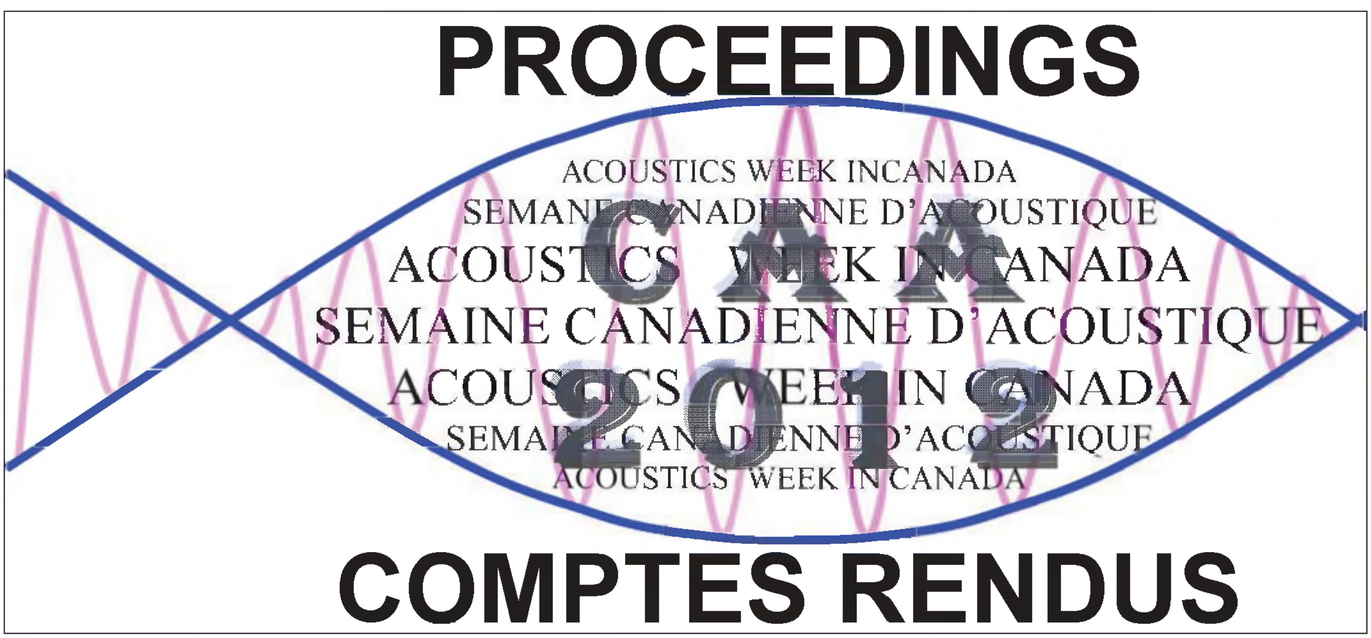
ISSN 0711-6659

canadian acoustics
acoustique canadienne

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

SEPTEMBER 2012 SEPTEMBRE 2012
Volume 40 -- Number 3 Volume 40 -- Numéro 3



canadian acoustics

acoustique canadienne

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

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

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

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

CANADIAN ACOUSTICS is published four times a year - in March, June, September and December. The deadline for submission of material is the first day of the month preceeding the issue month. Copyright on articles is held by the author(s), who should be contacted regarding reproduction. Annual subscription: $35 (student); $80 (individual, institution); $350 (sustaining - see back cover). Back issues (when available) may be obtained from the CAA Secretary - price $20 including postage. Advertisement prices: $350 (full page- $1200 for four issues); $200 (half page - $700 for four issues); $150 (quarter page - $500 for four issues). Contact the Associate Editor (advertising) to place advertisements. Canadian Publication Mail Product Sales Agreement No. 0557188.

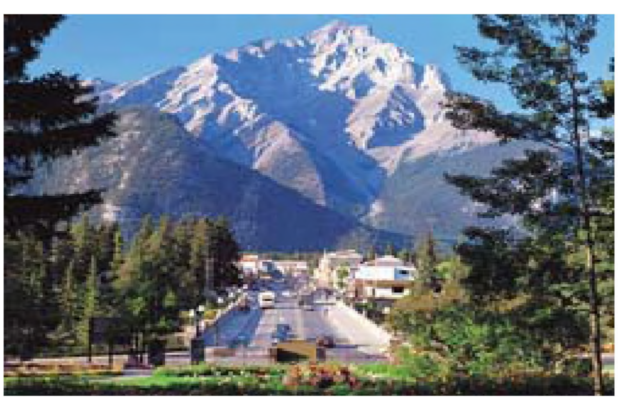
ACOUSTIQUE CANADIENNE est publiée quatre fois par année - en mars, juin, septembre et décembre. La date de tombée pour la soumission de matériel est fixée au premier jour du mois précédant la publication d'un numéro donné. Les droits d'auteur d'un article appartiennent à (aux) auteur(s). Toute demande de reproduction doit leur être acheminée. Abonnement annuel: $35 (étudiant); $80 (individuel, société); $350 (soutien - voir la couverture arrière). D'anciens numéros (non-épuisés) peuvent être obtenus du Secrétaire de l'ACA - prix: $20 (affranchissement inclus). Prix d'annonces publicitaires: $350 (page pleine - $1200 pour quatre publications); $200 (demi page - $700 pour quatre publications); $150 (quart de page - $500 pour quatre publications). Contacter le rédacteur associé (publicité) afin de placer des annonces. Société canadienne des postes - Envois de publications canadiennes - Numéro de convention 0557188.


Acoustics Week in Canada
Semaine Canadienne d'Acoustique

Banff Park Lodge
Banff AB, 10-12 October/octobre 2012

Organizing Committee / Comité d'organisation

Conference Chair / Président: *Stan Dosso*
Technical Chair / Directeur Technique: *Roberto Racca*
Accounting and Registration /Trésorerie et inscription: *Clair Wakefield*
Accounting and Registration /Trésorerie et inscription: *Lori Robson*
Exhibit and Sponsors / Exposition et Commandite: *Lisa Cooper*
Website / Site internet: *Brendan Rideout*
Students Awards / Etudiante Prix: *Michael Wilmut*




Sponsors

Exhibitors

GUEST EDITORIAL / EDITORIAL INVITÉ

On behalf of the Conference Organizing Committee, it is our pleasure to invite/welcome you to beautiful Banff, Alberta, for Acoustics Week in Canada 2012, the annual conference of the Canadian Acoustical Association (Wednesday to Friday, Oct. 10-12). The exceptional Rocky Mountain setting of the town, within Banff National Park (a UNESCO World Heritage Site), will make this a unique conference for the CAA. In keeping with the mountain grandeur of Banff, the Conference theme is Sound and the Natural World.

The meeting will be held at the Banff Park Lodge, which offers state-of-the-art conference facilities in a mountain-lodge ambience and a quiet location on the banks of the glacier-fed Bow River, yet just two blocks from the Banff dining/shopping/entertainment district.

The technical program includes close to 100 contributed talks covering all areas of acoustics and running from Wednesday morning to Friday noon (the technical content of the conference is summarized in the pages of this Proceedings issue). We also have great keynote talks to start each day, including Colleen Reichmuth, Adventures in bioacoustics: Sound reception, production, and perception in amphibious marine mammals; Frank Russo, From sound waves to brain waves and points in between; and Christian Giguère and Chantal Laroche, Auditory awareness and fitness-for-duty in the noisy workplace. The Acoustical Standards Committee Meeting will be held on Wednesday, and Thursday will include an all-day Exhibition of acoustical equipment, products, and services, with nearly 20 companies taking part.

A Welcome Reception will be held Wednesday evening at the Banff Ave. Brewing Company, featuring drinks and snacks, brewery tours, a games room, and comfortable space to mingle. The Conference Banquet will be held Thursday evening at the Banff Park Lodge and include fine dining with good company and the CAA Awards Ceremony.

We thank the members of the Organizing Committee for their dedicated and enthusiastic efforts, including Clair Wakefield and Lori Robson (Accounting and Registration), Lisa Cooper (Exhibits and Sponsors), Brendan Rideout (Web site), and Michael Wilmut (Student Awards and Travel).

We look forward to seeing you in Banff!

Stan Dosso, Conference Chair
University of Victoria

Roberto Racca, Technical Chair
JASCO Applied Sciences

Au nom du comité organisateur, nous sommes heureux de vous inviter/accueillir dans la magnifique ville de Banff, en Alberta, pour la semaine canadienne d'acoustique 2012, la conférence annuelle de l'Association Canadienne d'Acoustique (du mercredi au vendredi 10-12 octobre). L'emplacement de la ville parmi les exceptionnelles montagnes Rocheuses, dans le parc national Banff (un site du patrimoine mondial de l'UNESCO), rendront cette conférence unique pour l'ACA. En lien avec la splendeur naturelle de Banff, le thème de la conférence est Le son et le monde naturel.

La rencontre se déroulera au Banff Park Lodge, qui met à disposition des aménagements idéaux pour une conférence dans une atmosphère montagnarde et un endroit reposant au bord de la rivière Bow, mais à deux blocs seulement du quartier de Banff regroupant restaurants, magasins et divertissement.

Le programme technique inclut presqu'une centaine de présentations orales portant sur tous les domaines de l'acoustique et se tenant du mercredi matin au vendredi midi (le contenu technique de la conférence est résumé dans les pages de cette édition de Proceedings). De plus, des séances plénières prendront place à chaque début de journée, incluant Colleen Reichmuth, Adventures in bioacoustics: Sound reception, production, and perception in amphibious marine mammals; Frank Russo, From sound waves to brain waves and points in between; et Christian Giguère et Chantal Laroche, Auditory awareness and fitness-for-duty in the noisy workplace. La réunion du Comité des Standards en Acoustique se déroulera le mercredi, et presque 20 compagnies tiendront une exposition d'équipement, de produits et de services en acoustique durant toute la journée du jeudi.

