrum of the subject that is related to the stimulus will be manifest at Fm, and may also appear at its harmonics [6]. Since amplitudes of ASSR are quite well correlated with the level of stimulation, it may be possible to measure the

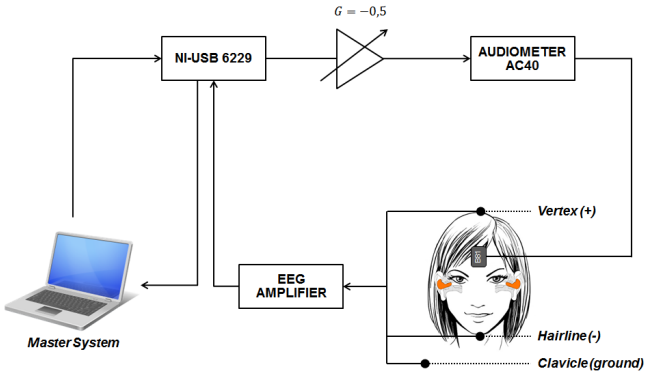


Figure 1: Overview of the experimental setup of ASSR recording. All components are monitored by a single PC. The stimulation signals from the DA output of the NI-USB 6229 board are attenuated by an operational amplifier with a gain of -0.5, so that they may be delivered to the "tape input" of the audiometer, which enables the operator to adjust the levels of stimuli delivered by the bone vibrator. In parallel, ASSRs are scalp-recorded on the electrodes (placed between vertex (+) and hairline (-), with clavicle as a ground) and are then amplified by an EEG amplifier, before reaching the AD input of the data acquisition board. Data is processed online through the MASTER SYSTEM software.

* m.olivier.valentin@gmail.com
† frederic.laville@etsmtl.ca

Stimulus	F_c	Stimulation levels	F_m	AM%	FM%
#1	250Hz	0 to 20 dB_{HL} (10 dB step)	40Hz	100 %	25%
#2	500Hz	20 to 40 dB_{HL} (10 dB step)			

Table 1: Characteristics of the amplitude-modulated tones used in the ASSR experiment. The same stimulation levels were used for both normal and occluded condition.

occlusion effect by recording ASSRs using both normal and occluded conditions. ASSR recordings and stimuli generations were conducted by using the LabVIEW™ based "MASTER SYSTEM™" Rotman Research software. Stimulus characteristics are reported in Table 1. The setup used for the recording of ASSR is illustrated on Figure 1.

Results and discussion

ASSR-based "physiological" OEs were calculated as the average difference between the occluded stimulation levels and their unoccluded equivalents that would be obtained for the same amplitudes, which were calculated using the linear least-squares regression performed on unoccluded data. An example is illustrated in Figure 2, on the plot of the 250 Hz amplitudes for subject #1.

Physiological OEs estimates were expected to be different from psychophysical OEs, because electrophysiological assessments of hearing threshold are not biased by the low-frequency masking effect. However, results suggest that the effect of low-frequency masking may not be as large an influence as previously assumed for frequencies below 500 Hz (87.5% of subjects at 500Hz, and 25% of subjects at 250 Hz had physiological OEs that were greater than psychophysical OEs).

Although ASSRs have been adapted in the past for measuring the threshold to bone-conduction stimuli [7], no study has considered using ASSRs as a method to evaluate the OE induced by wearing HPD. The present study seeks to ascertain whether it is possible to objectively measure the OE of HPD using ASSRs collected in the same subject both with and without protectors. The results are encouraging: we successfully measured the OE in every volunteer who participated but further research, using an extended frequency range, should be done to explore this hypothesis.

Acknowledgments

The authors wish to express their appreciation to 3M Company (St. Paul, Minnesota, USA) for providing the foam earplugs and to NSERC (Natural Sciences and Engineering Research Council of Canada) for their support to this project.

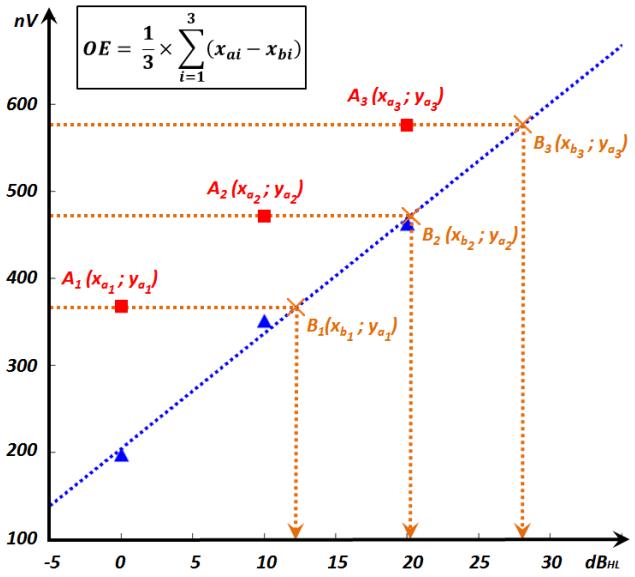


Figure 2: Example of calculation of the physiological OE from ASSR amplitudes as a function of stimulation intensity at 250 Hz, in the case of subject #1. Occluded condition results are represented by the red "■" markers and unoccluded condition results by the blue "▲" markers. The blue dotted straight line was obtained from linear last squares regression for the unoccluded condition.

References

[1] WHO, World Health Organization. Guidelines for community noise. Edited by B. Berglund, T. Lindvall, and DH Schwela, 1999.

[2] A. Smith. The fifteenth most serious health problem in the WHO perspective. Presentation to IFHOH World Congress, Helsinki, July 2004.

[3] E.H. Berger. The naked truth about NRRs. E-A-$RLOG_{20}$, 1993.

[4] O. Valentin, M.S. John, F. Laville. Use of auditory steady-state responses in measuring the attenuation of hearing protection devices. Canadian Acoustics - Acoustique Canadienne. 42(3), 96-97, 2014.

[5] B. Håkansson. The balanced electromagnetic separation transducer: A new bone conduction transducer. J. Acoust. Soc. Am. 113 (2), 818-825, 2003

[6] M.S. John, A. Dimitrijevic, T.W. Picton. Efficient stimuli for evoking auditory steady-state responses. Ear & Hearing. 24, 406–423, 2003.

[7] S.A. Small, D.R. Stapells. Multiple auditory steady-state response thresholds to bone-conduction stimuli in young infants with normal hearing. Ear & Hearing. 27, 219–228, 2006.

Comfort From Using Foam Earplugs – Project For A Study

Alberto Behar, Frank Russo

Comfort is an important characteristic of hearing protectors, as important as the sound attenuation. A bibliographical study was performed by Behar and Segu (submitted) examining research on comfort from hearing protectors published during the last 25 years. The study was a background document for developing a procedure for ranking comfort of hearing protectors on the basis of physical characteristics. The paper's main recommendation was work on only one type of protector. (Prof. S. Gerges, Universidade Federal de Santa Catarina, Brazil is studying comfort from ear muffs exclusively). The current project will focus on comfort from foam earplugs exclusively. Twenty participants will evaluate twelve types of foam earplugs. Each participant will be asked to assess the comfort using a visual analog scale. In addition, wax molds will be obtained from the ears of all participants. These molds will be digitized to obtain the shape and size of the ear canal. The physical characteristics of the earplugs will also be measured (density, stiffness, diameter, etc...) Finally, a correlation will be sought between the physical characteristics of the plugs and the comfort experienced by the participants. A multiple regression approach will be used to assess the influence of physical measures including stiffness, size, density, shape, and material on comfort ratings obtained across individuals with ear canals of varying diameter. The models will be realized as formulae specifying proportion of variance accounted for by each factor, the weight of the factor and the directionality. This formula could then be used to classify any foam plug that may be developed after completion of this study.

MUSICAL ACOUSTICS - ACOUSTIQUE MUSICALE

SPECTRUM ANALYSIS AND DIRECTIVITY PATTERN OF A TRANSDUCER-DRIVEN CONCH SHELL

Rasoul Morteza Pouraghdam [*1] **and Rama B. Bhat** [†1]
[1]Department of Mechanical and Industrial Engineering, Concordia University, Montreal, Quebec, Canada H3G 1M8

1 Introduction

Conch shells were and are still being used in different cultures around the world, such as in India, Greece and Japan. They served as fog warning devices, but also were used for coded communication purposes, to signify auspicious occasions and as a musical instrument, solo or as part of an ensemble. In this paper, we conduct a geometric analysis of the conch shell spiral cavity with the help of X-ray tomography scans and by approaches used in past literature. Moreover we study the directivity pattern of the conch shell in two excitation cases : 1. Loudspeaker 2. Electro-pneumatic transducer (EPT). Finally we examine the sound spectra in both previous excitation cases, as well as for lip-excitation.

2 Experimental Study

2.1 Geometrical Analysis

It is a difficult task to come up with a single geometry that matches all conch shell spiral cavities. Taylor, Prasad and Bhat [1] obtained X-ray tomography scans of a *Turbinella Pyrum* and showed that the shell's cavity approximates a conical spiral growing around a central stem with a cross-section area varying in both axial and transverse directions. Rath and Naik [2] found the geometry of a Nautilus shell to be a golden spiral, and discovered Fibonacci patterns inside a conch shell with the proper parameter selection. In our sample, we took X-ray tomography scans of a Turbinella Pyrum conch shell. Our measurements showed that none of the shell spiral geometries resembled a golden spiral. We also followed the same measurement procedure done by Rath and Naik and found Fibonacci patterns in our conch shell sample (table 1). We finally applied the same measurement approach done by Bhat [3] in order to find the equivalent acoustical straight tube to a given conch shell spiral cavity. We found a 55 cm conical tube to approximate our Turbinella Pyrum shell inner cavity. The resonances of a 55 cm conical tube and the lip-driven Turbinella Pyrum matched with minimal deviations. The result for the Turbinella Pyrum is displayed in table 2 (tagged as estimation).

Distance between points	Green Points	Pink Points
1 & 2	23 cm	12 cm
3 & 4	40 cm	22 cm
5 & 6	63 cm	34 cm

Table 1: The Fibonacci pattern is clearly present with the correct selection of parameters, as originally demonstrated by Rath and Naik.

∗. r_ morte@encs.concordia.ca
†. rama.bhat@concordia.ca

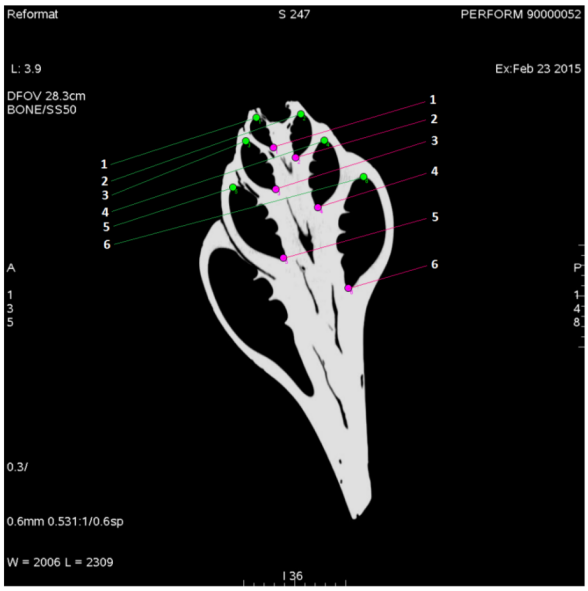


Figure 1: X-ray scan of our Turbinella Pyrum with selection points.

2.2 Spectrum Analysis

Taylor, Prasad and Bhat [1] also carried out a spectrum analysis of a conch shell. They found that when the shell is excited by lip-vibration, the sound spectrum contains harmonics that decrease smoothly in magnitude. Their results confirmed Bhat's [3] earlier findings. A skilled brass player could generate different notes by exciting different frequencies of the shell's cavity with the proper lip tension adjustment. This has been clearly displayed by the famous American trombonist Steve Turre who is also a conch shell virtuoso.

The sound spectrum of our lip-excited conch shell also displayed a fundamental frequency and several harmonics. Moreover, we measured the spectra in two other excitation cases as well. All of the measurements were done in an anechoic chamber and the results are presented in table 2.

Loudspeaker Sine Sweep Measurement

The sine sweep measurement was done by attaching a small loudspeaker to the throat of the conch shell. We set the frequency span to 100 - 2000 Hz with a time span of 1 second. Figure 2 shows the spectrum of the loudspeaker-driven Turbinella Pyrum.

The output spectra of the conch shell displayed harmonic resonances of a fundamental frequency, but slightly shifted which is acceptable as the shell's geometry is not perfect. The results verify that the shells are approximately wrapped conical tubes around a central stem, since it is a known fact

that resonances inside a conical tube are harmonics of a fundamental mode.

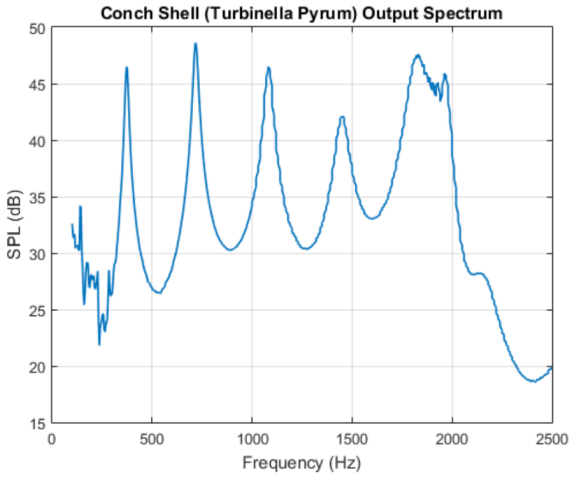


Figure 2: The Turbinella Pyrum resonances at 372, 716, 1080, 1450 and 1830 Hz

Electro-Pneumatic Transducer (EPT)

In this setup, the Turbinella Pyrum was driven by an EPT operating at 52 Hz. This was an attempt to see how the shell would respond to an excitation different than a sine sweep, and how closely it simulates lip-excitation. The output spectrum displayed $3k + 4$ multiples of the driver frequency (52 Hz) where $k = 0, 1, 2, ...$ and although a clear pattern is found, not much can be deduced due to the non-linear nature of the excitation.

Loudspeaker	EPT	Lip-vibration	Estimation
372	52	315	312
716	208	626	624
1.08K	364	944	936
1.45K	520	1.26K	1.25K
1.83K	676	1.57K	1.56K

Table 2: Resonance modes of the Turbinella Pyrum in different excitation cases (in Hertz). The estimated resonance frequencies correspond to the theoretical resonances of an equivalent conical (straight) tube to the shell's spiral cavity.

2.3 Directivity-Pattern

We carried out directivity measurements of the loudspeaker and EPT-driven Turbinella Pyrum in both horizontal and vertical alignments. The measurements were done in an anechoic chamber where the shell was mounted on a tripod, positioned at the center of the room. The loudspeaker was glued to the throat of the shell and a microphone positioned in front of the mouth of the shell at a distance of 1.35 m was used to capture the output.

The loudspeaker was driven at a given frequency and the shell was rotated at 60° interval about both its longitudinal and transverse axes to find the 2-D 360° directivity pattern in both vertical and horizontal alignments of the shell. The same configuration was used for the EPT-driven shell.

Based on our results, the shell radiates sound uniformly in space at frequencies near the shell's cavity resonance. It is also seen that the shell's alignment makes negligible difference.

Degree	LS-H	LS-V	EPT-H	EPT-V
0	73.7	73.9	55.1	55.1
60	73.1	72.8	53.4	53.5
120	73.1	72.2	53.0	52.5
180	73.6	72.8	53.9	53.0
240	74.6	73.6	55.2	54.3
300	74.5	73.6	53.9	53.2

Table 3: The captured output magnitude (in dB) for the loudspeaker (LS) and EPT excitations at 310 Hz in both horizontal (H) and vertical (V) alignments.

3 Conclusions

In this study we confirmed that by using X-ray scans of a conch shell, the spiral inner cavity could be approximated by a straight conical tube, originally suggested by Bhat [1,2]. We also carried out sine sweep measurement of the conch shell and found the resonance modes to fall in integer multiples of a fundamental frequency with minimal deviation. The lip-driven shells display perfect harmonics and only some of the harmonics are excited when driven by an Electro-pneumatic transducer (EPT). Finally we carried out directivity measurements of a conch shell in both vertical and horizontal alignments, excited by a loudspeaker and an EPT. The results indicated that the shell is essentially an omnidirectional source when excited near its cavity resonance frequency.

Acknowledgments

We would like to highly thank Chetas Shah and Pratik Desai for having carried out the directivity and the Electro-pneumatic transducer spectrum measurements in the anechoic chamber.

References

[1] M.G. Prasad, L.R. Taylor and R.B. Bhat. Geometrical modeling and spectral analysis of a conch shell trumpet. *Third International Congress on Air And Structure-Borne Sound And Vibration*, June 1994.

[2] S.K. Rath and P.C. Naik. Fibonacci structure in conch shell. *Current Science*, 88(4) :555–557, February 2005.

[3] R.B. Bhat. Acoustics of conch shells. *Journal of Sound and Vibration*, 157(1) :190–191, 1992.

THE SOURCE DILEMMA: PERCEPTUAL COHERENCE AND THE CONTINUUM OF CONSONANCE-DISSONANCE

Tanor Bonin and Daniel Smilek
Department of Psychology, University of Waterloo
200 University Avenue W., Waterloo, ON N2M 5G1 tanorbonin@gmail.com

1. Introduction

1.1. Overview

Here we present counterintuitive datasets of music listeners' perception of consonance-dissonance in a variety of tonal and atonal musical contexts. Specifically, we present evidence that listeners consonance-dissonance ratings are influenced by psychoacoustic factors beyond the tonal structures of music. For example, participants rated atonal music as more consonant under particular spatial orientations or timbrel distributions of the sounds sources. In addition, we provide evidence that consonant and dissonant musics entail differential cognitive loads. These datasets resulted from various investigations of the *source dilemma hypothesis,* a psychophysical model of consonance-dissonance perception developed by the first author that describes dissonance as an emergent property of a *source dilemma,* where musical sounds produce incoherent auditory percepts, and consonance as an emergent property of *source transparency,* when musical sounds produce coherent auditory percepts.

1.2. Aims

The present report addresses two primary research interests. First, we sought to replicate previous findings (see Bonin, 2014) that timbrel and spatial coherence influence the consonance and dissonance of musical sounds. In doing so we hoped to extend the generality of this observation by using both a novel stimulus set and an alternative participant response method. Second, we investigated a corollary prediction of the source dilemma hypothesis: If dissonance accompanies incoherent auditory perception, then one might expect dissonant music to require greater cognitive processing than consonant, perceptually coherent music.

We tested these predictions with three experiments that measured participants' affective and cognitive responses to musical stimuli differing in their perceptual coherence. In each experiment, we selectively manipulated one of the music's harmonic, timbrel, or spatial characteristics to create coherent and incoherent musical counterparts. We expected that the incoherent musical stimuli would be perceived as more dissonant and interfere to a greater extent with performance on a concurrent cognitively demanding task than its coherent musical counterpart.

2. Method

Each of the three experiments consisted of three blocks. In each block, participants completed a cognitively demanding task (the visual 2-back task; 2500ms SOA, 20% target rate) presented on a 24" Phillips 244E monitor (1920x1080, 80pt Helvetica font) while listening concurrently over stereo headphones (Sony MDR-MA100) to no music, coherent music, or incoherent music. The no music condition served as a practice block and was always presented first, while the order of the music blocks were counterbalanced between participants.

The musical stimuli were derived as "coherent" and "incoherent" counterparts that differed only in the psychoacoustic parameter of interest for that experiment. In Experiment 1, participants (n=48) listened to no music, tonal (harmonic), or atonal (inharmonic) music. In Experiment 2, participants (n=75) listened to no music, atonal music exhibiting timbrel coherence (two segregated timbres), or atonal music exhibiting timbrel incoherence (one fused timbre). Finally, in Experiment 3, participants (n=75) listened to no music, atonal music exhibiting spatial coherence (two segregated spatial locations), or atonal music exhibiting spatial incoherence (one fused spatial location). Participants' performance on the concurrent 2-back task was analyzed both in terms of accuracy and response times. After each block in which music was presented, we asked participants to rate the music they had just heard in terms of "pleasantness," "unpleasantness," "consonance," and "dissonance" on a 7-point Likert scale.

3. Results

As expected, perceptual coherence predicted the listeners' consonance-dissonance ratings of the musical stimuli in all three experiments. Regarding our second hypothesis, the cognitive interference data revealed an unexpected effect of our psychoacoustic manipulations and provided only partial support for our predictions. Consistent with our predictions, we found in Experiment 1 that incoherent atonal music interfered to a greater extent with cognitive performance on the 2-back task than coherent tonal music. Surprisingly, however, we found in Experiments 2 and 3 that the timbrelly and spatially coherent atonal musics imposed greater cognitive loads on the listener than their incoherent counterparts despite the fact that they were rated as less dissonant and unpleasant.

In Experiment 1, participants rated the atonal music as more dissonant, more unpleasant, less consonant and less pleasant than the tonal music (all $t(1,47) > 2.5$, all $p < 0.01$). Furthermore, consistent with our second hypothesis, atonal music also led to slower response times ($t(1,47) = 5.7$, $p < 0.001$) and less accurate responses ($t(1,47) = 2.8$, $p < 0.01$) on the concurrent cognitively demanding task compared to tonal music. These results are consistent with the findings of Masataka and Perlovsky (2013) that dissonant music leads to slower and less accurate performance on incongruent Stroop trials compared to consonant music, but are at odds with the Bodner, Gilboa and Amir (2007)

findings that dissonant music enhances performance on cognitively demanding tasks relative to consonant music.

In Experiment 2, participants rated the atonal music exhibiting timbrel coherence as less dissonant and less unpleasant than the atonal music exhibiting timbrel incoherence (all $t(1,74) > 2.2$, all $p < 0.05$). Surprisingly, however, the atonal music exhibiting timbrel coherence interfered to a greater extent with performance on the concurrent 2-back task than did the atonal music exhibiting timbrel incoherence, eliciting slower ($F(1,74) = 16.203$, $p < 0.0001$) and less accurate ($t(1,74) = 5.435$, $p < 0.0001$) responses.

Mirroring the effects observed in Experiment 2, participants in Experiment 3 rated the atonal music exhibiting spatial coherence as less dissonant and less unpleasant than the atonal music exhibiting spatial incoherence (all $t(1,74) > 2.05$, all $p < 0.02$). However, again in contrast with our predictions, the spatially coherent atonal music led to slower ($F(1,74) = 17.01$, $p < 0.0001$) and less accurate ($t(1,74) = 5.45$, $p < 0.0001$) response times on the cognitively demanding 2-back task than did the spatially incoherent atonal music.

The unexpected cognitive interference effects we observed between the timbrelly and spatially coherent stimuli and their incoherent counterparts produce several interesting implications. First, they are in contrast to the broad prediction of the source dilemma hypothesis that more dissonant music should generally require greater cognitive processing than less dissonant music. Such results indicate that participants found the perceptually coherent musics more consonant and enjoyable despite the fact that the brain was working harder to process them, and indicate the need for more specific predictions about the relationship between cognitive processes and dissonance phenomenology (it is worth reiterating here that we *did* find the predicted interference effect using a harmonic coherence manipulation). Thus, in conjunction with those from Experiment 1, these results demonstrate that the affect induced by musical stimuli is not sufficiently predicted by the cognitive processing that the stimuli require (i.e., it is not simply the case that atonal musics induce negative affect because they are difficult to process, or vice versa).

With these conclusions in mind, the source of this additional cognitive load in the timbrelly and spatially coherent atonal musics in Experiments 2 and 3 remains to be determined. One possibility is that, in order to maintain the reduced dissonance and unpleasantness in the segregated timbrel and spatial conditions, listeners must actively maintain the segregation of the independent musical streams, producing a divided attention requirement that increases the cognitive demand of these stimuli. This divided attention requirement could ostensibly overshadow the more nuanced differences in cognitive demand produced as a function of perceptual coherence, and would explain why we observed our expected results as a function of harmonicity in Experiment 1, where the coherent (fused) harmonic condition did not require active maintenance of stream segregation, eliminating this potential source of cognitive interference.

Consistent with this possibility is a recent publication from Demany, Erviti and Semal (2015), demonstrating that attention can be divided between segregated musical tone streams, and that response sensitivity declines as divided attention requirements increase. Future empirical studies could address this issue by investigating the interactive influence between perceptual coherence and divided attention on cognitive processing demands.

4. Conclusion

We conclude that perceptual coherence readily predicts the perception of dissonance across a variety of psychoacoustic manipulations. Based on our results, however, the relationship between consonance-dissonance and cognitive processing demands remains unclear and requires further empirical investigation. Our results provide what we hope are compelling insights regarding the consideration of multidimensional psychoacoustics in dissonance research.

References

Bodner, E., Gilboa, A., & Amir, D. (2007). The unexpected side-effects of dissonance. Psychology of Music, 35(2), 286-305.

Bonin, T. (2014). The source dilemma: A psychophysical account of dissonance perception in music (Undergraduate dissertation, McMaster University Hamilton).

Bregman, A. S. (1994). Auditory scene analysis: The perceptual organization of sound. MIT press.

Demany, L., Erviti, M., & Semal, C. (2015). Auditory attention is divisible: Segregated tone streams can be tracked simultaneously. Journal of Experimental Psychology: Human Perception and Performance, 41(2), 356.

Masataka, N., & Perlovsky, L. I. (2013). Cognitive interference can be mitigated by consonant music and facilitated by dissonant music. Scientific Reports 3, Article number: 2028 (2013) doi: 10.1038/srep02028.

Acknowledgements

This work is funded by an NSERC CGS-M grant to TB and an NSERC Discovery Grant to DS.

Performance On 2 Tasks Of The Airs Test Battery Of Singing Skills In Persons With Cochlear Implants

Derek Hughes, Bing-Yi Pan, Annabel J. Cohen

Cochlear implants aim to maximize speech understanding, but to what extent do implants support musical activities? Here we focus on singing in 4 adults (age 38, 50, 59, 73 years) with cochlear implants who carried out the AIRS Test Battery of Singing Skills (ATBSS) (Cohen, et al., 2009). The ATBSS taps 7 singing skills including singing a familiar song, singing a favourite song, learning a new song, repeating melodic elements, and creating a new song. It also measures several verbal skills. The test was administered twice to 2 individuals and once to the others. Participants also completed a biographical questionnaire. One focus of analysis was singing the familiar son–in particular the 10 repetitions of the key-note (tonic). Variability (SD) in pitch of the 10 sung tonics ranged from 1.37 to 5.88 semitones, higher than for persons with normal hearing (0.55 semitones; obtained in a separate study of 20 normal-hearing non-musicians). Mean error of the sung tonic ranged from 0.29 to 8.06 semitones, for best to poorest performance (normal hearing mean was 1.58). Analysis of contour accuracy of 3-note sequences (triads) revealed a range of systematic success or failure across participants. Music played a role in the lives of each participant- one currently sang in a choir; another played guitar. The gap between the ability to sing with normal accuracy and the desire to make music highlights the importance of additional research to characterize further the singing abilities of those who have cochlear implants. Whether the current mismatch problems could be overcome through training remains an important question. Cohen, A. J., Armstrong, V., Lannan, M. & Coady, J. (2009). A protocol for cross-cultural research on acquisition of singing. Neurosciences and Music III-Disorders and Plasticity: Annals of the New York Academy of Science, 1169, 112-115.

Identification Of Thin Slices Of Music By University Students In Pei

Jessica Lynn McKellar, Annabel Cohen

Young people can often identify the artist and title of a large repertoire of popular music, but how much acoustical information is needed to do so? Krumhansl (2010) showed with audio excerpts (thin slices of music) as short as 400 ms, students performed at a level of 25% correct.The 23 students in her study however were from Cornell University and, though they were not majoring in music, had approximately 10.3 years on average of total formal music instruction. The present study explored whether this identification ability would generalize to students who had much less music training and were from a small Eastern Canadian university. Participants were 17 students (mean age 23.71 years, SD = 4.28) who had only 2.34 (3.5) years musical training. Participants were presented with 60 400-ms excerpts of 30 songs (each with 2 excerpts) popular between 1960 to 2015. They were asked to identify the artist name and title of the work. Sixty-one per cent of the songs from the Krumhansl set were included. After this, they were presented with 15-sec excerpts of these 30 pieces and asked the same questions about them, to determine general knowledge of the songs. For the 15-sec excerpts, on average 72.94% (0.12) of the titles and 57.25% (0.17) of the artists were identified. Titles were significantly more easily identified than were artists, $F(1,16)=9.32$, p less than.01. For the very short excerpts, the mean title identification was 33.24% (.10) and artist identification was 29.51% (.12), both of which were significantly different from 0, $t(16) = 13.14$ and 9.93 respectively, p's less than .001. This study replicates the basic finding of Krumhansl for this population and serves to illustrate the young person's enormous storage capacity for information about popular music and its accessibility.

PHYSICAL ACOUSTICS / ULTRASOUND - ACOUSTIQUE PHYSIQUE / ULTRASONS

OPTICAL COHERENCE TOMOGRAPHY FOR CLINICAL OTOLOGY

Dan MacDougall [*1], **Thomas Landry** [†1,2], **Manohar Bance** [‡1,2], **Jeremy Brown** [§1], **and Robert Adamson** [¶1]
[1]School of Biomedical Engineering, Dalhousie University, Halifax NS, Canada
[2]Department of Surgery, Dalhousie University, Halifax NS, Canada

1 Introduction

Pathology of the human middle ear (ME) remains a challenging diagnostic problem for clinical otologists. The ME, an air-filled cavity located medially to the tympanic membrane (TM) at the end of the external ear canal, contains the chain of millimeter-sized, bony ossicles (malleus, incus, stapes) that conduct sound energy to the inner ear. Pathology of the middle ear causes conductive hearing losses which can arise from a variety of disorders (e.g. otosclerotic fixation, trauma, cholesteatoma, otitis media, erosion), many of which can be treated surgically. However, the currently available suite of imaging and diagnostic tools usually requires exploratory surgery through removal of the TM and direct microscopy of the ME for confident diagnosis before a treatment solution is chosen.

Optical Coherence Tomography (OCT) is an interferometric optical imaging technique that can be used to produce depth resolved images in biological samples [1] and can be thought of as the optical analog to ultrasonic imaging. By probing a tissue sample with a spatially coherent light source and interfering the the back-scattered light with a reference beam, both the structure and dynamics of the sample can be inferred using a variety of interferometric techniques (i.e. phase-sensitive forms of time-domain OCT, spectral-domain OCT, swept-source OCT). Since the TM is thin ($\approx$ 100-300 μm) and translucent, infrared light is able to penetrate it and allow non-invasive imaging of the middle ear with OCT. The potential for OCT in the middle ear has been previously recognized, for example in [2–4], but the technology has yet to be optimized for clinical use in otology where it could provide diagnostic capabilities unavailable with any existing tools.

2 Methods

2.1 SS-OCT engine design

We built a phase-sensitive swept-source OCT (SS-OCT) system and optimized it for real-time imaging of the human middle-ear, *in vivo*, with Doppler vibrometry capabilities. Our design is shown diagrammatically in Figure 1. Our infrared light source is an akinetically-tuned (i.e. no moving parts) swept-laser (SLE-101, Insight Photonics, 40 nm bandwidth centered at 1550 nm). A custom-built Mach-Zehnder fiber interferometer directs 90 % of the light to the sample where it is focused to a narrow beam, and recombines the collected backscatter with the remaining 10% for balanced detection. A custom, FPGA-based controller board was designed to synchronize all timing aspects of the system, including the control of two-axis scanning galvanometer mirrors and the generation of an acoustic frequency signal used to drive a speaker and excite the sample during acquisition.

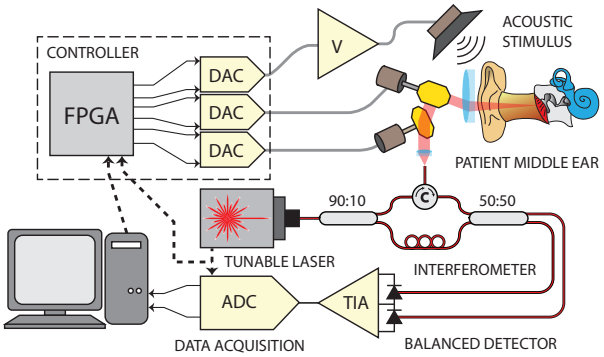


Figure 1: Diagram of the swept-source OCT system designed and built for real-time clinical imaging of middle ear anatomy and function

In our system, the laser performs linear optical-frequency sweeps at a repetition rate of $\approx$100 kHz. The interference pattern generated during each sweep is sampled ($\approx$400 MSPS @12-bit resolution) and by the nature of SS-OCT, the sample's reflection profile along the beam, called an A-line, is computed as its Discrete-Fourier-Transform (DFT). Structural 2D and 3D brightness-mode (B-Mode) images of the sample are generated from the magnitude of a number of A-lines. Simultaneously, since changes in position of a reflecting structure produce proportional phase shifts in its back-scattered light, sample dynamics along the beam path are extracted from the phase differences between consecutively acquired A-lines. In our system, by applying a pure-tone acoustic stimulus to the sample, the induced motions and resulting phase changes impart a frequency signature that is exploited to perform Doppler-Mode Vibrography and extract functional information about the response of middle ear structureto sound stimulus using Fourier analysis.

In order to make the imaging engine suitable for clinical use, real-time processing, display and control were necessary. A graphical user interface (GUI) written in C++ using the Qt framework was developed to interface with the system hardware and manage patient data. Real-time signal processing of the interferometric data was implemented on a GPU with Nvidia's parallel programming platform, CUDA, with cuFFT and several custom-written execution kernels that perform raw-data conditioning, windowing and dispersion compensation, B-Mode averaging, and Doppler-Mode calculations.

*. dan.macdougall@dal.ca
†. tglandry@dal.ca
‡. m.bance@dal.ca
§. j.brown@dal.ca
¶. rob.adamson@dal.ca

2.2 Optimization for clinical use in patients

The clinical adoption of OCT into otology will be made easier if it can provide new diagnostic capabilities without disturbing the existing clinical workflow. In a modern clinic, typically a patient is asked to lie down while a surgical microscope is used to look down the ear canal for visible abnormalities. A speculum is used to straighten the natural curvature of the ear canal and gain line-of-sight to the TM and ME. Our OCT imaging system was designed to provide images without disrupting this workflow.

We outfitted an existing articulated surgical microscope mount with custom-designed OCT scanning optics using off-the-shelf parts. The scanner images through a rigidly mounted speculum of standard size with a 4.5 mm aperture. A field-of-view wide enough to accomodate the entire TM's width, $\approx 10\ m$m, is achieved by intentionally introducing fish-eye distortion to the optics' scan pattern.

The OCT engine described above was integrated into a mobile cart to operate the system alongside the articulated scanning head. The entire assembly is located in a dedicated testing room within the ENT clinic over a gurney for patients to lie on during examination.

3 Results

3.1 Preliminary imaging and validation

Figure 2 shows a screen capture from the GUI during real-time imaging (20 frames-per-second) in the clinic. The large field of view in both the lateral and depth directions ($\approx$ 12 mm x 10 mm on center, respectively) is a key feature that makes our system well-suited to the middle ear.

While the Doppler vibrographic imaging capabilities are still being integrated into the GUI as a user-friendly real-time diagnostic tool, we have previously confirmed our system's ability to measure nanometer-scale vibration of the ossicles in bench-top experiments using human cadaver specimens.

4 Discussion

Across the several design iterations and cadaver experiments leading to the current clinical system, the benefit of wide-field imaging became apparent. While excellent images are easily obtained using simple optics in cadaver specimens that have had the bony ear canal drilled away to expose the TM, the resulting limited field-of-view in intact ears makes the identification of anatomical structures and landmarks difficult. Our wide-field optics' ability to provide images of the entire middle ear simultaneously has met with favour from our clinical partners who have found operating the system on volunteers from within our research group to be very natural. We expect the technology to allow better decision making and surgical treatment planning, to provide better post-operative monitoring, and to improve the clinican's ability to recommend against surgery altogether where it is counterindicated.

We are currently in the process of training our clinical partners on the use of the system and ethics approval has been obtained to commence imaging on patient volunteers.

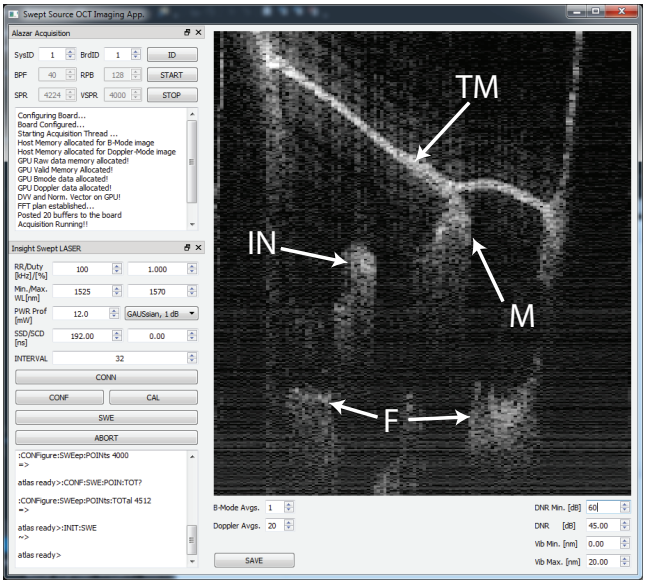


Figure 2: Screen capture from the real-time imaging GUI during *in vivo* imaging. The entire tympanic membrane (TM), the malleus handle (M), the end of the long process of the incus (I), and the floor of the middle ear (F) are all captured in the same image acquired down the intact ear canal

5 Conclusions

We have presented our novel phase-sensitive SS-OCT middle ear imaging system in its final design stages as we prepare to deploy it in clinical trials. We have also outlined the role OCT will likely play within the existing clinical workflow and where it is likely to impact middle ear diagnostics.

Acknowledgments

We would like to thank Colibri Technologies for providing the computer cart used to house our system in-clinic and NSERC and ACOA for project funding.

References

[1] J.A. Izatt and M.A. Choma. *Optical Coherence Tomography Technology and Applications*, chapter Theory of Optical Coherence Tomography. Springer, 2008.

[2] Pitris C, Saunders KT, Fujimoto JG, and Brezinski ME. High-resolution imaging of the middle ear with optical coherence tomography : A feasibility study. *Archives of Otolaryngology–Head & Neck Surgery*, 127(6) :637–642, 2001.

[3] Ernest W Chang, Jeffrey T Cheng, Christof Röösli, James B Kobler, John J Rosowski, and Seok Hyun Yun. Simultaneous 3d imaging of sound-induced motions of the tympanic membrane and middle ear ossicles. *Hearing research*, 304 :49–56, 2013.

[4] Hrebesh M Subhash, Anh Nguyen-Huynh, Ruikang K Wang, Steven L Jacques, Niloy Choudhury, and Alfred L Nuttall. Feasibility of spectral-domain phase-sensitive optical coherence tomography for middle ear vibrometry. *Journal of biomedical optics*, 17(6) :0605051–0605053, 2012.

ULTRASONIC DRY COUPLING THROUGH TISSUE

Jeremy Norman *[1]
[1]SENSE Lab, Dalhousie School of Biomedical Engineering

1 Introduction

Ultrasonic transcutaneous energy transmission (UTET) is a promising new technology for delivering electrical power to active biomedical implants. A key piece of technology required to make UTET devices viable is a "dry" acoustic coupling that can be worn daily for long periods of time by a patient. The dry coupling must provide a low-reflectance, low-loss, transmission pathway for acoustic energy. For other ultrasound applications such as diagnostic imaging, a coupling gel is typically used to acoustically interface a transducer with the patient. However, in a more permanent application like UTET coupling gel is likely dry out or wash away over time and so a solid-state solution is needed.

In order to reduce reflectance, the dry coupling must have an acoustic impedance as close to that of tissue as possible since this difference between acoustic impedances of two material at an interface is responsible for reflected sound according to equation 1.

$$R = \frac{z_2 - z_1}{z_2 + z_1} \quad \textbf{[1]}$$

Acoustic impedance is the product of a material's density and its speed of sound.

$$z = \rho * c \quad \textbf{[2]}$$

By manipulating either of these two properties, the acoustic impedance of a material can be controlled.

2 Method

Samples to be tested were manufactured from Dow Corning Sylgard 184. This kit provides a base with a crosslinker. The idea is to manipulate the acoustic impedance by varying either the speed of sound or the density of the sample. The more crosslinked the silicone elastomer, the higher the sample's bulk modulus which will increase the speed of sound in the material according to the Newton-Laplace equation.

$$c = \sqrt{\frac{K}{\rho}} \quad \textbf{[3]}$$

Where c is the speed of sound, K is the adiabatic bulk modulus, and ρ is the density. Density is increased by mixing ZnO powder into the silicone during the curing process.

* jeremy.norman@dal.ca

Samples were made with the kit's suggested 10:1 base:crosslinker ratio, as well as a 5:1 and 2:1 ratios. The loaded samples were produced by adding the needed ZnO mass to achieve the desired density. Samples with targeted densities of 1500, 1300 and 1100 kg/m^3 were produced. This results I samples with acoustic impedances ranging between water and unloaded silicone.

The samples were cast in a simple cylinder with tape on the inside to assist in delamination. In order to produce a flat surface on the top of the sample a plug was immersed in the silicone, otherwise surface tension in the silicone would form the top side into a convex surface.