Une réception d'accueil prendra place le mercredi soir à la brasserie Banff Ave. Brewing Company, et mettra à disposition des boissons et collations, des visites guidées de la brasserie, une salle de jeux, et un endroit propice aux conversations. Le banquet de la conférence se tiendra le jeudi soir au Banff Park Lodge et inclura un repas gastronomique en bonne compagnie, ainsi que la cérémonie de remise des prix ACA.

Nous tenons à remercier les membres du comité organisateur pour tous leurs efforts et leur enthousiasme, incluant Clair Wakefield et Lori Robson (comptabilité et inscription), Lisa Cooper (exposition et commandites), Brendan Rideout (site internet) et Michael Wilmut (prix étudiants et subventions de voyage).

Nous nous réjouissons de vous voir à Banff!

Stan Dosso, Président de la conférence
University of Victoria

Roberto Racca, Directeur scientifique
JASCO Applied Sciences

Table of Contents/Tables des matières

DAY ONE	WEDNESDAY 10 OCT 2012 REGISTRATION OPEN 8:30 Alpine Meadows		
9:00–9:10	Welcome		
9:10–10:10	Keynote Talk: Colleen Reichmuth (Room: Summit) *Adventures in bioacoustics: Sound reception, production & perception in amphibious marine mammals*		
10:10–10:30	COFFEE BREAK		
	ROOM: SUMMIT	ROOM: ASSINIBOINE	ROOM: CASTLE
10:30–12:10	*SPEECH COMMUNICATION*	*BIOACOUSTICS*	*NOISE & NOISE CONTROL*
12:10–1:20	LUNCH (Glacier Salon)		
1:20–3:00	*SPEECH COMMUNICATION*	*BIOACOUSTICS / UNDERWATER ACOUSTICS*	*NOISE & NOISE CONTROL*
3:00–3:20	COFFEE BREAK		
3:20–4:20	*SPEECH COMMUNICATION*	*UNDERWATER ACOUSTICS*	*NOISE & NOISE CONTROL*
4:30–9:30	Acoustics Standards Committee Meeting (All welcome)		
6:00–8:00	WELCOME RECEPTION (Banff Ave. Brewing Company)		

DAY TWO	THURSDAY 11 OCT 2012 EXHIBITION: 10:00–5:00 Alpine Meadows		
9:00–10:00	Keynote Talk: Frank Russo (Room: Summit) *From sound waves to brain waves and points in between*		
10:00–10:40	COFFEE BREAK		
	ROOM: SUMMIT	ROOM: ASSINIBOINE	ROOM: CASTLE
10:40–12:20	*SPEECH COMMUNICATION*	*UNDERWATER ACOUSTICS*	*NOISE & NOISE CONTROL / MUSICAL ACOUSTICS*
12:20–1:20	LUNCH (Glacier Salon)		
1:20–2:40	*SPEECH COMMUNICATION*	*UNDERWATER ACOUSTICS*	*ARCHITECTURAL ACOUSTICS*
2:40–3:20	COFFEE BREAK		
3:20–4:40		*HEARING SCIENCES*	*ARCHITECTURAL ACOUSTICS*
5:00–6:00	CAA ANNUAL GENERAL MEETING (Room: Castle)		
6:30–9:00	AWC 2010 BANQUET & AWARDS (Room: Summit)		

DAY THREE	FRIDAY 12 OCT 2012		
9:00–10:00	Keynote Talk: Christian Giguère and Chantal Laroche (Room: Summit) *Auditory awareness and fitness-for-duty in the noisy workplace*		
10:00–10:20	COFFEE BREAK		
	ROOM: SUMMIT	ROOM: ASSINIBOINE	ROOM: CASTLE
10:20–12:20	*PHYSIO- & PSYCHO-ACOUSTICS / VIBRATION, ENGINEERING & PHYSICAL ACOUSTICS*	*HEARING SCIENCES*	*ARCHITECTURAL ACOUSTICS / ACOUSTICAL STANDARDS*
12:20–13:30	LUNCH & STUDENT PRESENTATION AWARDS (Glacier Salon)		

PERCEPTION OF SPEAKER AGE IN CHILDREN'S VOICES

Peter F. Assmann[1], Terrance M. Nearey[2]
[1]School of Behavioral and Brain Sciences, The University of Texas at Dallas, Richardson TX 75083-0688 USA
[2] Department of Linguistics, University of Alberta, Edmonton, Alberta, Canada T6G 2E7

1. INTRODUCTION

Children's speech differs from adult speech in several respects. Their vocal tracts are shorter, leading to higher formant frequencies, and their larynges are smaller, resulting in higher average fundamental frequency (f0). These properties are linked to phonetic properties of speech but also affect indexical properties (such as the age, sex, and size of the speaker). As part of a larger study to investigate the interaction between indexical and phonetic properties we presented a sample of speech sounds spoken by children ranging in age from 5 to 18 years to adult listeners and asked them to judge the age of the speaker.

1.1. Background

Previous studies have shown that listeners can judge the age of a speaker fairly accurately in adult voices (Ptacek and Sander, 1966), but few studies have examined children's voices. One exception is a recent study by Amir et al. (2012) who investigated age and gender recognition of a sample of Hebrew-speaking children between 8 and 18 years. Age recognition accuracy (±2 years) was fairly low (40% for sentences, 35% for isolated vowels) with the lowest performance for the oldest group where there was a tendency to systematically underestimate the perceived age in female voices. Gender recognition was fairly accurate (85% for sentences, 78% for vowels) but again showed lower accuracy for older girls compared to older boys.

1.2. Research questions

The aim of the present study was to replicate and extend the research reported by Amir et al. (2011) with a sample of American English speaking children to answer the following questions: (1) *Does knowing the sex of the speaker help determine their age*? We previously found (Assmann and Nearey, 2011) that providing information about the age of the speaker provides a small benefit, in some conditions, for determining whether the speaker is male or female. Here we ask if knowledge of the speaker's sex also provides a benefit when listeners are asked to judge the speaker's age. (2) *To what extent is perceived age influenced by context*? Previous studies (Hillenbrand and Clark, 2009; Assmann and Nearey, 2011) have shown that listeners can identify the sex of the speaker more accurately from sentences than from single syllables. Amir et al. (2011) found higher accuracy for the perception of age in children's voices based on complete sentences rather than isolated vowels. In the present study we compared isolated /hVd/ syllables with the same syllables embedded in a fixed carrier sentence ("Please say the word _____ again").

2. METHODS

Stimuli

The stimuli were recorded syllables and sentences drawn from a vowel database (Assmann et al., 2008) of 208 children ranging in age from 5 to 18 years. In the syllable condition, 140 speakers (5 boys and 5 girls at each age level) contributed 3 syllables (heed, hod, and who'd) for a total of 420 stimuli. In the sentence condition (tested with a separate group of listeners) a subset of 84 of speakers was included, for a total of 252 stimuli (3 boys and 3 girls at each age level, each speaking the same 3 syllables in a carrier sentence) to keep the experiment to a reasonable length.

Participants

Separate sets of 20 listeners completed the syllable and sentence conditions; of these, 10 were provided with age information prior to responding, and 10 were not. The listeners were undergraduate students at the University of Texas at Dallas, native speakers of American English with normal hearing who received experimental credits for their participation. Prior to the experiment they completed a hearing screen and a questionnaire to provide information along their age, sex, younger siblings and exposure to children's voices on a daily basis.

Procedure

Stimuli were presented monaurally using earphones with Tucker-Davis System 3 and RP2.1 hardware. All stimulus conditions were randomly interspersed. Listeners used a graphical slider to register their estimate of the speaker's age, after which they checked one of five buttons indicating their confidence level. The experiment was self-paced, with an optional break in the middle, and lasted about 50 minutes.

3. RESULTS

Age judgments showed a fairly close match to chronological age across the age range for boys, but progressively underestimated chronological age for older girls, with the discrepancy reaching nearly four years on average by age 18. This is consistent with findings reported by Amir et al. (2012). Informing listeners about the sex of the speaker reduced the discrepancy slightly, although this information did not lead to substantially improved age estimation. Figure 1 shows that perceived age judgments (pooled across talkers and listeners) were well fit by a quadratic function (R^2=0.93) although age judgments for boys' voices were close to a linear function of age. When /hVd/ syllables were presented in a carrier sentence, perceived age was closer to chronological age and the degree of underestimation in older girls' voices was reduced, though not eliminated entirely.

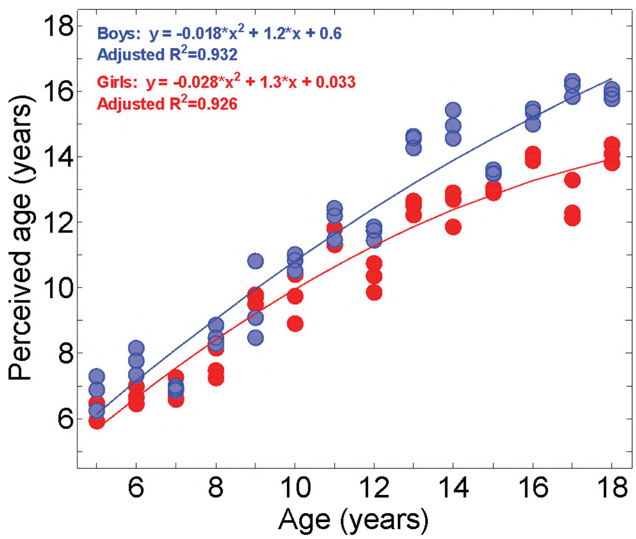


Figure 1. Perceived versus chronological age (/hVd/ syllables).

Accuracy of perceived age was quantified in terms of the absolute deviation of perceived age from chronological age. Analysis of variance revealed a significant effect of context, with more accurate age estimates from sentences compared to syllables, $F(1,40)=4.99$; $p<.05$. There was a significant interaction of context, sex and age, $F(13,520)=8.97$; $p<.01$, reflecting an increased benefit of context for older girls whose ages were often underestimated. Gender information did not lead to an overall improvement, but the interaction with sex, age and context was significant, $F(13,520)=2.92$; $p<.01$, with slightly improved estimates for some age/sex combinations when gender information and context were both available.

4. DISCUSSION AND CONCLUSIONS

The results provide further support for several findings reported by Amir et al. (2012). First, age estimates are more accurate for boys' voices than for those of girls. Second, there is a consistent trend for listeners to underestimate the ages of older girls. It is unlikely that the discrepancy is caused by a tendency for listeners to avoid the extreme responses on the slider, since the pattern is present only at the high end of the scale and only for girl's voices. Moreover, Amir et al. found a similar pattern when listeners were instructed to respond by selecting one of six response buttons representing the age ranges of the speakers.

The pattern in Figure 1 suggests that listeners can gauge the age of children's voices with relatively high accuracy. However, it should be noted that each data point is averaged across 5 talkers and 24 listeners. The inclusion of talkers and listeners introduces additional sources of variability in the data. Fitting a quadratic function to the entire dataset provides similar coefficients but reduced R^2 (0.66 for boys and 0.56 for girls). We are currently investigating acoustic factors that might account for variability across talkers and demographic factors – for example, familiarity and exposure to children's voices at different ages – to account for listener variability.

Previous findings have indicated that mean f0 and formant frequencies as well as durational properties provide important cues for the perception of speaker age (Linville, & Fisher, 1985; Harnberger et al., 2006). Studies with vocoded speech have indicated that upward scaling of the frequencies of the formants in a male voice, when coupled with an appropriate increase in f0, can raise the probability that the voice will be perceived as female. Similarly, downward scaling of the formants and f0 in a female voice raises the probability that the voice will be perceived as male. However, a change in F0 or formants alone is generally not sufficient to produce a compelling conversion of speaker sex (Hillenbrand & Clark, 2009). We are currently investigating whether a similar pattern holds for the perception of speaker age, and we are investigating properties of children's voices that might help to explain the discrepancy between perceived and chronological age in female voices and the effects of context.

REFERENCES

Amir, O., Engel, M., Shabtai, E., & Amir, N. (2012). *Identification of children's gender and age by listeners.* Journal of Voice 26(3), 313-21. Epub 2011 Aug 12.

Assmann, P.F., Nearey T.M. & Bharadwaj, S. (2008). *Analysis and classification of a vowel database*. Canadian Acoustics 36(3), 148-149.

Assmann P.F. & Nearey T.M. (2011). *Perception of speaker sex in children's voices*. Journal of the Acoustical Society of America 130(4), 2446(A).

Harnsberger, J. D., Shrivastav, R., Brown, Jr., W. S., Rothman, H. & Hollien, H. (2006). *Speaking rate and fundamental frequency as speech cues to perceived age*. Journal of Voice 22, 58-69.

Hillenbrand, J. M. & Clark, M. J. (2009). *The role of f0 and formant frequencies in distinguishing the voices of men and women*. Attention, Perception, & Psychophysics 71, 1150-1166.

Linville, S. E., and Fisher, H. B. (1985). *Acoustic characteristics of perceived versus actual vocal age in controlled phonation by adult females* Journal of the Acoustical Society of America 78, 40–8.

Ptacek, P. H., & Sander, E. K. (1966). *Age recognition from voice*. Journal of Speech and Hearing Research 9(2), 273-277.

ACKNOWLEDGEMENTS

Work supported by the National Science Foundation, Grant No. 1124479. Thanks to Daniel Hubbard and Shaikat Hossain for their assistance in data collection and analysis.

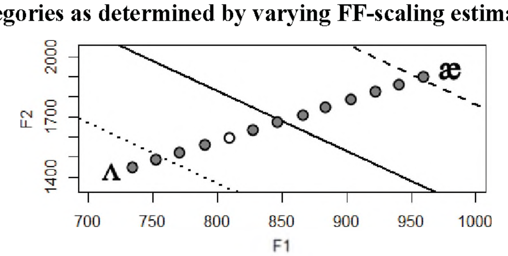
THE ASSOCIATION BETWEEN SPEAKER-DEPENDENT FORMANT SPACE ESTIMATES AND PERCEIVED VOWEL QUALITY

[1]Santiago Barreda, and Terrance M. Nearey
Dept. of Linguistics, University of Alberta, 4-32 Assiniboia Hall, Canada, T6G 2E7 [1]sbarreda@ualberta.ca

1. INTRODUCTION

The term *normalization* will be used here to denote the mechanism listeners use to accommodate between-speaker variation so that they can better identify vowels. Many theories of normalization suggest that vowels are interpreted relative to a representation of a speaker-dependent formant space, rather than being interpreted in an absolute way (Ladefoged & Broadbent 1957, Nearey 1989). Using a representation of a speaker's formant space, vowel sounds may be compared to the expected pattern for each vowel category for that speaker. There is good evidence that the formant spaces of speakers of the same language vary primarily by a single multiplicative parameter (Nearey 1978, Turner et al. 2009). This parameter is closely related to speaker vocal tract length and will be referred to as FF-scaling. Differences in speaker-dependent formant space estimates will be discussed in terms of differing FF-scaling estimates. Speakers with a relatively higher FF-scaling produce higher formant frequencies (FFs) overall than speakers with lower FF-scalings. As a result, when a speaker has a relatively higher FF-scaling, the expected FFs for all vowel categories increase, as do the FFs associated with the boundaries between any given set of vowel categories.

Fig. 1. Points indicate steps along a vowel continuum ranging from a less open vowel (/ʌ/) to a more open vowel (/æ/). Diagonal lines indicate boundaries between vowel categories as determined by varying FF-scaling estimates.



In experiments that involve vowel continua, different apparent speaker FF-scalings can result in vowel category shifts. Consider the simplified situation (ignoring F3 and higher formants) depicted in Figure 1. The dotted line represents the boundary between /ʌ/ and /æ/ when the speaker has a low FF-scaling. If this is the case, then the majority of vowels on the continuum will be identified as instances of /æ/, since they fall above the boundary. On the other hand, if the speaker has a high FF-scaling (as indicated by the dashed line), the majority of vowels along the continuum will be heard as instances of /ʌ/.

If normalization is driven by a process in which listeners estimate the location of the speaker's formant space (based on the apparent FF-scaling of the speaker), the interpretation of a sound with a given set of FFs (represented by a fixed point in Figure 1) should be determined by the FF-scaling estimate arrived at by the listener. Furthermore, the shift in vowel quality should be predictable based on the FF-scaling estimate. In the example given in Figure 1, the white point on the continuum would be more likely to be identified as an /æ/ when FF-scaling estimates were relatively low, indicating an inverse relationship between vowel openness and FF-scaling estimates.

In the experiment to be outlined here, listeners were first trained to report apparent speaker FF-scaling using the training method outlined in Barreda & Nearey (2011). This training method uses voices which vary in f0 and average FFs to teach listeners to report the psychoacoustic quality associated with higher overall FFs. Since higher FFs are associated with a higher FF-scaling, this response variable should be correlated with the apparent FF-scaling of the speaker and, consequently, with the speaker-dependent formant-space estimate arrived at by the speaker. After training, listeners performed a perceptual task similar to that in Barreda & Nearey (2012), in which they were presented with isolated vowel stimuli and were asked to indicate, on each trial: 1) The category of the vowel, 2) The gender of the apparent speaker and 3) Their FF-scaling estimate.

2. METHOD

Participants were 25 native speakers of Canadian English from the University of Alberta. Participants were drawn from a participant pool in which undergraduate linguistics students take part in experiments in exchange for partial course credit.

During the training phase, listeners learned to report apparent FF-scaling using the training method outlined in Barreda & Nearey (2011). After training, listeners proceeded to a testing phase. During the testing phase, listeners were presented with fully-randomized, isolated-vowel stimuli. For each vowel, listeners were asked to indicate the vowel category the stimulus belonged to (either /ʌ/ or /æ/) and the apparent gender of the speaker. Listeners were also asked to indicate the FF-scaling of the speaker using a discrete, 5-point scale, with higher values indicating a higher FF-scaling. Listeners heard each unique vowel stimulus 6 times, resulting in a maximum of 270 responses collected from each listener.

The testing stimuli consisted of a five-step F1-F2 continuum which spanned from FFs roughly appropriate for the /ʌ/ of an adult male to the /æ/ of an adult female. The vowels of the continuum varied in terms of increasing openness, with increasing F1 and F2 frequencies generally resulting in the perception of a more open vowel. The third point in the continuum had FFs which were appropriate for an /æ/ when produced by an adult male or an /ʌ/ when produced by an adult female. Since F1 and F2 frequencies are perfectly correlated, this factor will simply be referred to as F1. The FF values of each step along the continuum is provided in Table I. Each point along the continuum was combined with 3 f0 values (140 Hz, 198 Hz, and 280 Hz), and three F3 values (2475 Hz, 2774 Hz, 3109 Hz), resulting in 45 unique vowel stimuli. All vowels had f0s which decreased linearly by 10% from the start to the end of the vowel, and were 200 ms in duration.

Table 1. Formant frequencies of stimulus vowels.

Step #	1	2	3	4	5
F1	718	775	838	905	977
F2	1422	1536	1659	1792	1935

3. RESULTS

To confirm that listeners were reporting apparent FF-scaling in a consistent manner based on the stimulus properties, a linear model was fit to the pooled data across all participants in which reported FF-scaling was the dependent variable. Stimulus F1, F3 and f0 were the independent variables, and all were coded as continuous covariates. This model explained 18% of the variance in reported FF-scaling, with F1 accounting for 67.8%, f0 accounting for 28.1%, and F3 accounting for only 0.2% of the explained variance.

Table II. Results of significance tests carried out on the within-participant logistic regression coefficients.

Coefficient	Mean	t(24)	p
F1	3.45	16.6	< 0.001
F3	-1.83	11.3	< 0.001
f0	-0.74	5.9	< 0.001
Maleness	0.47	2.6	0.015
FF-scaling	-0.20	2.9	0.009

To investigate the relationship between vowel openness and apparent speaker FF-scaling, a two-stage (Lorch & Myers 1990) logistic regression analysis was carried out. A logistic regression model was fit to the data collected from each participant individually. In each case, vowel openness was the dependent variable, where responses of /æ/ were coded as 1 and responses of /ʌ/ were coded as 0. The stimulus properties F1 step, and F3 and f0 level were coded as continuous covariates. The response variable reported speaker gender was coded as a dummy variable, while reported speaker FF-scaling was coded as a continuous covariate. A series of independent-sample t-tests were carried out on the coefficients collected from all participants to see which independent variables significantly affect perceived vowel openness. The results of this are presented in Table II.