In order to analyse the samples, a pulse-echo technique was used to determine the amplitudes of echoes reflected off the sample's front and back face and the time of flight between echoes. The pulse was created using a signal generator, sent into the transducer, through deionized water into the sample, reflected off a glass slide, amplified, and then to an oscilloscope.

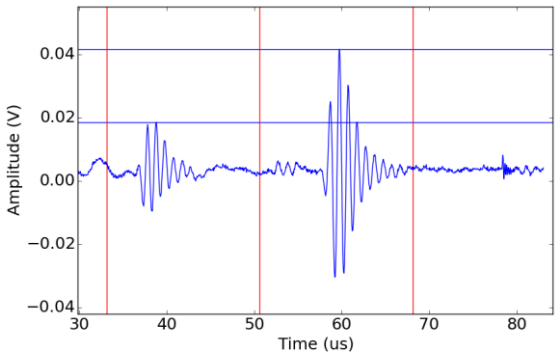


Figure 1: Front and Back Face Echoes of ZnO Loaded Sample with Maxima

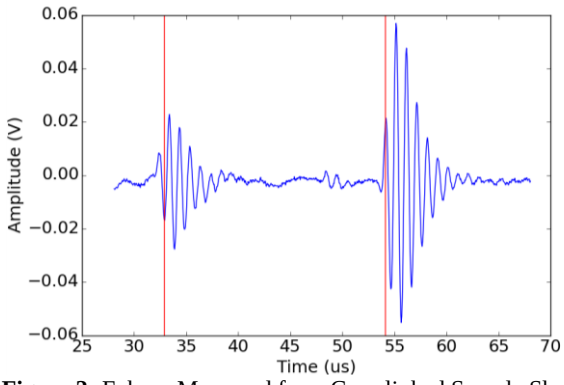


Figure 2: Echoes Measured from Crosslinked Sample Showing Points Used to Determine Time of Flight

The pulse must travel through the length of the material twice before the 2nd echo reaches the transducer behind the first echo. The speed of sound can be found by dividing twice the length of the sample by the difference in time between the start of each echo.

3 Results and Discussion

The speed of sound of the samples with varying amounts of crosslinker is shown in Figure 3. Increasing the amount of crosslinker has a negligible effect on the speed of sound. IT is likely that the kit's suggested 10:1 ratio of base to crosslinker is the point where the silicone becomes fully crosslinked and no amount of additional crosslinking agent will further stiffen the material. It is likely that reducing the amount of crosslinker will result in a less stiff silicone, however, this is not useful for increasing acoustic impedance.

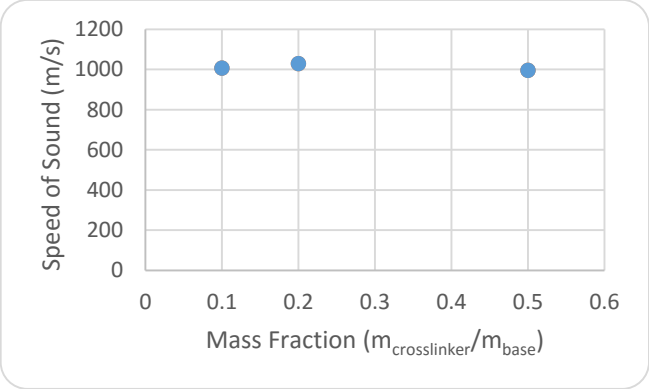


Figure 3: Speed of Sound in Silicone with Increasing Amounts of Crosslinker

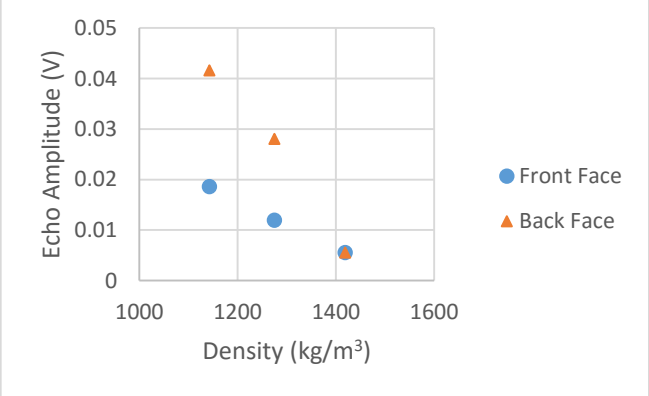


Figure 4: Maximum Amplitude of Front and Back Face Echo vs Composite Density

Figure 4 shows the maximum amplitudes from the echoes off the front and back faces of the samples loaded with ZnO. The back face echoes are much stronger because the sample was placed on a glass slide which has a very high acoustic impedance and therefore reflected nearly all the sound that came from the sample which had a much lower impedance. The sample and the water had much more closely matched acoustic impedances which is why a much weaker echo was reflected back to the transducer. It is evident that, as the samples increase in density due to ZnO loading, the reflected pulse off the front face becomes weaker. This suggests that ZnO loading increases does in fact increase the acoustic impedance as the samples with more loading are more closely matched to the water.

Figure 4 also shows a drawback however, the more closely matched, higher density samples should have increasingly strong echoes off the back face as more sound is transmitted through rather than being reflected back. The reason the back face echo is decreasing is due to attenuation introduced by the addition of the ZnO powder. The small granules of ZnO are large enough to interfere with the sound and create losses due to internal reflections within the sample. This puts an upper limit on the thickness that the dry coupling layer can be as well as the amount of loading that can be done.

4 Conclusion

Silicone as a material for coupling ultrasound to tissue is a viable material. In order to optimise the material to get the maximum amount of sound energy delivered into living tissue, losses from reflectance and attenuation must be minimized. This work has shown that for Sylgard 184, the maximum speed of sound is approximately 1000m/s and this cannot by surpassed by adding more crosslinking agent during curing.

It has also been demonstrated that ZnO powder can increase the acoustic impedance of the silicone by forming a composite material with a higher density than the silicone alone. There is a drawback in that the composite is much more attenuating and this attenuation increases as the concentration of the ZnO power increases in the composite. Attenuation losses can be mitigated by making the coupling layer as thin as possible as attenuation is also a function of the length of material that the sound must travel through. This suggests that an optimal solution may exist consisting of a thin layer of ZnO loaded silicone that balances losses due to reflection with those due to attenuation.

TIME-COURSE CHARACTERIZATION OF HIGH-INTENSITY FOCUSED ULTRASOUND EXPOSURE USING B-MODE ULTRASOUND IMAGING

Matthew Jahns [1], Rob Adamson [1]
[1]Department of Biomedical Engineering, Dalhousie University, Halifax, Nova Scotia, Canada

1 Introduction

High-Intensity Focused Ultrasound (HIFU) is an emerging non-invasive surgical technique. HIFU systems are generally designed to deliver a fixed amount of energy to tissue, but because the tissue thermal response is variable across patients and tissue type, this approach can often lead to underdosing or overdosing. In order to improve the dosing precision in HIFU we are investigating real time monitoring of the energy deposition process through high-frequency diagnostic ultrasound imaging. Ultimately the development of monitoring techniques will allow real-time adjustment of energy to obtain optimal treatment. We present results from an exploratory study into the transient response of HIFU exposure in phantoms and animal tissue. The transient response observed in the B-mode ultrasound images is due to a combination of the thermal-acoustic lens effect, thermal expansion of the medium and cavitation. These preliminary experiments will inform the development of new image analysis techniques for determining dosage received in tissue during HIFU exposure.

2 Method

2.1 HIFU System

HIFU deposition was performed in phantoms and chicken breast using a commercial HIFU system across a range of energy levels at a frequency of 4MHz. HIFU deposition points had an exposure duration of 40ms and a focal spot diameter of 100µm. Energy levels of 0.75J and 1.2J were used, corresponding to an intensity level of approximately 60kW/cm^2 and 95kW/cm^2 respectively.

2.2 Imaging system

Imaging was performed with a Visual Sonics Vevo 2100 50 MHz ultrasound imaging system with 40µm resolution and a maximum B-mode frame rate of 1000Hz. The imaging plane was taken at an angle 105 degrees from the axis of HIFU deposition.

2.3 Overview of Experiments

Temperature recordings were obtained in the region surrounding the HIFU focal spot during deposition in a polybutadiene rubber phantom material. Recordings were obtained at a sampling interval of 35ms using a 200µm diameter, type T thermocouples (Omega HYP-0).

To investigate the effect of a HIFU focal spot on the resulting ultrasound image, a 21-gauge needle was inserted into the polybutadiene rubber phantom material such that its edge was approximately 2mm distal to the HIFU focal spot in the imaging plane. The effect of HIFU deposition on the appearance of the straight and stiff metal needle as observed in the acquired ultrasound image was analyzed.

Lastly, HIFU deposition was performed in chicken breast. The acquired time-course images of the HIFU focal spot were analyzed using ImageJ software to obtain intensity decay plots of characteristic features.

3 Results

3.1 Thermocouple Measurement

An average of two thermocouple measurements obtained in the polybutadiene rubber phantom material before, during, and after HIFU exposure at 1.2J is shown in Figure 1. A peak temperature greater than 90°C was reached at the end of the 40ms HIFU exposure period. After HIFU was turned off, measured temperatures decreased as a result of diffusion.

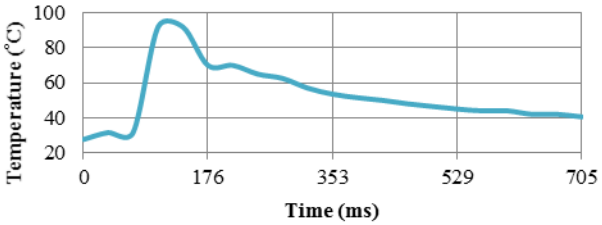


Figure 1: Temperature profile of HIFU deposition in phantom material with values prior to, during and after exposure as obtained by thermocouple measurements at the HIFU focal spot.

3.2 HIFU Deposition in Phantom Material

The time course of the ultrasound images indicated that shadowing distal to the focal spot temporarily prevented the needle from being observed. The extent of this shadowing was dependent upon the energy level. In Table 1, the maximum deflection of the needle was measured when its edge reappeared in the image.

Energy Level (J)	**0.75**	**1.20**
Shadow Width immediately post HIFU (mm)	1.0	1.45
Time of Max Deflection post HIFU (ms)	676	907
Max Deflection Width (mm)	2.2	2.32
Max Deflection Height (mm)	0.12	0.18
Deflection Height 4 seconds post HIFU (mm)	0.03	0.05

Table 1: Resulting deflection of the embedded needle as observed in the ultrasound image following HIFU deposition

The results indicate that changing the energy level of deposition will alter the observed curvature of the needle. The observed deflection of the needle is not permanent and 4 seconds post HIFU the needle has almost returned to its original straight line form. An abbreviated time course of the ultrasound images is shown in Figure 2.

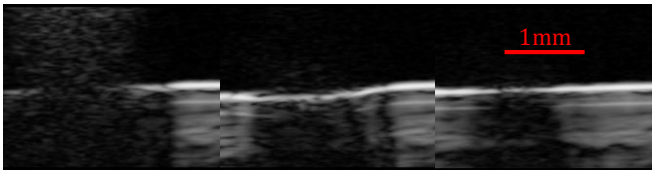


Figure 2: Effect of HIFU deposition in the phantom material on a straight metal needle: (left) immediately post HIFU, (middle) time of max deflection, (right) 4 seconds post HIFU

3.3 HIFU Deposition in Chicken

The deposition of HIFU in chicken breast at an energy level of 1.2J resulted in the formation of a hyperechoic region and signs of shadowing at distal locations in the captured B-mode ultrasound images. As shown in figure 3, these characteristic features were transient over time.

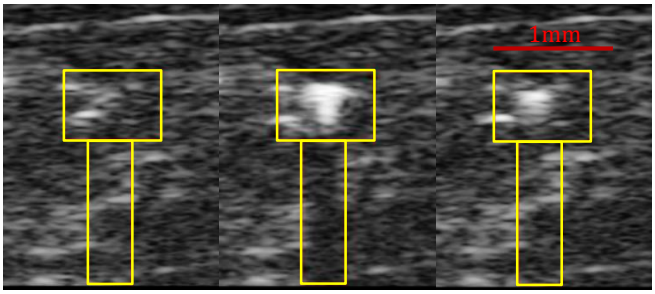


Figure 3: HIFU deposition in chicken breast as observed with B-mode ultrasound images: (left) prior to HIFU, (middle) immediately post HIFU, (right) 0.5 seconds post HIFU. Top box shows hyperechoic region and bottom box shows shadowing.

Referring to figure 4, immediately after HIFU exposure the mean greyscale value of the hyperechoic region in the ultrasound image was 55% greater than pre-exposure levels. A higher greyscale value corresponds to a whiter area in the ultrasound image. Within 10ms after HIFU exposure, the mean reduced to a steady state that was 30% above pre-exposure values. The tissue in the hyperechoic region appeared to shrink and contract as the tissue cooled, which could be the effect of thermal expansion in the tissue. The mean greyscale value of the shadow region rose logarithmically towards pre-exposure levels. In the first 40ms post exposure, the mean of the shadow region increased from 65% to 90% of pre-exposure greyscale values.

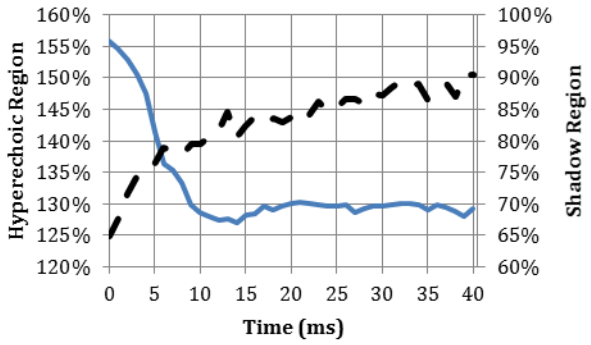


Figure 4: Mean greyscale value of regions in ultrasound images following HIFU exposure in chicken breast relative to pre-exposure values. Solid line (left axis) represents hyperechoic region and dashed line (right axis) represents shadow region.

4 Discussion

The high temporospatial resolution of the Vevo 2100 allows for a unique time-course characterization of the material response to HIFU exposure. Gudar et al [1], has also performed high frequency ultrasound imaging of HIFU exposure, however they only were able to achieve a M-Mode frame rate of 100Hz. The achieved B-Mode frame rate of 1000Hz in our study revealed transient behaviour within a 40ms observation window post HIFU exposure.

Other studies commonly use lower intensities maintained for longer exposure periods [2]. A threshold intensity of 1kW/cm^2 produced a hyperechoic spot in less than 1 second. Increasing the intensity to 8 kW/cm^2 for 5 seconds formed visible lesions within tissue. The effect of using higher intensities (>60 kW/cm^2) for shorter durations (40ms), as in our experiments, still remains to be fully investigated. It is believed that the hyperechoic region is from cavitation producing vapour bubbles in the material [2]. These bubbles may also cause the temporary shadowing observed throughout our experiments.

It is important to note the difference in time scale between the temperature and RF intensity decay plots. At present the quick transient behaviour of the hyperecho observed in the chicken breast experiment cannot be due solely to a reduction in temperature. Our measured peak temperatures exceeded 90°C near the HIFU focus and may indicate thermal cavitation. It is possible that following HIFU exposure, the quick collapse of microbubbles formed from cavitation could cause the observed changes in the first 40ms of ultrasound images.

In the phantom material, the HIFU focal spot appears to act like a converging lens due to the distortion of the embedded needle. The local increase in temperature modifies the speed of sound and the local refractive index of the medium. If the heated region is spherical, one could adapt optical lens equations to compute a focal length of the resulting thermal-acoustic lens. By analyzing the observed needle deflections, the radius of curvature may align with the focal length of the thermal lens. This lensing effect could provide a method of monitoring HIFU dosage.

5 Conclusion

This exploratory study has identified phenomena that could be useful in real-time monitoring of high-intensity focused ultrasound with high-speed B-mode imaging. The characteristic features observed were the thermal-acoustic lens effect, hyperechoic decay and shadow decay rates. With the development of appropriate image analysis techniques these may be useful in measuring HIFU dose.

References

[1] Gudur et al. High-frequency rapid B-mode ultrasound imaging for real-time monitoring of lesion formation and gas body activity during high-intensity focused ultrasound ablation. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 59:1687-1699, 2012.

[2] Vaezy et al. Real-time visualization of high-intensity focused ultrasound treatment using ultrasound imaging. *Ultrasound in Medicine & Biology*, 27:33-42, 2001

SIMULATION OF FRESNEL BASED BEAM FOCUSING AND STEERING FOR A CROSSED ELECTRODE ARRAY

K. Latham [*1], R. Zemp [†2], J. Brown [‡1]
[1]School of Biomedical Engineering, Dalhousie University, Halifax, Canada
[2]Electrical and Computer Engineering, University of Alberta, Edmonton, Canada

1 Introduction

Fresnel based beamforming has been investigated for use as a steerable lens for 3D imaging arrays. Acoustic waves, like light, can be focused using a Fresnel lens or zone plate approach. Fresnel zone plates are capable of producing a tight focus, especially when using large aperture. A zone plate is made up of rings or strips of alternating transmissive and opaque regions. The waves diffract around the opaque zones and, because of the specific spacing, constructively interfere around the focus. In a phase zone plate both zones transmit the wave, however, there is a phase reversal for every other zone. This type of plate has the advantage of greater efficiency [1] and is a good approach for passive ultrasound focusing.

Implementing a Fresnel zone plate approach in an ultrasound transducer requires control of the pulse polarity. Arrays built on either electrostrictive ceramics or CMUTs are appropriate for this approach because the response is controlled by a DC bias. Electrostrictive ceramics such as PMN-PT (lead magnesium niobate-lead titanate) ceramic can be used as the array substrate in place of conventional piezoelectrics. This type of material is only piezoelectrically active while a bias voltage is applied and the response is tunable with the amplitude of the bias voltage. When no voltage applied the transducer, the response is negligible, and when a DC bias is applied, the phase of the acoustic wave produced is quantized to either +90 or -90 degrees, depending on whether the bias is positive or negative. Array elements defined on an electrostrictive substrate can therefore be addressed individually and in parallel. This allows for reconfigurable Fresnel zone plates to be created by varying biasing patterns.

Typical linear-phased arrays use an acoustic lens to improve the elevational slice resolution (thickness) of the image. If an elevation lens could be reconfigured to steer to moderate angles a 3D volumetric image could be captured without adding additional beamforming channels and only moderately increasing the number of electrical connections. This can be accomplished by replacing the mechanical acoustic lens with an electrically reconfigurable lens, which approximates a Fresnel lens. A Fresnel lens is created by applying the appropriate pattern of positive and negative biases along the elevation direction of the array which determine the polarity of the pulses from each element. In implementation, the array would have a set of bottom electrodes running orthogonal to the top electrodes, similar to a crossed electrode array [2]. The bottom electrodes provide the active lens control in the elevation plane and the result is an array that requires approximately 2N (or ~128) electrical connections. The alternative approach to capturing 3D image volumes is to use a 2D array with N^2 (or ~4096) elements. The low beamforming complexity and minimal electrical connectivity used in our approach provide an enormous practical advantage over conventional 2D array designs.

2 Methods

The bias value for each element is calculated by considering the geometric path length between the element and the focus. The relative phase delay for that element is given by (1).

$$\varphi = 2\pi \left[z - \sqrt{x^2 + y^2 + z^2} \right] / \lambda \qquad (1)$$

Where x, y and z are the coordinates of the desired focus relative to the array element and λ is the wavelength of the centre frequency of the excitation pulse in the medium. The sign of the bias (S_{bias}) is given by (2).

$$S_{bias} = sign[mod(\varphi + offset, -2\pi) - \pi] \qquad (2)$$

Here we are approximating the relative phase delay for each element as the portion that falls within a single wavelength (quantized as shifted by $+\pi/2$ or $-\pi/2$ radians). This models the purely transparent regions and the pulse inverted regions of the zone plate. An offset phase can be added in the calculation that shifts the reference phase of the center element [3], [4]. Consequently, there is not one unique Fresnel pattern for a given focal point. The pattern can be chosen to optimize different beam shapes (e.g. main lobe width, side lobe level, sensitivity).

Field II [5], [6] was used to simulate the Fresnel aperture as a steerable lens. The Fresnel aperture was simulated by setting the apodization values to correspond to the sign of the bias for each element as described above. These simulations were completed for a 40MHz array with λ pitch, 64 azimuthal elements and 64 elevational elements. The peak absolute pressure is plotted in the radiation patterns.

3 Results

The most promising results were obtained by switching the side of the array that receives the bias and the side of the array that carries the signals between transmit and receive events. This technique is advantageous since it produces equivalent beam profiles in both the azimuthal and elevational directions. In practice, the Fresnel aperture will focus in azimuth on transmit while the elevational elements are beamformed traditionally. Between transmit and receive events the signals switch sides. The biases are then applied in elevation and dynamic receive beamforming can be completed in azimuth. A two-way focus is achieved.

[*] kate.latham@dal.ca [‡] j.brown@dal.ca
[†] rzemp@ualberta.ca

Representative radiation patterns and beam profiles simulating this technique are shown in Figure 1. The results were compared with a 4096 element (64x64) fully sampled and beamformed array. The -6dB beamwidths using the Fresnel aperture were 102, 116 and 140 µm when steered to 0, 15 and 25 degrees respectively. For comparison, the 4096 element array had simulated beamwidths of 89, 92 and 96 µm for the same set of steering angles. The side lobe levels raised by approximately 15dB using the Fresnel approach. Sensitivity remained mostly unaffected as the received energy amplitude is comparable.

4 Discussion

The Fresnel aperture simulations show that a two-way focus can be achieved in each imaging plane with beamwidths comparable to a conventionally beamformed 2D array. This steerable lens technique could generate 3D images with only 64 signal channels and 64 bias channels. The array performance in the simulated angle range is comparable to that of a conventional 2D array that requires 4096 channels.

Switching the bias dimension with the signal dimension between transmit and receive minimizes the negative effect of the approximations made with the Fresnel aperture. Using this approach the beamwidths in azimuth and elevation are equivalent. Additional improvements can be made by carefully choosing the Fresnel pattern. As mentioned previously, a Fresnel pattern for a given focal location is not unique. The beam characteristics can be improved for a given focal point by choosing the optimal Fresnel pattern.

The theory for the Fresnel zone plate is based on continuous wave operation. There are challenges when applying the technique to pulsed ultrasound imaging. The most obvious is a degradation in pulse bandwidth at the expense of lateral resolution at the focus. This is especially detrimental when steering to wide angles where the path length differences are larger. A novel technique has been developed that preserves the pulse bandwidth at larger steering angles, correcting for the beamforming errors inherent in the Fresnel approach.

5 Conclusion

Fresnel based focusing and steering is an interesting approach to ultrasound imaging in three dimensions. The simulations in this study show the potential performance of a steerable lens technique that limits the complications in both array fabrication and beamforming electronics. The planned future work for this project includes validating these simulations with experimental results using an electrostrictive array.

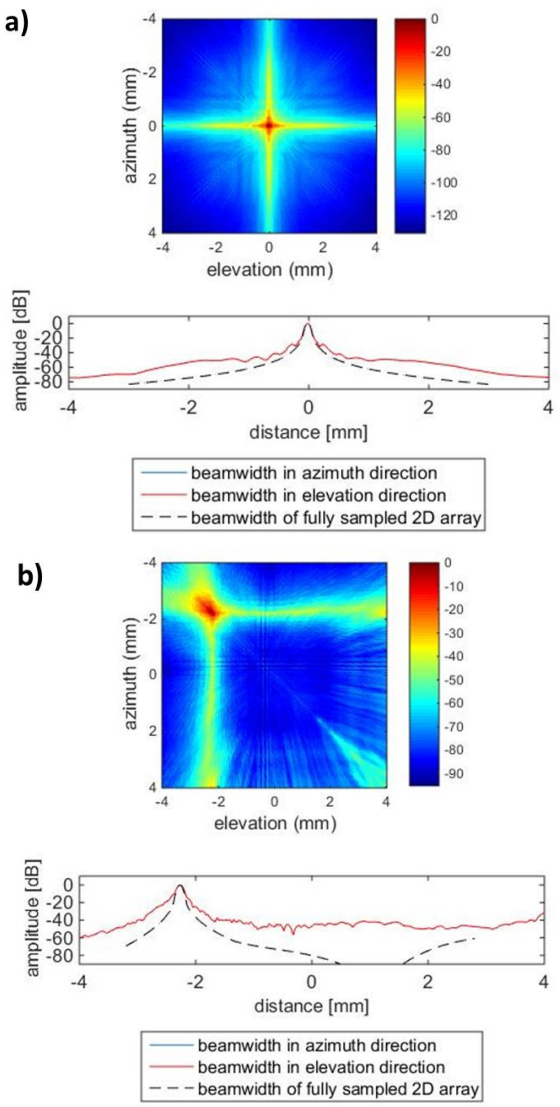


Figure 1: Two-way radiation pattern in the elevation-azimuth plane and beam profile for the array using a steerable Fresnel aperture focused a) on axis and b) to 20 degrees in both planes.

References

[1] H. D. Hristov, *Fresnel Zones in Wireless Links, Zone Plate Lenses and Antennas*. Boston: Artech House, 2000.

[2] C. E. Morton and G. R. Lockwood, "Theoretical assessment of a crossed electrode 2-D array for 3-D imaging," in *Ultrasonics, 2003 IEEE Symposium on*, 2003, pp. 968–971.

[3] C. Emery, R. Lahman, and J. Kook, "Therapeutic and diagnostic electrostrictor ultrasonic arrays," WO 2008/033528 A22008.

[4] C. Daft, P. Wagner, S. Panda, and I. Ladabaum, "Elevation beam profile control with bias polarity patterns applied to microfabricated ultrasound transducers," in *IEEE Symposium on Ultrasonics, 2003*, 2003, pp. 1578–1581.

[5] J. A. Jensen, "Field : A Program for Simulating Ultrasound Systems," *Pap. Present. 10th Nord. Conf. Biomed. Imaging. Pap. Publ. Med. Biol. Eng. Comput.*, vol. 34, no. Supplement 1 Part 1, pp. 351–353, 1996.

[6] J. A. Jensen and N. B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 39, no. 2, pp. 262–267, 1992.

Ultrasonics Investigations Of Molecular Interactions In Polymer Solutions Of Poly Propylene Glycol- 400 And N-Alkanols

Devinder Pal Singh, Arun Upmanyu

Using experimental ultrasonic velocities and densities of binary solutions of Poly propylene glycol (PPG 400) in ethanol, 1-propanol and 1-butanol at 30°C, various acoustical parameters such as molar volume, available volume, free volume, van der Waal's constants, Wada's constant, internal pressure, coefficient of viscosity, surface tension and cohesive energy have been determined. Several thermo-dynamical parameters such as isothermal compressibility, ratio of specific heats, collision factor, space filling factor, geometrical volume, thermal expansion coefficient, index of refraction, Eykman's constant, molar refraction, relaxation time and pseudo-Gruneisen parameter have been evaluated. The obtained results are compared with experimental values wherever available. The variation of these parameters with change in mole fraction of the PPG 400 in the binary solution, is investigated to understand the nature, type and strength of molecular interactions prevalent in these systems. For several acoustical and thermo-dynamical parameters, excess parametric values and percentage deviations have also been studied to confirm the results arrived at earlier. Using various empirical relations such as ideal mixing relation, Nomoto's relation, Van Dael and Vangeel relation, impedance dependence relation, Rao's specific velocity and Jungie's relation, ultrasonic velocities are calculated and the validity of these relations is assessed.

PSYCHOLOGICAL ACOUSTICS - PSYCHOACOUSTIQUE

THE EFFECT OF VOCAL EMOTION IDENTIFICATION ON WORD REPETITION AND RECALL ACCURACY IN NOISE

Sylvia Mancini[1], Kate Dupuis[1] & M. Kathleen Pichora-Fuller[1]
[1]Department of Psychology, University of Toronto Mississauga

1 Introduction

In everyday life, facts and events associated with strong emotions are better remembered than those lacking rich emotional content [1]. In particular, information that is associated with the emotion of fear has been found to be highly memorable [2]. Little is known about whether or not there is a similar effect of emotion on memory for speech spoken with emotion, in particular fear. In a prior study, we investigated if there would be such an effect if the semantic content of the spoken information is neutral, but it is spoken with vocal emotion and heard in noise. In the current study, we added an emotion identification task to investigate if the effect of vocal emotion on memory would be enhanced. We tested young adults' recognition and recall of semantically neutral words spoken to portray fear, sadness, neutral and pleasant surprise in a noisy background in two conditions, one with and one without an additional vocal emotion identification task. It was hypothesized that there would be an effect of vocal emotion on repetition and recall accuracy, with performance being better for words portraying arousing emotions (fear and pleasant surprise) than for words spoken in a neutral voice or with sad vocal emotion, and that the effect would be greater for repetition and recall accuracy when listeners completed the emotion identification, and.

2 Method

2.1 Participants

Participants were 12 undergraduate students from the University of Toronto Mississauga (M = 18.08 years, SD = 0.67, 67% female) who were enrolled in PSY100. All were native speakers of Canadian English who reported being in good health. They had pure-tone audiometric thresholds ≤ 25 dB HL from 250 to 3000 Hz, and no significant inter-aural differences in thresholds.

2.2 Design

All participants completed a demographic questionnaire and nine measures were taken to categorize their characteristics, including two measures of hearing, six measures of cognition and one measure of emotion. All participants completed the same experimental procedures which yielded three key outcome measures, including (1) accuracy of word repetition, (2) accuracy of emotion identification, and (3) accuracy of word recall.

Note that the design followed that used in an earlier study [3] so that it would be possible to compare the current findings to previous findings. The key difference between the previous experiment conducted in our lab and the current experiment is that in the present experiment participants completed the additional emotion identification task following each sentence. Thus, the procedure used in the current experiment is distinguished in its label of WARRMNIE by adding "IE" for "identification of emotion" to the label of WARRMN used to describe the procedure of the previous study.

2.3 Experimental Measures

Working Auditory Recognition and Recall Memory in Noise with Identification of Emotion (WARRMNIE). The 12-talker babble stimuli from the Speech Perception in Noise test were used as background noise. Participants listened to 100 sentences that were spoken in one of four different emotion conditions: fear, sadness, neutral, or pleasant surprise. All words were semantically neutral. The sentences all began with "Say the word" and then a semantically neutral monosyllabic target word was presented. After each sentence, participants repeated the target, identified the emotion in which it was spoken (1 = fear, 2 = neutral, 3 = pleasant surprise, 4 = sad), judged whether the word begins with the first or second half of the alphabet (first = A-M, Second = N-Z) as an unrelated additional processing task to tax working memory. Sentences were presented in recall set sizes of 2, 4, 6, and 8 items, with 5 trials at each set size. After each set, participants recalled as many words as possible that had been heard in the set.

3 Results

3.1 Repetition

An ANOVA was conducted with word repetition accuracy as the dependent variable and Emotion (fear, neutral, pleasant surprise, sadness) and Setsize (2, 4, 6, 8) as within-subjects factors. As seen in Figure 1, there was a significant main effect of Emotion on repetition, $F(3, 33) = 7.574$, $p < .0005$, with word repetition accuracy being poorer for sad compared to fear ($t(11) = 4.78$, $p < .0005$) and neutral ($t(11)3.1$, $p = .003$). There was no main effect of Setsize on repetition accuracy, $p > .05$, nor was there a significant interaction ($p > .05$).

3.2 Recall

An ANOVA was conducted with recall as the dependent variable and Emotion (fear, neutral, pleasant surprise, sadness) and Setsize (2, 4, 6, 8) as the within-subject factors. As seen in Figure 2, results indicated a significant main effect of Emotion on recall, $F(3, 33) = 3.419$, $p = .03$, with a significant difference in performance for words spoken to portray sadness compared to those spoken to portray fear ($t(11) = 2.04$, $p = .03$). There was also a significant main effect of Setsize on recall, $F(1,11) = 26.23$, $p = .000$. All setsizes were significantly different from each other ($ps < .05$), with recall scores decreasing as the setsize increased.

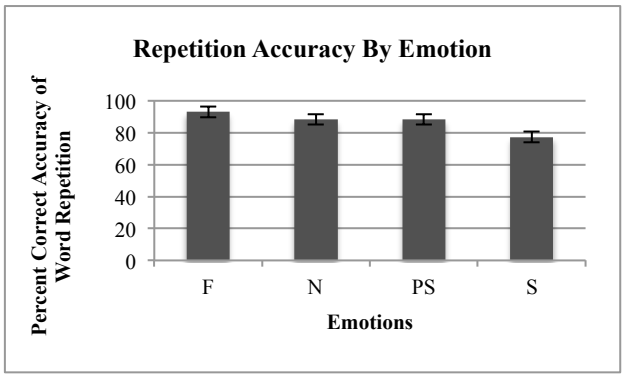


Figure 1. Repetition accuracy by Emotion (Fear, Neutral, Pleasant Surprise, Sad). Error bars are standard deviations.

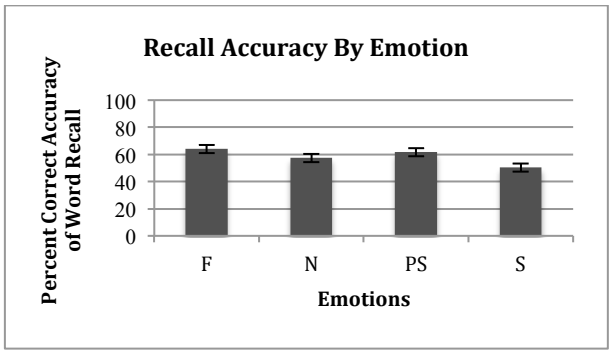


Figure 2. Recall accuracy by Emotion (Fear, Neutral, Pleasant Surprise, Sad). Error bars are standard deviations.

3.3 Emotion Identification

An ANOVA was conducted with emotion identification accuracy as the dependent variable and Emotion (fear, neutral, pleasant surprise, sadness) and Setsize (2, 4, 6, and 8) as the within-subject factors. The ANOVA revealed a significant main effect of Emotion ($F(3,33) = 8.76$, $p < .0005$), and a significant main effect of Setsize ($F(1,11) = 12.24$, $p = .005$), but no significant interaction ($p > .05$). Post hoc comparisons indicated that words spoken to portray pleasant surprise and sadness were identified significantly better compared to words spoken with the other emotions *ps* < .05. (See figure 3).

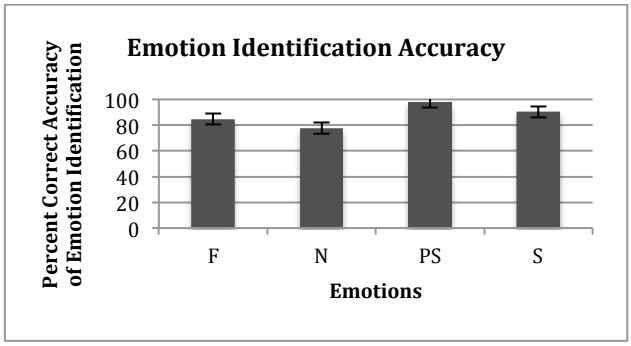


Figure 3. Emotion identification accuracy for the four emotions tested. Error bars are standard deviations.

To see if there was an association between the accuracy of repeating words and the accuracy of later recalling the words, a 2x2 chi-square test of independence was conducted. Results indicated that recall accuracy was not dependent on repetition accuracy, $\chi^2 = 1.75$, $p = .17$. Similarly, we tested if there was an association between correctly identifying the emotion and repeating the word. Results indicated that word repetition was not dependent on emotion identification accuracy, $\chi^2 = 1.47$, $p = .23$. Finally, we tested if there was an association between correctly identifying the emotion and recalling the word. Results indicated that recall was not dependent on emotion repetition accuracy, $\chi^2 = 0.0009$, $p = .98$.

3.4 Comparisons to Prior Study

Repetition. An ANOVA to examine the accuracy of word repetition was conducted with Experiment (WARRMNIE, WARRMN) as a between-subjects factor, and Emotion (fear, neutral, pleasant surprise, sadness) as the within-subject factor. Results indicated a significant main effect of Experiment, $F(1,22) = 15.45$, $p < .0005$. There was also a main effect of Emotion $F(3,66) = 11.962$, $p = .000$, but there was no significant interaction, $p = .08$.

Recall. The effect of emotion on recall accuracy did not differ significantly between the prior WARRMN and current WARRMNIE experiments. This was confirmed with an ANOVA for recall accuracy, with Experiment (WARRMN and WARRMNIE) as the between-subjects factor and Emotion (fear, sad, neutral, pleasant surprise) as the within-subject factor. Results showed that there was a significant main effect of Emotion on recall $F(3, 66) = 8.946$, $p < .0005$, but no significant effect of Experiment or an interaction between Experiment and Emotion (*ps* > .05).

4 Discussion

Taken together, the present experiment confirmed the expected effects of vocal emotion on the accuracy or word repetition and recall, but these effects were not enhanced when an emotion identification task was added in an attempt to focus more of the listeners' attention on the vocal emotion of the talker. Indeed, adding the additional task reduced overall performance on the repetition task, likely because processing was increased.

Acknowledgements

A special thank you to Huiwen Goy and Jenna Pattison for their generosity in assisting with this project.

References

[1] Kensinger, E., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory and Cognition,31*(8), 1169-1180.

[2] Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, 5, 73- 99. doi:10.1016/0010-0277(77)90018-X.

[3] Pichora-Fuller, M.K., Dupuis, K., & Smith, S. (in press). Effect of vocal emotion on word recall. *Experimental Aging Research*.

SPEECH SCIENCES - SCIENCES DE LA PAROLE

RECORDING AND REPRODUCING SPEECH AIRFLOW OUTSIDE THE MOUTH

Donald Derrick [*1], **Tom De Rybel** [†1], **and Romain Fiasson** [‡1]
[1]University of Canterbury, New Zealand Institute of Language, Brain, and Behaviour

1 Introduction

When we speak, air leaves the mouth and nose with time- and space-varying airflow, air pressure, directionality, and turbulence. Reproduction of which requires an ability to both estimate the airflow from speech and control a source of artificial airflow. Here we present a system that produces artificial airflow with a similar intensity envelope to that produced during speech, as measured 1 cm away from the lips.

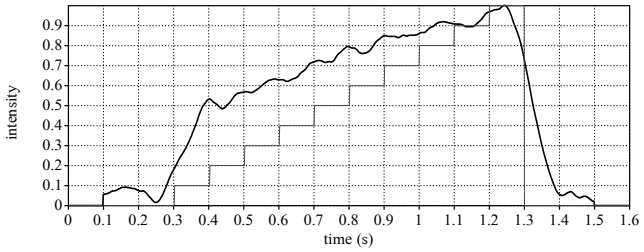


Figure 1: Relationship between direct voltage stimulation of air-pump, 10% increments every 100 ms. Input voltage (thin line). Five Hz low-pass filtered relative air-flow rate (bold line)

Researchers studying the effects of airflow on speech perception began with simple, solenoid-operated, compressor-based airflow device [1]. This system demonstrated that correctly time-aligned airflow enhanced the perception of aspirated stop consonants in syllables (/pa/, /ta/) produced in noisy environments. But our goal is to enhance speech perception during continuous speech, which requires a system able to produce both time and intensity-varying airflow. Our current system uses Murata's *microblower*, a 20x20x1.85mm piezoelectric air pump with up to 0.8 l/m flow, max 19.38 cm/H_2O pressure, and approximately 30ms 5-95% intensity rise time, allowing artificial approximation of continuously-varying airflow in speech [2]. This system is controlled using a circuit that converts a voltage output from a sound-card's right channel into a signal suitable to drive the air-pump, and routes the left channel to both headphone speakers, allowing direct presentation of both audio and airflow to the listener. The system allowed us to uncover speech perception enhancement in fricatives [3] and affricates [4], as well as stops.

However, the control system in use did not allow for a full range of airflow from zero to maximum, as the pump has a minimum airflow output of about 0.4 l/m when active, and exhibited other non-linearities. These effects are largely due to the electro-mechanical nature of the pump itself. Figure 1 shows the relationship between direct input and airflow output from our system. The solution presented here allows control of airflow from 0 to 0.8 l/m of airflow, as directly estimated during speech production.

∗. donald.derrick@canterbury.ac.nz
†. tom.derybel@canterbury.ac.nz
‡. romain.fiasson@canterbury.ac.nz

2 Methods and Results

To begin, we made an airflow estimator system that does not interfere with speech production. The system consists of a ping pong ball mounted on a carbon fiber lever that flexes a thin, flat, carbon-fibre member to which two strain gauges are affixed on opposite sides. These strain gages are part of a sinusoidal-driven bridge, the output of which, after amplification, is an amplitude-modulated sine wave that is can be recorded using a standard XLR based pre-amplifier into a computer. This allows us to bypass the DC blocking circuitry of sound cards for the very low frequency information we intend to measure. The sine wave is then demodulated using an envelope detector and lowpass filtered (100 Hz) in custom software, and a 11.35 Hz resonance is notch-filtered out, along with the second and third harmonic of that resonance. These are due to the mechanical design of the sensor.