4. DISCUSSION

As seen in Table II, reported FF-scaling has a significant negative effect on vowel openness. This means that for a given vowel sound, when listeners reported a higher FF-scaling, they were less likely to hear an open vowel. This negative association exists despite the fact that vowel openness and reported FF-scaling have a positive marginal relationship. For example, listeners heard /æ/ in 42% of cases when they reported the lowest FF-scaling level, and in 76% of cases when they reported the highest FF-scaling level. As discussed in the introduction, this counter-intuitive result is what would be expected if listeners were normalizing vowels based on speaker-dependent formant space estimates (driven by FF-scaling estimates). Although, in general, vowels with higher FFs will be perceived as indicating a higher FF-scaling *and* a more open vowel, after controlling for stimulus FFs (i.e., considering a fixed continuum point in Figure 1), a higher FF-scaling estimate results in the perception of fewer open vowels overall.

If the association between vowel openness and reported FF-scaling were simply a result of formant estimation errors on the part of the listener, we would expect a positive relation between vowel openness and reported FF-scaling. For example, for a vowel with a given set of FFs, relative to a fixed boundary, in cases where listeners overestimated the FFs, they would be more likely to hear a more open vowel *and* they would be more likely to report a higher FF-scaling.

REFERENCES

Barreda, S. & Nearey, T. (2011). Training listeners to report fundamental frequency and formant range information independently. The 161th meeting of the Acoustical Society of America, Seattle, WA.

Barreda, S. & T.M. Nearey. (2012). The direct and indirect roles of fundamental frequency in vowel perception. *Journal of the Acoustical Society of America* 131: 466-477.

Ladefoged, P., and Broadbent, D. E. (1957). “Information conveyed by vowels,”. J. Acoust. Soc. Am. 29, 98–104.

Lorch, R. F., and Myers, J. L. (1990). “Regression analyses of repeated measures data in cognitive research,” J. Exp. Psychol. Learn. Mem. Cogn. 16, 149–157.

Nearey, T. M. (1978). Phonetic Feature Systems for Vowels. PhD thesis, Indiana University Linguistics Club.

Nearey, T. M. (1989). “Static, dynamic, and relational properties in vowel perception,” J. Acoust. Soc. Am. 85, 2088–2113.

Turner, R. E., Al-Hames, M. A., Smith, D. R. R., Kawahara, H., Irino, T., and Patterson, R. D. (2006). Vowel normalisation: Time-domain processing of the internal dynamics of speech. in Dynamics of Speech Production and Perception, edited by P. Divenyi. Amsterdam: IOS Press. pp. 153-170.

TO REDUCE OR NOT TO REDUCE : EVIDENCE FROM SENĆOŦEN STORYTELLING

Sonya Bird, Ewa Czaykowska-Higgins, and Janet Leonard
Dept. of Linguistics, University of Victoria, PO Box 3045, Victoria BC, V8T3K5 sbird@uvic.ca

1. INTRODUCTION

It is a well-known fact that articulatory, and consequently acoustic, events are compressed in fluent speech; a process known as ‘reduction’ (Johnson 2004). Research has shown that when it comes to reduction, not all segments are equally affected; for example when reduction occurs at fast speech rates, effects are often greater on vowels than on consonants (Gay 1981). This paper reports on a preliminary investigation of reduction in the speech of a single fluent SENĆOŦEN speaker. The focus is on /vowel-ʔ-ə/ vs. other /vowel-ʔ-vowel/ sequences; we show that while the former reduce to a single lengthened vowel, the latter do not. As a whole, results support previous claims that /ə/ in Salish is phonologically and phonetically distinct from full vowels (Czaykowska-Higgins & Kinkade 1998), and that reduction is sensitive to the particular vowels involved, affecting some but not others.

2. METHOD

The dataset for the study consisted of 130 words extracted from a SENĆOŦEN story told by a fluent speaker, recorded in the 1970s. These words fell into three sets, corresponding to three different kinds of sequences: Set 1 consisted of 58 /vʔə/ sequences; of these, 52 were /eʔə/, hence the focus on these sequences in the acoustic analysis (see below). Set 2 consisted of 12 other /vʔv/ sequences; as is clear from the token count, these occurred relatively rarely in the story. Set 3 consisted of 60 tokens of /e/; these provided a baseline against which to compare the acoustic properties of /eʔə/. Table 1 provides a summary of the dataset.

Table 1. Tokens analyzed (target sequence in bold)

Set	Sequence	#	Example	Gloss
Set 1	/eʔə/	52	/l**eʔə**/	‘there’
	/iʔə/	3	/netʃt**iʔə**s/	‘different
	/ɑʔə/	3	/qʷ**ɑʔə**ŋ/	‘water’
Set 2	/eʔi/	8	/tʃ**eʔi**/	‘work’
	/eʔu/	2	/j**eʔu**/	‘went’
	/iʔe/	2	/t**iʔe**/	‘this’
Set 3	/e/	60	/məqsten/	‘everything’

The dataset included a number of words which were repeated multiple times (e.g. seven of the eight /eʔi/ sequences are from different repetitions of the word /tʃeʔi/ ‘work’). Multiple repetitions of a single word were treated as separate items in the qualitative analysis (3.1) because there was no easy way of averaging across them; in the quantitative analysis (3.2), they were aggregated and treated as a single item, as long as they were consistent in terms of stress and position (see 3.2 below for details).

3. RESULTS

3.1 Qualitative analysis

All /vʔv/ sequences (58 /vʔə/ and 12 other /vʔv/) were first transcribed based on auditory analysis, to determine (impressionistically) to what extent they were reduced. Table 2 summarizes the results.

Table 2. Transcriptions by sequence type

Set	Sequence	Transcription (#)	**Most common**
Set 1 /vʔə/ (58)	/eʔə/ (52)	[eː] (47); [e] (1); [eʔə] (2)	**[vː]**
	/iʔə/ (3)	[ije] (2); [eː] (1)	
	/ɑʔə/ (3)	[ɑː] (2); [aʔə] (1)	
Set 2 /vʔv/ (12)	/eʔi/ (8)	[eʔei̯] (7); [ei̯] (1)	**[vʔv]** or **[vjv]**
	/eʔu/ (2)	[eju]	
	/iʔe/ (2)	[iː] (1); [ije] (1)	

In general, transcriptions reflect the fact that /vʔə/ tends to reduce to [vː] while other /vʔv/ sequences tend either not to reduce, or to reduce to [vjv]; the latter case is interesting and may have to do with the phonological status of /ʔ/ in these words (underlying vs. derived from glottalized /j’/), but will not be further discussed here. Focusing on /vʔə/ vs. other /vʔv/ sequences, it is interesting to note that while the seven repetitions of /tʃeʔi/ (‘work’) *do* seem to reduce to varying degrees, none of them reduce to the extent that they lose the glottal stop entirely, as do the vast majority of /vʔə/ sequences.

3.2 Quantitative analysis

Based on the finding that /vʔə/ tends to reduce to [vː], a subset of these sequences - /eʔə/ ones - were analyzed in Praat in terms of: a) duration, b) vowel quality (F1 and F2 at 25% and 75% into the vowel), and c) glottalization (jitter, spectral tilt, amplitude dip, and pitch dip during the target interval). Acoustic analysis was limited to /eʔə/ sequences for two reasons: 1) they were by far the most common /vʔə/ sequence and so provided a unified set for analysis, and 2) the resulting [e:] could easily be compared to the underlying SENĆOTEN /e/ vowel, which also occurred relatively frequently in the story. As mentioned above, the set of words used in this study included a number of repetitions. As it turned out, the 60 /e/ tokens came from a much more varied set of words than did the 52 /eʔə/ tokens, which were extracted from a relatively small set of frequently repeated function words. Repetitions were aggregated only if stress (stressed vs. unstressed) and position (final vs. non-final) were consistent, leading to the analysis of 50 /e/ items and 22 /eʔə/ items. A series of two-factor between-items ANOVAs was used to investigate acoustic differences between underlying /eʔə/ and /e/; the two factors were sequence (/e/ vs. /eʔə/) and position (final vs. non-final). Position was included because the correlates of phonemic glottalization are sometimes confounded with those of prosodic (utterance-final) position.

The primary difference between /eʔə/ and /e/ was in term of duration: the main effect of sequence was significant $F(1, 71) = 50.38$; $p < 0.001$), with /eʔə/ almost twice as long (238ms) as /e/ (130ms). The effect of position was not significant, and neither was the interaction. This durational difference confirmed the auditory analysis (see 3.1), in which /eʔə/ was transcribed as [e:] and /e/ as [e].

Although less salient auditorily, /eʔə/ and /e/ also differed in vowel quality, particularly in terms of F2: /eʔə/ had a significantly lower F2 than did /e/ at both 25% ($F(1,71 = 10.10$, $p < 0.01$) and 75% ($F(1,71 = 16.73$, $p < 0.001$) into the vowel - see Figure 2. F1 was significantly higher in /eʔə/ than in /e/ only at 25% into the vowel ($F(1,71 = 4.09$, $p < 0.05$). Together, F1 and F2 values indicate that /eʔə/ is more retracted and slightly lower than /e/. Interestingly, both /eʔə/ and /e/ are realized in the range of [ɛ] (Kent & Read 2002), a lower and laxer version of the mid-front vowel previously documented in SENĆOTEN (Montler 1986).

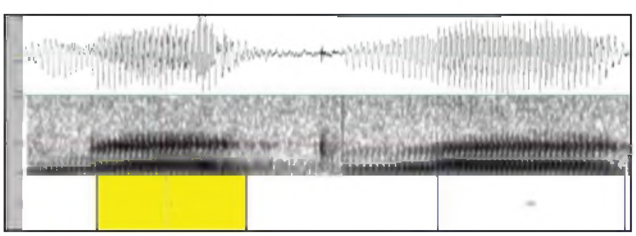


Figure 2. /e/ in /net/ 'name' vs. /eʔə/ in /leʔə/ 'there' (/e/ is shaded).