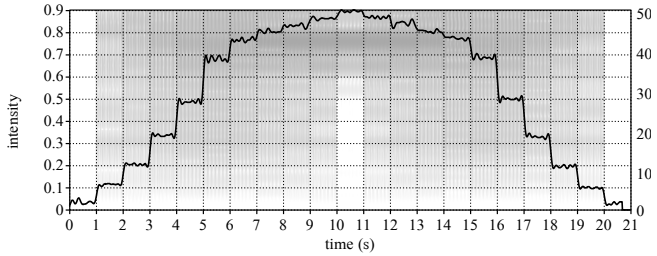


Figure 2: Relative airflow output (bold line) from linear 10 ms period-modulated 12kHz sine-waves sent to the air-pump, stepped at 10% intensity increments (blended background)

To overcome the non-linearities in the air pump output, we first had to overcome the DC blocking behaviour of the sound card. A 12 kHz sinusoidal carrier was chosen for this purpose. The air pump drive circuitry performs an implicit, envelope detection-based demodulation. Second, a form of period modulation was implemented to overcome the 0.4 l/m minimum flow rate. The modulation controls the number of 12 kHz periods during a fixed, 10 ms time window. By driving the pump with these bursts, and relying on the non-negligible quality factor of the piezo resonator in the pump to smooth the air flow, lower flow rates are obtained. Through experimentally-varying the width of these 12 kHz bursts, a 10 ms (100 Hz) burst width was found optimum, corresponding with the 30 ms step response of the *microblower*. We then generated a 21 second pyramid step signal with 1 second long 10% steps from 0 to 100% of signal strength and back. The period-modulated signal was then passed to the air pump. The audio output was recorded with a Sennheiser MKH-146 microphone, and the airflow output recorded using the air-

flow estimator. The output of one of the eight recordings can be seen in Figure 2. Although we could cover a much wider range of air flows, some non-linear behaviour remained.

To resolve these nonlinearities, we recorded 8 instances of the output of this 21 second pyramid, measuring the intensities of each step, and fitting the data to polynomial regressions. We found that a 5^{th} order regression was best. The transformed modulation was then applied to a new 21 second pyramid, which was then sent to the air pump. An example of the much more linear output can be seen in Figure 3.

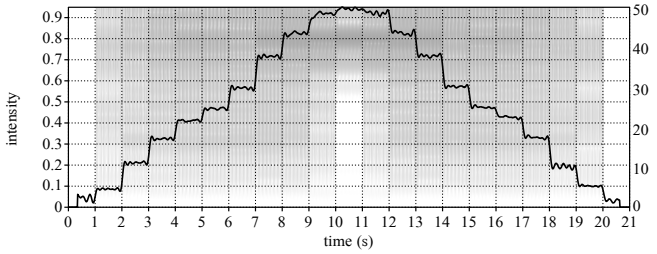


Figure 3: Relative airflow output (bold line) from 5th order polynomial fitted 10ms period-modulated 12kHz sine-waves sent to the air-pump, stepped at 10% intensity increments (blended background)

Our measurements also showed a 30 ms delay from the signal to air-flow from our pump. We used this information to realign, transform, and period-modulate our recorded airflow from speech to generate the signal to drive the air pump. We then played-back a long section of speech, recording the headphone and airflow output. The results show we can reasonably reproduce the envelope of airflow intensity from complex speech acts, such as the word "six" as seen in Figure 4.

3 Discussion

The results show that it is possible to reasonably reproduce the airflow envelope of speech using a low-cost and compact system. However, we do not know how linear the estimates from the airflow estimator are, and instead are relying on the fact that we are using the same measurement system for recording both speech airflow and artificial airflow.

The airflow estimator also has resonances and inertia that cause large impulses of air to appear to have a longer duration than they do. The issue can be seen in Figure 4 after the larger airflows from /s/ and /ks/ where there appears to be a second airflow burst that is produced by the under-damped resonances in the estimator instead of artificial airflow.

The results presented are also scaled, and the actual flow measurement intensity is 1/10th of that in speech. But despite this tiny airflow intensity, our research had already demonstrated that the airflow can alter speech perception, and, from experience, it is quite easy to feel and notice the airflow from this system. During experiments, the full-on flow from the air pump was even deemed uncomfortable by participants. Current research is using this new control system to test for speech perception enhancement of whole sentences.

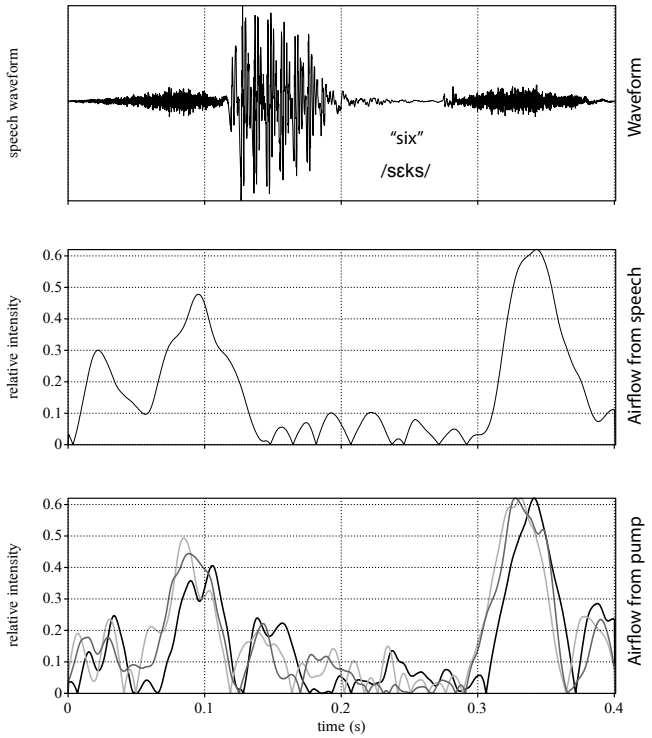


Figure 4: NZE "six" speech waveform (top). Scaled speech airflow from speaker (middle). Three overlapped tokens of scaled speech airflow from air pump generated from modulated input (bottom)

Acknowledgements

Research was funded by two New Zealand Ministry of Business, Innovation and Employment (MBIE) grants for Aerotactile Enhancement of Speech Perception.

References

[1] Bryan Gick and Donald Derrick. Aero-tactile integration in speech perception. *Nature*, 462 :502–504, 26 November 2009. doi :10.1038/nature08572.

[2] Donald Derrick, Tom De Rybel, Greg. A. O'beirne, and Jennifer Hay. Listen with your skin : Aerotak speech perception enhancement system. In *Proceedings of the 15th Annual Conference of the International Speech Communication Association (INTERSPEECH 2014)*, pages 1484–1485, Singapore, 14-18 September 2014.

[3] Donald Derrick, Greg A. O'Beirne, Tom De Rybel, and Jennifer Hay. Aero-tactile integration in fricatives : Converting audio to air flow information for speech perception enhancement. In *Proceedings of the 15th Annual Conference of the International Speech Communication Association (INTERSPEECH 2014)*, pages 2580–2584, Singapore, September 14-18 2014.

[4] Donald Derrick, Greg A. O'Beirne, Tom De Rybel, Jennifer Hay, and Romain Fiasson. Aero-tactile integration in mandarin syllables. In Prep.

[5] Polynomial regression data fit - aracnoid.com, July 2015.

THE TIMING OF ACCENTUAL PHRASES IN READ AND SPONTANEOUS SPEECH: DATA FROM ACADIAN FRENCH

Wladyslaw Cichocki
University of New Brunswick, Fredericton NB, E3B5A3, CANADA

1 Introduction

In recent years, researchers have developed rhythm metrics to capture patterns of timing in speech. Based on acoustic measures of vocalic and consonantal intervals, these metrics focus on variability in the durations of these segmental intervals; certain metrics (pairwise variability measures) measure the variability of pairs of immediately adjacent intervals while other metrics (interval measures) quantify variation of intervals over an entire utterance. These rhythm metrics have been applied in different areas of research including automatic speech recognition, dialectology, disordered speech, language typology and second language phonology.

Some researchers have looked beyond measures of vocalic and consonantal intervals to examine duration variability in larger units such as syllables and feet [1] Their work is motivated by a conception of rhythm that rests on notions such as grouping and patterns of prominence. In this paper, we measure the variability of durations of units known as accentual phrases (APs). Also refereed to as interstress units, *pieds accentuels*, *groupes rythmiques*, *mots prosodiques* and *syntagmes accentuels*, APs are groups of syllables that are demarcated by a stress. In French, these groups are generally realized by a sequence of unstressed syllables that ends in a stressed syllable; they tend to correspond to short syntactic units. APs are also basic units in the analysis of French intonation [2]. Classical phonetic studies of duration in French – see the summary in [3] – show that length of APs varies according to speech style.

The main elicitation technique in rhythm research has been reading, which is used in order to control for phonotactic effects on the stability of rhythmic measures. Of the small number of studies that include spontaneous speech, several show that spontaneous speech has greater variability in interval duration variation than reading [4, 5]; however, a few researchers find no between-style timing differences [6].

The goal of the present study is to apply rhythm metrics based on AP durations to speech data from a dialect of Acadian French spoken in New Brunswick (Canada). We compare speech from two elicitation methods – reading and spontaneous storytelling – and examine how well rhythm metrics based on AP durational variability can distinguish between the two styles. Our hypothesis is that spontaneous speech will show more variability (i.e. higher scores) than read speech.

2 Methodology

2.1 Speakers and speech materials

The corpus consists of recordings by 12 native speakers of Acadian French from the municipality of Tracadie in northeastern New Brunswick [7]. The speakers are stratified by age and gender. There are three age groups, each with four speakers: younger (mean age: 23.5 years), middle-aged (41.6 years) and older adults (76.3 years); each age group has an equal number of males and females.

Speech materials are from sociolinguistic interviews conducted according to the protocol of the *Phonologie du français contemporain* project (www.projet-pfc.net), which compares dialects of French spoken around the world. Among the interview tasks are: reading a passage and telling a story about a personal event. The personal stories were told spontaneously with no preparation and, because the interviews were carried out by a member of the speaker's family, these stories come close to representing an informal or vernacular speech style. The total corpus of speech studied is reasonably large: about 34 minutes (almost 3 minutes per speaker).

2.2 Procedure and measurements

Ten (out of a total of 21) sentences were extracted from the reading passage. Similary, between 2 and 3 minutes (per speaker) of the spontaneous speech were divided into utterances that were of varying lengths. All pauses, hesitations, and false starts were excluded from analysis. The speech data from both styles were segmented, first automatically with *EasyAlign* and then manually, into vocalic and consonantal intervals using *Praat* and following generally accepted segmentation criteria.

The task of identifying the APs in the reading and spontaneous story data was carried out by two native speakers of Acadian French. The transcribers were instructed to divide the speech into 'chunks' that 'are spoken together as a unit'; the transcribers also determined whether the boundary between these units is major or minor. Over 2,300 APs were identified: 1,112 in the reading texts and 1,237 in the spontaneous speech.

Durations of APs (in ms) were extracted using a *Praat* script. To calculate rhythm metrics for the APs for each speaker for both speech styles, we used the same formulas as for duration-based rhythm metrics. Two pairwise variability indexes – nPVI-AP, rPVI-AP – and two interval-based indexes – deltaVP, VarcoVP – were calculated.

* cicho@unb.ca

3 Results

Articulation rates, which exclude pauses and hesitations, tend to be slightly higher (i.e. there is a faster tempo) in spontaneous speech than in the reading style (5.2 vs. 4.8 syllables per second; $p < .171$), although this difference is not statistically significant in the data studied here. Mean AP durations are significantly shorter in spontaneous speech than in reading (667.3 vs. 824.0 ms; $p < .0001$).

Descriptive statistics for the AP-based rhythm metrics are given in Table 1. For normalized measures, variability of AP durations is significantly greater in spontaneous speech (i.e. higher scores) than in read speech (based on repeated measures ANOVAs). No effect is found for non-normalized measures.

	reading	spontaneous	*p*-level
nPVI-AP	48.3 (4.1)	61.8 (8.0)	.002
rPVI-AP	399.0 (59.8)	399.5 (55.4)	.993
VarcoAP	43.8 (3.8)	52.0 (4.2)	.002
deltaAP	361.9 (50.6)	343.1 (34.5)	.270

Table 1: Means and standard deviations of four rhythm metrics for two styles. Statistical significance is shown in the final column.

To determine which metrics contribute most to distinguishing between styles, we carried out linear discrimination analyses. Best discrimination was obtained with the two PVI-based metrics, achieving 91.3% correct classification. Figure 1 presents the plot of speaker scores.

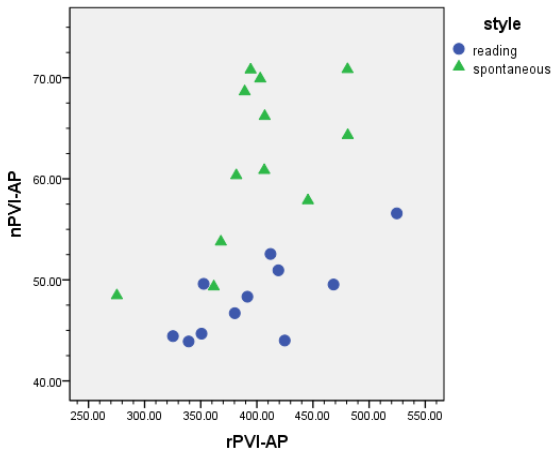


Figure 1: Distribution of speakers in the (rPVI-AP, nPVI-AP) plane by style.

Noticeable in Figure 1, for both normalized (nPVI) and non-normalized (rPVI) scores, are the large between-speaker variations in reading and in spontaneous styles. A similar spread of scores occurs in the (deltaAP, VarcoAP) plane. Strong correlations exist between the two normalized, ((nPVI-AP, VarcoAP) r=.853), and the two non-normalized, ((rPVI-AP, deltaAP) r=.844), metric scores.

4 Discussion and Conclusion

The results confirm three patterns found in studies of rhythm that use segmental-based metrics. First, spontaneous speech has more variability than read speech, as shown by the higher normalized metric scores. While this may appear to be an expected result, the existence of stylistic differences has been demonstrated in only a small number of studies. Second, there is considerable between-speaker variation in both speech styles. This source of variability remains a challenging shortcoming for metric-based research on speech timing. Finally, normalized metrics are better than non-normalized metrics in differentiating between the speech varieties.

The present study demonstrates that AP-based rhythm metrics can characterize stylistic differences in French. Followup (phonetic, phonological, syntactic) research is needed to uncover the nature of this variability. In addition, a study by [8] shows that, when considered with articulation rate, AP-based metrics discriminate among national, regional and contact varieties of French. Both results invite more extensive applications of these metrics to other varieties of French.

One implication of this study is that AP-based rhythm metrics contribute to a framework for the comparison of read and spontaneous speech materials. More generally, this research supports the claim that rhythm metrics can be useful in characterizing interesting patterns of speech timing that are located at levels above vocalic and consonantal segments.

Acknowledgments

Many thanks to Mélissa Chiasson, Stéphanie Duguay and Yves Perreault for their help with transcriptions and programming. The author gratefully acknowledges support from the Canada Foundation for Innovation and the Social Sciences and Humanities Research Council of Canada.

References

[1] F. Nolan & E.L. Asu. The pairwise variability index and coexisting rhythms in language. *Phonetica*, 66, 64-77, 2009.
[2] S-A. Jun & C. Fougeron. Realizations of Accentual Phrase in French intonation. *Probus*, 14, 147-172, 2002.
[3] P.R. Léon. *Phonétisme et prononciations du français*. 5th edition, Paris, Armand Colin, 2007.
[4] A. Arvaniti. The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics* 40, 351-373, 2012.
[5] T. Meisenburg. "Southern vibes? On rhythmic features of (Midi) French. *Language Sciences*, 39, 167-177, 2013.
[6] V. Dellwo, A. Leeman & M-J. Kolly. The recognition of read and spontaneous speech in local vernacular: The case of Zurich German. *Journal of Phonetics*, 48, 13-28, 2015.
[7] W. Cichocki. The PFC-Tracadie Corpus: A variety of Acadian French spoken in northeastern New Brunswick (Canada). In A. Barysevich, A. D'Arcy & D. Heap (eds) *Proceedings of Methods XIV. Papers from the 14th International Conference on Methods in Dialectology*, Bern, Peter Lang, 149-157, 2013.
[8] N. Obin, M. Avanzi, G. Bordal & A. Bardiaux. "À la recherche des temps perdus": Variations sur le rythme en français. *Actes des 29èmes Journées d'Étude sur la Parole*, Grenoble, 65-72, 2012.

STARTLE-INDUCED EVIDENCE FOR MULTI-DIMENSIONAL INFORMATION IN SPEECH PLANS

Chenhao Chiu [*1] **and Bryan Gick** [†1,2]
[1]Dept. of Linguistics, University of British Columbia, 2613 West Mall, Vancouver, BC, Canada, V6T1Z4
[2]Haskins Laboratories, 300 George St., New Haven, CT, USA 06511

1 Introduction

In order to uncover some of the details in speech plans, a design using a startling auditory stimulus (SAS ; $>$ 120dB) has been applied to prepared speech production [1]. Stevenson et al. [1] showed that a syllable as long as a CV sequence (i.e., /ba/) can be prepared in advance and subject to rapid execution when elicited by a SAS. Although little or no sensory feedback may be available at such a short latency, the prepared syllable is performed as intended without any interruption to the synergistic system. When elicited by a SAS, additional lip compression at the beginning of the response was observed. It is not clear whether the observed lip compression is a part of speech plans (i.e., speech-specific) or a generic effect induced by a SAS. This study tackles this question by comparing SAS-induced spoken speech with mouthed speech and non-speech lip movements.

Compared with spoken speech, mouthed speech has shorter word duration, reduced EMG amplitude, and overall hypoarticulation in lip movements [2, 3]. However, the most general statement of the difference between spoken and mouthed speech is that mouthed speech differs from vocalized speech not merely in that it lacks vocalization but also, as with internal speech, in that it entirely lacks air movement. Different aerodynamic requirements for different speech tasks may induce different orofacial muscle activity and kinematics [4]. The experiment presented here was conducted to test the hypothesis that multi-dimensional information is pretuned in speech plans. Information such as aerodynamics is reflected in tasks with different requirements. It is predicted that lip compression will vary under different aerodynamic conditions ; specifically, more and larger lip compression is expected in vocalized speech than in silent speech since intraoral pressure is greater in the former, but not so much in the latter.

2 Method

2.1 Participants

Data were collected and analyzed from the same nine participants reported in [1] (3 male and 6 female ; Mean = 23 years, SD = 4.2 years). Participants were all native speakers of North American English. Prior to the testing, participants signed an informed consent form and were naïve to the hypothesis under investigation. The experiment was conducted following the ethical guidelines established by the University of British Columbia.

∗. chenhao@alumni.ubc.ca
†. gick@mail.ubc.ca

2.2 Apparatus, task, and procedures

The same apparatus and recording equipment from [1] were applied. Two tasks were designed in blocks. Throughout the testing session, participants were asked to start with their mouths closed in a relaxed posture without compressing their lips during the preparation. The first block *Non-speech* condition required participants to respond to an acoustic stimulus by opening their mouths. In the second block, participants were instructed to respond to the acoustic stimulus by mouthing a silent articulation of /ba/ (the *Mouthed* condition). Kinematic data of the spoken syllable /ba/ from [1] were included and analyzed.

The testing procedures are identical to those reported in [1]. The testing block consisted of twenty control trials and five startle trials. The five startle trials were presented pseudo-randomly such that the first trial was never a startle trial, nor were there two consecutive startle trials.

2.3 Data reduction, dependent measures, and analyses

Two conditions and twenty-five trials in each condition for nine participants yielded 450 trials in total. A total of 40 of the 450 trials were excluded from analysis for the following reasons : anticipation (4 trials ; 0.89%), data loss (1 trial ; 0.2%), false startle (6 trials, 1.3%), hesitation (5 trials, 1.1%), startle indicator with late RT ($>$120 ms ; 5 trials ; 1.1%), and outlier filtering by subject (19 trials, 4.2%), where "outlier" was defined as a voluntary onset of 2 SD or more away from the mean. The remaining 410 trials were examined and further analyzed.

The trials were also analyzed for lip compression. In each experimental trial, lower lip movements were carefully examined in terms of their displacement trajectories. The initiation of the lower lip movement was recorded as the voluntary movement onset. Since the lips began in a mouth-closed position, after the initiation of the voluntary movement, the upward movement observed in the lower lip is understood as lip compression. Conversely, direct downward movement of the lower lip at the onset of voluntary movement indicates an absence of lip compression. The number of trials with and without lip compression was calculated. Compression displacement (i.e., the vertical height between voluntary movement onset and lower lip opening onset) was also measured and analyzed. Compression displacement data was analyzed by 2 $\times$ 3 ANOVA. Only trials that exhibited lip compression were included for the analyses here. In order to perform a 2 $\times$ 3 ANOVA, missing values were filled in by a constant. Missing data was treated as random and substituted with the mean

across participants in the same condition.

3 Results

The frequency of lip compression in the current design was calculated, with the raw count of trials showing lip compression presented in Table 1. As revealed, for Spoken responses, lip compression occurred in 91% of control responses and 87% of startle responses. Lip compression in Mouthed and Non-speech responses did not occur as frequently as in Spoken responses. The frequency of lip compression across control and startle was highest for Spoken, followed by Mouthed, and then Non-speech. For Spoken responses, 91% of trials showed lip compression, whereas for Mouthed responses, the overall occurrence of lip compression was 56%. Only 22% of overall Non-speech responses exhibited lip compression. Note that a decrease in the occurrence of lip compression from control to startle trials was also observed in Mouthed responses whereas an increase was observed in Non-speech responses.

	Spoken	Mouthed	Non-speech
Control	145 (91%)	104 (63%)	23 (14%)
Startle	26 (87%)	10 (24%)	22 (55%)
TOTAL	171 (90%)	114 (56%)	45 (22%)

Table 1: Numbers of trials with lip compression across conditions. Percentages were calculated from the trials included in analyses.

Compression displacement (as measured by the vertical height between lower lip voluntary onset and lower lip opening onset) across conditions is summarized in Table 2.

	Spoken	Mouthed	Non-speech
Control	0.96 (0.73)	0.30 (0.41)	0.21 (0.12)
Startle	1.38 (0.95)	0.48 (0.42)	0.48 (0.50)

Table 2: Mean compression displacement across conditions (mm); standard deviations in parentheses.

As the table shows, the largest compression displacement was observed in the Spoken condition, followed by the Mouthed condition, and then the Non-speech condition. Across all three conditions, more compression displacement was observed for startle responses than for control responses. Results revealed a main effect for Condition ($F(2, 16) = 12.46$, $p < .01$, $\eta_p^2 = 0.61$). No effect for Stimulus ($p = 0.06$) or Condition $\times$ Stimulus interaction ($p = 0.47$) was noted. Post-hoc analyses found significant differences between Spoken and Non-speech ($p = 0.027$) and between Spoken and Mouthed ($p = .006$), but not between Mouthed and Non-speech ($p = 1$).

4 Discussion

Considering both lip compression frequency and amplitude, Spoken responses were significantly larger than both Mouthed and Non-speech responses. For Spoken responses, a SAS did not appear to have an impact on the occurrence frequency of lip compression. Frequent lip compression observed in Spoken responses supports the view that such lip movements are part of the plan for a Spoken task. While non-startled Mouthed responses revealed less lip compression (both frequency and amplitude) than Spoken responses did overall, the percentage of Mouthed responses with lip compression decreased even further in SAS-elicited trials. Compared with Mouthed responses, a much lower lip compression frequency in Non-speech responses suggests that non-speech lip movement does not require lip compression as part of the preparatory movement. Lip compression is performed in anticipation of speech behaviours, we suggest possibly in anticipation of the need to contain an increase in intraoral pressure associated with an oral stop.

As a SAS reliably elicits prepared movements, these results suggest that lip compression may not be pre-tuned for Mouthed speech tasks. If lip compression were part of the speech-related motor plan, we should have observed a comparable frequency and magnitude of compression in both Mouthed and Spoken responses. A more general implication is that speech plans may include multi-dimensional information other than simple kinematic trajectories, i.e., we speculate that different levels of intraoral pressure may yield differentiated lip compression in the response. Investigation of intraoral pressure will call for further investigation.

Acknowledgments

This work was supported by NSERC and the NIH Grant DC-02717 to Haskins Laboratories. We thank the Artisynth Opal-Model Group, and also Dr. Ian Franks, Dr. Romeo Chua, Dr. Dana Maslovat, and Dr. Andrew Stevenson from the Motor Control and Learning Laboratories and at UBC.

References

[1] A. J. Stevenson, C. Chiu, D. Maslovat, R. Chua, B. Gick, J-S. Blouin, and I. M. Franks. Cortical involvement in the StartReact effect. *Neuroscience*, 269 :21–34, 2014.

[2] L. Crevier-Buchman, C. Gendrot, B. Denby, C. Pillot-Loiseau, P. Roussel, A. Colazo-Simon, and G. Dreyfus. Articulatory strategies for lip and tongue movements in silent versus vocalized speech. In *Proceedings of the 17th International Congress of Phonetic Science*, pages 1–4, 2011.

[3] M. Janke, M. Wand, and T. Schultz. Impact of lack of acoustic feedback in emg-based silent speech recognition. In *Interspeech*, pages 2686–2689, 2010.

[4] B. Gick, C. Chiu, C. Flynn, I. Stavness, N. Francis, and S. Fels. Producing whole speech events : Differential facial stiffness across the labial stops. In *Proceedings of Acoustics 2012 Hong Kong*, pages 2pSC1, 6pp. The Hong Kong Institute of Acoustics (HKIOA), 2012.

CONTRASTIVE TONGUE SHAPE OF THREE SIBILANT FRICATIVES IN TAIWAN MANDARIN READ SPEECH

Masaki Noguchi [*1], **Chenhao Chiu** [†1], **Po-Chun Wei** [‡1], **and Noriko Yamane** [§1]
[1]University of British Columbia

1 Introduction

This study investigates the production of three sibilant fricatives in Taiwan Mandarin, dental [s], retroflex [ʂ], and alveolopalatal [ɕ], using ultrasound recording. Previous studies have pointed out that some speakers of Taiwan Mandarin lose the contrast between dental and retroflex sibilants in connected speech [1,2]. In the literature on speech production, it has been well recognized that factors like speaking rate and speaking style significantly affect articulatory movements in such a way that fast speech or casual speech tends to result in the undershooting of articulatory targets and the loss of phonetic contrasts [3]. In this study, we investigated whether the loss of the contrast between dental and retroflex sibilants in Taiwan Mandarin would be conditioned by speaking style. We examined whether the contrast between dental [s] and retroflex [ʂ] in Taiwan Mandarin was maintained in the production of citation forms in read speech.

2 Methods

2.1 Participants

Two native speakers of Mandarin from Taiwan (one female and one male) participated in the study. Both of them self-reported living in Taiwan until adolescence. Importantly, they had participated in our previous study and had shown a significant overlap in tongue shape between dental [s] and retroflex [ʂ], when followed by the vowel /a/ [2].

2.2 Recording

Participants read 36 nonce words visually presented in written form on a computer screen. Half of the words are shown in Table 1. Each item comprises three syllables, and the second syllable contained one of three sibilant fricatives (dental [s], retroflex [ʂ], alveolopalatal [ɕ]) followed by one of three vowels (/a/, /ɨ/, /o/) with either tone 1 or tone 4. The items were divided into two lists according to tone category. In each list, the items were presented in random order. Participants read each list seven times. Ultrasound and audio recordings were made during the production.

2.3 Measurements

The ultrasound recordings were analyzed for each participant. For each token of production, a frame of ultrasound video was extracted from the mid-point of the target fricative sound. EdgeTrak was used to trace tongue shape in the ultrasound still image [4], and SSANOVA was used to compare tongue shapes [5].

Table 1: Items

	Tone	/a/	/ɨ/	/o/
s	1	加撒湖	八司館	八搜宮
	4	帕薩碑	霸四亭	駕嗽館
ʂ	1	搭沙堂	搭師碑	拉收亭
	4	法剎湖	納市樓	納受道
ɕ	1	搭瞎館	發吸園	發休道
	4	大夏堂	納戲園	法秀橋

2.4 Results

Figures 1 - 3 compare tongue shapes by fricative type for each vowel context. Both speakers showed clearly non-overlapping tongue shapes for three fricatives in all vowel contexts. However, they differed from each other in the way they made the contrast between dental [s] and retroflex [ʂ].

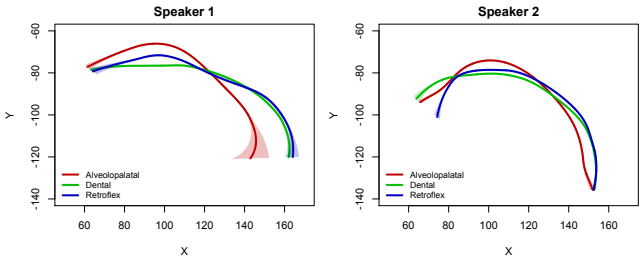


Figure 1: Three sibilants in the /a/ context

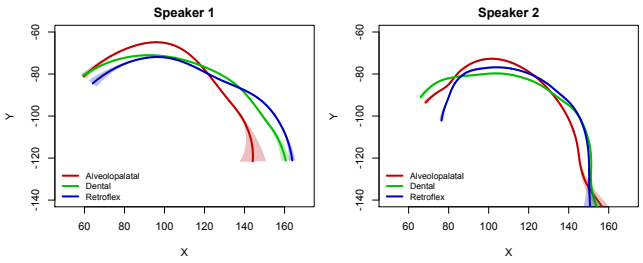


Figure 2: Three sibilants in the /ɨ/ context

Speaker 1 made the contrast through differences in tongue tip position the tongue body height in the /a/ context, but through differences in tongue tip position and tongue root position in the /ɨ/ and the /o/ contexts. Speaker 2 did not show such variation ; the contrast was made through differences in tongue tip position and tongue body height.

∗. mskngch@alumni.ubc.ca
†. chenhao@alumni.ubc.ca
‡. xyzgrace@alumni.ubc.ca
§. noriko.yamane@outlook.com

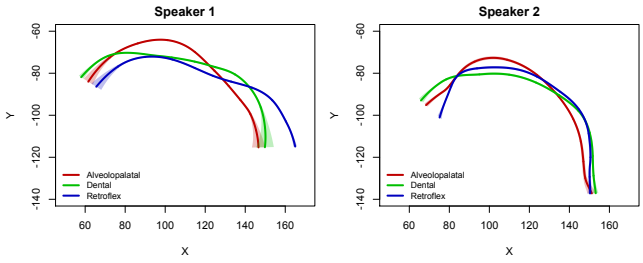


Figure 3: Three sibilants in the /o/ context

Figures 4 - 6 compare tongue shapes by vowel context for each fricative type. As revealed in Figure 4, Speaker 1 showed a vowel-dependent variation in the production of dental /s/. The tongue body was lower and the tongue root was more retracted in the /a/ context than in the other two contexts. This was probably due to coarticulation between the fricative and the following vowels ; the tongue height of the following vowels affected the height of the tongue body of the preceding fricative.

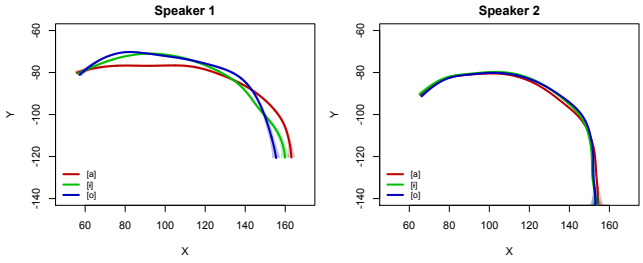


Figure 4: Dental [s] across three vowel contexts

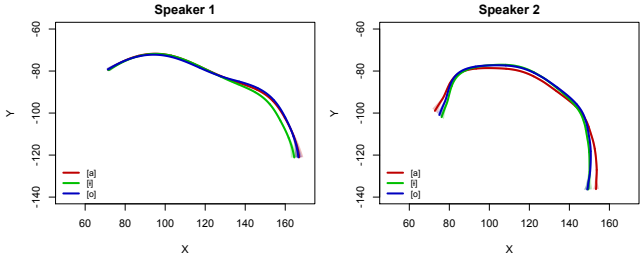


Figure 5: Retroflex [ʂ] across three vowel contexts

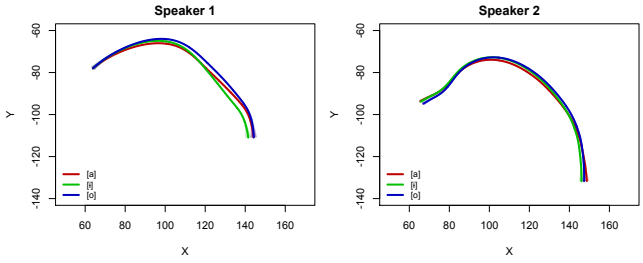


Figure 6: Alveolopalatal [ɕ] across three vowel contexts

Both speakers exhibited consistent tongue shapes for retroflex [ʂ] and alveolopalatal [ɕ] across vowel contexts (Figures 5 - 6). Tongue shape for retroflex [ʂ] was characterized by the lowering of the tongue tip. This gesture probably contributed to the formation of a large front cavity, an articulatory target for the production of retroflex sibilants in Mandarin. Tongue shape for alveolopalatal [ɕ] was characterized by the raising of the tongue body and the advancement of the tongue root. These gestures probably contributed to the formation of a long and narrow constriction channel from the alveolar to palatal regions, an articulatory target for the production of alveolopalatal sibilants in Mandarin.

3 Discussion

Speakers in this study showed clearly distinctive tongue shapes for all three sibilant fricatives across vowel contexts. These speakers had been known to show a significant overlap in tongue shape between dental [s] and retroflex [ʂ] in connected speech [2]. Therefore, the results of this study suggest that the loss of the contrast partly depends on speaking style. It is more likely to happen in connected speech than in read speech. The results of this study also revealed an interesting vowel-dependent variation in the production of dental [s] but not retroflex [ʂ] and alveolopalatal [ɕ]. The reason why this variation happened only in the production of dental [s] is probably that the production of the dental fricative requires the active involvement of the tongue tip alone, while the production of post-alveolar fricatives requires the active involvement of the tongue as a whole. This means that for the production of the dental fricative the tongue body can remain more susceptible to coarticulation with the following vowels but for the production of post-alveolar fricatives the tongue body does not show the same susceptibility to coarticulation.

References

[1] Y-H Chang. *Variability in cross-dialectal production and perception of contrasting phonemes : the case of the alveolar-retroflex contrast in Beijing and Taiwan Mandarin*. PhD thesis, University of Illinois at Urbana-Champaign, 2013.

[2] M Noguchi, C Chiu, P-C Wei, and N Yamane. *Uncovering sibilant fricative merger in Taiwan Mandarin : Evidence from ultrasound imaging and acoustics*, 2015. Poster presented at the Annual meeting of the Linguistic Society of America, Portland, OR.

[3] E Farnetani and D Recasens. Coarticulation and connected speech processes. In William J Hardcastle, John Laver, and Fiona E Gibbon, editors, *The handbook of phonetic sciences*, pages 371–404. Blackwell Oxford, 1997.

[4] M Li, C Kambhamettu, and M Stone. Automatic contour tracking in ultrasound images. *Clinical linguistics &amp ; phonetics*, 19(6-7) :545–554, 2005.

[5] L Davidson. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variancea). *The Journal of the Acoustical Society of America*, 120(1) :407–415, 2006.

COMMUNICATION BETWEEN NATIVE AND NON-NATIVE SPEAKERS OF ENGLISH IN NOISE

Ann Nakashima*, Sharon M. Abel and Ingrid Smith
Defence Research and Development Canada, Toronto Research Centre
1133 Sheppard Ave West, Toronto, ON M3K 2C9

1 Introduction

Communication between multilingual speakers in noisy environments can be problematic for both civilian and military operations. Pilots and air-traffic controllers communicate in high-workload situations, often in their second language [1]. For military personnel, battlefield communication in multi-national operations can be further complicated by extreme noise exposure from armoured vehicles, weapons and aircraft flyovers. The wearing of tactical communication and protection systems (TCAPS) provides users with hearing protection and integrated radio communication, but could interfere with face-to-face communication. To date, there have been few studies of the communication effectiveness of non-native speakers in noise wearing hearing protection devices (HPDs), particularly among those with a range of linguistic backgrounds. With two official languages in Canada and a culturally diverse population in the Canadian Armed Forces, this study of communication between native and non-native speakers was conducted in the interest of improving battlefield communication.

2 Methods

2.1 Participants

Ethics approval was obtained from the Human Research Ethics Committee (HREC) of Defence Research and Development Canada (DRDC). Twenty-four normal-hearing men and women, military and civilian, participated. The average age was 33.9±9.3 years (range: 21 to 53). Half of each gender subgroup of 12 were native monolingual English speakers (NA) and half were non-native speakers (NN) who learned English after the age of 10 years.

2.2 Test Facility and Materials

Responses on the language experience and proficiency questionnaire (LEAP-Q) [2] were used to confirm group assignment. The experimental sessions were conducted in the noise simulation facility at DRDC, Toronto Research Centre, which is a large, semi-reverberant room (10.55 x 6.10 x 3.05m^3). Two tests of speech intelligibility were presented: the modified rhyme test (MRT) [3] and the speech perception in noise test (SPIN) [4].

2.3 Experimental Protocol

Each participant completed two experimental sessions in same gender pairs, in which they alternated as a talker and listener. The NA participants were paired with another NA in one session and an NN in the other session. Similarly, the NN participants were paired with an NN in one session and an NA in the other. Thus there were four groups of talker-listener pairs: NA-NA, NA-NN, NN-NA and NN-NN. The linguistic backgrounds of the NN-NN pairs were mismatched to avoid a possible interlanguage intelligibility benefit [5].

The MRT and SPIN tests were administered for two modes of communication: face-to-face (F2F) and radio. In the F2F condition, the talker and listener were seated facing each other at the ends of a two-meter long table. An earmuff-style TCAPS device (3M™ Peltor™ LiteCom Plus [3M, St. Paul, MN]) was worn by each participant with the radio turned off. Participants were instructed to maintain a "normal" voice level of 55 dBA; this was practiced during a training run held prior to the experimental sessions. The background noise was a recording of a light-armoured vehicle (LAV III) driving on a highway. Three levels of the background noise were used: 55, 60 and 65 dBA, giving approximate speech-to-noise ratios of 0, -5 and -10dB, respectively. In the radio condition, a visual barrier was placed in the middle of the table so that the participants could not lip-read. The headset was turned on and used for communication. The talker used the push-to-talk mode to transmit the MRT word ("the word is ___") or the SPIN sentence to the listener. The background noise was presented at 80 dBA. Each participant pair alternated being a talker and a listener for the radio condition, and the three F2F conditions.

3 Results

3.1 Language Experience

Based on the responses to the LEAP-Q, the average age of English acquisition for the NN group was 10.9±3.7 years and the number of years of schooling or work in English was 13.7±6.7 years. Their self-ratings of English speaking, understanding and reading comprehension were 7.8±1.2, 8.6±0.8 and 8.8±0.7, respectively, out of a possible 10. The reported first languages included Chinese, Serbian, Spanish, French and Russian.

3.2 F2F Condition

The results of the MRT for the three background noise levels in the F2F condition are shown in Figure 1. The MRT results (combined for NA and NN listeners) are shown as percentage of correct responses, separated by trials

*ann.nakashima@drdc-rddc.gc.ca

with contrasting initial and final consonants. An analysis of variance (ANOVA) showed main effects of background noise level ($p<0.001$), talker (NA versus NN; $p<0.04$), and consonant ($p<0.001$). The interaction of talker by consonant was significant ($p<0.03$).

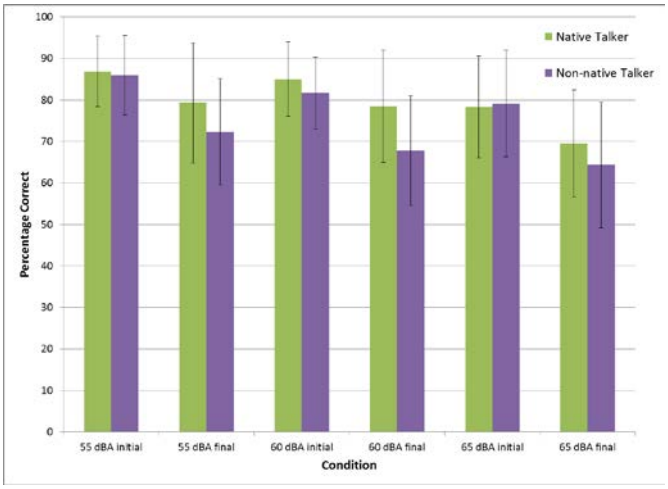


Figure 1: Mean percentage correct for MRT F2F and standard deviation, shown separately for initial and final consonants.

The results for the SPIN (combined for NA and NN listeners) are shown in Figure 2, separately for sentences with high and low contextual cues. An ANOVA showed main effects of background noise level ($p<0.001$) and contextual cues ($p<0.001$) and a between-subjects effect of listener ($p<0.006$).

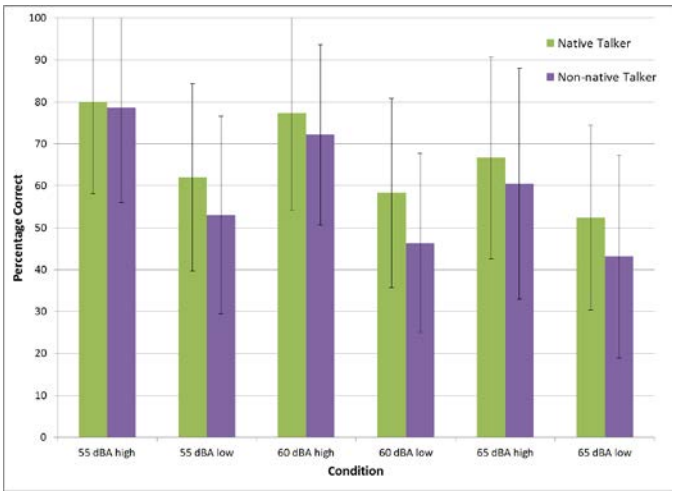


Figure 2: Mean percentage correct for the SPIN F2F and standard deviation, shown separately for low and high contextual cues.

3.3 Radio Condition

The results of the MRT and SPIN for the radio condition are summarized in Table 1. For the MRT, an ANOVA showed main effects of talker ($p<0.003$) and consonant position ($p<0.001$). For the SPIN, there were main effects of talker ($p<0.02$), contextual cues ($p<0.001$) and a between-subjects effect of listener ($p<0.01$).

Table 1: Mean percentage correct and standard deviation for the MRT (by initial and final contrasting consonant) and SPIN (by high and low contextual cues) for the radio condition.

	Native Talker	Non-Native Talker
MRT initial	90.7 ± 9.2	83.5 ± 10.3
MRT final	83.7 ± 11.7	71.5 ± 17.9
SPIN high	92.5 ± 10.2	84.8 ± 17.6
SPIN low	75.8 ± 16.4	67.2 ± 22.5

In all F2F and radio conditions, NN listeners performed more poorly overall on the SPIN than NA listeners.

4 Discussion

Based on the LEAP-Q responses for self-rating of English proficiency and years of education and work experience in English, the NN group can be described as highly proficient. Regardless, NN listeners performed more poorly than NA listeners on the SPIN test in both the F2F and radio conditions, but not on the MRT. This outcome may have been due to the fact for the SPIN, test listeners must recognize the final word, while for the MRT, response alternatives are provided. This suggests that it may be beneficial for NN workers to communicate with a limited vocabulary. Both NN and NA listeners had difficulty understanding NN talkers. Our future studies will investigate the benefits of linguistic training options in a military context. Previous studies have shown that listeners can be trained to improve speech understanding in noise [6].

References

[1] Farris C, Trofimovich P, Segalowitz N. and Gatbonton E. (2008). Air traffic communications in a second language: Implications of cognitive factors for training and assessment. TESOL Quarterly, 42(3):397-410.

[2] Marian V, Blumenfeld HK, and Kaushanskaya M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. J Speech Lang Hear Res, 50(4):940-967.

[3] Bell DW, Kruel EJ and Nixon JC. (1972). Reliability of the modified rhyme test for hearing. J Speech Hear Res 15(2):287-295.

[4] Kalikow DN, Stevens KN, Elliott LL. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. J Acoust Soc Am 61:1337-1351.

[5] Bent, T. and Bradlow, A.R. (2003). The interlanguage speech intelligibility benefit. J Acoust Soc Am 114(3):1600-1610.

[6] Song, J.H., Skoe, E., Banai, K. and Kraus, N. (2011). Training to improve hearing speech in noise: Biological mechanisms. Cereb Cortex 22(1):1180-1190.

CATEGORICAL PERCEPTION OF POSTALVEOLAR SIBILANTS BY TAIWAN AND BEIJING MANDARIN SPEAKERS

Masaki Noguchi *[1] **and Carla Hudson Kam** †[1]
[1]University of British Columbia

1 Introduction

This study addresses the question of how dialect-specific phonetic variation affects speech perception. Previous studies have pointed out that the contrast between dental and retroflex sibilants tends to be lost in connected speech in Taiwan Mandarin, while the contrast is maintained in Beijing Mandarin [1, 2]. A study on speech perception has demonstrated that, concomitantly, Taiwan Mandarin speakers have a broader/weaker category for the retroflex sibilants relative to the dental sibilants [1]. In this study, we tested whether the perception of retroflex sibilants by Taiwan Mandarin speakers relative to the other class of sibilants in the inventory of Mandarin phonemes is also affected. Specifically, we compared Taiwan and Beijing Mandarin speakers on their perception of two postalveolar sibilants, retroflex [ʂ] and alveolopalatal [ɕ].

2 Methods

2.1 Design

We used an ABX discrimination task as well as an identification task to test the categorical perception of the contrast between retroflex [ʂ] and alveolopalatal [ɕ]. Stimuli were drawn from a 10-step continuum from retroflex [ʂa] to alveolopalatal [ɕa]. In the ABX discrimination task, participants compared two test items that were separated by one step (e.g. step 1 vs. step 3). In the identification task, participants were asked to label a single test item either as retroflex or alveolopalatal.

2.2 Participants

10 Taiwan Mandarin speakers and 7 Beijing Mandarin speakers participated in the study. All were living in or visiting Vancouver at the time of the study. All participants in the Taiwan Mandarin group self-reported living in Taiwan until adolescence. Similarly, all participants in the Beijing group self-reported living in Beijing until adolescence. Participants were paid $5 for their participation.

2.3 Stimuli

The 10-step continuum was constructed by resynthesizing naturally produced tokens of [ʂa] and [ɕa]. The syllables were broken into frication noise and vowel portions, and separate continua were constructed for each. Then, the two continua were combined together to create a syllable continuum. For the frication noise, after the duration and the intensity were normalized, the PSOLA method was used to interpolate a stepwise transition from [ʂ] to [ɕ]. Figure 1 shows the LPC spectra of frication noise at steps 1, 4, 7, and 10: from step 1 to step 10, the weight of the spectra gradually shifts towards higher frequencies. For the vowel, after duration and intensity were normalized, STRAIGHT was used to interpolate a stepwise transition from [(ʂ)a] to [(ɕ)a]. Figure 2 shows the first two formants of steps 1, 4, 7, and 10, measured at 10 ms and 100 ms: from step 1 to step 10, the F1 onset gradually goes down and the F2 onset gradually goes up.

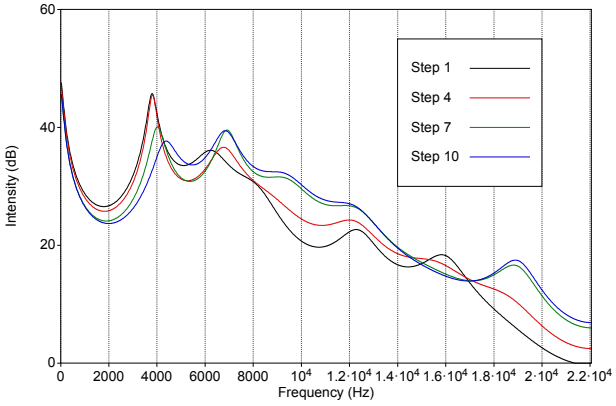


Figure 1: LPC spectra of resynthesized frication noise

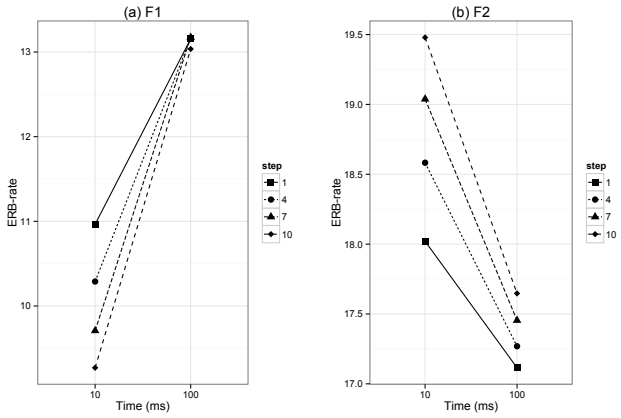


Figure 2: Formant transition of resynthesized vowel

2.4 Procedure

All participants did the ABX discrimination task first and the identification task second. Instructions were given in Mandarin in written form. For the ABX discrimination task, participants heard three syllables and were asked to decide whether the last syllable was identical to the first one or the second one. There were eight blocks in this task. Each block contained 32 trials (8 ABX triads in 4 different orders). For

∗. mskngch@alumni.ubc.ca
†. Carla.HudsonKam@ubc.ca

the identification task, participants heard a single syllable and were asked to identify it as either retroflex [ʂa] or alveolopalatal [ɕa]. Category labels were shown in *Zhuyin* for Taiwan Mandarin speakers and *Pinyin* for Beijing Mandarin speakers. There were eight blocks in the task, each consisting of 10 trials, one per step on the continuum. Trials within a block were presented in random order.

3 Results

For the discrimination task, responses to each unique triad were converted into sensitivity scores (d prime). Figure 3 shows the mean d prime scores by language group. Participants in both groups show the highest sensitivity for the trials comparing step 5 and step 7. A repeated-measures ANOVA was conducted with language group as a between-participant factor and pair as a within-participant factor. The analysis yielded a significant effect of pair ($F(7, 105)=21.45$, $p<0.001$) but not language group ($F(1, 15)=1.68$, $p=0.21$). There was no interaction between the two factors ($F(7, 105)=1.02$, $p=0.41$).

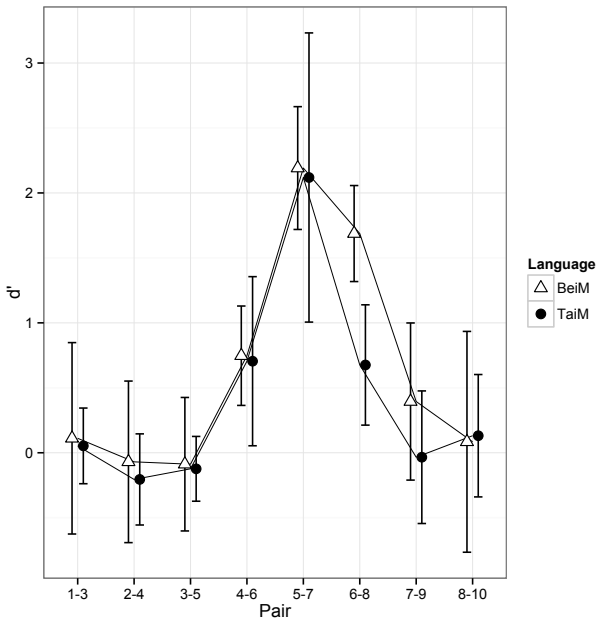


Figure 3: d prime (with 95% CI)

For the identification task, the proportion of /ɕa/ responses was calculated for each step on the continuum. Figure 4 shows the mean proportion of /ɕa/ responses by language group. For both groups, the proportion of /ɕa/ responses is close to 0 in steps 1-5, jumps to around .5 at step 6, then increases to close to 1 in steps 7-10, indicating a perceptual boundary between retroflex [ʂa] and alveolopalatal [ɕa] at step 6. Note that both Taiwan and Beijing Mandarin speakers show the same identification curve.

4 Discussion

The results of this study show that both Taiwan and Beijing Mandarin speakers perceive the contrast between retroflex [ʂa] and alveolopalatal [ɕa] in a categorical fashion, and there is no difference between the two language groups in where they place the boundary. It has been claimed that Taiwan Mandarin speakers have broader/weaker category for retroflex sibilants because of the merger between retroflex and dental sibilants in their dialects [1]. In contrast to previous studies showing a perceptual effect of this merger, the present study demonstrates that the weak category does not affect the perception of retroflex [ʂ] relative to alveolopalatal [ɕ]. A possible explanation for the seemingly conflicting results is that Mandarin speakers rely on different perceptual cues when they process different classes of sibilants. For example, studies have demonstrated that Mandarin speakers rely on frication noise to process dental and retroflex sibilants, but formant transitions to process alveolopaltal sibilants [3]. Therefore, it is possible that the merger affects the processing of cues from frication noise, but not from formant transitions. More research is needed to establish whether this is actually the case.

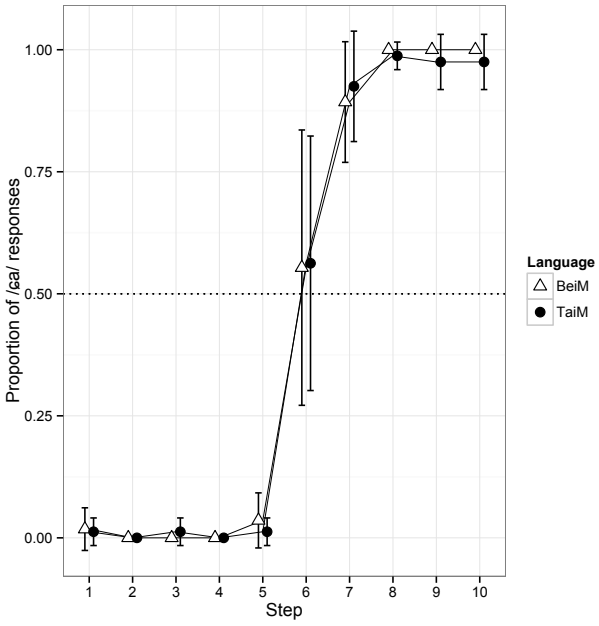


Figure 4: Proportion of /ɕa/ responses (with 95% CI)

References

[1] Y-H Chang. *Variability in cross-dialectal production and perception of contrasting phonemes : the case of the alveolar-retroflex contrast in Beijing and Taiwan Mandarin*. PhD thesis, University of Illinois at Urbana-Champaign, 2013.

[2] M Noguchi, C Chiu, P-C Wei, and N Yamane. *Uncovering sibilant fricative merger in Taiwan Mandarin : Evidence from ultrasound imaging and acoustics*, 2015. Poster presented at the Annual meeting of the Linguistic Society of America, Portland, OR.

[3] C Chiu. Attentional weighting of polish and taiwanese mandarin sibilant perception. In *Proceedings of the 2010 Canadian Linguistics Association Annual Conference*, 2010.

RHYTHM METRICS OF SPONTANEOUS SPEECH AND ACCENT IN JAPANESE AND ENGLISH

Yoichi Mukai [*1] **and Benjamin V. Tucker** [†1]
[1]Dept. of Linguistics, University of Alberta, Alberta, Canada, T6G 2E7

1 Introduction

The rhythm class hypothesis is an attempt to classify the human impression of language rhythm based on the concept of isochrony, which posits that language rhythm underlies the regulation of a particular unit re-occurring at regular intervals [1]. The present study provides an experimental evaluation of rhythm metrics between read and spontaneous speech in Japanese and English.

Researchers, such as Ramus et al. [2] and Grabe & Low [3], have proposed several rhythm metrics by adapting Dauer's [4] notion that phonological and phonetic factors modulating structures of syllables contribute to the perception of different language rhythms. Two types of rhythm metrics have been proposed, 'global' and 'local' metrics. Global metrics reveal the degree of overall durational variability of segments in utterances, e.g., %V—the relative proportion of vocalic intervals, and VarcoV—the standard deviation of vocalic intervals divided by the mean of vocalic intervals and multiplied by 100 [5]. Local metrics, on the other hand, capture durational differences between consecutive vocalic and consonant intervals, such as nPVI_V—the mean of the differences between successive vocalic intervals divided by their sum and multiplied by 100, see Eq. (1), and rPVI_C—the mean of differences between successive consonant intervals, see Eq. (2) [3].

$$\text{nPVI} = \sum_{m-1}^{k=1} |\frac{d_k - d_{k+1}}{(d_k + d_{k+1}/2)}|/(m-1) \quad (1)$$

$$\text{rPVI} = \sum_{m-1}^{k=1} |d_k - d_{k+1}|/(m-1) \quad (2)$$

Researchers have also applied these metrics to measure the influence of first language (L1) rhythm on second language (L2), and suggest that VarcoV and nPVI_V are particularly useful in predicting impressionistic judgements of accentedness [6].

Researchers, however, have identified inconsistencies in the classification of language rhythms across studies [3]. One possible cause is the type of speech collected (e.g., whether speech is spontaneously produced or carefully read). Though a few studies have compared read sentences to spontaneous speech using these metrics [7], Japanese has not yet been investigated in this respect. We predict that these metrics will differentiate read and spontaneous speech, with spontaneous speech falling on the faster side of each metric by showing a smaller proportion of vowels (lower %V) and more variability in the duration of vocalic intervals (higher VarcoV) than read sentences. Also, that spontaneous speech will allow for better discrimination of language groups by showing that L2 English falls between L1 English and L1 Japanese.

∗. mukai@ualberta.ca
†. bvtucker@ualberta.ca

2 Method

Spontaneous speech was collected from two existing data sets, The Wildcat Corpus [8] and unpublished data from Warner [9]. Forty-eight utterances were retrieved from three native speakers of Japanese (one male ; three utterances × three speakers = nine utterances), English (three males ; three utterances × three speakers = nine utterances), and six English-speaking Japanese subjects (one male ; five utterances × six speakers = 30 utterances). All were manually labeled using Praat [10], and segmented following Arvaniti [7]. Praat was then used to calculate %V, VarcoV, rPVI_C, and speech rate. We compare the present research to Grenon & White's [11] data (read speech) for L1 and L2 rhythm in Canadian English and Japanese. The resulting data were modeled, with linear mixed- effects regression using the lme4 package [12] in R to investigate the relationship between the metrics and the language groups across the speech types. The metrics and speech rate were entered as separate dependent variables, and the language groups and speech types were treated as fixed effects, with subjects (speakers) as random effects.

3 Results

Table 1 shows the means and standard errors of the metric scores and speech rates in the L1 English, L1 Japanese, and L2 English for the current study. Figure 1 illustrates the comparison of %V and VarcoV between the current study (spontaneous speech) and Grenon & White (read speech) [11]. A comparison of the present study to the Grenon & White's [11] data shows that the two data sets are significantly different for all the metrics except rPVI_C (%V (*t*=-2.69), VarcoV(*t*=2.27), Speech Rate(*t*=-5.04)). Analysis of %V for the spontaneous speech data indicates that the vowel for L1 English takes a significantly larger portion of the utterance than L1 Japanese (*t*=2.66). A similar result is also found for rPVI_C (*t*=2.924)

Groups	L1EN	L1JP	L2EN
Metrics			
%V	39.87 (1.19)	49.47 (2.34)	44.98 (1.57)
VarcoV	52.90 (5.26)	65.84 (5.86)	57.33 (3.22)
rPVI_C	83.18 (12.27)	42.53 (2.57)	68.39 (5.11)
SRate	3.82 (0.20)	4.72 (0.34)	4.11 (0.19)

Table 1: The means and standard errors of the metric scores and speech rates (SRate) in L1 English (L1EN), L1 Japanese (L1JP), and L2 English (L2EN).

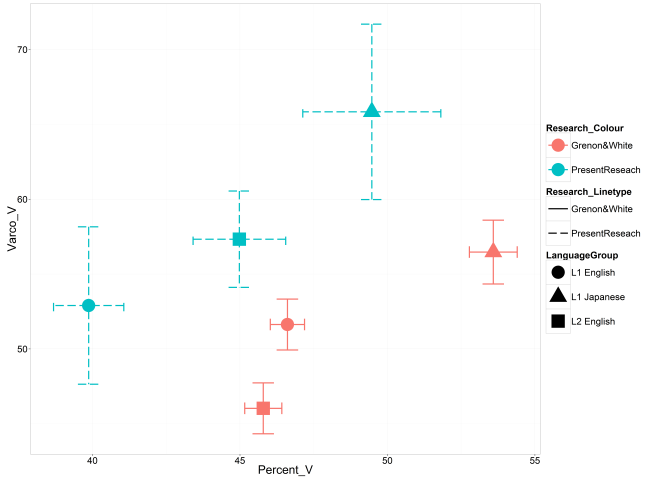


Figure 1: The plot shows the means and standard errors of the %V and VarcoV scores in L1 English, L1 Japanese, and L2 English speakers both in the present study (spontaneous speech) and Grenon & White (read speech) [11].

which indicates that there is more variability in the successive consonant intervals for L1 English as compared to L1 Japanese. Additionally, we find that L2 English significantly differs from L1 Japanese for rPVI_C (t=2.356). Analysis of VarcoV and Speech Rate do not show any significant differences between language groups (though some comparisons are trending toward significance).

4 Discussion

The variability and lack of some significant effects found in the present study may be due to the limited data size (we have only 48 utterances) rather than speech type differences. However, statistical analysis reveals that there are some significant differences between the metrics of the two speech types. As predicted, there are significant differences between spontaneous speech and read speech. Further, we find that spontaneous speech better discriminates the language groups by showing L2 English between L1 English and L1 Japanese, as predicted. It is possible that the read speech causes L2 English speakers to overemphasize English rhythm causing the L2 English not to fall between the two L1s. The difference may also be due to the differing L2 fluency levels. We predicted that spontaneous speech would fall on the faster side of each metric. However, the spontaneous data is produced at a slower speech rate than the read sentences, and it is produced with a smaller proportion of vowels and more variability in the duration of vocalic intervals. This is likely because subjects in Grenon & White [11] practiced reading their sentences before recording, and the familiarity with the sentences encouraged them to read at a high rate, but maintained the proportion of vowels and variability in the duration of vocalic intervals. It is noteworthy that L2 English significantly differs from L1 Japanese for rPVI_C, suggesting that spontaneous L2 English is similar to L1 English. In other words, their L2 speech is rhythmically less accented in terms of rPVI_C.

5 Conclusions

This study provides an experimental evaluation of rhythm metrics between read and spontaneous speech in Japanese and English. Statistical analysis indicates that these metrics are sufficient for distinguishing between the two speech types. Additionally, our findings indicate that L2 speakers do fall between L1s in spontaneous speech. Further examination is required with a larger sample of data to further investigate the details of linguistic isochrony and its relationship to impressionistic observations of language rhythm.

Acknowledgments

We would like to thank Laurence White and Izabelle Grenon for generously sharing their data with us. We would also like thank Natasha Warner for sharing her spontaneous data, and Volker Dellwo for sharing his Praat scripts.

References

[1] David Abercrombie et al. *Elements of general phonetics*, volume 203. Edinburgh University Press Edinburgh, 1967.

[2] Franck Ramus, Marina Nespor, and Jacques Mehler. Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3) :265–292, 1999.

[3] Esther Grabe and Ee Ling Low. Durational variability in speech and the rhythm class hypothesis. *Papers in laboratory phonology*, 7(515-546), 2002.

[4] Rebecca M Dauer. Stress-timing and syllable-timing reanalyzed. *Journal of phonetics*, 1983.

[5] Volker Dellwo. Rhythm and speech rate : A variation coefficient for deltac. *Language and language-processing*, pages 231–241, 2006.

[6] Laurence White and Sven L Mattys. Calibrating rhythm : First language and second language studies. *Journal of Phonetics*, 35(4) :501–522, 2007.

[7] Amalia Arvaniti. The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, 40(3) :351–373, 2012.

[8] Kristin J Van Engen, Melissa Baese-Berk, Rachel E Baker, Arim Choi, Midam Kim, and Ann R Bradlow. The wildcat corpus of native-and foreign-accented english : Communicative efficiency across conversational dyads with varying language alignment profiles. *Language and speech*, 53(4) :510–540, 2010.

[9] Natasha Warner. Speech reduction across languages and dialects, 2015.

[10] Paul Boersma and David Weenink. Praat, a system for doing phonetics by computer. 2001.

[11] Izabelle Grenon and Laurence White. Acquiring rhythm : A comparison of l1 and l2 speakers of canadian english and japanese. In *Proceedings of the 32nd Boston University conference on language development*, pages 155–166, 2008.

[12] Douglas Bates, Martin Maechler, and Ben Bolker. lme4 : Linear mixed-effects models using s4 classes. 2012.

THE EFFECTS OF DURATION ON HUMAN PROCESSING OF REDUCED SPEECH

Dylan Bernhard[1] and Benjamin V. Tucker[2]

[1]Department of Linguistics, University of Alberta, Edmonton, AB, Canada T6G 2E7, dbernhar@ualberta.ca

[2]Department of Linguistics, University of Alberta, Edmonton, AB, Canada T6G 2E7, bvtucker@ualberta.ca

1 Introduction

For most listeners, parsing of reduced speech is required on a daily basis. When humans produce speech in a fast manner or casual context, there is a tendency to delete, assimilate, and generally weaken the phonemes that are being produced. The resulting acoustic signal can vary by a large degree from what speakers would consider a canonical pronunciation [3]. While studies have been performed to identify exactly how humans can translate from reduced forms to standardized internal representations, we are still far from a comprehensive model to explain all of the aspects of our understanding of how listeners parse a reduced speech input.

Past studies have shown that listeners use a combination of syntax, semantics, rate of speech, and proximal phonetic cues to aid in the processing of reduced speech [4, 5, 6, 9]. This present study investigates the degree to which duration affects human processing of reduced speech stimuli, with the hope of aiding in the building of a comprehensive model of how humans perceive reduced speech.

2 Method

Responses were gathered from 51 participants, 2 sets of responses had to be excluded from the analysis due to lost or damaged data. Participants in this study were students from introductory level Linguistics classes at the University of Alberta, and received partial course credit for participation.

This study consisted of cloze tasks [8] with an auditory component, where sections of reduced speech corresponding to between one and five words were removed from spontaneous utterances (frames) and replaced with silence. Seventy-two frames consisting of phrase length utterances with varying degrees of reduction were extracted from recorded audio of a young adult Western Canadian English speaker [7]. In each frame, a target of 1 to 5 words in length was extracted and replaced with silence to form the original gap stimuli. This set of 72 original gap length stimuli was then manipulated into two additional sets. In the first, the original stimuli were manipulated to have silent gaps with twice the original length. In a similar way, the second set was manipulated to consist of frames with half-length silent gaps. Two controls were also prepared: one where only a visual cloze was presented, and another where the participant heard the full audio from the original frame.

Participants were presented with a visual cloze where the target had been replaced with ten underscores, and a corresponding audio. They were then asked to type in the standard English orthography what word or words they thought best fit in the gap. Each participant was presented with the stimuli in five blocks. The first contained all of the visual cloze stimuli in a randomized order. Blocks two through four were then run, with each consisting of a set of seventy-two audible cloze stimuli with varying gap conditions (a balanced set of 24 short, original, and long gaps). Each of these blocks only contained one of each frame. The final block consisted of the full audio stimuli. This setup ensured that each participant was given an opportunity to respond to every stimulus that had been created for this study.

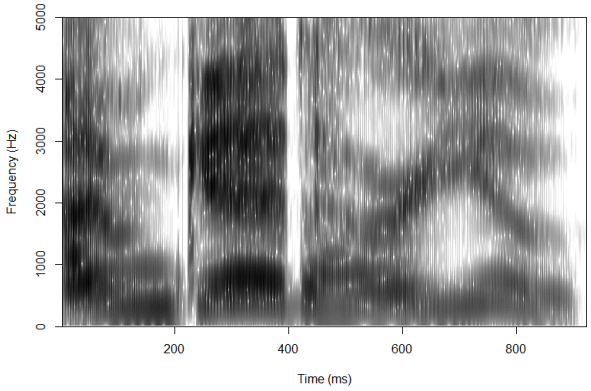


Figure 1: Example original stimulus

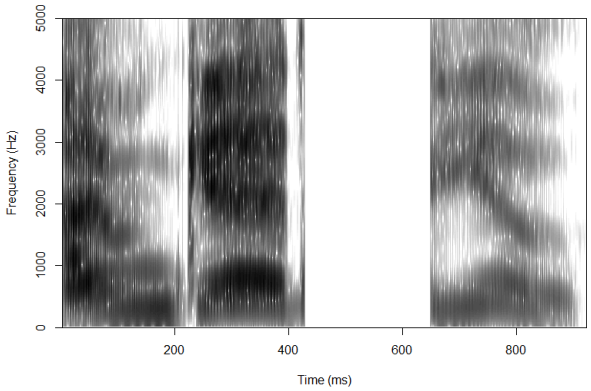


Figure 2: Example auditory cloze stimulus (original gap length)

3 Results

Responses were graded as correct or incorrect based upon what words were part of the original utterance. In trials where the participant did not give a response, the trial was marked as incorrect, as this implies that the information given was not enough to properly identify any possible response. Contractions were marked as their constituent words (e.g., "didn't" becomes "did not").

With only the visual information given in the Visual Cloze task, the rate of correct responses was low, at 8.4%. Given more auditory data, participants became approximately twice as good at the task, with the short gap, original gap, and long gap response rates being 15.8%, 16.7%, and 17.7% respectively. Finally, with the Full Audio task, participants' responses were 79.6% correct.

A logistic linear mixed-effects regression using the lme4 package [1] in R was performed to analyze the significance of the difference between the different tasks (Visual Cloze, Auditory Cloze and Full Audio). A first simple effects analysis with subjects and items as random-effects factors comparing the three different tasks finds that responses to the Visual Cloze task are significantly less accurate than the Auditory Cloze ($p<0.001$) and Full Audio ($p<0.001$). We also found that the responses to the Auditory Cloze were significantly less accurate then to the Full Audio ($p<0.001$). A second analysis was performed on the items for the Auditory Cloze task to investigate differences between the three manipulation types (Original, Short, Long). In this analysis, the responses from the Visual Cloze task were used as a way to measure the predictability of each item. To create this variable we calculated the average response for each item in the Visual Cloze task and used this value as a predictor (vcScore) in the second model. In this model Type of Manipulation and vcScore were our independent variables with subjects and items as random-effects factors. A summary of the model is found in Table 1.

	Estimate	S. Error	z value	Pr(>|z|)
Intercept	-4.47	0.213	-21.02	< 2e-16 ***
Type: Long	0.22	0.082	2.71	0.00672 **
Type: Orig	0.10	0.083	1.22	0.22235
vcScore	11.78	0.748	15.75	< 2e-16 ***

Table 1: Results of linear mixed-effects regression for the auditory cloze manipulations

In this analysis we find that vcScore is a significant predictor, indicating that as the predictability increases the responses are more accurate. We also see a significant different between the Long and the Short manipulations, indication that participants were significantly more accurate for the Long manipulation than the Short manipulation.

4 Discussion

We find that the increase in correct response rate between the visual cloze control and the auditory cloze conditions supports the claim that there are proximal phonetic cues that humans use when parsing speech [7].

With regard to the control condition which presented participants with the full audio frame and with no silent gap, the rate of correct responses is lower than might be expected. At approximately 80%, it would seem that some factor from our stimuli is hindering complete recognition of the target. This is interesting because there was no indication in the original recording the other interlocutor had any difficulty understanding the conversation. This is likely an effect of degree of reduction; more reduced targets are harder to parse than those closer to their canonical pronunciation. This combined with the fact that the participants don't have the full conversational context make it a more difficult task.

The results of this study also indicate that there is an effect of duration on human speech processing of spontaneous speech. However, there does appear to be a trend in the data that the Long stimuli were easier to predict than the Original duration, and that Short stimuli were the most difficult, even though the only significant comparison is between the Long and Short items. This difference suggests that the increased processing time offered by a larger gap is aiding in word prediction.

It is likely that the listeners make predictions with regard to the expected duration when processing speech [2]. This would support a speech processing model that contains activation and competition in the mental lexicon. The reason for shorter gaps being harder to predict correctly is that the correct response is removed from the pool of possible candidates due to the gap being too short to fit it.

The current experimental results suggest that the inclusion of a measure of reduction would benefit the analysis. Measures of reduction have proved only marginally predictive in previous research [6] but may be useful here. It may also be fruitful to investigate the response length. Items that may have been marked incorrect might indicate a correlation between the duration manipulation and response length, such that participants provide shorter responses for the shorter duration and longer responses for the longer silence duration manipulation.

Overall, the results of this study indicate that listeners are sensitive to manipulations of duration in the reconstruction of spontaneous speech. In combination with the results of other studies investigating speech perception, we may yet discover aspects which assist speech recognition technologies to recognize more spontaneous speech.

References

[1] Bates D, Maechler M, Bolker B and Walker S (2014). _lme4: Linear mixed-effects models using Eigen and S4_. R package version 1.1-7, <URL: http://CRAN.R-project.org/package=lme4>.

[2] Dilley, L. C., & Pitt, M. (2010). Altering context speech rate can cause words to appear or disappear. *Psychological Science*, 21(11). 1664-1670.

[3] Dilts, P. (2013). *Modeling phonetic reduction in a corpus of spoken english using random forests and mixed-effects regression*. Thesis, Department of Linguistics, University of Alberta, Edmonton, AB.

[4] Ernestus, M., Baayen, H., & Schreuder, R. (2002). The recognition of reduced word forms. *Brain and Language, 81*(1), 162-173.

[5] Mehta, G., & Cutler, A. (1988). Detection of target phonemes in spontaneous and read speech. *Language and Speech. 31*(2), 135-156.

[6] Pickett, J., & Pollack, I. (1963). Intelligibility of excerpts from fluent speech: effects of rate of utterance and duration of excerpt. *Language and Speech 6*(3), 151-164.

[7] Podlubny, R. (2013). *Acoustic Decomposition: The Roles of Duration and Intensity in Spoken Language Processing*. Honors Thesis, Department of Linguistics, University of Alberta, Edmonton, AB.

[8] Taylor, W. L. (1953). "Cloze Procedure": a new tool for measuring readability. *Journalism Quarterly, 30*(4), 415-33.

[9] van de Ven, M., Tucker, B. V., & Ernestus, M. (2011). Semantic context effects in the comprehension of reduced pronunciation variants. *Mem. & Cog., 39*(7), 1301-1316.

VOWELS SPACES AND REDUCTION IN PLAINS CREE

A.G. Harrigan * **and Benjamin V. Tucker** †
University of Alberta, 116 St & 85 Ave, Edmonton, AB T6G 2R3

1 Introduction

Plains Cree is an Algonquian language [1] spoken by approximately 35,000 individuals [2]. Plains Cree can be divided into a series of areal dialects that exhibit grammatical and phonetic differences. The present work focuses on the community of Maskwacîs, Alberta and investigates the acoustic characteristics of vowels and phonetic reduction in the community.

2 Background

According to Thunder [3], Plains Cree vowels come in pairs of short and long, orthographically distinguished by a circumflex to mark length: a â, i î, o ô. The mid-high front vowel <ê>, which is written with or without a circumflex, is always considered long. Despite indicating that the vowels sound different, Thunder [3] differentiates vowels in a pair primarily through duration.

Muehlbauer [4] phonetically documents the vowel space of three speakers from various Cree communities using archival recordings. He reports that short and long vowels differ in both duration and quality. Muehlbauer [4] measures the F1 and F2 of vowels occurring between /k/ and /t/. This acoustic vowel plot indicates that, while front vowels overlap some, the vowel pairs are qualitatively different.

Relatively little research has also been conducted on Plains Cree reduction. Russel [5] focused on the reduction of a vowel when preceding another across a word boundary. Russell uses the orthographic form as an indication of the pre-reduced form. Investigating instances of various sandhi (in the form of $(V_1)\#V_2$) and comparing the F1 and F2 of V_2, vowel averages, Russell determines the effect reduced V_1 has on the remaining V_2. Indeed, most remaining vowels show diphthongization in their onset when compared to the average vowel, indicating that "reduced" vowels are present.

3 Methodology

We first investigate the acoustic characteristics of the vowels in Maskwacîs Cree. This will provide a baseline for comparison for determining what sort of reduction is occurring elsewhere for the vowels. Syllables with a glide onset (/j/ and /w/) were selected for the study of reduction. This restriction was based on a pilot study, which identified these syllables as the most likely candidates for reduction.

3.1 Speakers

Three fluent female speakers from Maskwacîs, ages 50-75, participated in the recordings. The speakers all work as language educators.

3.2 Vowel Acoustics Items

Following Muehlbauer [4], we created a list of words containing vowels in various positions. Target words contained vowels in word initial (before either /t/ or /s/) stressed and unstressed positions, as well as interconsonantally (between two /s/ phonemes) in a stressed position. English translations of targets were used for elicitations.

3.3 Reduction Items

Focusing on the reduction of words with glide onsets, syllables were selected from word initial, medial, and final positions (where possible). Elicitations followed the vowel acoustics items and we used the English translation of the word for elicitation. Similar to Russell [3], we used the Standard Roman Orthographic representation of a word as a pre-reduced/canonical form.

3.4 Recordings

Recordings were conducted in Maskwacîs, AB in a quiet room using a portable recorder (Korg MR2000) and a head mounted omnidirectional microphone (Countryman E6). Mark-up, transcription and measurement were performed using Praat [6].

4 Results and Discussion

4.1 Vowel space

We find the vowel space of Maskwacîs Plains Cree shows some overlap for the front vowels, in accord with Muehlbauer [4]. However, there are qualitative differences distinguishing the front vowels. We find substantial overlap for the back and low vowel pairs, illustrated in Figure 1. Orthographically long vowels are indicated with a double vowel in all figures.

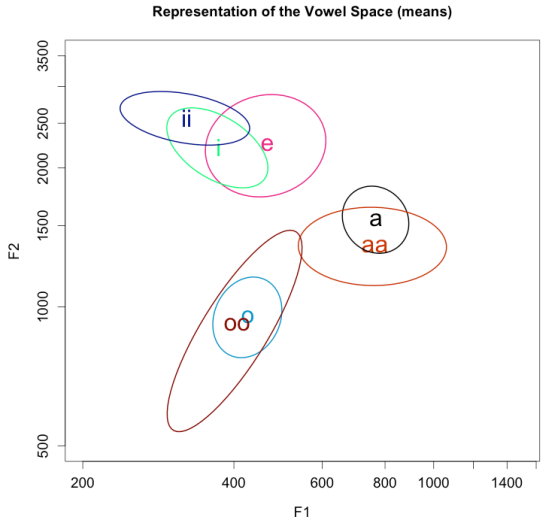


Figure 1: Vowel Space of Maskwacîs Cree for three female speakers (log Hz).

* galvin@ualberta.ca
† bvtucker@ualberta.ca

The short low vowel generally has a higher F2 than the long vowel. We also note the overlap of <o> and <oo>, with the long vowel's standard deviation almost entirely encompassing the short vowel's. In fact, each long/short vowel pairing indicates some overlap. This may explain the expected duration distinction between long and short vowels, seen in Table 1.

Vowel	a	aa	i	Ii	o	oo	e
Duration (ms)	83	157	97	122	122	157	203

Table 1: Mean duration of vowels across participants.

As in Muehlbauer [7], stress seemed to have no effect on vowel space or duration.

4.2 Reduction

Looking towards reduction, we identify three main types: fusion of /j/ near high front vowels, assimilation of /i/ to [u] when preceding a /w/, and the fusion of /w/ when near high back vowels.

Dealing first with /j/ reduction near high vowels, we look to Figure 3. Because they're phonetically similar, the reduction of /j/ and /i/ is not surprising. In Figure 3 we see an example of this reduction, leading the /iji/ cluster to surface as [iː]. Interestingly, the resulting [iː] is longer (102 ms) than the speaker's average /i/ (83 ms), indicating an instance of sound fusion, rather than deletion. Similarly, we see the fusion of /ɑji/ into [ɑɪ].

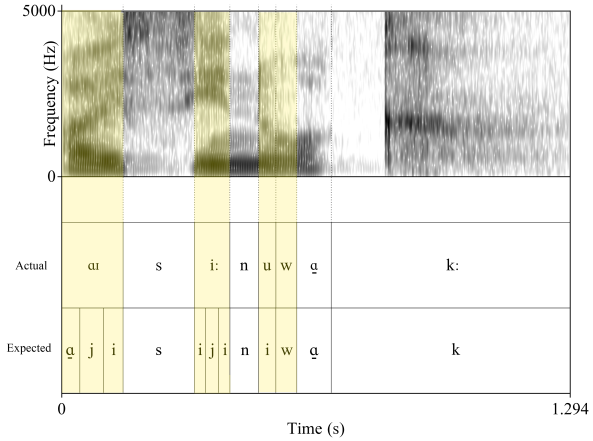


Figure 3: Reduction of /iji/ to [iː] and /iw/ to [uw] in the word <ayisiyiniwak> (meaning 'People').

Figure 3 also illustrates the second form of reduction, where find the change of /iw/ to /uw/. This can be seen as an assimilation of the high front /i/ to the back-round /u/. Although this sort of assimilation is mentioned by Thunder [3], the /w/ is generally lost, rather than left intact as in Figure 3.

The last reduction form concerns /w/ near mid and back vowels. We find two instances of this reduction illustrated in Figure 4. First is the reduction of /ɑ̱wo:/ to [oː]. The resulting /oː/ has a duration of 184 ms, just slightly less than the average /oː/ of 187 ms. Because of this slight difference, this reduction might be one of sound deletion rather than fusion.

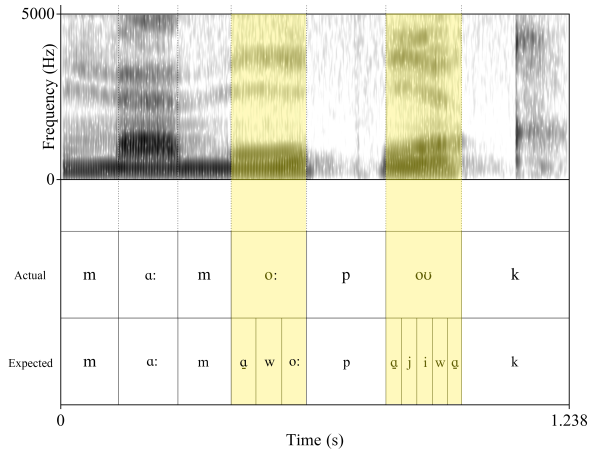


Figure 4: Reduction of /ɑ̱wo:/ to [oː] and /ɑ̱jiwɑ̱/ to [oʊ] in the word <maamawoo piwak> (meaning 'They come together).