There was little consistent evidence of any underlying glottalization in /eʔə/ sequences. Of the acoustic measurements taken, only pitch and amplitude dips showed effects. These were calculated by subtracting the minimum pitch/amplitude from the maximum within the target interval (/e/ or /eʔə/). For both pitch and amplitude, there was an interaction between sequence and position, with dips significantly greater in /eʔə/ than in /e/ in utterance final position only (pitch: $F(1,20) = 14.43$, $p<0.01$; amplitude: $F(1,20) = 11.65$, $p<0.01$). Table 3 summarizes these results.

Table 3. Pitch and amplitude dips

Sequence	Pitch dip (Hz)	Amplitude dip (dB)
/e/	29 (14)	8 (3)
/eʔə/	33 (20)	9 (2)

4. DISCUSSION

Overall, results showed two things: 1) /vʔə/ sequences tended to reduce to [v:] while other /vʔv/ sequences did not; 2) reduced /eʔə/ sequences were distinguishable from underlying /e/ in duration and to a lesser extent in vowel quality, but not (consistently) in degree of glottalization. The pronunciation of /vʔə/ as [v:] can be viewed as a more extreme version of "schwa assimilation across glottal stop", which has previously been reported in SENĆOŦEN (Montler 1986, p. 29). The fact that /eʔə/ sequences exhibited greater reduction effects than other /vʔv/ sequences, and also that they showed greater dips in pitch and amplitude utterance-finally than did /e/ is possibly related to the fact that /eʔə/ sequences were extracted primarily from function words, whereas other /vʔv/ sequences and /e/ were extracted from a more varied set of words. It may be the case that function words are more prone to a range of prosodic effects than are content words, a tendency that could prove useful as a diagnostic for teasing apart different word classes in Salish languages (Czaykowska-Higgins & Kinkade 1998).

REFERENCES

Czaykowska-Higgins, E & D. Kinkade (1998). Salish languages and linguistics. Berlin/New York: Mouton de Gruyter.

Gay, T. (1981). Mechanisms in the control of speech rate. Phonetica, 38, 148-158.

Johnson, K. (2004). Massive reduction in conversational American English. In Spontaneous Speech: Data and analysis – Proceedings of the 1st session of the 10th International Symposium, ed. Yoneyama, K. & Maekawa, K. Tokyo, Japan: The National International Institute for Japanese Language, 29-54.

Kent R. & C. Read (2002) Acoustic analysis of speech, second edition. Albany NY: Singular; Thomson Learning.

Montler, T. (1986). An outline of the morphology and phonology of Saanich, North Straits Salish. University of Montana Working Papers in Linguistics 4.

ACKNOWLEDGEMENTS

Many thanks to our SENĆOŦEN-speaking collaborators in transcribing and translating the story we used for this study, and to the Jacobs Fund for funding the project.

PROSODIC PHRASING IN NXAʔAMXCÍN (SALISH) DECLARATIVE CLAUSES

Marion Caldecott[1] and Ewa Czaykowka-Higgins[2]
[1]Dept. of Linguistics, Simon Fraser University, 8888 University Dr. Burnaby, BC, Canada, V5A 1S6
[2]Department of Linguistics, University of Victoria, PO Box 3045, Victoria, BC, Canada, V8W 3P4

1. INTRODUCTION

This paper reports results from a pilot study of the phonetic correlates of declarative clauses from a narrative in Nxaʔamxcín (Salish). There is a small but growing body of research describing the properties of prosodic phrasing in Salish languages. While seminal research on prosodic phrasing in English and Japanese is based on targeted elicitations [9], [11], work on prosody in endangered languages often comes from varying sources. In the case of Salish languages some of this research is based on examples elicited to answer specific phonetic questions associated with intonation and phasing [5], [6], [7], some is based on examples that were elicited for syntactic rather than prosodic purposes [4], some on conversational data [7], and some on analysis of connected speech found in a narrative text [1], [2], [3]. In addition, there are differences in the types of correlates examined for different phrase/clause types in Salish: e.g., while [3] uses pauses as cues in narratives, [7] does not consider pauses in elicited sentences. Thjs pilot study examines the phonetic correlates of prosodic phrases in Nxaʔamxcín, an Interior Salish language whose prosodic phrasing has not been studied previously, to determine the extent to which data from this language is consistent with results from other Salish languages. Specifically, we examine whether i) lexical prominence is associated with a pitch peak and ii) the boundaries of prosodic phrases are correlated with boundary tones, pitch resets and pauses.

2. METHOD

2.1 Materials

The analysis was conducted on a 30-minute narrative x̣áʕx̣aʕ st̓ámkaʔs "Crow's Daughter" told by a fluent storyteller in her 70s. It was recorded in 1990 on a Marantz 430 cassette recorder, using an AKG D320B unidirectional microphone, and digitized in 2002. The narrative consists of approximately 500 lines of text, each representing approximately one clause.

To make the data directly comparable to that used in earlier studies, two phrase-types were initially selected for analysis: the first 13 random phrases of the text, uncontrolled for word order, and predicate-initial (V-initial) declarative clauses, reflecting the most common non-focus word order [1,12]. The text contained only 13 such clauses (including 2 from the first 13), all with an overt Subject and/or Object. The scope was then widened to include all k^{w}aʔ-initial phrases, whose form differs minimally from V-initial sentences. This resulted in an additional 14 phrases, 9 in which k^{w}aʔ is a conjunction, 4 where its function is that of a discourse linking particle and one where it is ambiguous. 7 phrases were excluded due to mistranscriptions, misspeaks or rhetorical lengthening. The total number of phrases analysed was 32.

2.2 Acoustical analysis procedure

The data were analyzed using [10] on a Macbook Pro computer. The measurements made are comparable to those in the earlier work on Salish languages cited above. The maximum F0 of stressed, phrase-initial and -final vowels of the target phrase, as well as the final vowel of the preceding phrase were measured. If pitch was higher than that of either surrounding vowel, it was considered a pitch peak. Low tone was assigned to initial/final vowels if the pitch was i) lower than the following or preceding vowel, respectively. Pitch resets were calculated by subtracting the pitch of a phrase-initial vowel from that of the final vowel of the preceding phrase. In addition, the length of pauses before each phrase was calculated. As this is a pilot study no statistical analyses were performed.

3. RESULTS

73% of primary-stress vowels are correlated with pitch peaks, as seen in Figure 1:

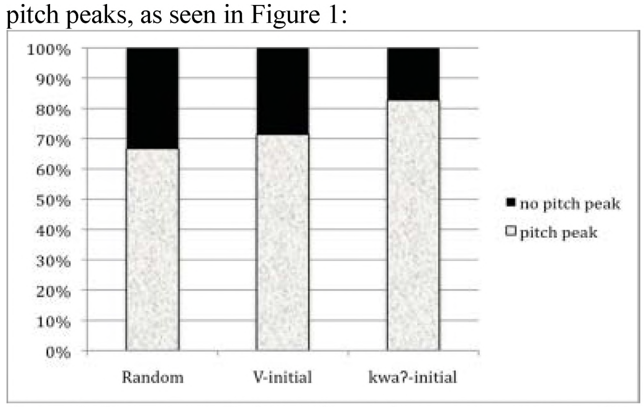


Figure 1: Percentage of tokens where primary stress was correlated with a pitch peak.

Phrase boundaries were correlated with a L% initial and final tone, provided the boundary vowels were not stressed. 89% of phrases with unstressed boundary vowels were associated with a L%, while 91% of phrases with a stressed boundary vowel showed no rise or fall. Verb-initial phrases had predominantly initial stressed vowels, while k^{w}aʔ-initial phrases were rarely initially-stressed.

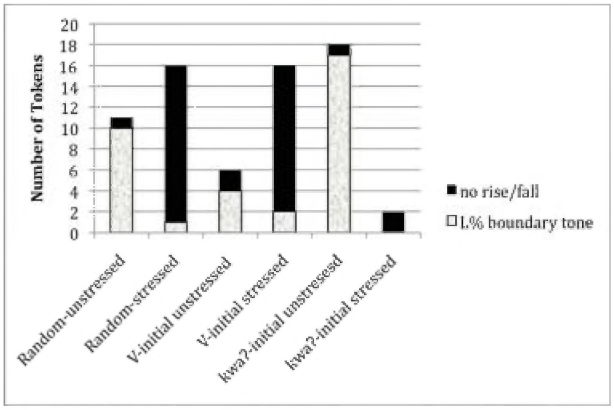


Figure 2: Boundary tones by phrase type and vowel type.

Pitch resets at boundaries occur 69% of the time in random phrases, 50% of the time in V(S)(O) structures, and never in structures beginning with kʷaʔ.

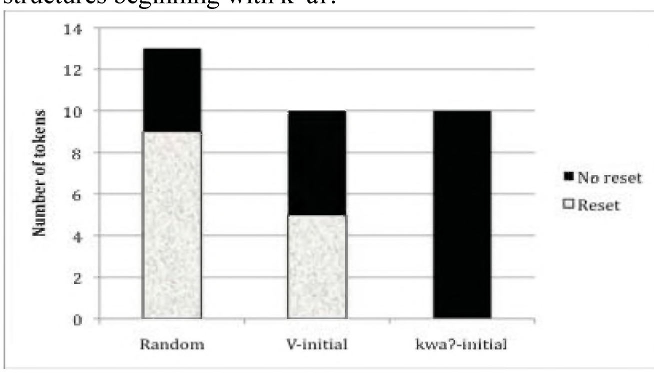


Figure 3: Pitch resets by phrase type

In V(S)(O) structures, pauses delimit phrases 70% of the time. *kʷaʔ*-conjunction cases have no preceding pause, but *kʷaʔ*-discourse particle cases do. The pauses in the *kʷaʔ*-initial phrases in Fig. 4 represent only the discourse particles.