Finally is the complex reduction of /ɑ̱jiwɑ̱/ to [oʊ]. Because of the multiple forms of reduction happening, it's difficult to determine the order of reduction. It seems likely that /ɑ̱ji/ reduced into something resemebling /iː/, which was assimilated by the /w/ into a /o/; Then, /w/ may have fused with the /ɑ̱/ to create a lower, unrounded /ʊ/.

5 Conclusion

We find that the Maskwacîs Plains Cree vowel space exhibits some overlap in the vowel space but we also find indication of qualitative and quantitative differences between long and short vowels. Reduction of /j/ and /w/ onset syllables is prevalent in the speech examined. More research is needed to determine further instances of reduction and help identify more possible causes of reduction.

Acknowledgments

We thank the speakers for their time, and the University of Alberta Faculty of Arts Roger S. Smith award for funding a pilot of this study.

References

[1] Wolvengrey, A. E. 2011. Semantic and pragmatic functions in Plains Cree syntax. LOT. Accessed November 6, 2014. < http://dare.uva.nl/record/1/342704>

[2] Ethnologue. 2015. Cree, Plain. Web. Accessed March 26, 2015 < http://www.ethnologue.com/language/crk>

[3] Thunder, Dorothy. 2012. *Introductory Cree: Part 1*. University of Alberta: Faculty of Native Studies.

[4] Muehlbauer, Jeffery. 2012. Vowel spaces in Plains Cree. *Journal of the International Phonetic Association*, 42, 91–105

[5] Russell, Kevin. 2008. Sandhi in Plains Cree. *Journal of Phonetics.* 36.3: 450–464

[6] Boersma, Paul & Weenink, David (2015). Praat: doing phonetics by computer [Computer program]. Version 5.4.16, retrieved 16 August 2015 from http://www.praat.org/

[7] Muehlbauer, Jeffery. 2006. Pitch as accent in Plains Cree. In Wolfart Christoph H. (ed.), *Actes du Trente-septie`me Congre`s des Algonquinistes*, 229–268. Winnipeg, MB: University of Manitoba Press.

TOWARDS CONVERGENCE OF METHODS FOR SPEECH AND SIGN SEGMENTATION

André Xavier [1], Oksana Tkachman [1], Bryan Gick [12]
[1]Department of Linguistics, University of British Columbia, Vancouver, BC, Canada
[2]Haskins Laboratories, New Haven, CT, USA

1 Présentation (Français)

Comme la parole, déterminer la limite d'un signe dans un flot de signes n'est pas une tâche facile. Pour certains chercheurs, le signe commence lorsque la main atteint l'emplacement où le signe est produit et termine lorsque la main quitte ce lieu [1-3]. Par conséquent, ces chercheurs laissent les mouvements transitoires hors des limites du signe. D'autres chercheurs pensent que les transitions doivent être partiellement ou entièrement considérées comme partie du signe, parce que (1) certaines caractéristiques articulatoires du signe sont visibles même avant ou encore après le signe est produit, et (2) les percepteurs sont capables de deviner les signes avec seulement l'information transmis dans les transitions [4-8]. L'étude présente utilise les données de video du langue des signes brésilienne (Libras; Xavier 2014) pour évaluer de façon critique les critères traditionnellement utilisés pour délimiter les éléments lexicaux dans le flot de signes. Utilisant la théorie de commande de moteur, nous proposons un traitement unifié de la production de langue, quelle que soit sa modalité.

2 Presentation (English)

Similar to speech, determining sign boundaries in a signing stream is not an easy task. For some researchers, the sign starts when the hand reaches the position where the sign is produced and ends when the hand leaves that position [1-3]. Thus, these researchers leave transitional movements out of the sign limits. Other researchers think that transitions should be partially or entirely regarded as part of the sign, because (1) some articulatory features of the sign are visible even before or still after the sign is produced, and (2) perceivers are able to guess signs solely drawing on information conveyed during transitions [4-8]. The present study uses video data of Brazilian Sign Language (Libras; Xavier 2014) to critically evaluate the criteria traditionally used to delimit lexical items in the sign stream. Using motor control theory, we propose a unified treatment of language production, regardless of its modality.

3 The study

3.1 Background

According to the traditional view, the location of the hand in sign language delimits the boundaries of a sign: the sign begins when the hand reaches the position where the sign is initially or entirely produced, and ends when the hand leaves that position [1-3]. This is also how signs are usually depicted in dictionaries, where only features crucial for sign recognition are depicted [7-8]. As a result, the *transitional movements*, that is, movements produced between signs to place the hand at a certain location or to move it to the rest position [10], are left out as irrelevant which results in multiple gaps in the signing stream.

The alternative view claims that the signs last longer, from the moment when some articulatory features can be recognized and until these features disappear in the transitional movements, and signers can guess the sign based solely on these features of the transitional movements [6-8]. Jantunen (2010b) also shows that mouthing can also indicate sign boundaries [11]. Gesture scholars [4] go even further and analyze the gesture/sign as production phases of preparation, stroke, and retraction.

3.2 Method and analysis

We used video data of Brazilian Sign Language (Libras) [9] to critically evaluate the traditional criteria of sign segmentation. The data have been analyzed and annotated using ELAN software [12].

We follow the alternative view [4, 6-8] in considering transitional movements as parts of the signs, but we further refine this claim by taking into account speech motor control theory, also successfully applied to spoken language production analysis [13]. In this approach, articulatory movement coincides with the end of an articulatory task, after which the articulator loses the control over the movement. Thus, when the hand has clearly accomplished its task, the sign ends. Therefore, the sign consists of the preparation and stroke phases only. We keep this distinction based on the qualitative differences between transitions and signs demonstrated in [6]. Retractions are annotated but not considered part of the sign.

This approach makes delimiting of sign boundaries easier compared to Jantunen's proposal [6-8], because coarticulation effects make it hard to distinguish between articulatory features of one sign and of the following sign, which makes determining a boundary between them problematic. In our transcriptions, the signing stream is broken into a sequence of preparation and stroke phases with no gaps between signs. We place the sign boundary at the frame where the hand starts to exhibit at least one of the features of the sign, which we take as the beginning of the sign production. Determining the beginning of the sign production is easier when the sign is at the onset of an utterance than when it follows some other sign. In the latter case, we place the sign boundary on the frame that

immediately follows the last frame of the stroke, similar to the beginning of retraction phases.

3.3 Preliminary results and discussion

The preliminary results show that the adopted approach is effective in determining sign boundaries as well as enables us to observe hand behaviour left out in the traditional approach to sign segmentation. Here we give examples of coarticulatory effects in one- and two-handed signs.

In two-handed signs, the dominant hand of a sign in the utterance onset position starts its preparation before the nondominand hand; however, both hands start the stroke synchronously. On the other hand, if the two-handed sign follows another two-handed sign, the hands act synchronically both in preparation and stroke phases. If a one-handed sign follows the two-handed sign, however, the nondominant hand of the two-handed sign moves back to the rest position before the dominant hand has finished its stroke. This effect is not observed if the two-handed sign is utterance-final; both hands produce the stroke and retract synchronously.

We also found some counter-evidence to Jantunen's claim [8] that mouthing coincides with sign boundaries: in some cases mouthing preceded or lags behind the preparation phase of the hand(s), or lasts across more than one sign.

4 Conclusion

Comparing natural human languages of different modalities is important in order to determine which aspects of those languages are modality-specific, and which are characteristic of all human languages. Likewise, the strength of a theory can be evaluated by applying it not only to data for which the theory was proposed, but also to additional data that were not originally meant to be accounted for by the theory. In this study we have successfully applied speech motor control theory to sign language data and the preliminary results show that methods used by speech researchers to delimit units in the speech stream are a good fit for delimiting units in the sign stream as well. By analyzing signs as a combination of preparation and stroke phases we not only were successful in identifying sign boundaries without leaving any gaps in the annotation, but also were able to capture coarticulation effects left out in the traditional approach to sign segmentation. This study points to the possibility of a unified account of articulatory movements of human language, regardless of its modality.

Remerciements/Acknowledgments

This project was funded by NSERC. We wish to express our gratitude to Adriana Dias for her help with data gathering and Emily Sadlier-Brown for her help with manuscript preparation.

References

[1] O. Crasborn, and I. Zwitserlood. *Annotation of the Video Data in the* "Corpus NGT." Department of Linguistics and Centre for Language Studies, Radboud University Nijmegen, the Netherlands. 2008. http://hdl.handle.net/1839/00-0000-0000-000A-3F63-4

[2] T. Johnston. Guidelines for Annotation of the Video Data in the Auslan Corpus. Department of Linguistics, Macquarie University, Sydney,Australia. 2009. http://media.auslan.org.au/media/upload/attachments/Annotation Guidelines_Auslan_CorpusT5.pdf (accessed January 15, 2010).

[3] R. E. Johnson, and S. K. Liddell. Toward a Phonetic Representation of Signs: Sequentiality and Contrast. *Sign Language Studies* 11: 241–74, 2011.

[4] S. Kita, I. Van Gijn, and H. Van der Hulst. Movement phases in signs and co-speech gestures, and their transcription by human coders. In *Gesture and sign language in human-computer interaction*, pp. 23-35. Springer Berlin Heidelberg, 1998.

[5] J. Bressem, and S. H. Ladewig. Rethinking gesture phases: Articulatory features of gestural movement?. *Semiotica*, 184: 53-91, 2011.

[6] T. Jantunen. On the role of transitions in signed language (or: What's wrong with the sign?). Paper presented at the 10th conference on the Theoretical Issues in Sign Language Research (TISLR10), Purdue University, West Lafayette, Indiana (USA), September 30–October 2. 2010.

[7] T. Jantunen. Signs and transitions: Do they differ phonetically and does it matter? *Sign Language Studies* 13(2): 211–237, 2013.

[8] T. Jantunen. How long is the sign? *Linguistics* 53(1): 93-124, 2015.

[9] Xavier, André Nogueira. "Uma ou duas? Eis a questão! Um estudo do parâmetro número de mãos na produção de sinais da língua brasileira de sinais (libras)". 158 f. Tese (Doutorado em Linguística). Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2014.

[10] S. K. Liddell, and R. E. Johnson. American sign language: The phonological base. *Sign language studies*, *64*(1): 195-277, 1989.

[11] T. Jantunen. A comparison of two linguistic sign identification methods. In Philippe Dreuw, Eleni Efthimiou, Thomas Hanke, Trevor Johnston, Gregorio Martínez Ruiz & Adam Schembri (eds.), *Proceedings of the 4th workshop on the representation and processing of sign languages: Corpora and sign language technologies* (organized as a part of the 7th Language resources and evaluation conference LREC at Valletta, Malta, May 22–23, 2010), 129–132. Paris: ELRA. 2010b

[12] Technical Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, ELAN – Linguistic Annotator (Version 4.6.2) [Computer Program]. First retrieved November 1, 2013, from http://www.lat-mpi.eu/tools/elan/

[13] C. P. Browman, and L. Goldstein. Articulatory phonology: An overview. *Phonetica*, *49*(3-4), 155-180, 1992.

ACOUSTIC AND ARTICULATORY QUALITIES OF SMILED SPEECH

Megan Keough, Avery Ozburn, Elise Kedersha McClay, Michael David Schwan, Murray Schellenberg, Samuel Akinbo, Bryan Gick*
University of British Columbia

1 Introduction

Studies have shown that listeners can use acoustics to identify smiled speech [1,2,3], which is characterised by increased amplitude, higher f0 [1], and some increases in formants [3,4]. Though smiling changes vocal tract shape [5], little is known about its effect on articulation, and how those articulatory changes affect the acoustics, nor whether results from single-speaker studies such as [4] apply to a larger population. Our study addresses this gap through a multi-speaker production experiment examining the articulation and acoustics of smiled and neutral English vowels, focusing on formant values, lip spreading, lip protrusion, lip angle, and larynx height. Facial positions were tracked following [4], and larynx height was measured with ultrasound following [6]. We hypothesize that smiled speech, compared to neutral speech, is characterised by higher f0, higher formant frequencies, raised larynx, and spread lips with corners turned up.

2 Methodology

Results are from 10 native English speakers (5 male, 5 female). Three target words (key, caw, coo) were used to elicit the vowels /i/, /a/, and /u/ in the carrier sentence 'I got a_toy', and were presented in E-Prime [7] alongside an image of a toy. Stimuli were grouped into 8 blocks of randomized all-neutral or all-smiled sentences, with 3 repetitions per sentence per block; there was also a training block. Neutral and smiled blocks alternated, with first block type randomly assigned. To remind participants of the target facial expression, stimulus slides were followed by images of neutral or smiling children from [8]. 16 dot stickers were placed at anthropometrically-defined points on participants' faces (Fig. 1).

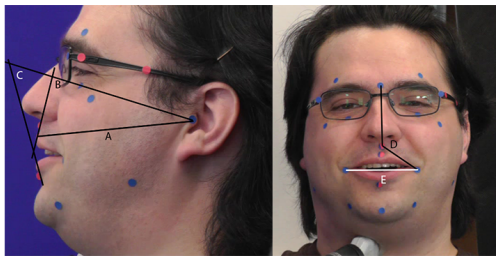


Figure 1: Anthropometric dots. A=Lip Protrusion Length (LPL), B=Upper Protrusion Angle (UPA), C=Lower Protrusion Angle (LPA), D=Lip Corner Angle (LPA), E=Lip Spread Length (LPL)

An Aloka Pro-Sound SSD 5000 ultrasound machine with an Aloka UST-9119-3.5 convex transducer (pulse frequency 3.5MHz, field of view 120º) collected larynx images. Following [6], the probe was positioned manually against the right thyroid lamina near the laryngeal prominence; it was kept stable with a mechanical arm. The audio (digitized at 44.1kHz) was recorded using a Sennheiser MKH 8060 shotgun microphone. Two Panasonic HC-V700 cameras recorded facial video.

Target vowels from the audio were delineated by hand. Praat [9] scripts marked the vowel midpoints, extracting midpoint f0, F1, and F2 values. A JPEG from each vowel midpoint was extracted from all videos. Facial markers were used to measure distances (in pixels) and angles in ImageJ [10]. For ultrasound images, annotators measured the distance from the most stable lamina point to the edge of the frame and converted to centimetres using the pixel length of the ultrasound's 10cm ruler.

3 Results and discussion

Measures were normalized within-speaker using a z-score transform. Data was subset by vowel, and for each (normalized) measure, a mixed effects model was run (using the lme4 package [11] in R [12]) with the measure as a dependent variable, a fixed effect of condition (smiled vs. neutral) and a random effect for subject with a random slope for condition.

3.1 Acoustic results

As Fig. 2 shows, smiling significantly ($|t| \geq 2$) raised f0 for all three vowels ($t = 6.13, 6.19, 5.97$ for *caw, coo, key*), an expected result based on previous literature [1,3,13].

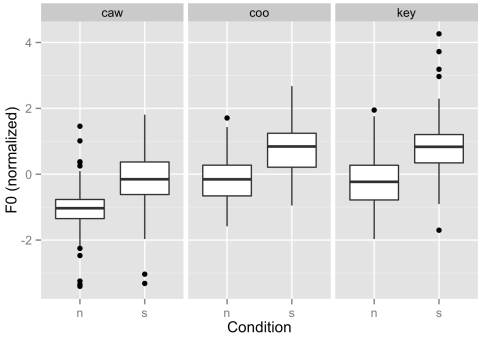


Figure 2: Normalized F0 for each vowel in neutral (n) and smiling (s) conditions.

* gick@mail.ubc.ca

However, while [1] found that all formant frequencies were higher in smiled speech, we found that smiling only significantly raised F1 and F2 for *caw* (t = 3.19, 3.26 respectively); F1 and F2 results for the other vowels were not significant. Our data thus align more with [13], who found that f0 was one of the most important factors in perceiving smiling. Further, we would expect /a/ to be affected most by smiling, since only in /a/ are the lips phonetically unconstrained; in /i/, they are spread, and in /u/, rounded. However, these results do not support findings of increased formants for only /i/ in German [4].

3.2 Larynx Measurements

The results showed no significant effect of smiling on larynx height. This is one way in which our results depart from [13], which found larynx height and f0 to be the main predictors of speech being perceived as smiled. Also, since we saw a significant effect in f0, and since larynx height is a main influence on f0 [14], the lack of larynx height difference in smiling is notable.

3.3 Articulatory Results

Lips were significantly more spread in smiled than in neutral speech (t = 12.64, 4.75, 18.08 for *caw*, *coo*, *key*). LCA was smaller for all three vowels (t = -9.79, -8.13, -7.22 for *caw*, *key*, *coo*), meaning lip corners were raised. This indicates that participants smiled, with lips spread and lip corners turned up. LPL was significantly less for smiled speech for *caw* (t = -2.88) and *key* (t = -5.21), but not for *coo*. Similarly, UPA and LPA were significantly smaller and larger respectively for *caw* and *key* (t = -2.98, 2.28 for *caw*, t = -5.98, t = 3.49), but not significant for *coo*. The lack of protrusion effect for /u/ suggests that lip spreading and rounding do not conflict. These results diverge from [4], which showed a suppression in the rounding of /u/ in German, possibly due to differences between German and English /u/. UPA and LPA measurements are novel to this study; they correspond to the standard LPL, suggesting these angles are informative and potentially useful.

4 Conclusion

The effects of smiling observed here on articulation were lip spreading and lip corner raising, and raised f0 for all three vowels, but raised F1 and F2 only for /a/. Larynx height was not significantly affected, contrary to findings in [13].

Acknowledgments

We thank: Molly Babel, Kyra Borland-Walker, Amon Ge, Thomas J. Heins, Michael McAuliffe, and Nick Romero. This project has been made possible by a NSERC grant to the final author. Errors are our own.

References

[1] Tartter, V. C. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech. Perception & Psychophysics, 27 (1), 2427. doi:10.3758/Bf03199901

[2] Quené, H., Semin, G., & Foroni, F. (2012). Audible smiles and frowns affect speech comprehension. *Speech Communication*, 54 (7), 917922. doi:10.1016/j.specom.2012.03.004

[3] Torre, I. (2014). Production and perception of smiling voice. Proceedings of the first Postgraduate and Academic Researchers in Linguistics at York (PARLAY 2013) conference.

[4] Fagel, S. (2010). Effects of smiling on articulation: Lips, larynx, and acoustics. In A. Esposito, N. Campbell, C. Vogel, A. Hussain, & A. Nijholt (Eds.), Development of Multimodal Interfaces: Active Listening and Synchrony Lecture Notes in Computer Science Volume 5967 (294303). Berlin: Springer.

[5] Shor, R. E. (1978). The production and judgment of smile magnitude. *The Journal of General Psychology*, 98(1), 79-96. doi: 10.1080/00221309.1978.9920859.

[6] Moisik, S. R., Lin, H., & Esling, J. H. (2014). A study of laryngeal gestures in Mandarin citation tones using simultaneous laryngoscopy and laryngeal ultrasound (SLLUS). *Journal of the International Phonetic Association*, 44(01), 21-58.

[7] Psychology Software Tools, Inc. (E-Prime 2.0). (2012). Retrieved from http://pstnet.com

[8] Egger, H. L., Pine, D. S., Nelson, E., Leibenluft, E., Ernst, M., Towbin, K. E., & Angold, A. (2011). The NIMH Child Emotional Faces Picture Set (NIMH-ChEFS): a new set of children's facial emotion stimuli. *Int. J. Methods Psychiatr. Res*., 20: 145–156. doi: 10.1002/mpr.343

[9] Boersma, P. & Weenink, D. (2015). Praat: doing phonetics by computer. Version 5.4.06, Available from http://praat.org/

[10] Rasband, W. S. (2014) ImageJ. Available from http://imagej.nih.gov/ij/

[11] Bates D., Maechler M., Bolker B., and Walker S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7, http://CRAN.R-project.org/package=lme4.

[12] R Development Core Team. R: A language and environment for statistical computing. R Foundation for Satistical Computing, Vienna, Austria. *http://www.R-project.org*. 2008.

[13] Lasarcyk, E. & Trouvain, J. (2008). Spread lips + raised larynx + higher f0 = smiled speech? - An articulatory synthesis approach. *Proceedings of the 8th International Seminar on Speech Production*, Strasbourg, France, 8-12 December 2008, pp. 345- 348.

[14] Honda, K., Hirai, H., Masaki, S., & Shimada, Y. (1999). Role of vertical larynx movement and cervical lordosis in FO control. *Language and Speech*, 42(4), 401.

ULTRASOUND-ENHANCED MULTIMODAL APPROACHES TO PRONUNCIATION TEACHING AND LEARNING

Jennifer Abel, Blake Allen, Strang Burton, Misuzu Kazama, Bosung Kim, Masaki Noguchi, Asami Tsuda, Noriko Yamane, and Bryan Gick
University of British Columbia.

1 Introduction

Pronunciation is an integral part of communication, as it directly affects speakers' communicative competence and performance, and ultimately their self-confidence and social interaction. Second language (L2) pronunciation is one of the most challenging skills to master for adult learners. Explicit pronunciation instruction from language instructors is often unavailable due to limited class time; even when time is available, instructors often lack knowledge of effective pronunciation teaching and learning methods. Imitating native speakers' utterances can be done outside of the classroom, but the absence of feedback makes it difficult for learners to improve their skills.

In an effort to improve pronunciation instruction, the Department of Linguistics and the Japanese language program in the Department of Asian Studies at the University of British Columbia began a collaboration in 2014 to develop new multimodal approaches to pronunciation teaching and learning. Japanese is the largest language program at UBC, with more than 1,500 students enrolled every year, and is known to be the most diverse in terms of learners' language backgrounds. The project has developed online resources to allow learners of Japanese to improve their pronunciation, as well as to allow Linguistics students to better understand sound production. The project's key technological innovation is the use of ultrasound overlay videos, which combine mid-sagittal ultrasound images of tongue movement in speech with external profile views of a speaker's head to allow learners to visualize speech production [1]. A strong focus to this point has been on the pronunciation of Japanese sounds, with instruction incorporating explicit awareness of tongue movements and insights from articulatory phonology, including the interference of L1 phonology on L2 learning [3].

2 Methods

Ultrasound of native speakers of Japanese and of English was recorded using an Aloka ProSound SSD-5000 system, and the exterior video was recorded using a JVC camcorder (GZ-E300AU). Both recordings were made at 30 frames per second. The exterior video showed the left profile of the speaker's head. A clapper was used to generate an audio alignment point.

The ultrasound overlay videos were created from raw footage using a four-step process. First, the ultrasound and exterior video were trimmed using Adobe Premiere to ensure alignment. Next, all elements of the ultrasound image aside from the tongue were manually erased using Adobe After Effects. The brightness of the tongue was increased, and the colour was changed from white to a shade of pink (colour #DE8887 in Adobe After Effects) to more closely resemble the human tongue. Then, the erased ultrasound image was overlaid on the exterior face video using Adobe After Effects. Scaling of the two sources was achieved by ensuring that the shadow of the probe in the ultrasound image is the same width as the top of the probe in the exterior video. The results of this process are exemplified in Figure 1.

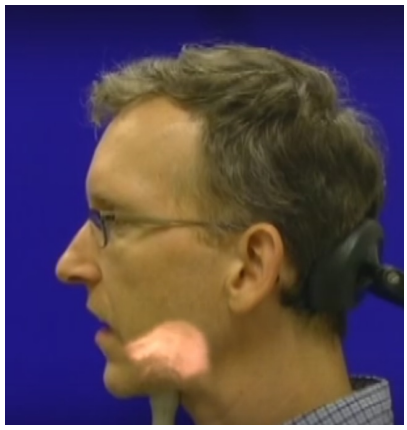


Figure 1. Ultrasound overlay video frame of [χ]

3 Results

The videos are available to the public through the eNunciate website (http://enunciate.arts.ubc.ca/), and are licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. The videos are categorized into 'Linguistics' and 'Japanese' content, although all pages are open to all students and instructors.

3.1 Linguistics content

The Linguistics pages feature ultrasound overlay videos of canonical examples of the sounds of the world's languages produced basic contexts: in the [_a] and [a_a] positions for consonants, and in isolation or in [C_C] contexts for vowels. Freeze frames are inserted in these videos to capture key moments in the articulation (e.g., the stop closure in a stop articulation), and beginning and end titles are inserted. These videos can be accessed through interactive IPA consonant and vowel charts. In addition to the ultrasound overlay videos, videos introducing the use of ultrasound in linguistics and the basics of vowel and consonant articulation are available. To date, these videos have been used mainly in introductory Linguistics courses as in-class

demonstration materials or as optional out-of-class resources.

3.2 Japanese content

The Japanese pages include instructional and exercise videos for Japanese pronunciation teaching and learning. These videos incorporate narration, cartoons, and animations in addition to ultrasound overlay elements, and are augmented with quizzes to allow students to reinforce what they have learned using the videos. The videos are grouped into three categories: introductory, which includes introductions to Japanese sounds and to basic phonetic concepts; 'challenging sounds', which features videos focusing on problems that L2 learners from different language backgrounds may encounter; and intonation.

These videos have been used in both introductory and upper-year Japanese classes as pronunciation teaching and assessment tools. In the introductory classes, students are expected to watch particular 'challenging sounds' or intonation videos as preparation for in-class quizzes. In the upper-year class in which the materials were used (JAPN 301 at UBC), the students were encouraged to use the videos to improve their fluency when reading aloud. Instructors and students have anecdotally reported satisfaction with the resources; full-scale effectiveness and satisfaction data collection is in progress.

4 Discussion: developments in progress

4.1 Biofeedback tools

As part of our plan to use biofeedback to facilitate L2 pronunciation learning, we are developing two tools that will allow learners to visualize their speech in real-time. The first tool is an interactive tongue visualizer, which will automate creation of the type of ultrasound overlay videos described in section 2 based on ultrasound and video feeds of a speaker producing sounds in real time. Development of this tool is still in the early stages. The visualizer will be implemented at a physical location ("Pronunciation Station") at UBC, and will be equipped with a CHISON ECO 1 portable ultrasound with a 6.0MHz D6C12L Transvaginal probe.

The second tool is a web app that allows visualization of speech pitch contours based on open-source audio manipulation software [4]. Students will be able to visually compare the pitch of their own pronunciations of L2 speech with pitch contours produced by native speakers, allowing them to improve their L2 intonation, pitch accent, and/or tone patterns. Students will also be able to save their pronunciations in order to track how their pronunciation improves over time. Currently, we have implemented a basic web interface for the app, along with an implementation of the standard autocorrelation algorithm for inferring pitch from sound signals. Our immediate goals are 1) to refine the pitch detection algorithm to allow it to better compensate for the rapid pitch variations in natural speech, and 2) to build a more user-friendly interface.

4.2 Ultrasound training

To overcome the lack of a standardized procedure for the teaching of L2 pronunciation with ultrasound imaging, we are developing guidelines based on the procedures previously used in the settings of L2 learning [1] and speech language pathology [2]. The guidelines target three consecutive days of teaching to allow teachers to use the Pronunciation Station: (1) initial evaluation of students' pronunciation, (2) training with ultrasound images as biovisual feedback, and (3) post-training evaluation of students' pronunciation. As a case study, we will implement the protocols in teaching Japanese pronunciation to native speakers of Korean, particularly focusing on the acquisition of the contrast between alveolar and alveo-palatal sibilants (e.g. [za] vs. [ʑa]), which is known to be especially difficult for Korean speakers. The effectiveness of the protocols will be discussed in light of the data collected at the Pronunciation Station.

5 Conclusion

As more research continues to support the positive effects of multimodal feedback via technologies such as ultrasound on L2 pronunciation training [e.g. 5], the importance of bringing these technologies into the classroom has become clear. We believe that the tools and techniques developed as part of the eNunciate project are helping to achieve that goal at UBC, and hope to extend them to other institutions and situations in the future.

Acknowledgments

This project is supported by a Flexible Learning Large Project Grant from the Teaching and Learning Enhancement Fund at the University of British Columbia. Many thanks to Joe D'Aquisto, Jonathan de Vries, Amir Entezaralmahdi, Lewis Haas, Tsuyoshi Hamanaka, Hisako Hayashi, Ross King, Andrea Lau, Yoshitaka Matsubara, Douglas Pulleyblank, Nicholas Romero, Hotze Rullmann, Murray Schellenberg, Joyce Tull, Martina Wiltschko, Jenny Wong, and Kazuhiro Yonemoto.

References

[1] Gick, B., et al. (2008). Ultrasound imaging applications in second language acquisition. In J. G. Hansen Edwards and M. L. Zampini (eds.), *Phonology and Second Language Acquisition* (pp. 309-322). Amsterdam: John Benjamins.

[2] Bernhardt, B., et al. (2005). Ultrasound in speech therapy with adolescents and adults. *Clinical Linguistics & Phonetics* 19.6-7: 605-617.

[3] Toda, T. (2003). *Second Language Speech Perception and Production: Acquisition of Phonological Contrasts in Japanese*. Lanham, Maryland: University Press of America.

[4] Cwilso/PitchDetect. (n.d.). Retrieved August 11, 2015 from https://github.com/cwilso/PitchDetect.

[5] Pillot-Loiseau, C., et al. (2015). French /y/-/u/ contrast in Japanese learners with/without ultrasound feedback: vowels, non-words and words. Paper presented at ICPhS 2015. Retrieved August 12, 2015 from http://www.icphs2015.info/pdfs/Papers/ICPHS0485.pdf.

Abstracts for Presentations without Proceedings Paper
Résumés des communications sans article

A Tool For Testing Phonetic Hypotheses In The Lab: The Prosodylab Experimenter

Michael Wagner, Erin Olson, Meghan Clayards

We describe a set of scripts that facilitate the testing of phonetic hypotheses in the lab, and exemplify them with a study of the realization of the voicing contrast in word-final obstruents in English. An important cue is the length of the preceding vowel, interpreted by Hayes (2007) and others as shortening preceding word-final voiceless obstruents. Two production experiments look at the phonetic realization of vowels preceding fricatives (6 triplets of the form bus/buzz/bun, 27 speakers) and stops (30 triplets of the form bit/bid/bin, 20 speakers), using scripts that randomize trials and maximize the distance between trials from the same item and condition, and order them such that both latin square and fully crossed analyses are possible. The segment durations were determined with forced-alignment with varying models (using scripts by Gorman et al. 2011), as well as manual RA annotations. Across different ways of annotating and analyzing the data we find similar results, thus validating more time-efficient automatic methods. The data suggest that (i) there is both shortening before voiceless and lengthening before voiced consonants; (ii) vowels preceding sonorants are of intermediate length; (iii) vowels preceding sonorants are more similar to vowels preceding voiceless fricatives than voiced ones, suggesting a primary effect lengthening before voiced fricatives (e.g., bus/bun/buz); but they are less similar to vowels preceding voiceless stops than voiced ones, suggesting a primary effect of shortening before voiceless stops (e.g., bit/bin/bid); (iv) there are substantial by-item differences depending on the individual vowel and consonant (Umeda 1975). Overall, the observed pattern speaks against reducing the effect to either lengthening or shortening, and suggest that the contrastiveness of an acoustic dimension matters for the phonetic realization of all sides of the contrast, as predicted by dispersion theories (Liljencrants & Lindblom 1972; Flemming 2004).

Using Automatic Alignment On Child Speech: Directions For Improvement

Meghan Clayards, Thea Knowles, Morgan Sonderegger, Michael Wagner, Kristine Onishi, Aparna Nadig

Phonetic analysis is labor intensive, limiting the amount of data that can be considered. Recently, automated techniques (e.g. forced alignment based on Automatic Speech Recognition - ASR) have emerged allowing for much larger-scale analysis. For adult speech, forced alignment can be accurate even when the phonetic transcription is automatically generated allowing for large-scale phonetic studies (e.g., Yuan & Liberman, 2009), however, such studies remain difficult for children's speech, where ASR methods perform much more poorly (Benzeghiba et al., 2007). We used a trainable forced aligner that performs well on adult speech (Gorman et al., 2011) to examine the effect of four factors on alignment accuracy with child speech: Factor 1: Dataset. Two datasets available on CHILDES varying in the number of children and type of speech (spontaneous versus picture naming) (McWhinney, 2000): Factor 2: Transcription type. Phonetic transcriptions were generated by hand (actual pronunciation) or automatically generated (canonical pronunciation from CMU dictionary (Weide, 1998). Factor 3: Training data. The aligner was trained using: Adult lab speech, the Specific dataset to-be-aligned, All child data, or All child data and adult lab speech. Factor 4: Type of segment. We examined whether the alignment worked better for certain segment types (voiceless stops, voiceless sibilants, or vowels). Automatically generated alignments were compared to hand segmentations. Results were generally better (1) with picture naming (2) when a rough phonetic transcription was available (3) when the training data included child data, versus adult data alone; (4) for stops, followed by sibilants and vowels. These four factors increase the utility of analyzing children's speech production using forced alignment potentially allowing researchers to ask questions that otherwise would require weeks or months of hand-segmentation.

Locus Equations As Proxies For Co-Articulation Lend Support To The Degree Of Articulatory Constraints Model

Sara Perillo, Hye-Young Bang, Meghan Clayards

The degree of articulatory constraints (DAC) model (Recasens, Pallarès, & Fontdevila, 1997) proposes that consonants involving the movement of the tongue dorsum are more resistant to coarticulation than those consonants that have a more fronted articulation. The present study aims to assess this claim using locus equation (LE) slopes as indicators of coarticulation. Participants were asked to produce V1(t).C1V2 sequences as part of two-word phrases in a scripted dialogue, where C1 is one of /p, t, s, ?/. LE were derived by measuring F2 at V2 onset and midpoint. Since LE slopes approaching 1 indicate high levels of coarticulation, it was hypothesized that those segments with the lowest DAC would have the steepest slopes, and vice versa. /p/ was predicted to have the lowest DAC and steepest slope, followed by /t/, /s/, and /?/. Results were highly consistent with these hypotheses, lending support to the DAC model. A secondary hypothesis assessed the effect of emphatically stressing C1 on the LE. Participants partook in a dialogue involving a "mishearing", which prompted them to repeat the original V1(t).C1V2 sequence. We expected participants to emphasize the misheard segment, which was either the target C1 (Prominent condition) or the preceding V1(t) (Control condition). It was predicted that prominence would reduce coarticulation, resulting in a downward shift in LE slopes relative to the Control condition. Our findings indicate that only the LE slopes of sibilants /s/ and /?/ were reduced under prominence as hypothesized, and that these reductions were statistically comparable. Results are thus consistent with the DAC model, since /s/ and /?/'s being less likely to co-articulate than /p/ and /t/ in the Prominent condition may due to their relatively large DAC values.

UNDERWATER ACOUSTICS - ACOUSTIQUE SOUS-MARINE

MANEUVERING VEHICLE LOCALIZATION WITH AN ACOUSTIC LONG BASELINE SYSTEM

Zhao Li [†1,2,3], Stan E. Dosso [3] and Dajun Sun [1,2]
[1]Science and Technology on Underwater Acoustic Laboratory, Harbin Engineering University, China;
[2]College of Underwater Acoustic Engineering, Harbin Engineering University, China;
[3]School of Earth and Ocean Sciences, University of Victoria, Canada.
[†]lizhao0517@gmail.com

1 Introduction

The localization error of the long baseline (LBL) acoustic localization system is sensitive to vehicle motion between the time the vehicle sends out an interrogation signal and the times of receptions for the acoustic replies from the various transponders. The static-vehicle model (SM) for localization does not take vehicle motion into consideration and calculate the one-way acoustic travel time from the vehicle to a transponder by halving the observed two-way travel-time from vehicle to the transponder and back. This is the reason why esult larger localization errors for a maneuvering vehicle are obtained in an actual LBL trial than the posterior uncertainty estimated by the posterior covariance matrix of the static localization model.

To address this problem, in this paper, we tried to use time corrections as additional unknown parameters estimated via Bayesian inversion to reduce the localization errors related to vehicle motion by the SM, rather than using Kalman filters. In this extended SM, referred as motion-compensated model (MM), vehicle interrogation coordinates, transponder coordinates, and time corrections are all estimated using a strongly underdetermined linearized Bayesian inversion, given reasonable prior information. An LBL field trial indicated that this MM approach is effective in compensating the influence of vehicle motion .

2 Method

In the MM inversion algorithm, the dependence of vehicle location (x,y,z), transponder location (X,Y,Z) and the one-way travel-time t^{obs} calculated by simply half the observed two-way travel-time can be written as

$$t\left(x_i, y_i, z_i, X_j, Y_j, Z_j, c\right) + \Delta t_{ij} = t_{ij}^{\text{obs}}, \qquad (1)$$

where Δt is the travel-time correction, which compensates the signal travel-time difference of interrogation and reception. The sequence number of the localization cycle is marked as i and j represents different transponders. The synthetic halved two-way travel time t is calculated with interrogation location, transponder location and SSP c via ray tracing method. With local linearization and prior information $\hat{\mathbf{m}}$, the Bayesian inversion solution of Eq.(1) is expressed as

$$\mathbf{m} = \hat{\mathbf{m}} + \left(\mathbf{J}^{\mathrm{T}}\mathbf{C}_D^{-1}\mathbf{J} + \mathbf{C}_M^{-1}\right)^{-1}\mathbf{J}^{\mathrm{T}}\mathbf{C}_D^{-1}\left(\mathbf{d}\text{-}\mathbf{J}\hat{\mathbf{m}}\right), \qquad (2)$$

where $\mathbf{m} = \left[x_i, y_i, z_i, X_j, Y_j, Z_j, \Delta t_{ij}\right]^{\mathrm{T}}$ is the unknown model parameters vector in this problem. $\mathbf{J}$ represents the Jacobian matrix consisted by the partial derivatives of t with respect to $\mathbf{m}$. $\mathbf{C}_D$ is the covariance matrix of the data, $\mathbf{C}_M$ is the covariance matrix of the prior model. As this approach is based on linearization, it is repeated iteratively until convergence (when model changes between iterations become insignificant).

Larger quantity of unknown parameter than data leads to a strongly under-determined inversion, which is regularized by including prior estimates (with Gaussian uncertainties) for all parameters. The key to this algorithm is setting the prior estimates for the travel-time corrections. This is carried out by first estimating the vehicle locations at the time instants the interrogation signals are sent by applying cubic-spline interpolation of the results of an initial localization inversion assuming a static vehicle, and then determining the corresponding travel-times by ray tracing.

3 Results

3.1 Lake Trial Procedure

A LBL field trial was carried out on Nov. 3rd, 2014, in Songhua Lake, Jilin, China. The lake trial was aimed to test a LBL acoustic localization system designed for AUV

Figure 1: Components of the LBL system tested in the field trial, including transducer, transponder, computer, signal processing box and power box (counter-clockwise from top left)

Localization. The LBL system is composed of four transponders (T1 to T4), a transceiver transducer, a processing box and a power box, shown in Fig. 1. The acoustic transducer of the transceiver was firmly installed on the starboard side of the boat by a triangular-prism-shaped steel frame. On top of the fame, the antenna of a mobile GPS station was fastened and was set in the real time kinematic mode. A gyroscope was installed between the frame and the transducer to measure yaw, pitch and roll of the boat. With the data from GPS and gyroscope, the geodetic coordinates of the transducer can be calculated with better than 4 cm precision. These geodetic coordinates are logged as references to evaluate the localization error of acoustic localization.

Transponders were deployed on the lakebed at about 35 m depth and are localized by the array element acoustic survey (AEL) with about 4 cm precision. When the AEL survey was finished, motion vehicle survey (MVL) were conducted in which the boat was treated as a substitute for an AUV, and moved back and forth within an about 300 m×300 m area centered on the average of transponder coordinates. The geometry of the transponder array and the survey tracks are shown in Fig. 2(a).

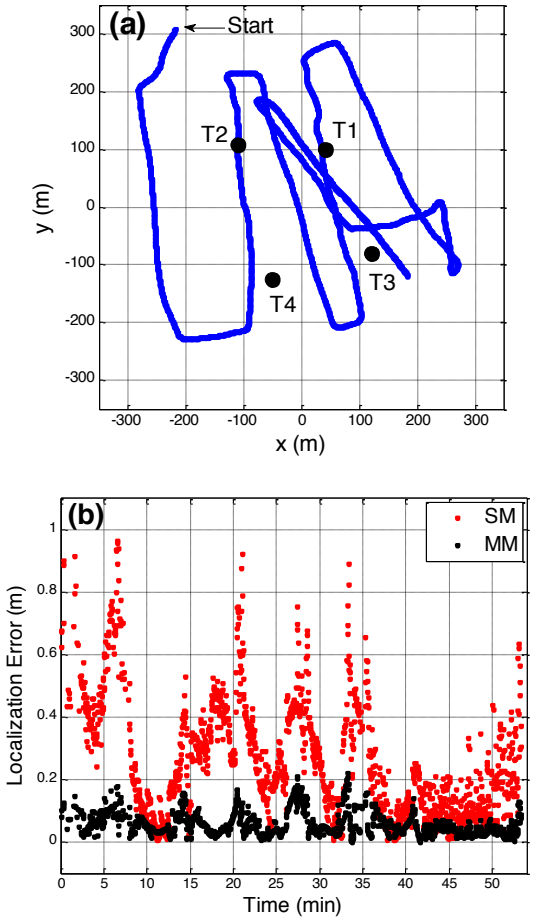


Figure 2: (a) Motion vehicle survey track and transponder locations. (b) Localization error comparison between SM inversion and MM inversion.