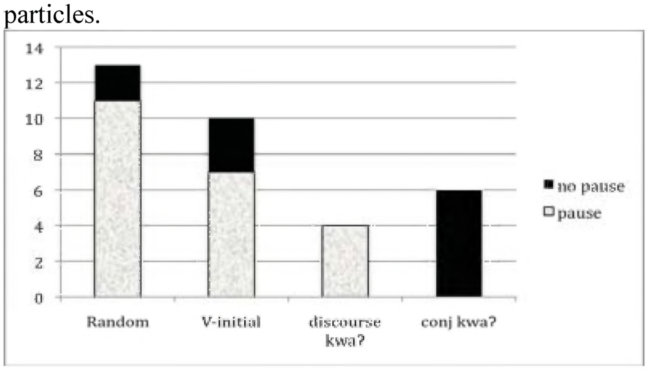


Figure 4: Tokens preceded by pauses by phrase-type

4. DISCUSSION

These results are partly consistent with observations made for other Salish languages. In Nxaʔamxcín, higher pitch is generally correlated with primary lexical stress, as in Nlhekepmxcin [7]. Declarative clauses are delimited by boundary tones, as is the case in St'át'imcets [5]. However, pitch resets are not consistent correlates of prosodic units in Nxaʔamxcín, unlike in Lushootseed [3] and Okanagan [1]. Even in phrases with initial stress (e.g. V-initial) where a pitch reset might be predicted, they occurred only 50% of the time. Lastly, pauses do not reliably delimit V(S)(O) phrases, which is not consistent with results found for Lushootseed and SENĆOTEN [3, 4, 8]. This result is surprising, given that pauses did delimit phrases where kʷaʔ acted as a discourse particle (and thus marked a new phrase), as opposed to when it functioned as a conjunction.

Our results, while preliminary are both comparable and vary systematically from the results found for other Salish languages. This suggests that while correlates for prosodic phrasing are selected from a universal set, the implementation is language-specific. A final point raised by our pilot study is the extent to which our results are directly comparable to those from studies using other types of data (elicited, conversational or narrative). The fact that our analysis produced comparable results to data elicited using other methods suggests that the source of the data may not have a strong effect on results. This would be an encouraging result for those communities with limited audio recordings.

REFERENCES

[1] Barthmaier, P. (2004). Intonation units in Okanagan. In D. Gerdts & L. Matthewson (Eds.), *Studies in Salish Linguistics in Honor of M. Dale Kinkade* (Vol. 17): UMOPL.
[2] Beck, D. (1999). Words and prosodic phrasing in Lushootseed. In T. A. Hall & U. Kleinhenz (Eds.), *Studies on the phonological word* (pp. 23-46). Amsterdam: Benjamins.
[3] Beck, D., & Bennett, D. (2007). Extending the Prosodic Hierarchy: Evidence from Lushootseed Narrative. *NWJL* 1: 1-34.
[4] Benner, A. (2006). The prosody of SENĆOTEN, a pilot study. Paper presented at *41st ICSNL*. UVIC.
[5] Caldecott, Marion (2009) Non-exhaustive parsing: Phonetic and phonological evidence from St'át'imcets. PhD. Dissertation. UBC.
[6] Jacobs, P. (2007). Intonation of yes/no questions in Sḵwx̱wú7mesh. In K. M. Johannsdottir & M. Oberg (Eds.), *ICSNL* 42 (Vol. 20, pp. 256-284). UBC Kelowna: UBCWPL.
[7] Koch, K. (2008). *Intonation and Focus in Nlhe7kepmxcin (Thompson River Salish)*. PhD Dissertation, UBC.
[8] Leonard, J. (2011) Prosodic units in SENCOTEN (North Straits Salish). UVIC, ms.
[9] Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation*. PhD Dissertation, MIT.
[10] Boersma, Paul & Weenink, David (2012). Praat: doing phonetics by computer [Computer program]. Version 5.1.44, retrieved July 2011 from http://www.praat.org/
[11] Selkirk, E. (1995). Sentence prosody: intonation, stress and phrasing. In J. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 550-569): Blackwell Reference.
[12] Willett, M.L. 2003. *A Grammatical Sketch of Nxa'amxcín (Moses-Columbia Salish)*. PhD dissertation, UVIC.

ACKNOWLEDGEMENTS

Lámlamt to Elizabeth Davis who told the story; to Colville Confederated Tribes for permission to work on the Reservation; and to UVIC-IRG and Jacobs Research Fund.

PRELIMINARY STATISTICAL PATTERN RECOGNITION METHODS IN THE STUDY OF VOWELS PRODUCED BY CHILDREN WITH AND WITHOUT SPEECH SOUND DISORDERS

Hyunju Chung[1], Terrance, M. Nearey[2], Megan Hodge[1], Karen, E. Pollock[1], and Benjamin, V. Tucker[2]
[1]Dept. of Speech Pathology & Audiology, University of Alberta, 8205 114St., Edmonton, AB, Canada, T6G 2G4
[2]Dept. of Linguistics, 4-32 Assiniboia Hall, Edmonton, AB, Canada, T6G 2E7

1. INTRODUCTION

The nature of vowel acquisition and vowel error patterns in young children, especially those with speech sound disorders (SSD), is not well understood. Previous studies of vowels produced by children with SSD of unknown origin (SSD-UNK) have shown that they demonstrate a similar order of acquisition and vowel error patterns as those of children with TDS (typically developing speech), but are slower in developing vowels and produce more vowel errors (e.g., Stoel-Gammon & Herrington, 1990). Previous studies have relied mainly on phonetic transcription to describe the characteristics of vowels produced by children with SSD. A few studies used acoustic analysis to investigate the vowels of children with SSD due to motor impairments. These studies have shown that children with cerebral palsy have smaller vowel space size than children with TDS, and that size of vowel space is correlated positively with their speech intelligibility (e.g., Higgins & Hodge, 2001). These studies, however, are based on the acoustic measurements of only 3 to 4 corner vowels. In addition, despite the proposed importance of spectral movement patterns in successful identification of vowels (e.g., Hillenbrand, & Nearey, 1999), only a few studies (e.g., Lee, 2009) have examined spectral movement patterns of vowels produced by young children with SSD.

In the current study, acoustic characteristics of each of the 10 English monophthong vowels were examined and compared to the same-aged children with and without SSD. For acoustic measurements, F1, F2, and fundamental frequency (F0), that incorporated spectral movement patterns, were obtained for the analysis. Using these measurements, statistical pattern recognition models (e.g., Thomson, Nearey, & Derwing, 2009) were used to examine whether sets of acoustic variables differentiate 1) vowels produced by children with and without SSD, 2) age subgroups within each child group, and 3) vowels that are identified accurately from those that are not.

2. METHOD

2.1 Participants

Adult participants were 15 women, ages 18 to 35 years. All were monolingual speakers of Western Canadian English with no history of speech delay or disorder. Two groups of children participated (TDS and SSD-UNK), with 3 children in each of four age groups (3, 4, 5, and 6-year-olds) in each group for a total of 24 children. All children were learning English as their first language and living in Western Canada. Parents of children in the TDS group reported no concerns about their child's speech and language development, and all children passed a standard speech and language screening measure. Children in the SSD group were receiving or on waiting lists for speech therapy.

2.2 Stimuli

The target words used in this study are a subset of words from the three *TOCS+* word lists (Hodge, Daniels, & Gotzke, 2009) (Table 1). Target vowels were 10 English monophthongs, [i, ɪ, e, ɛ, æ, ʌ, ɑ, o, ʊ, u].

Table 1. List of target words and vowels

Vowel	Target Words	Vowel	Target Words
/i/	bead, beat, bee, D, feet, tea, peep	/ɪ/	bit, fit, hid, sit
/e/	bait, pain	/ɛ/	bet, pen
/æ/	baa, bad, bag, fat, hat, pat, tap	/ʌ/	bud, bug, hut, pup, shut, tub
/ɑ/	Don, hot, jaw, paw, pop, pot, shot, top	/o/	cone, toe
/ʊ/	foot, hood, soot	/u/	boo, Pooh, shoot, hoot, suit, two, tube

2.3 Analysis

The boundaries for each vowel token were manually defined using *Praat* (Boersma & Weenink, 2012). A semi-automatic formant tracking program (Nearey *et al.*, 2002) created in MATLAB (7.8.0.347, R2009a) extracted vowel duration: it also extracted F1, F2, and F0 at 2ms steps over the entire duration of the vowel. For each vowel, the last 10ms or the earliest point where the amplitude falls 25dB below the peak has been trimmed. All F1 and F2 values were then log transformed. For F0, the median of the first half of the trimmed vowel was used. The log transformed F1 and F2 measured at 20% and 70% time points, median F0, and duration were used as input variables for the pattern recognition model. The acoustic measures of vowels produced by all speakers were used to train a linear discriminant analysis model. Predicted identification rates for each group of speakers were calculated using the resubstitution method that is the same data was used in training the model and predicting the classification.

3. RESULTS

The preliminary analyses indicated that adult vowels were classified with the highest accuracy (91.7%). Vowels of two groups of children were classified with similar accuracy (TDS - 80.6% and SSD-UNK - 74.9%), but at lower accuracy than for the adult vowels. Across age groups, vowels of the 6-year-olds were classified with higher accuracy than those of the younger age groups in each of the TDS and SSD-UNK group. Across all groups, /i/ was classified most accurately and the vowels /ɛ/ or /ʌ/ least accurately. The model predicted accuracies of 10 vowels of each speaker group are summarized in Table 2.

Table 2. Overall classification accuracy (%) for vowels produced by each speaker and age group.