3.2 Data Processing Results

The maneuvering surface boat is localized with two models: the static model (SM) that estimates vehicle and transponder locations ignoring vehicle motion; the motion-compensated model (MM) that estimates these locations and travel-time corrections. In this paper, localization errors in the trial are regarded as the difference between inversion results and the measured geodetic coordinates of the transducer.

The localization error curves versus time are shown in Fig. 2(b). The error of SM method varies invensively. The largest localization error is almost 1 m. Generally, large localization errors correspond to places where the boat is outside the transponder array. When the vehicle was inside the array, the maximum localization error is about 50 cm and the averaged error is 19±14 cm, which is significant larger than the expected 4 cm uncertainty estimated from the posterior matrix. In the motion-compensated inversion, the average localization error is 3.4±1.4 cm when the vehicle was inside the transponder array. Even when the vehicle moved outside the array, the error is still generally less than 20 cm. Therefore, the data processing results indicate that the MM method is helpful in compensate the localization error generated by vehicle motion.

4 Conclusion

This paper describes the compensation for vehicle motion during the interrogation-reception time interval between the vehicle and transponders in a long baseline (LBL) system lake trial without using Kalman filter. This compensation method is base on Bayesian inversion algorithm which includes travel-time corrections as additional unknown parameters with prior determined by interpolating the vehicle location at interrogation time instants using static-vehicle localization model (SM) results. Field trials results showed that the motion-compensated inversion (MM) reduced averaged acoustic localization errors of a surface vehicle to 3.4±1.4 cm, much accurate than the 19±14 cm error for the static inversion.

Acknowledgments

The authors thank Dr. Jucheng Zhang and his colleagues in Harbin Engineering University for designing the LBL localization system and the fieldwork.

References

[1] D. Thomson and S. E. Dosso, “AUV localization in an underwater acoustic positioning system,” *Canadian Acoustics*, vol. 40, no. 3, pp. 84–85, 2012.

[2] S. E. Dosso and G. R. Ebbeson, “Array element localization accuracy and survey design”, *Canadian Acoustic.*, vol. 34, no. 4, pp. 1–11, Mar. 2006.

[3] R. C. Aster, B. Borchers, and C. H. Thurber, “Bayesian Methods,” in Parameter estimation and inverse problems, 2nd. Waltham: Academic Press, 2013, pp. 253-268.

UNDERWATER NAVIGATION METHOD BASED ON SIDE-SCAN SONAR IMAGES

Ziqi Song [1], Adam Zielinski [2], and Hongyu Bian*[1]

[1] Acoustic Science and Technology Laboratory, College of Underwater Acoustic Engineering, Harbin Engineering University, Harbin, China

[2] Department of Electrical and Computer Engineering, University of Victoria, Victoria, Canada

1 Introduction

Underwater navigation is a critical technique for underwater platforms in both civil and military missions. In order to make autonomous underwater vehicle (AUV) capable of time-extensive submerged operations, navigation methods using geophysical characteristics have been proposed recently as a supplementary means to the inertial navigation system (INS) [1]. For an AUV equipped with side-scan sonars, image processing techniques can be applied for its navigation system.

Side-scan sonar images can provide a clear view of the sea floor in 2D with a relative higher resolution than multi-beam bathymetry sonar [2]. This equipment is often be employed for buried-object detection [3], seafloor sediments classification [4] and ocean environment monitoring [5] etc. Similar to the concept of terrain aided navigation [6], this paper investigated the possibility of using side-scan sonar and proper image analysis methods in the matching process of underwater positioning.

The remainder of this article describes the bispectrum of images and presents the simulation results about its robustness of rotation in underwater application.

2 Bispectrum of side-scan sonar images

2.1 High-order Cumulant

The theory of high-order statistics is developed from the theory of second order statistics (e.g. correlation function and power spectrum). It can reflect the phase information of signal which is absence in the theory of second order statistics. Benefit from this, high-order statistics can be employed to deal with non-minimum phase systems. Many applications based on this theory can now be found in the research of signal processing for radar, sonar, communication and fault diagnosis.

Let x be a random variable and $f(x)$ is its probability density function (PDF). The characteristic function of x, $\Phi(\omega)$, is define as the integration presented in (1).

$$\Phi(\omega) = \int_{-\infty}^{\infty} f(\mathrm{x})\, \mathrm{e}^{j\omega x}\, dx \tag{1}$$

In other words, the PDF of x equals to the Fourier transform (FT) of $f(x)$. $\Phi(\omega)$ is also called the first characteristic function of x.

Then, the second characteristic function of x, $\Psi(\omega)$, is defined as (2).

$$\Psi(\omega) = \ln[\Phi(\omega)] \tag{2}$$

Value of the k order derivative of $\Phi(\omega)$ at the origin, m_k, is the k order moment of x as shown in (3).

$$m_k = \Phi^{(k)}(\omega)\,|_{\omega=0} = E[x^k] = \int_{-\infty}^{\infty} x^k f(x)dx \tag{3}$$

Value of the k order derivative of $\Psi(\omega)$ at the origin, c_k, is the k order cumulant of x as shown in (4).

$$c_k = \Psi^{(k)}(\omega)\,|_{\omega=0} \tag{4}$$

Further, assume x obeys a Gaussian distribution $N(0,\sigma^2)$. Then the PDF of x will be like (5).

$$f(\mathrm{x}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \tag{5}$$

The first and second characteristic functions of x will then be (6) and (7) respectively.

$$\Phi(\omega) = \mathrm{e}^{-\frac{\sigma^2\omega^2}{2}} \tag{6}$$

$$\Psi(\omega) = \ln[\Phi(\omega)] = -\frac{\sigma^2\omega^2}{2} \tag{7}$$

Now every cumulant of x is available that is shown in (8).

$$\begin{cases} c_1 = 0 \\ c_2 = \sigma^2 \\ c_k = 0 \quad (k \geq 3) \end{cases} \tag{8}$$

As a result, high-order cumulant is capable of restraining the effect of Gaussian noise.

2.2 Bispectrum

As shown in 2.1, to any Gaussian process, its high-order cumulant is always zero while its high-order moment is not. This explains why high-order cumulant and its spectrum are more often used in scientific researches. Theoretically, high-order cumulant is able to restrain Gaussian colored noise completely. Spectrum of high-order cumulant is named high-order spectrum (HOS). Among them, three order and four order HOS, called bispectrum and trispectrum separately, are used most frequently. To reduce the computation cost, this paper selected the bispectrum of side-scan sonar images as the characteristic used in the following matching process.

Assume k order cumulant, $c_{kx}(\tau_1, \tau_2, \cdots, \tau_{k-1})$, is absolute summable ($\tau_{k-1}$ is the time delay between x_{k-1} and x_k), then the corresponding k order spectrum is defined as its k-1 dimensional Fourier transform presented in (9).

$$S_{kx}(f_1, f_2, \cdots, f_{k-1}) =$$
$$\sum_{\tau_1=-\infty}^{\infty} \cdots \sum_{\tau_{k-1}=-\infty}^{\infty} c_{kx}(\tau_1, \tau_2, \cdots, \tau_{k-1}) \quad (9)$$
$$\exp[-j2\pi(f_1\tau_1 + f_2\tau_2 + \cdots + f_{k-1}\tau_{k-1})]$$

Specifically, when $k = 3$, the definition of bispectrum is available in (10).

$$B_x(f_1, f_2) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} c_{3x}(\tau_1, \tau_2) e^{-j2\pi(f_1\tau_1 + f_2\tau_2)} \quad (10)$$

3 Estimation steps

The method that using bispectrum of side-scan sonar images for underwater navigation can be divided into the following steps:

Step1: A side-scan sonar image of a sea bottom is read as the reference map, such as Figure 1a. Size of the image should be determined by the requirement of resolution and confidence level of the certain task.

Step2: Get an real-time side-scan sonar image of the bottom, such as the region in the yellow dotted lines in Figure 1a. Calculate its bispectrum by (10).

Step3: Using the bispectrum from step 2 as an template (window) to scan the reference map for a match. Calculate the similarity between the bispectrums of the real-time image and the local area at each scanning point.

Step4: After completing the map scanning, the point that has the biggest similarity (such as the central point of the red region in Figure 1a) will be considered as the position estimation of the underwater platform.

4 Simulation results

Simulations were done in Matlab 2012a platform to investigate the performance of the proposed method. Two side-scan sonar images taken from different sorts of sea bottom were used as reference maps. It was assumed that necessary pre-processing to these images were completed as enhancement and the related methods will not be described in this paper.

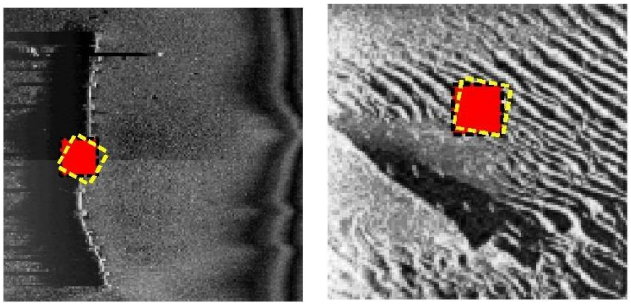
A local area of the reference map was picked out to synthesize the real-time image collected by a side-scan sonar installed on an autonomous underwater vehicle (AUV). Position of the AUV was assumed at the central point of this real-time image. Gaussian noise (signal to noise ratio is 10dB) was added to the real-time images in order to simulate the interference signal slipped into the sonar. The searching step is 2 pixels/move. Mean square deviation (MSD) was chosen as the tool to evaluate the similarity between the bispectrums of the real-time image and the map.

Figure 1 shows the matching results using different rotation of real-time images. Although the real-time images have direction differences against the reference map, the estimated regions (red) keep locating close to the true position (yellow). Horizontal and vertical errors in pixels are (2, 1) and (1, 1) in Figure 1(a) and Figure 1(b) respectively. This result shows good robustness of rotation.

5 Conclusion

Bispectrum of side-scan sonar images was investigated in this paper for its performance in underwater navigation. Bispectrum of synthesized side-scan sonar image was calculated as a template in the matching processing. Simulation results show that the proposed method has good robustness of image rotation.

To find a balance of the calculation time and the estimation reliability, more tests are needed before using this method in actual applications. Side-scan sonar images from sea trial will be processed with this method and a general regulation of template selection can be expected.



(a) (b)

Figure 1: Matching results with rotation in Gaussian noise background. (a) Rotation = 30 °. (b) Rotation = 10 °.

Acknowledgments

This work was funded by the Chinese Scholarship Council (Ziqi Song, grant number: 201406680029) and the Chinese Nature Science Foundation (Hongyu Bian, grant number: 41376102).

References

[1] Department of the Navy. The Navy Unmanned Undersea Vehicle (UUV) Master Plan. pp.1-11, 2000.

[2] Zhao, J., Wang A. and Guo J. Study on Fusion Method of the Block Image of MBS and SSS. *Geometrics and Information Science of Wuhan University*, 38(3), pp. 287-290, 2013.

[3] Chen X., Zhou L. Review of Current Status of Buried-object Detection Techniques. *Technical Acoustics*, 31(1), pp. 30-35, 2012.

[4] Shi D., Li Q., Fan X. and Huo G. Seafloor Sediments Classification of Side-scan Sonar Imagery in Fast Discrete Curvelet Transform Domain. *Journal of Applied Sciences – Electronics and Information Engineering*, 27(5), pp. 498-501, 2009.

[5] Yu J., Cheng E. Applications of Side-scan Sonar in Ocean Environment Monitoring. *Hydrographic Surveying and Charting*, 24(2), pp. 63-66, 2004.

[6] Se Hyun Yun, Wonhee Lee, Chan Gook Park. Covariance Calculation for Batch Processing Terrain Referenced Navigation. *Position, Location and Navigation Symposium - PLANS*, 2014 IEEE/ION, pp. 701-706, 2014.

A TOOLSET FOR MODELLING ANTHROPOGENIC UNDERWATER NOISE

Terry J. Deveau[*1]
[1] JASCO Applied Sciences, 202 – 32 Troop Ave., Dartmouth, NS, B3B 1Z1, Canada.

1 Introduction

Anthropogenic underwater noise is an increasing environmental concern. Accurate predictions of sound levels from such sources are required to estimate the impact on marine life that is exposed to it. JASCO has been developing software modelling tools for underwater noise exposure estimation for more than 30 years. Elements of this modelling process include estimation of the source levels, source spectra, and source radiation patterns; environmental characteristics of the underwater sound medium and the geoacoustics of the seabed; calculating the acoustic propagation loss; estimating the received levels, both in terms of rms Sound Pressure Level (SPL) and Sound Exposure Level (SEL); evaluating species-specific impact weighting; and compiling results into comprehensible summaries. These tools have been developed for accuracy in prediction and efficiency in computation, and have been used in projects for a wide range of international clients, both commercial and governmental. This paper presents an overview of the software toolset used in anthropogenic underwater noise and exposure modelling work being done by JASCO.

2 Discussion

2.1 Source Characteristics

The first step in anthropogenic underwater noise modelling is to characterize the acoustic source. Many studies involve multiple sources, and each source is treated individually first, with the combined received levels from all sources calculated only at the end. Each source is broadly characterized as either impulsive (e.g., pile-driving hammer impact, explosive detonation, or air-gun bubble pulse) or continuous (e.g., ships propeller cavitation noise, engine noise transmitted through the vessel hull, or rapid active sonar chirps). The long-distance acoustic projection of most underwater sources can be adequately characterized in terms of their source depth, third-octave frequency band levels, and the azimuthal beam pattern.

In particular, for modelling airgun array source levels and directivity, the Airgun Array Source Model (AASM) is employed. This model has been developed by JASCO's Alex MacGillivray [1,2]; it is based on the physics of the oscillation and radiation of airgun bubbles [3]. The model solves the set of parallel differential equations that govern bubble oscillations. AASM also accounts for non-linear pressure interactions between airguns, port throttling, bubble damping, and Generated Injection (GI) airgun behavior [4-6]. AASM includes four empirical parameters that are tuned so that the model output matches observed airgun behavior. The model parameters were fit to a large library of empirical airgun data using a "simulated annealing" global optimization algorithm [7].

[*] terry.deveau@jasco.com

2.2 Environmental Characteristics

Anthropogenic underwater acoustics emissions from the source propagate to distant points in the ocean as pressure waves carried by the sea water medium. This propagation typically involves some interaction with the sea surface and the sea bed, which can be visualized as some combination of reflection, transmission, absorption, and scattering. Acoustic energy transmitted into the sea bed can be carried by the sedimentary media and re-emerge into the water column at distant points. While being carried in the sediments, acoustic energy can be in the form of pressure waves, as in the sea water media, but also as transverse (shear) waves. A number of different underwater acoustic propagation models are available to account for these complex interactions, but before they are employed, the relevant environmental characteristics must be compiled. Some models have particular requirements, but in general the environmental parameters listed in Table 1 are always necessary.

Environmental parameter	Function of:	Coverage
Water depth	Horizontal position in meters	Entire area
Compressional-wave sound speed in water	Location, depth, time of year	One profile for the area
Sediment compressional-wave sound speed	Location, layer	A value for each layer
Sediment compressional-wave attenuation coefficient	Location, layer	″
Sediment density	Location, layer	″
Sediment layer thickness	Location, layer	″
Sediment shear wave sound speed	Location	One value for the area
Sediment shear wave attenuation coefficient	Location	″

Table 1: Minimum set of environmental parameters for underwater acoustic modelling.

2.3 Propagation Modelling

Underwater acoustic propagation models are generally divided into wave-based and ray-based categories. Wave-based models are better suited to modelling low-frequency underwater sound (typically below 2 kHz); using them at higher frequencies is possible but requires smaller step sizes and longer run times to achieve the desired accuracy and stability. Both types of model can be used to model the coherent pressure wave reaching the receiver as well as the average per pulse (or per time unit) received acoustic energy. The latter approach allows for some simplifying assumptions in the modelling technique and generally results in much faster modelling run times. Where estimates of both rms SPL and SEL are required at each receiver location, the

SEL can be calculated using the faster energy propagation modelling technique, and the rms SPL can be estimated from the calculated SEL by means of an empirical formula, or by running the full-wave simulation on a smaller number of receiver locations to construct an empirical conversion table that can then be interpolated for the bulk of the receiver locations. This is usually the approach used with JASCO's Marine Operations Noise Model (MONM). MONM is comprised of two modules, for high and low frequency regimes. Both modules account for horizontal directivity of the source, and full exposure from a direct acoustic wave as well as exposure from acoustic wave reflections.

At frequencies below 2 kHz, MONM computes acoustic propagation via a wide-angle parabolic equation solution to the acoustic wave equation [8] based on a version of the U.S. Naval Research Laboratory's Range-dependent Acoustic Model (RAM) that has been modified to account for an elastic seabed [9]. This method has been widely benchmarked and is a popular choice for underwater acoustics modelling [10]. MONM-RAM accounts for the additional reflection loss at the seabed due to partial conversion of incident compressional waves to shear waves at the seabed and sub-bottom interfaces, and it includes wave attenuations in all layers. MONM-RAM's predictions have been validated against experimental data in several underwater acoustic measurement programs conducted by JASCO [11-16].

At frequencies above 2 kHz, MONM employs the widely-used BELLHOP Gaussian beam ray-trace propagation model [17] and includes sound attenuation due to volumetric absorption at higher frequencies [18], which is significant above 5 kHz. In contrast to MONM-RAM, the geoacoustic input for MONM-BELLHOP consists of only one interface, the sea bottom. This limitation is acceptable because the sub-bottom layers have negligible impact on underwater sound propagation above 1 kHz. MONM-BELLHOP also takes account of the vertical directivity of the source beam pattern.

3 Conclusion

The foregoing is a brief summary of some of the main underwater acoustic modelling tools used at JASCO.

Acknowledgments

Thanks to Alex MacGillivray for providing the details of the AASM model background.

References

[1] Alexander O. MacGillivray and N. Ross Chapman, *Results from an acoustic modelling study of seismic airgun survey noise in Queen Charlotte Basin*, School of Earth and Ocean Sciences, University of Victoria, 2005. [link]

[2] Alexander Orion MacGillivray, *An Acoustic Modelling Study of Seismic Airgun Noise in Queen Charlotte Basin*, MSc Thesis, school of Earth and Ocean Sciences, University of Victoria, 2006. [link]

[3] A. Ziolkowski. A method for calculating the output pressure waveform from an airgun. *Geophysical Journal of the Royal Astronomical Society* 21(2):137-161, 1970.

[4] W.H. Dragoset. A comprehensive method for evaluating the design of airguns and airgun arrays. In: *16th Annual Proc. Offshore Tech. Conf.* 3:75–84, 1984.

[5] M. Laws, L. Hatton, and M. Haartsen. Computer modeling of clustered airguns. *First Break* 8:331–338, 1990.

[6] M. Landro. Modeling of GI gun signatures. *Geophysical Prospecting* 40:721–747, 1992.

[7] V. Černý. Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *Journal of Optimization Theory and Applications* 45:41-51,1985.

[8] M.D Collins. The split-step Padé solution for the parabolic equation method. *J. Acoust. Soc. Am.* 93(4):1736-1742, 1993.

[9] Y. Zhang and C. Tindle. Improved equivalent fluid approximations for a low shear speed ocean bottom. *J. Acoust. Soc. Am.* 98:3391-3396, 1995.

[10] M.D. Collins, R.J. Cederberg, D.B. King, and S.A. Chin-bing. Comparison of algorithms for solving parabolic wave equations. *J. Acoust. Soc. Am.* 100(1)178-182, 1996.

[11] D.E. Hannay and R.G. Racca. *Acoustic Model Validation*. Document 0000-S−90-04-T−7006-00-E, Revision 02. Technical report for Sakhalin Energy Investment Company Ltd., 2005. [link]

[12] L. Aerts, M. Blees, S. Blackwell, C. Greene, K. Kim, D. Hannay, and M. Austin. *Marine Mammal Monitoring and Mitigation During BP Liberty OBC Seismic Survey in Foggy Island Bay, Beaufort Sea, July-August 2008: 90-Day Report*. Document No. LGL Report P1011-1, 2008.

[13] D. Funk, D. Hannay, D. Ireland, R. Rodrigues, and W. Koski (eds.). *Marine Mammal Monitoring and Mitigation during Open Water Seismic Exploration by Shell Offshore Inc. in the Chukchi and Beaufort Seas, July–November 2007: 90-Day Report*. LGL Report P969-1, 2008. [link]

[14] D.S. Ireland, R. Rodrigues, D. Funk, W. Koski, and D. Hannay. *Marine Mammal Monitoring and Mitigation during Open Water Seismic Exploration by Shell Offshore Inc. in the Chukchi and Beaufort Seas, July–October 2008: 90-Day Report*. LGL Report P1049-1, 2009.

[15] C. O'Neill, D. Leary, and A. McCrodan. Sound Source Verification. (3) *In:* Blees, M.K., K.G. Hartin, D.S. Ireland, and D. Hannay (eds.). *Marine mammal monitoring and mitigation during open water seismic exploration by Statoil USA E&P Inc. in the Chukchi Sea, August-October 2010: 90-day report*. LGL Report P1119, pp 3-1 to 3-34, 2010.

[16] G. Warner, C. Erbe, and D. Hannay. 2010. Underwater sound measurements. (3) *In:* Reiser, C.M., D.W. Funk, R. Rodrigues, and D. Hannay (eds.). *Marine Mammal Monitoring and Mitigation during Open Water Shallow Hazards and Site Clearance Surveys by Shell Offshore Inc. in the Alaskan Chukchi Sea, July-October 2009: 90-Day Report*. LGL Report P1112-1,. pp 3-1 to 3-54. [link]

[17] M.B. Porter, and Y.-C. Liu. Finite-element ray tracing. *In*: Lee, D. and M.H. Schultz (eds.). *Proceedings of the International Conference on Theoretical and Computational Acoustics*. Volume 2. World Scientific Publishing Co. pp 947-956, 1994.

[18] F.H. Fisher and V.P. Simmons. Sound absorption in sea water. *J. Acoust. Soc. Am.* 62(3):558-564, 1977.

MODELLING EXPOSURE OF MARINE MAMMALS TO UNDERWATER NOISE FROM PULSED SOURCES IN LONG-DURATION SURVEYS

Mikhail M. Zykov[*1], Terry J. Deveau[†1], and David G. Zeddies [‡2]
[1]Jasco Applied Sciences, 202 – 32 Troop Ave., Dartmouth, NS, B3B 1Z1, Canada
[2]Jasco Applied Sciences, 2004 Coleridge Drive, Shady Side, MD, 20902, USA

1 Introduction

The current rate of geological and geophysical ocean exploration conducted through seismic surveys has brought about increased concern for marine species near such operations. Seismic surveys typically use sources capable of producing loud impulse sounds, and the surveys may run for weeks and cover hundreds of kilometers of acquisition track. Determining their potential effects on marine mammals requires an estimate of the sound levels to which those animals are exposed over prolonged periods. Approaches based on specific thresholds for instantaneous received levels have been commonly used, but more recent thinking has focused on cumulative exposure levels to provide accurate evaluation of the acoustic impact on marine mammal populations.

JASCO Applied Sciences has been involved in acoustic measurement and modelling for many years, and has built up an extensive set of source signatures and source modelling tools. JASCO has also developed and made use of acoustic wave propagation models and post-processing methods that have allowed us to provide accurate estimates of the acoustic field around a source. To estimate the impact of sound on a population of marine mammals, however, requires information on the probability of animals being exposed over time to given levels within the sound field. Such data can be obtained by combining the modelled acoustic field with animal movement patterns; in other words, departing from a static paradigm and considering the dynamic system of moving animals and moving sources. This approach entails a large amount of input data and considerable computer processing requirements.

The Effects of Sound on the Marine Environment (ESME) modelling package developed by the Office of Naval Research (ONR) and Boston University [1], is an open-source software that can be used to model the dynamic system of moving animals and sound sources. ESME integrates databases of acoustic environmental parameters, the Marine Mammal Movement and Behavior (3MB) model developed by Houser [2], and basic acoustic propagation tools (RAM-C, BELLHOP), within the structure of a graphical user interface (GUI). Our group began working with ESME as the operational framework for estimating sound exposure in a dynamic environment.

After initial alterations to the ESME code to incorporate JASCO's acoustic modelling tools for the estimation of the sound field, we were able to use ESME for small scale modelling projects. The biggest drawback of ESME, however, was performance. A 7-day modelling scenario with ~3000 simulated animals would take about 60 hours of computer time to execute. A project involving complex exposure scenarios played out over weeks or months would be prohibitive with ESME. We undertook therefore the development of a faster and more flexible system for combining simulated animal movement and acoustic fields.

2 Results

2.1 JEMS

The JASCO Exposure Modelling System (JEMS) was developed in the IDL programing language (Exelis, Herndon, VA) and, unlike ESME, uses a parallelized approach that allows independent processing blocks to execute on different machines. Each block has its own input and output and the interaction between blocks occurs through standardized interchange files. Specific project requirements are easily satisfied through the selection of processing blocks, including the seamless integration into JEMS of different modelling tools from JASCO's library.

The major inputs to JEMS are:

- acoustic field data,
- source pattern data, and
- animal movement data.

Execution of a processing block results in an output file that may be reused within or across different projects. Because each block runs separately, unit-level quality control procedures can be implemented by testing and verifying blocks before making them available for use.

2.2 Inputs

Acoustic data

The per-pulse acoustic field is modelled at selected locations (acoustic sampling points) within the exposure modelling area and passed to JEMS as a standard N × 2-D grid, where N is the number of radials modelled in a range-depth plane. The number of sampling points and their location are selected based on variation in the propagation environment (e.g., bathymetry and geoacoustic parameters) within the modelling region. Multiple acoustic data files can be included in a JEMS run to represent different sources, the same source at different locations, or different descriptions of a field (various metrics such as SEL, peak SPL, rms SPL, and frequency filtering in species-specific auditory bands).

Source pattern data

The pattern of the source movement is passed to JEMS in a table of coordinates indexed by time. Sources used for seismic exploration are often towed behind a vessel, so their geo-referenced sound field changes over time as the source

moves. Additional indices can be used to specify the acoustic field dataset to be used at each source location to reflect changing propagation conditions. JEMS includes a collection of scripts to produce specific track patterns commonly used in seismic surveys, such as "snake track", "race track", spiral, and coil. Arbitrary shot patterns can be defined, or shot points can be based on a time interval or spatial reference including water depth or azimuthal location. Stationary sources are defined using constants for the location coordinates. No limit is placed on the number of sources or patterns used in a JEMS run.

Animal movement data

JEMS can use animal movement data from models such as the Marine Mammal Movement and Behavior (3MB) model [2]. 3MB simulates realistic, species-specific movement of individual animal "agents", called animats. Individual animat tracks are logged by 3MB in a binary file as a triplet of coordinates (latitude, longitude, depth) and a specified time step. The size of the movement data files may reach several gigabytes for long simulation periods, but pre-loading of the file is not required. JEMS maintains a file pointer based on the time step to gain random access into the animal movement data file and extract the movement data needed at each time step.

2.3 Exposure modelling process

Source pattern data files are first loaded into memory along with the per-pulse acoustic field data. The process of exposure modelling is to step through the source data and load animal positions for the current time step. The received exposure level of every animat is logged for each time step, and the values are written into the exposure output file along with auxiliary information such as the source ID and coordinates, the coordinates of the animat, and the slant and horizontal ranges of the animat to the source.

2.4 Post-processing

The usefulness of this form of acoustic modelling is in determining the exposure probability function of a population of animals from which to estimate the potential impact. When many animats moving realistically are used to sample the sound fields of a survey, the probability of exposure for a real-world distribution of animals can be built up from the estimated received levels. In the post-processing stage, the exposure modelling output is analyzed to obtain datasets of acoustic metrics such as cumulative and maximum exposure for each animat, and fraction of animats exposed above specific thresholds.

3 Discussion

ESME is open-source code and can be freely modified, though revisions and additions of new features require the involvement of an experienced programmer and the use of specialized programming tools. ESME's built-in databases simplify the setting up of scenarios, but make including data from other sources a cumbersome process.

JEMS' design is based on a decentralized and modular approach. Ultimately, JEMS is a flexible infrastructure for convolving simulated animal movement data with simulated sound fields (from moving sources).

Advantages of the JEMS approach:

- processing speed;
- ease and flexibility to accommodate wide-ranging project requirements and incorporate various data types, such as new sources or motion patterns;
- natural entry points to perform quality control of all inputs using internal and external means;
- ability to reuse inputs that remain unchanged in subsequent scenarios;
- modelling scenarios set up through configuration files, making it easy to replicate scenarios or batch process multiple scenarios.

Shortfalls of JEMS still requiring development work, currently ongoing:

- can use only one CPU core per processing block (multiple instances of the code, however, can be executed on one machine if it is equipped with multiple cores);
- no feedback interaction with animal movement model, precluding simulation of aversion;
- no GUI interface.

4 Conclusion

The modularized approach used by JEMS allows greater flexibility in the implementation of new features and functionality. Changes to one component of the exposure modelling system do not require changes to the other components, as long as the same standard format is used to pass the data to the subsequent step. The input datasets required for the exposure modelling can be generated independently and in parallel on multiple computers. For subsequent scenario runs the input datasets that do not require changes can be reused without recalculation. With more than a 100 fold performance improvement over ESME, the approach is seen as more suitable for projects with a large number of exposure scenarios.

Acknowledgments

We thank Dorian Houser for assistance with 3MB, and members of David Mountain's lab at Boston University for helpful discussions of the ESME code.

References

[1] Shyu, H-J and Hillson, R. (2006) "A Software Workbench for Estimating the Effects of Cumulative Sound Exposure in Marine Mammals" IEEE J. Ocean. Eng., 31: 8-21

[2] Houser, D. S. (2006). "A method for modeling marine mammal movement and behavior for environmental impact assessment." IEEE J. Ocean. Eng. 31(1): 76-81.

PHYSICAL MECHANISMS UNDERLYING THE ACOUSTIC SIGNATURES OF BREAKING WAVES

Cameron Dallas[1] **and Cristina Tollefsen**[2]
[1,2]Defence Research and Development Canada - Atlantic, P.O. Box 1012, Dartmouth, NS, B2Y 3Z7
[2]cristina.tollefsen@drdc-rddc.gc.ca

1 Introduction

The characteristics of waves breaking along the shoreline have been extensively researched over the past century, including wave breaking motions and aftereffects [1], as well as the relationships among bathymetry, wave shape (plunging, spilling, or surging) and wave breaking location [2]. The resulting underwater sound generation has also been well studied and has been elucidated as the individual and collective oscillations of bubbles and bubble plumes produced by the breaking waves [3] [4] [5].

In spite of this large body of research, the airborne sound generated by breaking waves is not as well-understood. Bolin and Åbom [6] proposed a model in which sound from the underwater bubble clouds propagate through the air-water interface, but also concluded that the main airborne sound generation mechanism remains a mystery. The aim of this paper is to analyze acoustic data from breaking waves and determine the dominant airborne sound generating mechanisms.

2 Methods

Recordings were made of breaking waves at Osborne Head, Nova Scotia, Canada (44.612 N, 63.420 W) from June 1 to August 22, 2011. A microphone was situated on a cliff 6 m above a rocky beach which experiences abundant breaking wave activity. Five minutes of uncompressed audio data sampled at 25.6 kHz were recorded every half hour. A camera mounted on a nearby building acquired photographs of the surf zone every five minutes. Directional wave spectra and significant wave height (SWH) were measured at a location 800 m from the microphone site, while weather data were obtained from a weather buoy 12.5 km SE of the microphone site. Further details can be found in a previous paper [7].

Thirteen isolated plunging breaking wave events were selected from the dataset, for which the significant wave height ranged from 0.8 m to 1.08 m. The time of the wave break was determined by listening to the audio recordings and marking the time at which the water jet from the curling wave appeared to impact the water surface. All data analysis was performed in Python using the Scikits Audiolab library. The data were highpass filtered with a cutoff frequency of 20 Hz to remove the infrasound vibrations.

3 Results

Time series of the third-octave band levels were constructed by calculating the Fourier transform of the sound pressure signal in lengths of 0.4 seconds (10240 samples) with a Hann window and 2.5% overlap (256 samples) and then summing the energy in each band with centre frequencies ranging from 40 to 2500 Hz. Figure 1 is a time series of the third-octave band level in the 50-Hz band for each breaking wave event. The higher overall band levels correspond to periods with larger significant wave height. The plots of the other bands are similar to Figure 1, including the large variability in power ($\pm$10 dB) and waveform in the time surrounding the break.

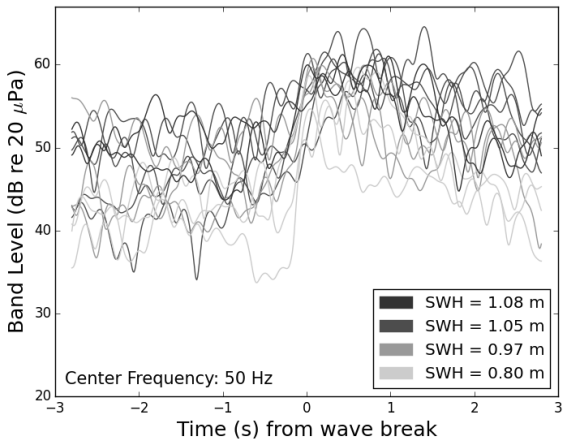


Figure 1: Time series of the third-octave band level for the 50-Hz band for each breaking wave in the data set. Greyscale of plot lines indicates significant wave height.

The band level change ΔBL was calculated for each wave and each third-octave band time series by manually marking the minimum BL_{min} and maximum BL_{max} band levels before and after the wave break and calculating their difference ($\Delta BL = BL_{max} - BL_{min}$). Figure 2 is a plot of the band level change averaged over all thirteen waves $\overline{\Delta BL}$ as a function of third-octave band centre frequency. As a wave breaks, the level in every band increases by 5 to 12 dB. At frequencies below 250 Hz, the band level change increases with decreasing frequency and reaches a peak of 12 dB at 50 Hz, while the band level change is constant (5-6 dB) for frequencies between 300 Hz and 800 Hz, and slightly larger (6-7 dB) for frequencies greater than 1000 Hz.

Plunging waves have a distinctive 'whoomphing' or bellowing sound that is absent when waves are merely spilling or surging. Figure 3 is a power spectrum for a plunging wave which exhibits this distinctive sound, computed for the 0.4 s immediately following the breaking, using a window length of 0.2 s and a 50% overlap. The spectrum contains first-order (and possibly second-order) harmonics of the prominent low-frequency maxima at 34 Hz, 48 Hz, and 78 Hz.

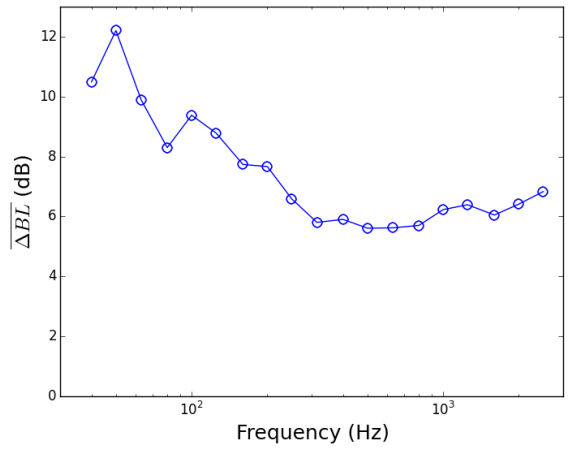


Figure 2: Average change in band level $\overline{\Delta BL}$ during the wave break as a function of third octave band centre frequency.

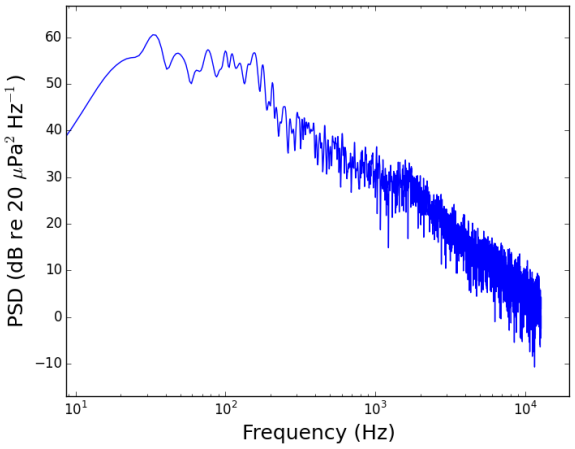


Figure 3: Power spectrum of the 0.4 s immediately following the plunging wave break.

4 Discussion

The mechanisms generating the airborne sounds are obscured to some extent by the observed variability : the band level increases differ in timing and magnitude among individual waves. The smaller peak in band level change centred at 100 Hz in Figure 2 may be caused by the plunging jet of water impacting the ocean surface. The same frequencies were found to be affected when Galbrun and Ali [8] measured the sound generated by small artificial waterfalls.

We postulate that the maximum in band level change near 50 Hz during the wave break is caused by the air column entrained by the plunging wave, which behaves like a hollow cylindrical pipe that is closed at one end. As a plunging wave breaks, the centrifugal force of the water will be briefly balanced by the pressure in the enclosed air tube [1]. The wave breaks in one direction, creating a bellows effect that sends a burst of air and aerated water through the opening in the side of the wave, much like the forcing in a flue organ pipe. The resonance frequency f_0 of a cylinder closed at one end is

$$f_0 = \frac{c}{4(L + 0.6R)} \quad (1)$$

where L and R are the length and radius of the cylinder, respectively, and c is the speed of sound in air [9]. The photographs were used to estimate the physical dimensions of representative breaking waves. Using values of $L = 1.0$ m and $R = 0.25$ m in Equation 1 yields a resonance frequency of $f_0 = 74.6$ Hz, which is typical of the sharp maxima in the observed spectra (e.g., Figure 3).

5 Conclusions

Recordings of airborne wave breaking sound demonstrated large variability in the signature of plunging breaking waves. Increases of 5-12 dB were observed in the third-octave band levels between 40 and 2500 Hz during the wave breaking. The two most prominent increases were postulated to be caused by the plunging jet impacting the water surface (100 Hz) and the resonance of the air cylinder entrained by the plunging wave (50-80 Hz). The resonance frequency of the wave modelled as a closed cylinder was comparable to the observed sharp frequency maxima in the spectrum.

References

[1] D.H. Peregrine. Breaking waves on beaches. *Ann, Rev. Fluid Mech*, 15 :149–178, 1983.

[2] J.A. Battjes. Surf similarity. In *Proceedings of the 14th Conference on Coastal Engineering*, Copenhagen, Denmark, 1974. American Society of Civil Engineers.

[3] G.B. Deane. Sound generation and air entrainment by breaking waves in the surf zone. *Journal of the Acoustical Society of America*, 102(5) :2671–2689, 1997.

[4] WK Melville, MR Loewen, and Eric Lamarre. Bubbles, noise and breaking waves : A review of laboratory experiments. In *Natural Physical Sources of Underwater Sound*, pages 483–501, Cambrideg, U.K., 1993.

[5] B.R. Kerman. Audio signature of a laboratory breaking wave. In *Sea Surface Sound*, pages 437–448, Lerici, Italy, 1988.

[6] Karl Bolin and Mats Åbom. Air-borne sound generated by sea waves. *Journal of the Acoustical Society of America*, 127(5) :2771–2779, 2010.

[7] Cristina Tollefsen and Brendan Byrne. Dependence of airborne surf noise on wave height. *Canadian Acoustics*, 39(3) :210–211, 2011.

[8] Laurent Galbrun and Tahrir T Ali. Acoustical and perceptual assessment of water sounds and their use over road traffic noise. *Journal of the Acoustical Society of America*, 133(1) :227–237, 2013.

[9] Charles Aaron Culver. *Musical acoustics*. McGraw-Hill, 1956.

HF SONAR PERFORMANCE IN WHITECAPS AND WAKES

Mark V. Trevorrow [1] and Svein Vagle [2]
[1]Defence R&D Canada Atlantic, Dartmouth, Nova Scotia
[2]Institute of Ocean Sciences, Sidney, British Columbia

1 Introduction

Small air bubbles present in the near the ocean surface are well known to interfere with the operation of High Frequency (HF, > 10 kHz) active sonars. These bubbles are naturally created by breaking surface waves (whitecaps) [1] and also generated by propeller and hull effects in a ship wake [2]. In both cases these bubbles are predicted to limit the performance of horizontally-oriented HF sonars operated from a ship [1,3]. Such sonars have been previously utilized for forward-looking obstacle avoidance and in-harbor surveillance purposes. There is renewed interest in such sonars for close-in ship defence purposes. In order to specifically assess HF sonar performance in the vicinity of a ship a sea-trial was conducted on the CCGS Vector April 8 to 14, 2015 in BC south coast waters.

2 Experimental Description

The CCGS Vector is 40 m length overall with displacement of 515 tonnes and top speed of 12 knots. These sea-trials took advantage of a retractable sonar strut on the ship's starboard side, allowing use of a HF multibeam sonar at a depth of 3.5 m in either a forward- or aft-looking horizontal orientation.

The forward-looking sonar tests examined the detectability of drifting near-surface targets at low and medium sea states. The targets were hollow steel spheres, 0.9 m diameter, ballasted to submerge them either 1 or 4 m below the surface. Sonar data was recorded as the ship made a series of close passes by the targets at differing approach headings relative to the wind/sea. The typical ship speed during these approach runs was 3 to 5 knots under sea state conditions ranging from 0 to 5. These tests were conducted in both Saanich Inlet and the southern Strait of Georgia in water depths of approximately 200 m.

The aft-looking sonar tests were designed to test target detection in the vicinity of the ship wake. In this case a newly designed acoustic target was towed at distances between 20 and 500 m behind the ship at speeds of 6 to 7.5 knots. The aft-looking sonar tests were conducted under calm conditions in Saanich Inlet.

A Kongsberg Mesotech SM-2000 system operating at 90 kHz was utilized for these sea-tests. This sonar provided 128 overlapping 1.5° beams covering an angular aperture of 90°. The vertical beam-width was 13.5°. These tests utilized sonar range settings between 200 m and 800 m, pulse lengths between 0.5 and 1 ms, and pulse intervals between 0.5 and 1 s. Ship position (GPS) and heading (gyrocompass) were recorded along with the raw sonar data. The present results are based on successive snapshots (images) of single ping data.

In order to provide *in situ* bubble density measurements, two acoustic resonator devices were deployed at 1.2 m and 4.7 m depth from a surface following float. These devices (described in [4]) were capable of detecting bubbles with radii between approximately 2 and 250 microns with a 1 s update rate.

3 Sonar Detection Performance

Based on a review of the sonar imagery there were a number of qualitative observations which apply to both the forward- and aft-looking sonar tests, specifically:

- Targets were generally detected and followed using their ping to ping consistency. On a single-ping basis targets were generally difficult or impossible to detect due to the presence of background clutter with similar intensity and spatial scale.
- The images exhibited significant small-scale variability from ping to ping. This might be best described as *fluctuations* or *scintillation*. Both target echoes and the background reverberation exhibited this variability.
- On occasions when the ship accelerated, significant acoustic interference attributed to propeller cavitation was observed. This disappeared when the ship speed reached a steady-state speed.

It was also found that the receiver bandwidth used in the SM-2000 was tightly coupled to the sonar pulse length; presumably set at minimum values in order to minimize electronic noise interference. Unfortunately in moving ship operations some parts of the sea surface backscatter reverberation or wake scattering were shifted out of the receiver band by the Doppler effect, reducing their apparent amplitude. This Doppler limitation was seen with the 600 m and 800 m sonar range settings at ship speeds above 5 knots.

3.1 Forward-Looking Drifting Target Tests

Forward-looking sonar data were collected in Saanich Inlet under relatively calm conditions on April 9th and 10th, and under medium sea states in the Strait of Georgia on April 11th and 12th. The targets were observed at significantly greater ranges in Saanich Inlet, due to the calmer wind conditions. Figure 1 shows an example target detection (small red dot) at 530 m range. This target could be clearly tracked inwards over a period of roughly 7 minutes as the ship approached. The large intense scattering features on the right side of the image is the remnant bubbly wake from a previous pass by the ship. While in this particular image the target has high signal strength relative to the local

background, there are other regions of higher sea-surface backscatter in the image against which the target would not be detectable. At higher sea states there was often a region of higher reverberation in the near-field zone, up to 180 m.

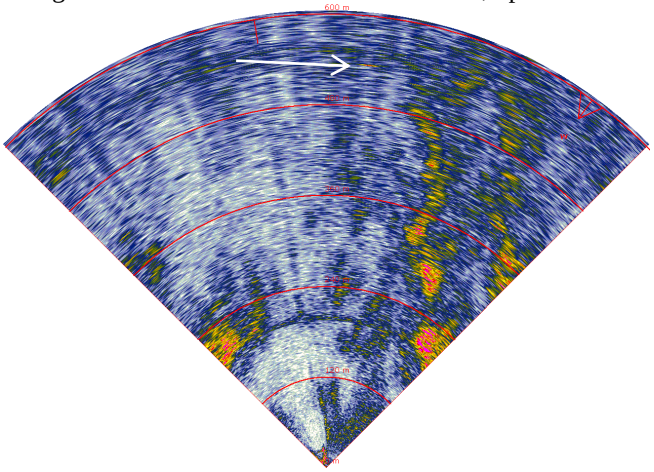


Figure 1: *Forward-looking HF sonar image to 600 m range, 1011h April 09th in Saanich Inlet. Range rings are in 120 m increments. Arrow denotes location of target sphere.*

Through visual examination of the sonar imagery, the target maximum detection ranges were extracted from the various runs. There were no cases where the targets were not detectable, although in higher sea-states the targets became difficult to track inside of roughly 100 to 150 m range. The average detection ranges were (mean ± standard deviation) 503 m ± 50 m and 274 m ± 85 m in Saanich Inlet and Strait of Georgia, respectively.

3.2 Aft-Looking Towed Target Tests

The aft-looking sonar tests were all conducted on April 13th under calm conditions in Saanich Inlet. Figure 2 shows an example of the towed target detected at approximately 150 m range. The sonar was inadvertently mounted with a 9° counter-clockwise heading offset, so that in the images the wake is apparently rotated 9° clockwise of the ship track. The ship speed at this time was 7.3 knots.

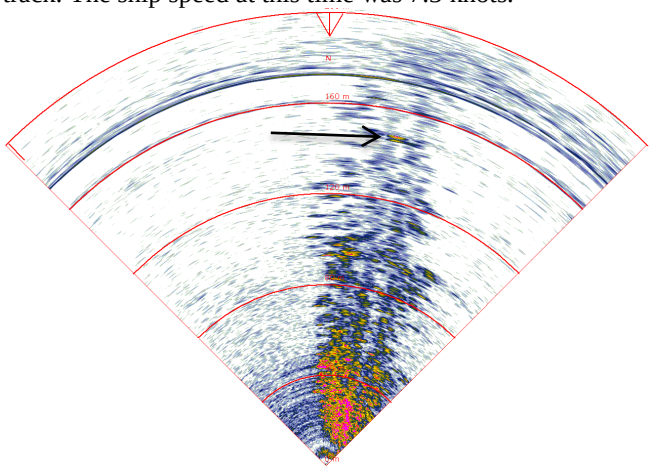


Figure 2: *Aft-looking HF sonar image to 200 m range, 1128h April 13th in Saanich Inlet. Range rings are 40 m increments. Arrow denotes location of towed target.*

At the time in Fig. 2 the towed target was located at a depth of roughly 12 m, well within the vertical aperture of the sonar beam but likely below the wake bubbles. In a reference frame fixed to the ship the wake backscattering appeared to translate away at the ship speed. The target was detectable by its apparent inward motion relative to the moving background. While the ship wake backscatter can be clearly seen in the *near-field* (< 50 m range), it generally faded in intensity at greater ranges presumably due to acoustic extinction by the near-field wake bubbles. This fading allowed relatively straightforward target detection at ranges from 50 m up to approximately 450 m. The target was often lost in the strong wake scattering in the near-field zone (< 50 m).

4 Summary Discussion

The HF multi-beam sonar functioned well in both the forward- and aft-looking target detection tests, producing a valuable data set covering a variety of sea states and near-surface sound-speed conditions. In the forward-looking tests targets were routinely observed at ranges up to 600 m in low sea-states, but the detection range was greatly reduced in higher sea states. While this suggests wind speed was the dominant factor in target detectability, the effect of near-surface sound speed gradients cannot be discounted. During the aft-looking sonar tests the towed target was detectable up to 500 m behind the ship at speeds from 5 to 7.5 knots. While this result is encouraging, the ship wake at the maximum speed (7.5 knots) was relatively weak. Greater wake width, depth, and densities have been previously observed with this same ship at speeds up to 12 knots [2]. If opportunities arise, further aft-looking wake tests should be conducted at higher ship speeds.

The next step in this work is to extract quantitative target Signal to Background Ratio (SBR) data from the various sonar tests. Here the acoustic background includes the combined effects of sea-surface reverberation, wake backscatter, and sonar self-noise. The variation of this SBR parameter against wind speed and measured bubble densities can then be used to validate sonar performance models. Acoustic calibration of this sonar will be required.

References

[1] Trevorrow, M., 2003. Measurements of near-surface bubble plumes in the open ocean with implications for high-frequency sonar performance, *J. Acoustical Soc. Am.* **114**(5), 2672-2684.

[2] Trevorrow, M., S. Vagle, and D. M. Farmer, 1994. Acoustic measurements of micro-bubbles within ship wakes, *J. Acoustical Soc. Am.* **95**(4), 1922-1930.

[3] Trevorrow, M., and Myers, V., 2012. Modeling obstacle avoidance sonar performance in the presence of near-surface bubbles, *11th European Conference on Underwater Acoustics*, 02 – 06 July, Edinburgh, UK.

[4] Farmer, D., Vagle, S., and Booth, A. D., 1998. A free-flooding acoustical resonator for measurement of bubble size distributions, *J. Atmos. Oceanic Technology* **15**(5), 1132-1146.

SENSITIVITY OF BELLHOP INTENSITY TO UNCERTAINTY IN SOUND SPEED

Diana McCammon *[1] and Sean Pecknold †[2]
[1]McCammon Acoustical Consulting, Black Rock, NS and Maritime Way Scientific, Ottawa, ON
[2]DRDC Atlantic Research Centre, Dartmouth, NS

1 Introduction

The Bellhop Gaussian beam solution formulated by Porter and Bucker [1] for acoustic propagation in an underwater environment is a complicated function of the environmental sound speed. In addition to the standard set of ray tracing differential equations, the incoherent intensity from the Bellhop algorithm is dependent upon a beamwidth factor obtained by the integration of a second pair of differential equations along the ray path. This factor can be expressed in closed form for linear constant gradient sound speed profiles. The derivatives of these equations with respect to (w.r.t) the sound speed in each layer express the sensitivity of the Gaussian Beam method to uncertainty in sound speed. Using these derivatives, the uncertainty in the transmission loss can be estimated from the uncertainty in the environment. In this paper, we present a pictorial display of the Bellhop intensity derivatives w.r.t. the sound speed.

2 Method

The Bellhop Gaussian Beam incoherent intensity can be written as

$$I(r,z) = \sum A \frac{c}{c_0} \frac{\exp(\frac{-n^2}{2q^2})}{q} \Re_B^2 \Re_S^2 \qquad (1)$$

where A is a constant, independent of beam number, that carries the cylindrical spreading and attenuation. The symbols $\Re$ contain the reflections coefficients from the boundaries (bottom and surface) and c is the sound speed The Gaussian factor depends upon n, the normal distance from the ray's position to the receiver location, and the beam's width q. The value of q satisfies the second system of differential equations that Bellhop traces and is capped when small. The function q is therefore an extremely important term in the amplitude of Bellhop.

When the sound speed layers are assumed to be linear with constant gradients, a closed-form expression can be found for q that is inversely proportional to the sine of the ray angle and the gradient of the layer. The function q can be differentiated with respect to the sound speed c in each layer, and therefore the intensity derivative dI/dc can be derived. It is found that one of the terms in this function is inversely proportional to $\sin^3\theta$, meaning that the sensitivity to the sound speed will be greatest as the beam is approaching a turning point.

The relation between the variance of the intensity σ_I^2 and the variance of the random sound speed uncertainty σ_c^2 is given in Papoulis [2] as

$$\sigma_I^2 = \left(\frac{dI}{dc}\right)^2 \sigma_c^2 \qquad (2)$$

This paper shows examples of the intensity derivatives' spatial distribution which highlight the regions where the uncertainty in sound speed would produce the highest uncertainty in the propagated intensity.

3 Results

3.1 Test environment

The sound speed profile used for these tests is shown in Fig 1. It is designed to provide structure to the acoustic field with both ducted and deep refracted rays. It features a surface duct as well as a sub-surface duct to illustrate the intensity and its derivative captured in each segment. The bottom is flat at 200m and the composition is hard sand to give good reflections. The frequency is 1000Hz, and about 400 beams are traced. The lowest SSP point, is defined well below the bottom depth.

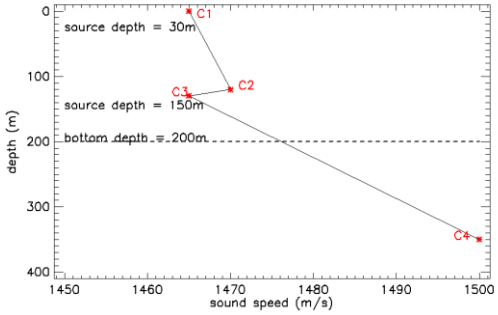


Figure 1 Sound speed profile defined by 4 points, c_1 thru c_4.

3.2 30 m source depth

The transmission loss at 1000Hz is shown in Fig 2 and reveals the expected ray concentrations in the surface duct as well as a pair of deep refracted rays and some leakage from the surface duct.

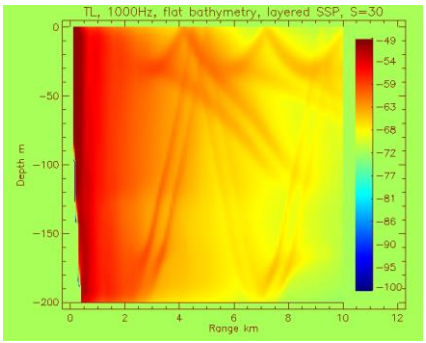


Figure 2 Full field transmission loss for 30m source.

Some of the dB intensity derivatives are shown in Fig 3. The upper plot is the derivative w.r.t. the surficial sound speed dI/dc_1. This plot is substantially the same for the derivative w.r.t. the second sound speed point dI/dc_2 (therefore not repeated here). The lower plot is the

derivative w.r.t. c_3 and this plot is also substantially the same for the last SSP point dI/dc_4 (not shown here).

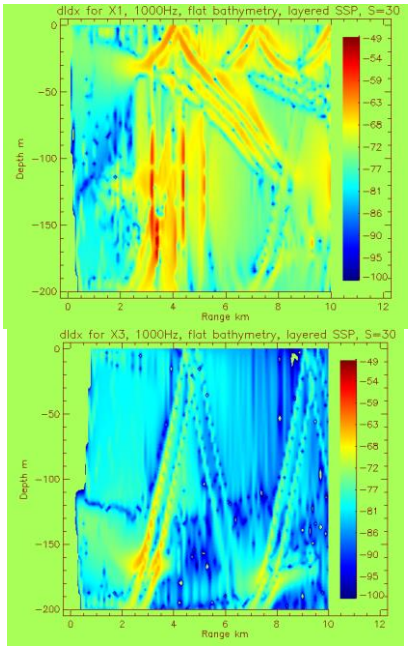


Figure 3 Full field intensity derivatives w.r.t. c_1 or c_2 (upper), c_3 or c_4 (lower).

3.3 150 m source depth

The transmission loss at 1000Hz for a 150m source (Fig. 4) shows the influence of the sub-surface duct as well as deeply cycled rays.

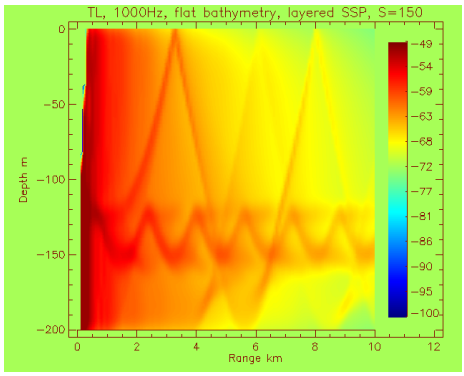
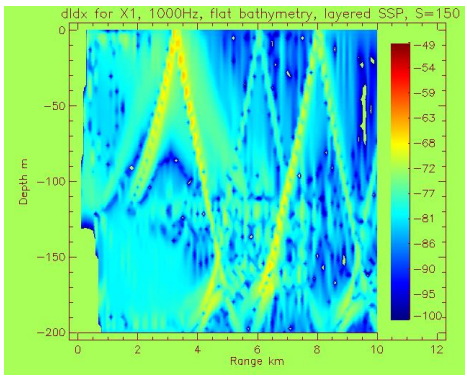


Figure 4 Full field transmission loss for 150m source.



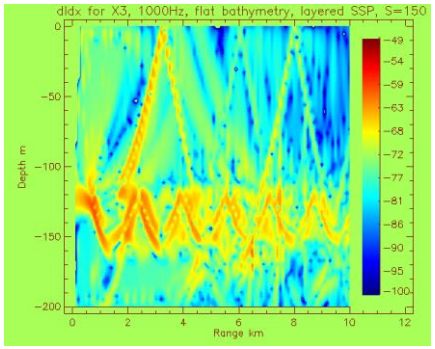


Figure 5 Full field intensity derivative w.r.t. c_1 or c_4 (bottom, previous column), c_2 or c_3 (above).

Some of the dB intensity derivatives are shown in Fig 5. The plot on the bottom of the previous column is the derivative w.r.t. the surficial sound speed dI/dc_1. This plot is substantially the same for the last sound speed point dI/dc_4 (not shown here). The plot above is the derivative w.r.t. c_3 and this plot is also substantially the same for the second SSP point dI/dc_2 (not shown here).

4 Discussion

A comparison of Fig 3 with Fig 2 and Fig 5 with Fig 4 shows that the intensity derivatives are largest along the edges of strong propagation intensity, particularly along caustic lines when the randomness is in the speeds defining the duct containing the source. The sound ducts that support strong propagation are also highly sensitive to changes in these speeds. The first set of intensity derivative plots (Fig 3) show strong sensitivity in dI/dc_1 and dI/dc_2 which define the surface duct in which the source is located, with much lower contributions from randomness in c_3 and c_4 because they lie outside the surface duct. Similarly, the second set of intensity derivative plots (Fig 5) show strong contributions from the derivative of the intensity w.r.t. c_2 and c_3 which define the sub-surface duct in which the source is located, but not as much sensitivity in dI/dc_1 or dI/dc_4. When the speed uncertainty is in the sound speed points outside the duct, only the deep cycled rays are affected.

5 Conclusion

Given the relationship between the variance of the intensity and the variance of the underlying sound speeds shown in Eq 2, the implication is that obtaining an accurate estimate for the speeds surrounding the source depth will be of most benefit in reducing the uncertainty in predicted transmission loss.

References

[1] M. Porter and H. Bucker Gaussian beam tracing for computing ocean acoustic fields, *J. Acoust Soc. Am*. 82 (4):1349, 1987.

[2] A. Papoulis, Probability, Random Variables, and Stochastic Processes, 2nd edition McGraw-Hill, 1984

MATCHED-FIELD SOURCE LOCALIZATION WITH NON-SYNCHRONIZED SENSOR ARRAYS

Stan E. Dosso
School of Earth and Ocean Sciences, University of Victoria
Victoria B.C. Canada V8W 3P6

1 Overview

This paper considers combining information from multiple hydrophone arrays in matched-field localization for an acoustic source in the ocean. Assuming each individual array is comprised of calibrated sensors which are synchronized in time, conventional matched-field methods can be applied for each array using the Bartlett (linear) processor and the resulting processors summed over arrays. However, if the relative calibration and/or time synchronization is known between some or all arrays, more informative multiple-array processors can be derived by maximum-likelihood methods. For example, if the relative calibration of the arrays is known, the observed amplitude variation of the received field between arrays provides additional information for source localization: the corresponding processor takes advantage of the fact that the source amplitude spectrum is the same (although unknown) for all arrays. Likewise, if synchronization between arrays is known, phase variation provides localization information. Various multiple-array processors are derived and evaluated in terms of the probability of correct localization from Monte-Carlo analyses. The analysis indicates that, dependent on specific array configurations, significant improvements in source localization performance can be achieved when including relative amplitude and/or phase information in multiple-array processors.

2 Theory

Consider measured complex (frequency-domain) acoustic data vectors $\mathbf{d}_{af}$ at N_a arrays (with H_a hydrophones each) and N_f frequencies due to a source at location $\mathbf{x} = (x, y, z)$. Assuming circularly-symmetric complex Gaussian-distributed errors with variances σ_f^2 and complex spectral strengths $\mathbf{S}$ (discussed below), the likelihood function is given by

$$L(\mathbf{x},\mathbf{S},\boldsymbol{\sigma}) = \prod_{f=1}^{N_f}\prod_{a=1}^{N_a}\frac{1}{\left(\pi\sigma_f^2\right)^{H_a}} \exp\left[-\left|\mathbf{d}_{af} - S e^{i\theta_{af}}\mathbf{d}_{af}(\mathbf{x})\right|^2 / \sigma_f^2\right], \quad (1)$$

where $\mathbf{d}_{af}(\mathbf{x})$ represents predicted acoustic fields computed by a numerical propagation model for a unit-amplitude, zero-phase source at location $\mathbf{x}$. The corresponding misfit (negative log-likelihood) function is given by

$$E(\mathbf{x},\mathbf{S},\boldsymbol{\sigma}) = \sum_{f=1}^{N_f}\sum_{a=1}^{N_a}\left|\mathbf{d}_{af} - S\mathbf{d}_{af}(\mathbf{x})\right|^2 / \sigma_f^2 + H_a \log_e \sigma_f^2. \quad (2)$$

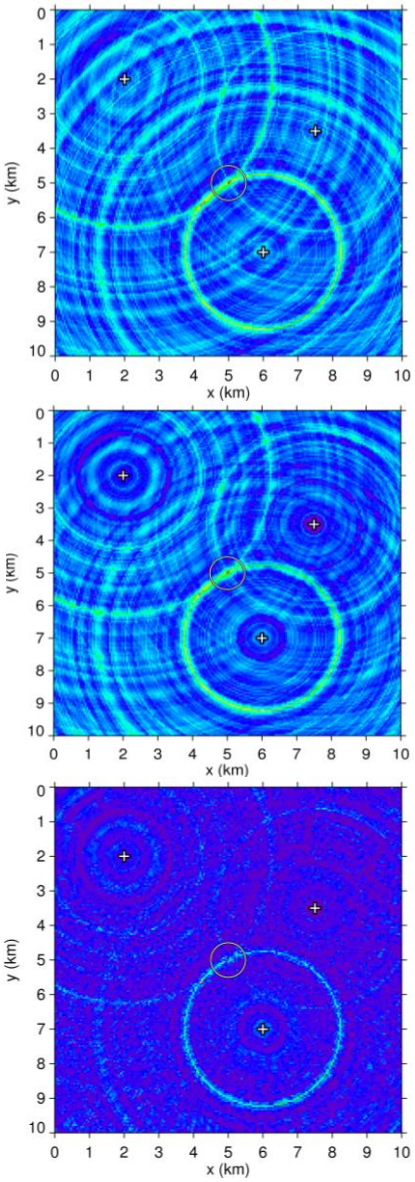


Figure 1: Comparison of horizontal (x-y) ambiguity surfaces (at the optimal source depth) for data from three vertical arrays (crosses) and the incoherent, amplitude-coherent, and coherent processors (top to bottom). The true source position is at (5,5) km, and white circles are centred at the maximum processor value.

Three cases consisting of different assumptions on the relative information available between arrays are considered. The first case assumes that the sensors of each individual array are calibrated and synchronized (have a

common clock), but that no information on calibration and timing is available between arrays. In this case the unknown spectral strength in Eqs. (1) and (2) is taken to be $S = A_{af} e^{i\theta_{af}}$ where A and θ represent source amplitude and phase, both of which vary with array (subscript a) and frequency (subscript f). Maximizing the likelihood over the unknown inter-array amplitude and phase factors as well as the noise variances (which are often unknown) by setting $\partial E / \partial A_{af} = \partial E / \partial \theta_{af} = \partial E / \partial \sigma_f = 0$ leads, after some algebra, to a misfit function (processor)

$$E(\mathbf{x}) = \sum_{f=1}^{N_f} \sum_{a=1}^{N_a} N_a \log_e \left[\left| \mathbf{d}_{af} \right|^2 - \frac{\left| \mathbf{d}_{af}(\mathbf{x})^\dagger \mathbf{d}_{af} \right|^2}{\left| \mathbf{d}_{af}(\mathbf{x}) \right|^2} \right] . \qquad (3)$$

The misfit E in Eq. (3) depends only on the source position $\mathbf{x}$ and the term in square brackets on the right is recognized as the incoherent Bartlett processor which is summed over arrays and frequencies.

The second case consists of relative knowledge of both sensor calibration (amplitude) and synchronization (phase) between arrays. In this case the unknown spectral strength is taken to be $S = A_f e^{i\theta_f}$ in which amplitude and phase depend on frequency but not array (no subscript a). Maximizing the likelihood over A_f, θ_f, and σ_f leads to a processor

$$E(\mathbf{x}) = \sum_{f=1}^{N_f} \left(\sum_{a=1}^{N_a} H_a \right) \log_e \left[\left| \sum_{a=1}^{N_a} \mathbf{d}_{af} \right|^2 - \frac{\left| \sum_{a=1}^{N_a} \mathbf{d}_{af}(\mathbf{x})^\dagger \mathbf{d}_{af} \right|^2}{\left| \sum_{a=1}^{N_a} \mathbf{d}_{af}(\mathbf{x}) \right|^2} \right], \qquad (4)$$

where the term in square brackets is recognized as the coherent Bartlett processor (over arrays and frequencies).

In many practical settings synchronization between arrays (i.e., a precise common clock) is not available, but the relative array calibrations are known. In this case the unknown spectral strength is taken to be $S = A_f e^{i\theta_{af}}$ in which the relative phase is unknown from array to array (subscript a on θ) but the relative amplitude is known. Maximizing the likelihood over A_f, θ_{af}, and σ_f leads to

$$E(\mathbf{x}) = \sum_{f=1}^{N_f} \left(\sum_{a=1}^{N_a} H_a \right) \log_e \left[\sum_{a=1}^{N_a} \left| \mathbf{d}_{af} \right|^2 - \frac{\left(\sum_{a=1}^{N_a} \left| \mathbf{d}_{af}(\mathbf{x})^\dagger \mathbf{d}_{af} \right| \right)^2}{\left| \sum_{a=1}^{N_a} \mathbf{d}_{af}(\mathbf{x}) \right|^2} \right]. \qquad (5)$$

In Eq. (5) the correlations between modelled and measured fields are summed incoherently (as in Eq. (3)), but the sum of the correlations is normalized by the summed amplitude of the modelled fields (as in Eq. (4)) to preserve relative information on source location contained in the measured field amplitudes from array to array (not included in Eq. (3) where the field at each array is normalized individually). Hence, this processor is refered to as "amplitude coherent."

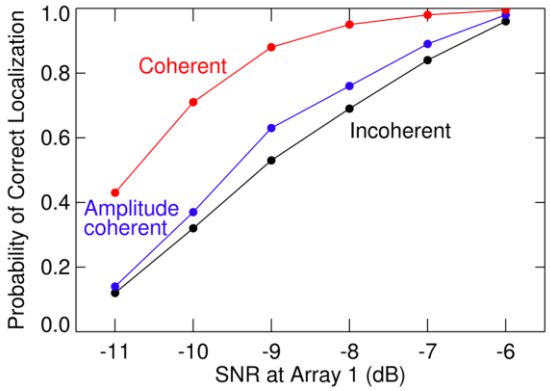


Figure 2: Comparison of probability of correct localization (from 500 random Monte Carlo trials) for the three processors as a function of SNR at Array 1 (upper left in Fig. 1).

3 Results

A (simulated) example of ambiguity surfaces for the three processors is shown in Fig. 1 for a source at (x,y) = (5,5) km and 50-m depth, and three vertical arrays of 24 sensors over a 100-m water column; five frequencies of 50, 100, 200, 300, 500 Hz are considered and the signal-to-noise ratios (SNR) are -5, -6.5, and -8.5 dB at the three arrays for all frequencies (arrays are numbered clockwise from the upper left in Fig. 1). This figure shows a substantial improvement in the ability to localize the source (relative to background processor values) for the coherent processor (bottom panel) over the incoherent processor (top panel); however, inter-array synchroniztaion required for the coherent processor is often unavailable. Figure 1 also shows a smaller but significant improvement over the incoherent processor due to the inclusion of relative calibration information (middle panel), which is frequently available.

The relative performance of the three processors is quantified using a Monte Carlo study in which 500 realizations of noisy data and random source locations were considered for the three processors, with correct localization defined to be within 200 m in x and y and 10 m in z of the true location. Figure 2 shows the probability of correct localization for the processors as a function of SNR at Array 1. While Fig. 2 shows the best performance at all SNRs for the coherent processor, the amplitude-coherent processor, which makes use of commonly-available information, is also significantly better than the incoherent processor.

4 Conclusion

Three maximum-likelihood processors were derived for matched-field source localization using multiple sensor arrays, differing in the level of knowledge of inter-array calibration and synchronization. The best results are obtained for the coherent processor based on full inter-array information. However, the (new) amplitude-coherent processor based on relative calibration knowledge (commonly available) also out-performs an incoherent processor which incorporates no inter-array information.

EFFICIENCY OF TRANS-DIMENSIONAL BAYESIAN INFERENCE FOR GEOACOUSTIC INVERSION

Stan E. Dosso
School of Earth and Ocean Sciences, University of Victoria, Victoria B.C. Canada V8W 3P6

1 Overview

This paper considers trans-dimensional (trans-D) Bayesian inference applied to a representative geoacoustic inverse problem of estimating seabed parameters from acoustic reflectivity measurements. Trans-D methods include model selection in inversion by sampling the posterior probability density (PPD) over models with differing numbers of parameters (dimensions) [1-3]. The approach is applied here to samplie over seabed geoacoustic models with a variable number of layers [2,3], providing seabed profile estimates with uncertainties that include the uncertainty in the model parameterization. However, trans-D sampling can be computationally intensive. Sampling efficiency is largely determined by the proposal schemes applied to generate perturbed values for existing parameters and for new parameters assigned to layers added to the model. Perturbations of existing parameters are considered in a principal-component (PC) space based on an eigenvector decomposition of the unit-lag parameter covariance matrix (computed from the history of sampled models, a diminishing adaptation) [3]. The relative efficiency of proposing new-layer parameters from the prior versus a Gaussian distribution focused near existing values (the common approch) is examined [3]. Parallel tempering [4], which employs a sequence of interacting samplers with successively-relaxed likelihoods, is also applied to increase the acceptance rate of new layers. The relative efficiency of various proposal schemes is compared through repeated inversions with a pragmatic convergence criterion [3].

2 Theory

Consider measured data $\mathbf{d}$ and a set of possible models $\{\mathbf{m}_k\}$ indexed by $k \in K$ (where K is a countable set); e.g., k can indicate the number of seabed interfaces in a geoacoustic model. Bayes' theorem for a hierarchical model including hyper-parameter k can be written [1]

$$P(k,\mathbf{m}_k \mid \mathbf{d}) = \frac{P(k)P(\mathbf{m}_k \mid k)P(\mathbf{d} \mid k,\mathbf{m}_k)}{\sum_{k' \in K}\int P(k')P(\mathbf{m}'_{k'} \mid k')P(\mathbf{d} \mid k',\mathbf{m}'_{k'})\,d\mathbf{m}'_{k'}}. \quad (1)$$

In Eq. (1), $P(k)$ and $P(\mathbf{m}_k|k)$ represent prior probability densities, and $P(\mathbf{d}|k,\mathbf{m}_k)$ is the conditional probability of $\mathbf{d}$ given k and $\mathbf{m}_k$, which, for (fixed) measured data is interpreted as the likelihood $L(k,\mathbf{m}_k)$. The PPD $P(k,\mathbf{m}_k|\mathbf{d})$ is defined over the trans-D model space defined by the union of all fixed-dimensional spaces spanned by K. A Markov chain which samples the trans-D PPD can be obtained via reversible-jump Markov-Chain Monte Carlo (rjMCMC) sampling [1] in which proposed state $(k',\mathbf{m}'_{k'})$ is generated from an existing state $(k,\mathbf{m}_k)$ via a proposal distribution $Q(k',\mathbf{m}'_{k'}|k,\mathbf{m}_k)$ and accepted with the Metropolis-Hastings-Green acceptance probability

$$A(k',\mathbf{m}'_{k'} \mid k,\mathbf{m}_k) = \min\left[1, \frac{Q(k,\mathbf{m}_k \mid k',\mathbf{m}'_{k'})}{Q(k',\mathbf{m}'_{k'} \mid k,\mathbf{m}_k)} \frac{P(k')P(\mathbf{m}'_{k'} \mid k')}{P(k)P(\mathbf{m}_k \mid k)} \frac{L^{1/T}(k',\mathbf{m}'_{k'})}{L^{1/T}(k,\mathbf{m}_k)} \mid \mathbf{J} \mid \right]. \quad (2)$$

In Eq. (2), T is referred to as the sampling temperature which relaxes the likelihood as used in parallel tempering (discussed below) and $|\mathbf{J}|$ is the determinant of the Jacobian matrix for transitions from $(k,\mathbf{m}_k)$ to $(k',\mathbf{m}'_{k'})$. For transitions which perturb the parameters of the present state but do not change model dimension (i.e., $k' = k$) it follows that $|\mathbf{J}| = 1$. The birth/death rjMCMC scheme [1,2] changes k by 1 (e.g., adds/deletes a single seabed interface) such that only the parameters added/deleted are changed, with proposed parameters in a birth step depending only on the parameters in the current state; in this case it also follows that $|\mathbf{J}| = 1$.

The rjMCMC algorithm draws dependent samples from the trans-D PPD such that representative parameter estimates and uncertainties (e.g., marginal probability densities) can be computed. However, the approach can be computationally intensive, with efficiency largely determined by the effectiveness of the proposal schemes employed for perturbing existing parameters in fixed dimensions and for adding new parameters in birth steps. Several approaches have been considered to improve sampling efficiency (described in detail in [3]). These include: (1) Parallel tempering [4], in which a sequence of interacting Markov chains are run at a range of temperatures $T \geq 1$ to successively relax the likelihood and improve the acceptance rate of birth and death steps. In this case either only the $T = 1$ chain is retained as parameter samples or all chains are retained and importance reweighting is applied to remove the sampling bias due to the non-unity sampling temperatures. (2) Drawing parameter perturbations in a PC parameter space determined by an eigenvector decomposition of the unit-lag covariance matrix estimated adaptively from the sampling history, effectively de-correlating parameters in the sampling. (3) Drawing new parameters in birth steps from their prior rather than from Gaussian distribution focussed near existing parameter values. Some of these approaches are considered in terms of sampling efficiency in this paper.

3 Results

Sampling approaches are considered here for the geoacoustic inverse problem of estimating profiles of seabed properties (sound speed, density, attenuation) from angle- and frequency-dependent measurements of acoustic reflection coefficients [2,3]. Figure 1 shows an example of the results

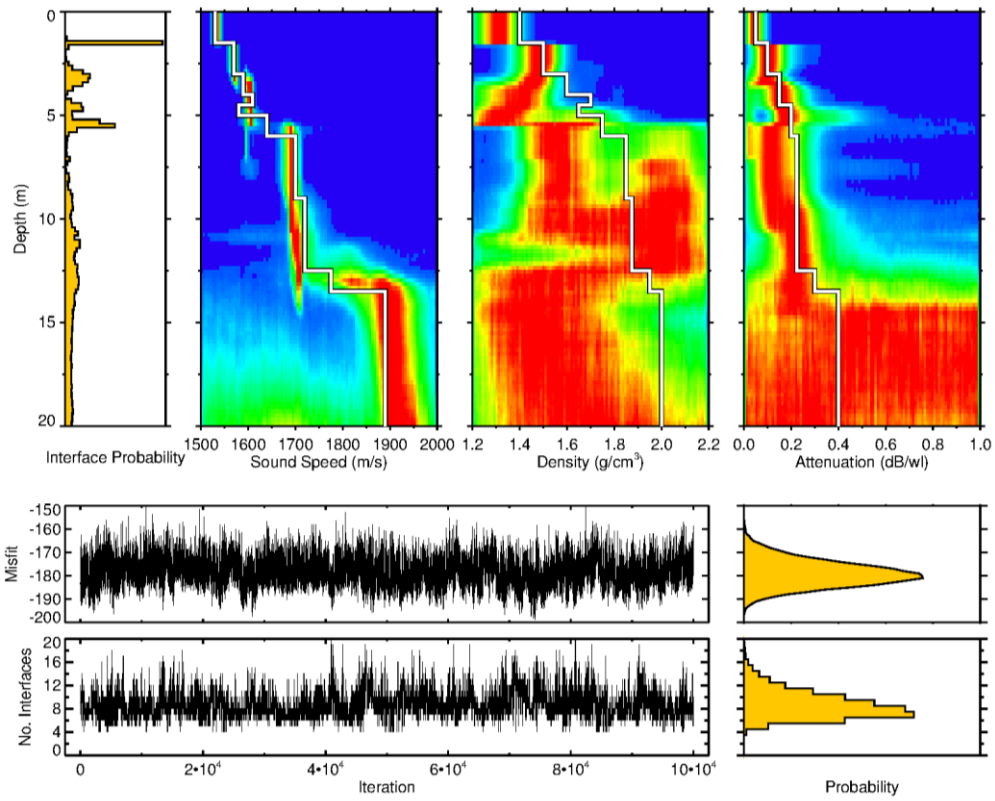


Figure 1: Example result of geoacoustics inversion results for (simulated) reflection coefficient data. The upper row, from left to right, shows marginal probability profiles of interface occurrence and of geoacoustic properties sound speed, density, and attenuation (solid lines indicate true values; hot colours indicate high probabilities). The middle and lower rows show the sampling history of the rjMCMC algorithm and marginal distributions for the misfit (negative log likelihood function) and the number of interfaces (the true number is 9).

of the trans-D geoacoustic inversion algorithm applied to (simulated) reflection coefficient data. To consider the efficiency of various approaches to sampling, a set of 10 inversions were run for each of four proposal schemes (PSs) with an automated convergence criterion based on examining the differences between parallel samplers [3]. PS1 consisted of PC parameter perturbations, geoacoustic parameters drawn from the prior in birth moves, and parallel tempering with 5 chains and $T_i = 1.25^{i-1}$. PS2 is the same as PS1 except that importance re-weighting is applied such that marginal profiles are computed using all parallel-tempering chains. PS3 is the same as PS1 except that parameter perturbations are drawn from a Gaussian distribution centred at current values with standard deviation equal to 1/10 of the prior-bound width. Finally, PS4 is the same as PS1 except that geoacoustic parameters in birth moves are drawn from a Gaussian distribution with standard deviation 1/30 of the prior-bound width and the number of parallel tempering chains is increased to 10 to help offset the resulting decrease in birth/death acceptance.

Figure 2 shows that proposal schemes PS1 and PS2 are similar in efficiency, indicating that retaining all samples in parallel tempering and applying importance re-weighting does not provide a significant benefit in sampling efficiency. However, both PS1 and PS2 are much more efficient than PS3 and PS4. This indicates that drawing new parameters in birth steps from the prior is more efficient than the commonly-used approach of applying a focused (Gaussian) proposal density. Mathematical reasons for this are discussed in [3], but are beyond the scope of this proceedings paper.

References

[1] P. J. Green, Trans-dimensional Markov chain Monte Carlo in highly structured stochastic systems. *Oxford Statistical Sciences Series* (Oxford: Oxford University Press), 179–98, 2003.

[2] J. Dettmer, S. E. Dosso, and C. W. Holland, Trans-dimensional geoacoustic inversion, *J. Acoust. Soc. Am.* 128:3393–405, 2010.

[3] S. E. Dosso, J. Dettmer, G. A. W. Steininger, and C. W. Holland, Efficient trans-dimensional inversion for geoacoustic parameter estimation. *Inverse Problems*, 30:1-29, 2014.

[4] S. E. Dosso, C. W. Holland, and M. Sambridge, Parallel tempering in strongly nonlinear geoacoustic inversion, *J. Acoust. Soc. Am*. 132:3030–3040, 2012.

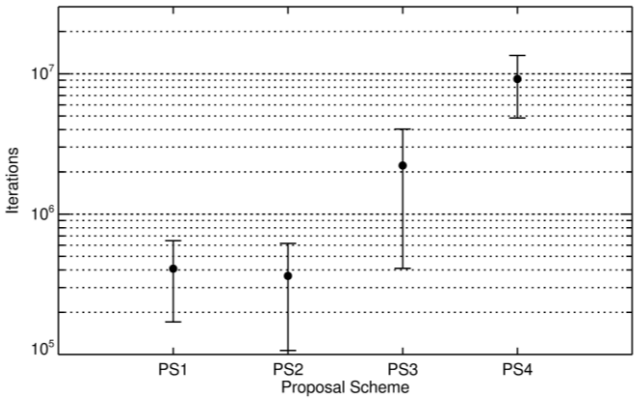


Figure 2: Comparison of the average number of iterations to convergence (with one standard-deviation error bars) for various proposal schemes (PSs) as described in the text.

A UNIVERSAL EXPRESSION FOR THE LOW-FREQUENCY ATTENUATION COEFFICIENT IN THE SURFACE DUCT

Rui Duan[1], N. Ross Chapman[†2], Kunde Yang[‡1] and Yuanliang Ma[1]

[1] Institute of Acoustic Engineering, Northwestern Polytechnical University, Xi'an, Shaanxi, P. R. China, 710072.

[2] School of Earth and Ocean Sciences, University of Victoria, P.O. Box 3055, Victoria, British Columbia V8W 3P6, Canada.