	Vowel Category									
Speaker Group	i	ɪ	e	ɛ	æ	ʌ	ɑ	o	ʊ	u
Adults	98	100	84	63	88	84	95	100	92	96
TDS 6yr	100	100	100	100	63	82	86	100	83	86
TDS 5yr	100	100	25	50	86	62	79	75	50	100
TDS 4yr	100	100	50	25	87	54	69	100	60	86
TDS 3yr	100	100	100	0	92	50	75	50	50	93
SSD-UNK 6yr	100	88	100	60	64	54	69	100	80	93
SSD-UNK 5yr	87	88	25	0	92	75	79	50	40	86
SSD-UNK 4yr	75	100	100	100	100	39	73	100	17	87
SSD-UNK 3yr	100	100	50	33	80	36	64	50	60	54

Vowels were better identified with two measurements (at 20% and 70%) of the vowel formant pattern than a single measurement (at 50%) for all groups. Classification scores were higher when all the acoustic variables were entered than when either duration or F0 was absent (Table 3). The result of the two-way ANOVA showed that each or the combination of acoustic measurements differs significantly by vowel type and speaker groups.

Table 3. Overall classification accuracy (%) of each speaker group by acoustic variables.

Speaker Group	Acoustic measures							
	duration F0		no duration F0		duration no F0		no duration no F0	
	20,70%	50%	20,70%	50%	20,70%	50%	20,70%	50%
Adults	91.7	85.5	87.3	83.4	91.1	85.4	85.9	81.9
TDS	80.6	77.3	79.0	73.5	80.6	77.3	78.7	73.0
SSD-UNK	74.9	67.5	73.3	68.8	73.9	66.1	72.5	67.2

4. DISCUSSION

The analysis using a pattern recognition model showed that vowels of adults were better identified than those of the child groups, as expected. The classification accuracy of the two child groups was not very different; accuracy of SSD-UNK was slightly lower than those of TDS group. An age difference in classification accuracy was also found between the oldest children (6-year-olds) and the younger ages. Vowels of all groups were more successfully classified with two measurements representing the formant movement patterns, than a single point measurement, and a combination of all acoustic measures than a single or a subset of measures. Regardless, some vowel categories were always classified with higher accuracy than others.

Our next steps include 1) addressing ways to minimize measurement errors (e.g., rechecking formant frequencies and F0 of poorly classified tokens) and 2) developing methods to compare the model predicted accuracy with the judged accuracy of vowels based on listener identification scores. Further testing of the model will follow, using the same acoustic measurements from additional children with and without SSD.

REFERENCES

Boersma, P. & Weenink, D.(2012). Praat: doing phonetics by computer [Computer program]. Version 5.03.02, retrieved 27 March 2012 from http://www.praat.org/.

Hillenbrand, J.M., & Nearey, T.M. (1999). Identification of resynthesized /hVd/ syllables: Effects of formant contour, *Journal of the Acoustical Society of America, 105 (6)*, 3509-3523.

Higgins, C. M., & Hodge, M. M. (2001). F2/F1 vowel quadrilateral area in young children with and without dysarthria. *Canadian Acoustics - Acoustique Canadienne, 29*(3), 66-67.

Hodge, M., Daniels, J., & Gotzke, C. L. (2009). TOCS+ *intelligibility measures* (version 5.3)[computer software]. Edmonton, AB: University of Alberta.

Lee, J. (2009). *Development of vowels and their relationship with speech intelligibility in children with cerebral palsy.* Unpublished doctoral dissertation, University of Wisconsin, Madison.

MATLAB (2010). Version 7.10.0. Natick, Massachusetts: The MathWorks Inc.

Nearey, T. M., Assmann, P. F. & Hillenbrand, J. M. (2002). Evaluation of a strategy for automatic formant tracking. *Journal of the Acoustical Society of America, 112,* 2323.

Stoel-Gammon, C. & Herrington, P. (1990). Vowel systems of normally developing and phonologically disordered children. *Clinical Linguistics and Phonetics, 4*, 145-160.

Thomson, R. I., Nearey, T. M., & Derwing, T. M. (2009). A modified statistical pattern recognition approach to measuring the crosslinguistic similarity of Mandarin and English vowels. *The Journal of the Acoustical Society of America, 126 (3)*, 1447-1460.

DENTALS ARE GRAVE

Darin Flynn[1] and Sean Fulop[2]

[1]Dept. of Linguistics, University of Calgary, 2500 University Dr NW, Calgary, Alberta, Canada T2N 1N4

[2]Dept. of Linguistics, California State University, 5245 N. Backer Ave. PB92, Fresno, CA, USA 93740-8001

1. INTRODUCTION

Phonetic features require either an articulatory or an acoustic basis. Defining a feature in an optimal fashion can lead to improved explanatory force concerning, for example, phonetically motivated sound change. This article highlights the increased explanation of certain auditorily based sound changes and assimilations, obtained by correcting the definition of the old-school feature [grave], and concomitant adjustments to the classification of segments. In particular, non-sibilant dentals must be [grave]. Like all coronals, dentals are considered [acute] in Jakobsonian taxonomy [1] et seq. However, their noise energy and their involvement in [flat] enhancement and assimilation suggest instead that they are [grave], like labials and velars.

2. CLASSIFICATION

With noise measurements alone, it is notoriously difficult to discriminate reliably between non-sibilant dental and labial consonants, as both present generally level spectra with no significant peaks.

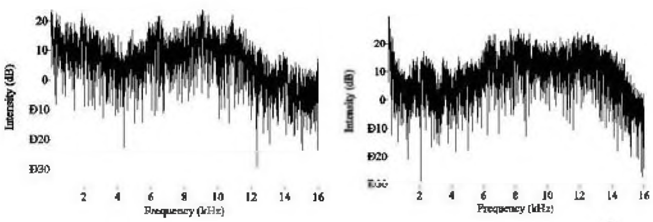


Fig. 1. Simple Fourier power spectra of [θ] in Slavey [t̪͡θʰah] 'carrot' and [f] in English 'fan' (rendered in Praat)

Indeed their noise energies are so similar that labials commonly substitute for dentals across languages. Table 1 showcases such substitution in the Slavey (Athabascan) dialect centered in Tulita, NT; cf. dentals in South Slavey (NT, AB).

Table 1. Dentals > labials in North Slavey, Tulita, NT [2]

Slavey	Tulita		Sl.	Tul.	
ʔeht̪͡θáa	ehpaː	'dryfish'	θę́	fẽʔ	'star'
-t̪͡θʰiʔ	-pʰiʔ	'head'	θa	fa	'sand'
t̪͡θ'ih	p'ih/p'ĩ́ɛ	'mosquito'	-ðáʔ	-va	'mouth'
-t̪͡θ'éhé	p'é/p'ɛ́h	'sinew'	-ðeʔ	-ve	'liver'

Since labials are always considered [grave], we can find no support at all for the claim that (inter)dentals are not [grave], let alone [acute]. Rather it seems clear that (inter)dentals can only be given the same value of [grave] as the labials from which they are so hard to distinguish. To quote Ladefoged and Maddieson: "It seems that in the case of the pairs [f, θ], and [v, ð] in English, the inconsistencies between speakers are so great that it may be profitless to try to characterize the acoustic spectra of the fricatives themselves." [3]

Moreover, given their rather level spectrum, labials (and dentals) cannot be [grave] in the sense of "having predominantly low frequency energy" [1]. Rather, given that an acoustic feature must really be an *auditory* feature, we propose to redefine [grave] as the *audible* presence of significant low frequency noise in a sound. In particular this means that the low frequency noise (< 2.5 kHz) must not be overshadowed by predominant high frequency noise (as in sibilants).

On this definition, [grave] applies equally to labials and dentals as it does to velars, which present a preponderance of noise in lower frequencies. This allows us to rationalize shifts not only between velars and labiodentals as in Table 2, but also between velars and dentals as in Table 3.

Table 2. [f] ~ [x] in Hare Slavey, Fort Good Hope, NT [2]

fori ~ xori	'quickly'	lifuʃé ~ lixuʃé	'fork'
lifótõ ~ lixótõ	'nine'	fawéhgewe ~ xawéhgewe	'Old Baldy'

Table 3. [θ, ð] > [x, ɣ] in South Slavey, Wrigley, NT [2]

Standard	Wrigley		St.	Wrg.	
θe-	xe-	PERF.	-ðáʔ	-ɣáʔ	'mouth'
θę́	xẽʔ	'star'	-ðéh	-ɣéʔ	'skin'

3. ENHANCEMENT

According to Jakobson et al. [1] [grave] is *enhanced* (cf. [4]) by another "low tonality" feature of vocoids, [flat], characterized by a downward shift of formants—particularly F_2. (Similarly, consonantal [acute] is enhanced by vocalic [sharp], an upward shift of formants.) Indeed, across languages, F_2 transitions tend to be lower or equal in dentals vs. alveolars [5]. This pattern is shown for American English in Table 4.

Table 4. Starting F2 values (Hz) for alveolars vs. dentals [6]

si	2050	zi	1950	di	2000	θi	1950	ði	1950
sæ	1700	zæ	1700	dæ	1750	θæ	1650	ðæ	1650
sɔ	1150	zɔ	1200	dɔ	1350	θɔ	1050	ðɔ	1150
su	1600	zu	1550	du	1700	θu	1600	ðu	1500

Like its consonantal counterpart [grave], [flat] has diverse articulatory exponents in speech: labialization, velarization, pharyngealization, and retroflexion. We present diachronic evidence that dentals —as [grave]— are enhanced acoustically by all such incongruent articulations.

3.1 Labialization and velarization

Table 5 illustrates that dental consonants, which remain in South Slavey, have evolved into labiovelars in the North Slavey dialect centered in Deline, NT. This sound change also occurred in Hare, another Slavey dialect of NT, in Tlicho (also NT), and word-finally in Gwich'in (YK). As predicted, the [grave] feature of dentals was enhanced by the [flat] feature of labialization and velarization (and the dental gesture was eventually lost).

Table 5. Dentals > labiovelars in North Slavey, Deline, NT [2]

Slavey	Deline		Slavey	Del.	
t̪͡θʰe	kʷʰe	'rock'	-t̪͡θ'ɛ́hɛ́	-kʷ'é	'sinew'
-t̪͡θ'iəne	kʷ'ɛnɛ́	'bone'	-ðé	-wéʔ	'liver'

3.2 Pharyngealization

Table 6 illustrates that the dentalized sibilants of Proto-Athabascan which survive in some speakers of South Dakelh (BC) have all become pharyngealized ("emphatics") in adjacent Tsilhqot'in. In the sibilants, the decidedly weak [grave] tonality of the dental gesture was enhanced and eventually replaced by the [flat] tonality of tongue root retraction.