1 Introduction

The surface duct is an effective channel for sound propagation near the ocean surface. At low frequency the region outside the surface duct, the geometrical shadow zone, can be reliably insonified by energy leakage from the surface duct. Although features of the duct arrival (arrivals generated by the surface duct) at low frequency (around or lower than the cut-off frequency of the duct) have been concluded by Weston *et al.*[1] and Avilov *et al.*[2] using geometric diffraction theory, an expression for the above-mentioned attenuation related to the frequency and the duct thickness is still a grey area. The sensitivities of the attenuation to these parameters have not been quantitatively analyzed in real ocean environments. This paper develops a universal expression for the attenuation in the surface duct through simulations and comparison with experimental data from a deep water environment.

2 Attenuation Coefficient

The transmission loss of the duct arrival at low frequency beyond a transition range can be fitted accurately by cylindrical spreading plus a constant attenuation. Therefore, the energy can be expressed as

$$E(r,d,f)=\frac{E_0(d,f,r_0)r_0}{r}\exp(-\beta(f)r), \qquad (1)$$

where r, d, f denote the hydrophone range, the hydrophone depth, and the frequency, respectively. $E_0(d,f,r_0)$ is the energy at the transition range r_0. $\beta(f)$ is the attenuation coefficient. If the source is inside the surface duct, the energy within the transition range decays following the spherical law[1]. When the source is outside the surface duct, as energy is diffracted into the surface duct, the energy within the transition range is complicated. This paper focuses on the attenuation $\beta(f)$ beyond the transition range, so the value of $E_0(d,f,r_0)$ is not necessary. The transmission loss (in dB) from Eq. (1) is

$$\mathrm{TL}=SL(f)-C(d,f)+10\log_{10}(r)+\alpha(f)r, \qquad (2)$$

where $SL(f)$ is the source level and $C(d,f)$ denotes $10\log_{10}(E_0(d,f,r_0)r_0)$. The attenuation coefficient in dB, $\alpha(f)$, equals $10\beta(f)\log_{10}(e)$. Hereafter, transmission loss compensating for the cylindrical spreading loss (net loss) is used for estimating $\alpha(f)$.

† chapman@uvic.ca

‡ ykdzym@nwpu.edu.cn

3 Experimental Data and Model Calibration

In the experimental area, the ocean bottom was almost flat and the ocean depth is about 3450 m. A vertical uniform linear array of 18 hydrophones spaced at 4 m was deployed near the ocean surface. The topmost hydrophone was 30 m below the ocean surface. Explosive sound sources were deployed every 1 km when the ship was moving away from the vertical array to a distance of 40 km. The preset explosion depth was 300 m and the mass of TNT in the explosive charges was 1 kg. A Conductivity-Temperature-Depth (CTD) sensor was used to measure the sound speed profile (SSP) at the array location before the track. Besides, five expendable bathythermographs (XBT) were deployed at ranges 1.6, 10, 20, 30 and 40 km to measure the range-dependent SSPs along the track. The surface duct thickness is 37.13 m at the origin and increases to 44.8 m at 10 km. Between 10 and 20 km, the thickness is steady but it falls to 36.02 m at 30 km.

Figure 1 (a) shows the net losses of the duct arrival received at the top-most hydrophone (30 m). The net losses are calculated from 300 Hz to 1100 Hz every 100 Hz. The fluctuations of the net loss are attributed to the fluctuations of the source levels and the range-dependent environment. The net losses at mid frequencies (600 Hz to 1100 Hz) have inflection points around 29.5 km, where a great increase of the attenuation occurs. The inflection point is more distinct when the net losses at 900 Hz of all the hydrophones are plotted in Fig. 1 (b). The jump of the attenuation indicates an abrupt decrease of surface duct thickness. Note that the measured thickness is thinnest at 30 km. As the attenuations from 10 km to 30 km are roughly range independent in Fig. 1 (a), it is assumed that the surface duct thickness is invariant within this range interval. This assumption partly agrees with the measured thicknesses at 10 km (44.8 m) and 20 km (44.2 m).

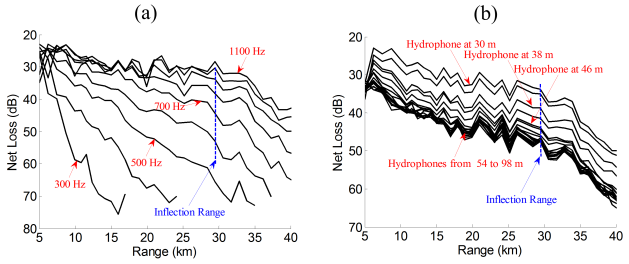


Figure 1: (a) Transmission loss versus range compensating for the cylindrical spreading loss (net loss) at the depth of 30 m. (b) Comparison between the net losses at 900 Hz of all the 18 hydrophones.

The black curve in Figure 2 denotes the average attenuation coefficients over all the 18 hydrophones. The net

losses from 10 km to 28.5 km are used to calculate the attenuation coefficients from 350 to 1150 Hz. The attenuation coefficient versus frequency is a curve of the exponential type. The decrease rate of the attenuation coefficient decreases with the increasing frequency. The red curve in Figure 2 displays the simulated results using the Range-dependent Acoustic Model (RAM) based on the parabolic equation model. The accuracy of the sound field calculated by the parabolic equation is sensitive to the depth step and range step, which are determined through a convergence test in this paper. The ocean environment for the simulation is range dependent and divided into three segments every 10 km. The SSPs at 0, 10 and 20 km are the ones from the measurements at the corresponding ranges. The net losses from 10 to 30 km are used to calculate the attenuation coefficients. In general, the simulations agree well with the measurements except that the measurements present fluctuation.

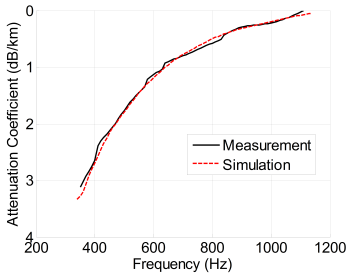


Figure 2: Extracted attenuations from experimental data in comparison with the simulated values using the range-dependent SSPs obtained in the experiment.

4 Universal expression for the attenuation coefficient

In our simulations, the sound speed gradient in the surface duct is fixed to be the hydrostatic one, 0.0167 s-1, and the sound speed below the surface duct is taken from the measurement at 10 km. Figure 3 (a) shows the frequency dependence of the attenuation coefficient, which is calculated for different surface duct thicknesses from 30 to 80 m. The attenuation coefficient is cut off at the critical frequency of the corresponding thickness except for the 30 m duct, whose critical frequency (about 2200 Hz) is above the maximum frequency of our simulation. The attenuation is very sensitive to the surface duct thickness. For example, the attenuation at 1 kHz for the 30 m duct is about twice that for the 35 duct. Note the cut-off frequency is proportional to the thickness to the power -1.5 in the bilinear SSP[3] and this power can change in real environments. Since the attenuation is similar to the cut-off frequency in terms of representing the ability of trapping energy, it is assumed that the attenuation coefficient is proportional to the thickness to the power b. Then the frequency f at the thickness h is scaled by $F = fh^b / 1000$, where F denotes the scaled frequency with the unit $kHz{\cdot}m^b$. Finally, the sum of squares of deviations (SSD) of all the attenuation coefficients which are plotted versus the scaled frequency is calculated at different b. It turns out that when b equals 1.77 the SSD is minimum. Fig. 3 (b) presents the attenuation coefficients versus the scaled frequency with the power b 1.77. Every curve is shown both on the bottom plane and in the three-dimension space. At the bottom plane, all the curves converge to a curve of the parabolic type. Discrepancies between these curves demonstrate that the thickness dependence of the attenuation is more complicated than a single term. However, this single term represents the major relation between the thickness and the attenuation.

Next, the attenuation coefficient is assumed as

$$\alpha(F) = \gamma \exp(\eta F) \quad dB/km \tag{3}$$

where, γ and η are unknown constants. The method of least squares was used in a regression analysis to find the approximate solution of Eq. (3). γ and η are 14.8 (14.79, 14.91) and -0.005164 (-0.005176, -0.005151), respectively. The values in parentheses are the 95% confidence bounds. Thus, an expression for the attenuation coefficient related to the frequency and the duct thickness is

$$\alpha(f,h) = 14.8 \exp(-5.164 \times 10^{-6} fh^{1.77}) \quad \text{dB/km}. \tag{4}$$

The grey solid lines in Fig. 3 (a) and grey surface in Fig. 3 (c) denote the attenuation coefficients calculated from Eq. (4). In general, the regressed coefficients agree well with the simulated values.

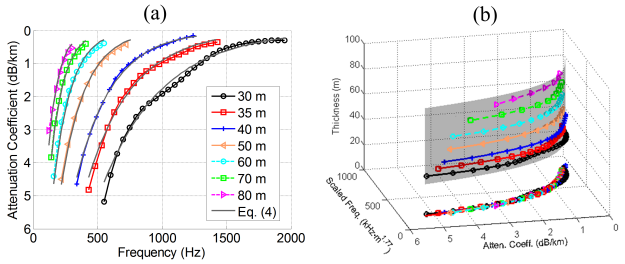


Figure 3: (a) Attenuations extracted from simulations at various surface-duct thicknesses. (b) The sum of squares of deviations (SSD) of all the attenuation coefficients, plotted versus the scaled frequency, at different power of the surface duct thickness.

5 Conclusion

This paper presented a universal expression for the attenuation coefficient of the duct-related energy at low frequency. A scaled frequency was proposed to combine the frequency and surface duct thickness. Then the attenuation was assumed to be exponentially dependent on the scaled frequency and a regression equation was obtained. The calculated attenuation agrees with the simulated value and the experimental measurement. This expression can be directed used in practice after the surface duct thickness has been acquired.

References

[1] D. Weston, C. Esmond, and A. Ferris, The duct leakage relation for the surface sound channel. J. Acoust. Soc. Am., 89: 156-164, 1991.

[2] K. Avilov, A. Mikryukov, and O. Popov, The Effect of the Subsurface Channel on the Insonification Rate of a Shadow Zone in a Deep Ocean. Acoustical Physics, 48: 656, 2002.

[3] F. M. Labianca, Normal modes, virtual modes, and alternative representations in the theory of surface-duct sound propagation. J. Acoust. Soc. Am., 53: 1137, 1973.

A BAFFIN BAY ACOUSTIC COMMUNICATION AND NAVIGATION SYSTEM

Eric Rehm [1], Walid Baccari [1] and Marcel Babin [1]
[1]Takuvik Joint International Laboratory, Laval University (Canada) - CNRS (France), Département de biologie et Québec-Océan, Université Laval, Québec, Québec, Canada

1 Introduction

In open water, autonomous floats and gliders can periodically surface and use GPS to determine their position. Iridium satellite communication can be used to send data back to land-based ocean data centers for analysis, research and operational decision-making. However, neither GPS nor satellite data communication systems function under the water. We propose a Baffin Bay Acoustic Navigation and Communication System (BBANC), which will use broadband low-frequency sources to provide basin-wide coverage of RAFOS-style signals for acoustic positioning of underwater assets. Low-frequency acoustic receivers would be collocated with the sources to enable ocean acoustic tomography for long-term study of heat content and currents. Finally, passive acoustic listening systems would enable the study of marine mammal communication and ambient noise from ships, sonar, ocean-based resource exploitation and ice dynamics, as well as gating acoustic source operation in the presence of marine mammals.

2 Motivation

Since 2002, the international Argo program has deployed a fleet of profiling floats, currently numbering over 3800, each measuring the temperature and salinity of the upper 2000 m of the ocean for up to 4 years. Over 30 countries, including Canada, have contributed to this program, yielding sustained real-time observations of the oceans that have revolutionized studies of global ocean circulation and the development and validation of climate models. Notice in Fig. 1**Error! Reference source not found.** that there are no Argo floats in the Arctic Ocean or Baffin Bay (red ellipse).

In open water, floats and gliders can periodically surface and use GPS to determine their position. Iridium satellite communication is used to send data back to land-based ocean data centers for analysis, research and operational decision-making. However, autonomous operation in the Arctic is not currently possible because ice cover prevents floats and gliders from positioning their measurements as well as relaying data to shore for long periods; neither GPS nor satellite data communications are possible under the water. As a result our current set of scientific observations of the Canadian Arctic is biased towards spring and summer, when it is navigable by ships and when autonomous platforms can safely surface. **To truly comprehend winter Arctic waters in the manner that has been achieved in the world's more temperate oceans, floats and gliders require additional capability for underwater geolocation and periodic communication.**

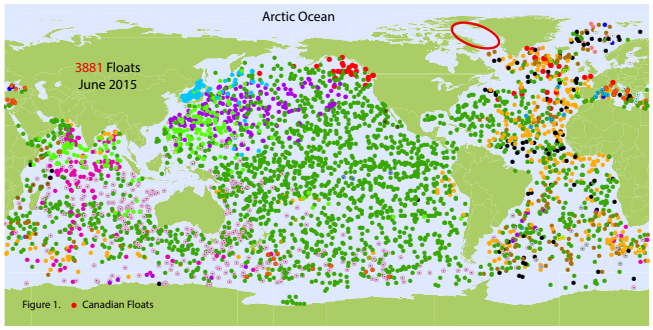


Figure 1 Argo floats.

Underwater geolocation and communication are possible using underwater sound. Refraction of sound within distinct ocean layers defined by their temperature, salinity and depth enables long distance propagation of sound deep in the ocean. RAFOS underwater positioning technology was developed in the 1980's, allowing autonomous scientific platforms to triangulate their position by listening to three or more underwater acoustic sources strategically-placed in these sound channels [1]. Since 2002, RAFOS-enabled platforms have been deployed in open and ice-covered high-latitude regions providing acoustic geolocation.

When ice-covered, Arctic waters poses special challenges for underwater acoustics due to the existence of a shallow sound channel with a highly reflective and scattering ice cover. This ice cover, along with the ambient noise created by ice formation, deformation, cracking and ridging, limits the maximum range for acoustic signals in the surface channel to a distance significantly less than that for signals in the deep acoustic channels. To meet these unique challenges of Arctic geolocation and communication, an international team of acousticians, autonomous platform developers, high-latitude oceanographers, and marine mammal researchers gathered at University of Washington in January 2005 for an NSF-sponsored "Acoustic Navigation and Communication for High-latitude Research" (ANCHOR) workshop [2]. They established international collaboration and developed an overarching system specification for the Arctic Ocean.

Following the workshop, several regional Arctic deployments of underwater acoustic positioning (Davis Strait, Fram Strait, Beaufort Sea) were achieved by a small number of partner nation efforts. Since 2006, UW-APL (USA) has completed 14 Seaglider missions across Davis Strait (~290 km) using acoustic geolocation [3], amounting to 1300 days of all-season operation (orange dots, Fig. 2). Since 2010, AWI (Germany) has enabled under-ice float

and glider missions with acoustic geolocation as part of the Fram Strait Observatory; acoustic sources are separated by 200-300 km [4]. **A permanent installation of moored acoustic sources for Arctic under-ice geolocation and communication on the scale of Baffin Bay (≥ 500 km source separation) has yet to be realized.**

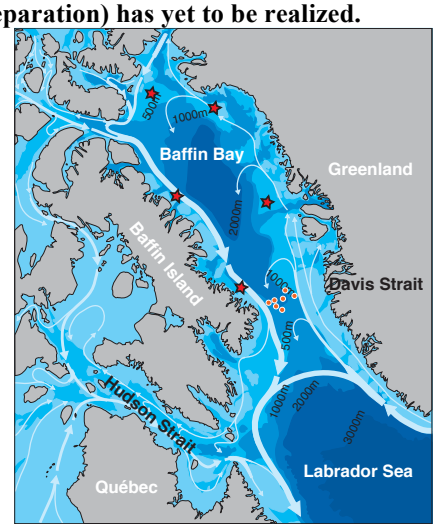


Figure 2 Baffin Bay and Davis Strait. White arrows indicate major circulation. Red stars indicate representative locations of possible acoustic sources, *only* for the purpose of comparing scale to existing 780 Hz Davis Strait RAFOS sources (orange dots). UW-APL gliders travel approximately East-West along a line of four moorings, with the three additional RAFOS moorings providing off-axis range measurements.

Characterization of acoustic propagation channels in Baffin Bay is incomplete, consisting of limited open-water results from Defence Canada (e.g., [5,6]) and marine mammal impact studies of shallow (<10 m) seismic air gun arrays (e.g., [7]) and core drilling operations for oil/gas exploration. Ice scattering losses are not well known for the largely first-year Arctic sea ice conditions of Baffin Bay [8].

3 Project Description

Our project consists of a feasibility study of a Baffin Bay Acoustic Navigation and Communication (**BBANC**) system, addressing the ANCHOR report recommendation of a pilot system "significantly larger" than current Arctic acoustic geolocation and navigation efforts. The benefits of this proposed system would be:

1. **Geolocation / "Underwater GPS":** Year-round, 24x7* underwater geolocation capability for autonomous platforms, supporting the study of the "unseen" winter arctic waters and springtime ice edge blooms. (*Gaps are possible due to maintenance and time and space outages or reduced duty cycle during whale migration, if necessary.)
2. **Data Communication:** Low data rate messaging capability for RAFOS system synchronization and autonomous platform command and control.
3. **Physical Oceanography:** Using co-sited acoustic receivers [9], acoustic tomography can "take the temperature" of Baffin Bay in less than one hour and repeat the measurements weekly on a year-round basis, enabling climate change research.
4. **Passive Acoustic Listening:** Passive acoustic recorders would enable the study of marine mammal communication and ambient noise from ships, sonar, resource exploitation, as well as wind and ice dynamics, as well as gating acoustic source operation in the presence of marine mammals.

The BBANC feasibility study will carry out underwater acoustic propagation and design modeling and simulation under partial and complete ice cover, exploring acoustic source type, location, frequency and transmission loss for basin-scale operation. Accomplishments to date include a large hydrographic database support range-dependent acoustic modeling across selected Baffin Bay sections. Ice cover thickness statistics along these sections has been derived from passive (SMOS) and active (CryoSat2) microwave satellite measurements of sea ice freeboard. An initial assessment of impacts on marine mammal communication will also be completed [10]. An international scientific network of underwater acoustics experts, marine mammal experts, and potential users is being established. We expect this study to lead to a field measurement program to verify and improve acoustic propagation modeling in an ice-covered Baffin Bay prior to implementation.

References

[1]Rossby et al. "The RAFOS system." Journal of Atmospheric and Oceanic Technology 3.4 (1986).

[2]"Acoustic Navigation and Communication for High-Latitude Ocean Research Workshop Report." University of Washington Applied Physics Laboratory (2008). http://anchor.apl.washington.edu

[3]Lee et al. "Long-Endurance, Ice-capable Autonomous Seagliders." AGU Fall Meeting Abstracts. Vol. 1 (2012).

[4]ACOBAR: Acoustic Technology for Observing the Interior of the Arctic Ocean, http://acobar.nersc.no/

[5]Leggat et al. "LNG Carrier Underwater Noise Study for Baffin Bay", (Report DREA/R/81/3) Defence Research Establishment Atlantic (1981).

[6]Desharnais. "Bottom loss in areas with ice-rafted sediments." Canadian Acoustics, 19.5 (1991).

[7]Wisniewska et al. Propagation of airgun pulses in Baffin Bay 2012. Aarhus University, DCE-Danish Centre for Environment and Energy, (2014).

[8]Bourke et al. "Sea ice thickness distribution in the Arctic Ocean." Cold Reg Sci Tech 13.3 (1987).

[9]Mikhalevsky. "Measuring heat in the Arctic Ocean with acoustic thermometry." Proceedings Underwater Acoustics and Measurements Conference, Kos, Greece (2011).

[10]Clark et al. "Acoustic masking in marine ecosystems: intuitions, analysis, and implication." Marine Ecology Progress Series 395 (2009).

MODELLING OF SEAFLOOR REVERBERATION IN NORTHERN GULF OF MEXICO SANDY SEDIMENT

Shannon Steele *[1] **and Sean Pecknold** †[2]
[1]Dalhousie University and Defence Research Development Canada-Atlantic
[2]Defence Research Development Canada-Atlanticc

1 Introduction

Underwater acoustics is an important tool for detecting and classifying targets such as submarines, mines, and fauna. Backscatter from the seafloor degrades the acoustic signal received from targets and thus modelling seafloor backscatter is an important component of active sonar. In this paper seabed roughness and cone pentrometer data taken during the TREX2013 sea trial will be used to produce a model of backscatter from sediment. The TREX sea trial was conducted in the Northern Gulf of Mexico, just off the coast of Panama City, Florida. The data was collected along two different tracks known as the main track(Figure 1) and the clutter track, which runs approximately perpendicular to the to the main track. The study area is composed of fine to medium grained quartz sand with a low carbonate content and patches of fine-grained sediment(mud and silt) [1,2]. The bathymetry of the TREX study area is characterized by large-scale north-south-trending sand ripples with amplitudes ranging from 1–3 m and wavelengths between 0.1 to 0.5 km [1, 3]. At high resolutions the acoustic scattering is considered nearly isotropic, however, on a broad scale it appears to show both gradual and abrupt changes in backscatter strength [4]. At high frequencies backscatter in the study area is mainly controlled by roughness spectra, however, at low and medium frequencies it is expected that volume scattering will also play a role in sediment backscatter. The most prominent volume scatterers in the study area are likely to be shell hash, hard packed sand, and other carbonate scatterers [5].

2 Methods

2.1 Roughness Spectra

The one dimensional roughness spectra for the multiple different sources of bathymetry collected during TREX were calculated in both the along track and across track directions. For each line a peridogram was calculated and then in each direction the peridograms were averaged. To test for isotropy a 2D FFT was also calculated.Two important parameters for inputs into backscattering models are spectral strength(ω_2) and spectral slope($-\gamma_2$).This can be modelled by the power law, which can be expressed as $W(\vec{k}) = \omega_2 k^{-\gamma_2}$ [6]. Since, in this study, the roughness was found to be roughly isotropic the spectral strength and slope can be determined from the 1D spectra, which is modelled as a power law of spectral strength (ω_1) and slope($-\gamma_1$). It was assumed that the measured roughness scaling continued to scales that affect the scattering.

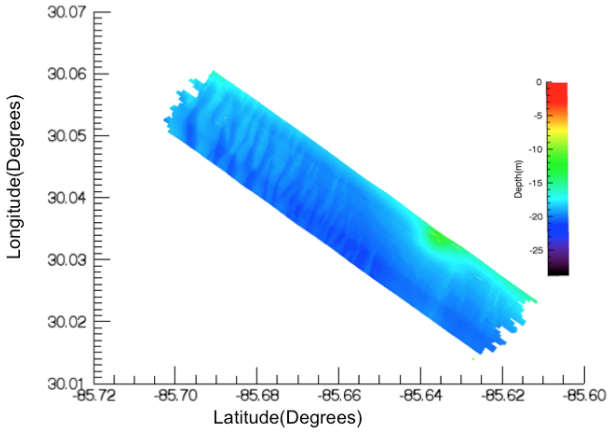


Figure 1: Image of main track bathymetry at 1m resolution

2.2 Scattering Model

The two-dimensional roughness spectra and geoacoustic parameters measured by the cone pentrometer were used to model acoustic backscatter as a function of Grazing Angle. Methods described in [7] were used to empirically estimate sound speed ratio(υ), density ratio(ρ), loss parameter(δ), and volume scattering parameter(σ_2) based on the grain sizes determined by the cone penetrometer. A composite roughness approximation (CRA) model [6,7] was used to determine the backscattering strength. In the CRA backscattering strength is defined as the product of roughness scattering and volume scattering.

3 Results

3.1 Roughness Spectra

For all sources of bathymetry the 2D FFTs show roughly isotropic spectral characteristics at the scale of 75m or less, and thus in terms of high to mid frequency acoustics the sediment can be considered isotropic. Shown in Figure 2 is the average one dimensional spectra across the north/south lines, east/west lines, and across all lines. The power law fit obtained from the 1D spectra is $\omega_1 = 5.716^{-8} m^{3-\gamma_1}$ and $\gamma_1 = 1.454$. The 2D spectral parameters were calculated to be $\omega_2 = 4.464^{-7} m^{4-\gamma_2}$ and $\gamma_2 = 2.454$.

3.2 Backscatter Model

The following geoacoustic parameters were obtained from the cone penetrometer data, and used as inputs into the scattering model : υ=1.23, ρ=1.46, σ_2=0.002, and δ=0.016. The CRA backscatter model has been plotted as a function of grazing angle along with Lambert's Law and a Lambert's Law

fit for sandy sediment (μ=-20.2) for comparison. Between 30° and 90° the model appears to be most predominantly controlled by volume scattering. At grazing angles below approximately 18° roughness scattering is the dominant control on scattering strength. In between 18° and 30° both types of scattering appear to contribute equally to the scattering strength. Near the critical angle (35.7°) the CRA demonstrates a steep slope, causing the backscattering strength to be much lower than what is predicted by Lambert's Law. Below 15° the CRA plot has a much smaller slope than predicted by Lambert's Law. The CRA predicts significantly lower backscatter than the sandy sediment value fit with μ=-20.2 measured previously at other sites.

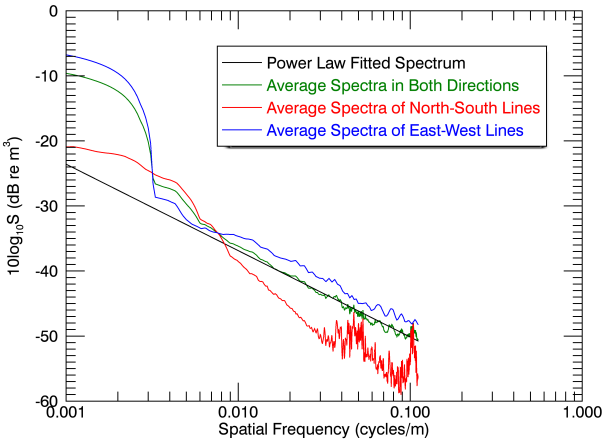


Figure 2: Average spectra and Power-Law fitted spectrum to average Spectra from both directions

4 Discussion

In the CRA the volume scattering coefficient is determined either empirically based on measured backscattering data or by estimation based on mean grain size. In this study, the volume parameter was determined based on mean grain size, however in many cases sediment grain size is not an accurate predictor of volume scattering [6, 7].In addition to this, the model also fails to consider volume scattering due to heterogeneity with depth as well as the presence of discrete scatterers [7]. Therefore it is possible that the CRA model predicts significantly lower backscattering strength than expected because volume scattering is not being fully accounted for.

5 Conclusions

The CRA model predicts that the sediment in the study area is not an ideal diffusely reflecting surface. Below the critical angle the CRA model predicts grazing angle dependency that is far different than what would be predicted by Lambert's Law. The significant difference in scattering strength predicted by the CRA model and the Lambert's law fit to sandy sediment indicates that predicting volume scattering with sediment grain size may not always be an effective way to model backscattering. Thus more research needs to be conducted to test and further develop models that account for volume heterogeneity (both at and below the sediment water interface) and discreet scatterers. In future work sub-bottom profiler data collected during TREX2013 will be used in order to quantify volume scattering.

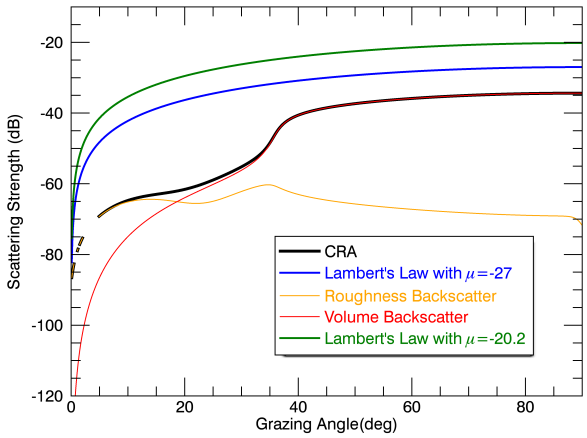


Figure 3: Plot comparing scattering strength as function of grazing angle for : the CRA model of TREX study area, Lambert's Law(μ=-27) and Lambert's Law fit for sandy sediment(μ=-20.2). Also shown is the seabed roughness scattering strength computed for the CRA and the volume scattering strength computed for the CRA

References

[1] Peter. Fleischer, William B. Sawyer, Hannelore Fiedler, and Ingo H. Stender. Spatial and temporal variability of a coarse-sand anomaly on a sandy inner shelf, northeastern gulf of mexico. *Geo-Marine Letters*, 16(3) :266–272, 1996.

[2] Larry J. Doyle and Thomas N. Sparks. Sediments of the Mississippi, Alabama, and Florida (MAFLA) continental shelf. *Journal of Sedimentary Petrology*, 50(3) :905–915, 1980.

[3] Steven J. Parker, Albert W. Shultz, and William W. Schroeder. Sediment Characteristics and Seafloor Topography of a Palimpsest Shelf, Mississippi-Alabama Continental Shelf. *SEPM*, 48, 1992.

[4] Kenneth S. Davis, Niall C. Slowey, Ingo H. Stender, Hannelore Fiedler, William R. Bryant, and Gunther Fechner. Acoustic backscatter and sediment textural properties of inner shelf sands, northeastern gulf of mexico. *Geo-Marine Letters*, 16(3) :273–278, 1996.

[5] John A. Goff. Reconnaissance marine geophysical survey for the shallow water acoustics program. 2013.

[6] D. Jackson and M. Richardson. *High-Frequency Seafloor Acoustics*. The Underwater Acoustics Series. Springer New York, 2007.

[7] APL-UW High-Frequency Ocean Environmental Acoustic Models Handbook,. Technical report, Dod, Defense Technical Information Center, 1995.

Abstracts for Presentations without Proceedings Paper
Résumés des communications sans article

Multistatic Sonar Operator Visualization Development Requirements

Mark A. Gammon, Christopher Strode

The operational community has been using multistatic sonar for numerous years yet the development of applications for operator visualization of multistatic sonar performance is lacking. Such applications may enhance the operator's ability to visualize the multistatic sonar's real time performance to optimize sensor employment. One of the key issues is the transition from legacy sonar monostatic performance acoustic prediction to multistatic sonar acoustic prediction. While many such applications have been developed, the majority of these pertain to the analysis of multistatic sonar and not the direct operator visualization of that performance. In order to be able to provide a completely effective system from the standpoint of both sensors and users, this challenge must be addressed. One of the key parameters is the ocean environment, in particular realistic modelling for the purpose of sonar performance prediction, yet numerous other factors will influence actual operations. Techniques have recently been field tested to provide the operator with Operational Analysis (OA) tools and advice, and further work in the area of near-real time Operator Visualization Development (OVD) is essential. A system engineering approach advocates an integrated system to provide the near-real time sonar performance prediction embedded in the sonar display. Such a system does not and will probably never fully allay the acoustic performance prediction problem given the uncertainty in acoustic analysis. However it must provide useful system performance information to the operator in order to optimize sensor employment. Hence, a new conceptual framework is required for Multistatic Sonar OVD (MSOVD).

The Canadian Acoustical Association - L'Association canadienne d'acoustique

CANADIAN ACOUSTICS TELEGRAM ANNOUNCEMENTS - ANNONCES TÉLÉGRAPHIQUES DE L'ACOUSTIQUE CANADIENNE

Tri-Agency Open Access Policy on Publications

The good news? Publication and copyright policies of Canadian Acoustics journal are fully compliant with these new rules! That's another good reason for reseachers to publish in Canadian Acoustics!

The Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Social Sciences and Humanities Research Council of Canada (SSHRC) are federal granting agencies that promote and support research, research training and innovation within Canada. As publicly funded organizations, the Agencies have a fundamental interest in promoting the availability of findings that result from the research they fund, including research publications and data, to the widest possible audience, and at the earliest possible opportunity. Societal advancement is made possible through widespread and barrier-free access to cutting-edge research and knowledge, enabling researchers, scholars, clinicians, policymakers, private sector and not-for-profit organizations and the public to use and build on this knowledge.According to a new policy, all grant recipients that were funded in whole or in part by NSERC or SSHRC for grants awarded May 1, 2015 and onwar (Januray 1, 2008 for CIHR) are required to ensure that any peer-reviewed journal publications arising from Agency-supported research are freely accessible within 12 months of publication.

May 28th 2015

Editorial Position Available for Canadian Acoustics

Our current deputy editor, Prof. Josée Lagacé, will be stepping out this summer and the position will become available September 1st, 2015. Many thanks from all the editorial board for her help during the 4 years spent on the editorial team.

The position of Deputy Editor of the Canadian Acoustics journal will become available on September 1st and represents a great opportunity for an academic or anyone interested in the editorial process of a peer-reviewed journal.The duties of the Deputy Editor are:to establish, together with the Editor-in-Chief, the editorial board of the journal,to organize annually, together with the editorial board of the journal, special issues of the journal,to ensure the fair and balanced language representation within the journal,to assist the Editor-in-Chief in the editorial tasks, especially in the enforcement of the Conflict of Interest journal policy.If you or someone you know may be interested, simply contact the Editor (jcaa@caa-aca.ca).

June 18th 2015

Acoustics Week in Canada Conference

ACOUSTICS WEEK IN CANADA 2015 will be held October 7-9, 2015, Halifax, Nova Scotia

See last call for paper at the end of the June 2015 journal issue

June 18th 2015

Michael R. Stinson was elected President-Elect of the Acoustical Society of America (ASA)

The Acoustical Society of America (ASA) has elected officers and Executive Council members in 2015 including President-Elect Michael R. Stinson, an old time member of the Canadian Acoustical Association.

Michael R. Stinson was elected President-Elect. His term began on 22 May 2015 and he will automatically assum the office of President on 27 May 2016. Mike Stinson served as Principal Research Officer at the National Research Council of Canada, in Ottawa where he is now Researcher Emeritus. Mike was awarded a Ph.D. in Physics from Queen's University in Kingston, Ontario, Canada. His research activities have spanned a broadrange of technical activities including studies of the acoustics of the human ear canal and middle ear which have led to advances in hearing aid design. His recent work has looked at propagation of infrasonic noise from wind turbines.

August 5th 2015

Looking for a job in Acoustics?

There are many job offers listed on the website of the Canadian Acoustical Association!

You can see them online, under http://www.caa-aca.ca/jobs/

August 5th 2015

Canadian Acoustical Association
Association canadienne d'acoustique

Application for Membership

CAA membership is open to all individuals who have an interest in acoustics. Annual dues total $100.00 for individual members and $50.00 for student members. This includes a subscription to ***Canadian Acoustics***, the journal of the Association, which is published 4 times/year, and voting privileges at the Annual General Meeting.

Subscriptions to *Canadian Acoustics* or Sustaining Subscriptions

Subscriptions to ***Canadian Acoustics*** are available to companies and institutions at a cost of $100.00 per year. Many organizations choose to become benefactors of the CAA by contributing as Sustaining Subscribers, paying $475.00 per year (no voting privileges at AGM). The list of Sustaining Subscribers is published in each issue of ***Canadian Acoustics*** and on the CAA website.

Please note that online payments will be accepted at http://jcaa.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 $ 100.00
[ ] CAA Student Membership $ 50.00

Corporate Subscriptions (4 issues/yr)
[ ] $100 including mailing in Canada
[ ] $108 including mailing to USA,
[ ] $115 including International mailing

[ ] Sustaining Subscription $475.00 (4 issues/yr)

Please note that the preferred method of payment is by credit card, online at http://jcaa.caa-aca.ca

For individuals or organizations wishing to pay by check, please register online at *http://jcaa.caa-aca.ca* and then mail your check to:

Executive Secretary, The Canadian Acoustical Association, PO Box 74068, Ottawa, Ontario, K1M 2H9, Canada

Association Canadienne d'Acoustique
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 100.00$ pour les membres individuels, et de 50.00$ 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 100.00$. Des compagnies et établissements préfèrent souvent la cotisation de membre bienfaiteur, de 475.00$ 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.

Notez que tous les paiements électroniques sont acceptés en ligne http://jcaa.caa-aca.ca

Pour correspondance administrative et financière:

Nom / Organisation _____

Adresse _____

Ville/Province _____ Code postal _____Pays _____

Téléphone _____ Téléc. _____ Courriel _____

Adresse postale pour la revue Acoustique Canadienne

Nom / Organisation _____

Adresse _____

Ville/Province _____ Code postal _____Pays _____

Cocher vos champs d'intérêt: (maximum 3)

1. Acoustique architecturale
2. Génie acoustique / Contrôle du bruit
3. Acoustique physique / Ultrasons
4. Acoustique musicale / Électro-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 100.00 $
[ ] Membre étudiant(e) 50.00 $

Abonnement institutionnel

[ ] 100 $ à l'intérieur du Canada
[ ] 108 $ vers les États-Unis
[ ] 115 $ tout autre envoi international

[ ] Abonnement de soutien 475.00 $ (comprend l'abonnement à *L'acoustique Canadienne*)

Merci de noter que le moyen de paiement privilégie est le paiement par carte crédit en ligne à *http://jcaa.caa-aca.ca*

Pour les individus ou les organisations qui préféreraient payer par chèque, l'inscription se fait en ligne à *http://jcaa.caa-aca.ca* puis le chèque peut être envoyé à :

Secrétaire exécutif, Association canadienne d'acoustique, BP 74068, Ottawa, Ontario, K1M 2H9, Canada

Canadian Acoustical Association
Association canadienne d'acoustique

CANADIAN ACOUSTICAL ASSOCIATION
ASSOCIATION CANADIENNE D'ACOUSTIQUE

SUSTAINING SUBSCRIBERS - ABONNÉS DE SOUTIEN

The Canadian Acoustical Association gratefully acknowledges the financial assistance of the Sustaining Subscribers listed below. Their annual donations (of $475.00 or more) enable the journal to be distributed to all at a reasonable cost.
L'Association Canadienne d'Acoustique tient à témoigner sa reconnaissance à l'égard de ses Abonnés de Soutien en publiant ci-dessous leur nom et leur adresse. En amortissants les coûts de publication et de distribution, les dons annuels (de 475.00$ et plus) rendent le journal accessible à tous les membres.

Acoustec Inc.
J.G. Migneron - 418-496-6600
courrier@acoustec.qc.ca
acoustec.qc.ca

ACOUSTIKALAB Inc.
Jean Laporte - 514-692-1147
jlaporte@acoustikalab.com
ACOUSTIKALAB.com

AECOM
Alex Dundon -
alex.dundon@aecom.com
aecom.com

Aercoustics Engineering Ltd.
Nicholas Sylvestre-Williams - (416) 249-3361
NicholasS@aercoustics.com
aercoustics.com

AMEC
Mr. Frank Babic - 905.568.2929
frank.babic@amec.com
amec.com

Bruel & Kjaer North America Inc.
Andrew Khoury - 514-695-8225
andrew.khoury@bksv.com
bksv.com

Dalimar Instruments Inc
Daniel Larose - 514-424-0033
daniel@dalimar.ca
www.dalimar.ca

dB Noise Reduction
Mehrzad Salkhordeh - 519-651-3330 x 220
mehrzad@dbnoisereduction.com
dbnoisereduction.com

FFA Consultants in Acoustics and Noise Control
Clifford Faszer - 403.508.4996
info@ffaacoustics.com
ffaacoustics.com

H.L. Blachford Ltd.
Duncan Spence - 905-823-3200
amsales@blachford.ca
blachford.ca

Hatch Associates Ltd.
Tim Kelsall - 905-403-3932
tkelsall@hatch.ca
hatch.ca

HGC Engineering Ltd.
Bill Gastmeier -
gastmeier@rogers.com
hgcengineering.com

Integral DX Engineering Ltd.
Greg Clunis - 613-761-1565
greg@integraldxengineering.ca
integraldxengineering.ca

Jade Acoustics Inc.
Ms. Dalila Giusti - 905-660-2444
dalila@jadeacoustics.com
jadeacoustics.com

MJM Conseillers en Acoustique Inc.
Michel Morin - 514-737-9811
mmorin@mjm.qc.ca
mjm.qc.ca

Peutz & Associés
Marc Asselineau - 33-1-45230500
m.asselineau@peutz.fr
peutz.fr

Pliteq Inc.
Wil Byrick - 416-449-0049
wbyrick@pliteq.com
pliteq.com

Pyrok Inc.
Howard Podolsky - 914-777-7770
mrpyrok@aol.com
pyrok.com

ROCKFON
Dr. Gary S. Madaras - 708.563.4548
gary.madaras@rockfon.com
rockfon.com

RWDI AIR Inc.
Mr. Peter VanDelden - 519-823-1311
peter.vandelden@rwdi.com
rwdi.com

Scantek Inc.
Steve Marshall - 1-410-290-7726
S.Marshall@scantekinc.com
scantekinc.com

Soft dB Inc.
André L'Espérance - 418-686-0993
info@softdb.com
softdb.com

Swallow Acoustic Consultants
John Swallow -
info@swallowacoustic.ca
swallowacoustic.ca

True Grit Consulting Ltd.
Ms Ina Chomyshyn - 807-626-5640
ina@tgcl.ca
tgcl.ca

Valcoustics Canada Ltd.
Al Lightstone - 905-764-5223
solutions@valcoustics.com
valcoustics.com

Vibra-Sonic Control
Mr. Chris N. Wolfe - 604-294-9495
chris@vibra-sonic.com
vibra-sonic.com

Wakefield Acoustics Ltd.
Mr. Clair W. Wakefield - 250-370-9302
clair@wakefieldacoustics.com
wakefieldacoustics.com

West Caldwell Calibration Labs
Mary Richard - 905-554-8005
wcclca@gmail.com
wccl.com