Table 6. Sibilants in Dakelh vs. Tsilhqot'in, BC [2]

Dakelh	Tsilhq.		Dakelh	Tsilhq.	
t͡s̪ʰet͡s̪ʰel	t͡s̪ʰĩ́ɬ	'axe'	jʌs̪	jəṣ	'snow'
t͡s̪ʰi	-t͡s̪ʰi	'head'	-jiz̪	-neẓ	'long'

Interestingly, Tsilhqot'in's own neighbor St'at'imcets Salish has pharyngealized coronal approximants /ẓ, ẓ'/ which are phonetically dental or interdental. (Arabic has a similar voiced continuant, called ð̣a:ʔ.)

3.3 Retroflexion

Retroflexion cannot enhance dentalization, as these gestures are incompatible. Revealingly, however, an interdental approximant /ð̣/ which occurs in disparate Philippine languages has evolved into a retroflex lateral /ɭ/ in Southern Kalinga, and a retroflex rhotic /ɻ/ in Madukayang Kalinga, Balangao, Mansaka and Upper Tanudan Kalinga. We assume that retroflexion came to substitute for interdentalization on the basis of a shared "low tonality": [flat] in /ɭ, ɻ/ and [grave] in /ð/. (An acoustic study of Kagayanen /ð̣/ confirms that it is not [flat]; its F_2 and F_3 are similar to those of an alveolar liquid [7].)

4. ASSIMILATION

That dentals are [grave] predicts that coronal consonants may become dental when released into a [flat] vowel or approximant. This is because "low tonality" in an approximant or vowel, viz. [flat], can be mistaken for "low tonality" in a preceding consonant, viz. [grave]—a kind of acoustic assimilation. This prediction is confirmed in the subsections below.

4.1 Back vowels

Table 7 illustrates that in the Australian language Lardil /t/ is realized as dental before /u, a/ (and as laminal-postalveolar before /i/). On our interpretation, the "low tonality" of [flat] in /u, a/ is assimilated into /t/ as [grave]/dental (and the "high tonality" of [sharp] in /i/ is assimilated into /t/ as [acute]/laminal-postalveolar).

Table 7. Coronal allophony in Lardil [4]

nom.	fut.	nonfut.	Acc.	
kal̠t̠it	kal̠t̠it̪-uɽ	kal̠t̠it̪-at	kal̠t̠it̠-in	'urine'
jaɽput	jaɽput̪-uɽ	jaɽput̪-at	jaɽput̠-in	'snake, bird'

4.2 Retracted vowels

A palatographic study of Kamwe (Afro-Asiatic) reveals that coronal consonants are alveolar or postalveolar when adjacent to advanced tongue root vowels, but dental when adjacent to retracted tongue root vowels [8]. A similar pattern occurs in Kalenjin (Nilo-Saharan) [9]. In our view, the "low tonality" of [flat] in retracted vowels, which may be considered pharyngealized, is assimilated into coronal consonants as [grave]/dental.

4.3 /ɻ, ɚ/

In Irish English, alveolar consonants can be realized as dental before /ɻ, ɚ/, which are retroflexed (and perhaps rounded) [10]. For instance, /t, d, n, l/ are dental in e.g. *train, spider, manner, pillar*. Again, on our interpretation, the "low tonality" of [flat] in retroflex/rounded /ɻ, ɚ/ is assimilated into coronal consonants as [grave]/dental (cf. [10]).

5. CONCLUSION

We have argued that the Jakobsonian feature [grave] does not require a predominance of low-frequency noise, but rather requires that the noise below 2.5 kHz is "sufficiently audible" owing to a lack of predominance of high-frequency noise. This effectively extends the reach of the feature, since all the noisy sounds which were classed as [grave] under the original definition still are—notably labials and velars. We have argued that non-sibilant dentals, too, are [grave]. On the one hand, their noise energy is very similar to that of labials. On the other hand, their interaction with the vocalic feature [flat] across languages strongly suggests that they bear the consonantal counterpart [grave].

REFERENCES

[1] R. Jakobson, G. Fant & M. Halle. (1952). *Preliminaries to speech analysis.* Cambridge, MA: MIT Press.

[2] D. Flynn [Howe] & S. Fulop. (2005). Acoustic features in Athabaskan. Paper presented at the Annual Meeting of the Linguistic Society of America, Oakland, CA, Jan. 9, 2005.

[3] P. Ladefoged & I. Maddieson. (1996). *The sounds of the world's languages*. Oxford, UK. Quotation from p. 173.

[4] K. N. Stevens, S. J. Keyser, & H. Kawasaki. (1986). Toward a phonetic and phonological theory of redundant features. In J. S. Perkell & D. H. Klatt (eds.), *Invariance and variability in speech processes*, 426-449. Hillsdale, NJ: Lawrence Erlbaum.

[5] S. N. Dart & P. Nihalani. (1999). The articulation of Malayalam coronal stops and nasals. *JIPA* 29, 129-142.

[6] J. Olive, A. Greenwood & J. Coleman. (1993). *Acoustics of American English speech: a dynamic approach*. New York: Springer-Verlag.

[7] K. Olson & J. Mielke. (2008). Acoustic properties of the interdental approximant in Kagayanen. *J Acoust Soc Am* 123(5), 3460.

[8] J. Laver. (1965). Some observations on alveolar and dental consonant-articulations in Higi. *Journal of West African Languages*, 2, 59-61.

[9] J. Local & K. Lodge. (2004). Some auditory and acoustic observations on the phonetics of [ATR] harmony in a speaker of a dialect of Kalenjin. *JIPA*, 34, 1-16.

[10] J. Harris. (2002). Dentals are dark: coronals in Irish English. Paper presented at the Ninth International Phonology Meeting, Vienna, Nov. 1-3, 2002.

ACKNOWLEDGEMENTS

Thanks to SSHRC (CE101-3663) and many consultants, e.g. Elsie Vital-Swanson, Valerie MacKeinzo (Deline, NT); Rosemary Andrew, Vivian Pellisey, Jessie Campbell (Tulita, NT); Violet Fabian (Hay River, NT); Elizabeth Enfield (Meander River, AB); Helen Metchooyaeah (Chateh, AB).

AEROTACTILE ACUITY AS A PREDICTOR OF SIBILANT CONTRAST

Naomi Francis[1], Jamie Ma[1], and Bryan Gick[1,2]
[1]Dept. of Linguistics, University of British Columbia, 2613 West Mall, Vancouver, BC, Canada, V6T1Z4 gick@mail.ubc.ca
[2]Haskins Laboratories, 300 George St., New Haven, CT, USA 06511

1. INTRODUCTION

Somatosensory acuity has been shown to correlate with produced acoustic contrast distance between /s/ and /ʃ/ (Ghosh et al., 2010). This effect is independent of auditory acuity, suggesting that speech sounds have independent somatosensory and auditory perceptual goals.

The term "somatosensory", however, applies to a broad range of different sensory modalities (Hsiao & Gomez-Ramirez, 2011). It is not known whether the different somatosenses contribute independently to determining production goals. One somatosense that has been shown to play an important role in speech, but that has an unknown effect on production goals, is the aerotactile sense (Gick & Derrick, 2009).

Aerodynamic properties have long been identified as defining characteristics of sibilant sounds, viz. the feature [strident] (Jakobson & Halle 1956). Indeed, the achievement of aerodynamic goals is often seen as the purpose of the tongue shapes adopted during the production of sibilants (Iskarous et al., 2011).

The present investigation replicates Ghosh et al.'s (2010) study of contact acuity, adding to this a comparison of the aerotactile modality. We hypothesize that aerotactile acuity is an independent predictor of produced contrast distance between sibilants.

2. METHOD

21 paid volunteers, all native speakers of English, participated in this study. Data were rejected from three subjects who scored below chance or reached ceiling, and from one subject because of technical problems; the data discussed in this paper therefore came from 17 participants, 10 male and 7 female, ranging in age from 20 to 58 years.

Participants performed three tasks, designed to test for each individual a) the contact acuity of the anterior tongue, b) the contrast distance – that is, the difference in spectral energy distribution – between /s/ and /ʃ/, and c) the aerotactile acuity of the anterior tongue. The contrast distance task was performed between the two acuity tasks. The order of the acuity tasks was reversed between participants so that half performed the contact acuity task first and half performed the aerotactile acuity task first; the order of tasks did not significantly affect results.

2.1 Methods for contact acuity

This portion of the experiment was a four-way forced-choice grating-orientation judgment task similar to the one described in Ghosh et al. (2010), with the following adjustments. Participants were seated comfortably with their chin in a chin rest, their eyes closed and their tongue flat and protruded. A domed probe with grooves spaced 1.75mm apart was pressed gently against the tongue with the grooves oriented at one of four different angles, exerting a pressure of approximately 94.3g/cm² for 500ms. Probe pressure was applied by a solenoid plunger activated by a sine wave signal for each trial. Participants responded with hand gestures indicating the orientation of the grooves and their responses were logged by the experimenter. Participants were free to take breaks during this phase of the experiment. There were a total of 40 trials, with each orientation appearing 10 times. The order of the trials was randomized.

2.2 Methods for sibilant contrast distance

This portion of the experiment was a production task replicating the methods of Ghosh et al. (2010) with the following adjustments. Participants were seated comfortably in front of a computer in a sound-proof booth and read sentences from the screen. The targets, "said" and "shed", were each repeated 15 times in a random order in the carrier phrase "He ___ a lot." These productions were recorded using the internal microphone of a Macintosh OS X computer at a sampling rate of 44100 Hz. The middle portion of each target word's sibilant was isolated in Praat. The centre of gravity, skewness and kurtosis of each sibilant production were calculated using Praat and the average value of each spectral moment was calculated for each sibilant. These values were used to calculate the average contrast distance (CD) as the Euclidean distance between /s/ and /ʃ/ in the three dimensional space defined by the three acoustic measures.

2.3 Methods for aerotactile acuity

This portion of the experiment was a two-way forced-choice task. Participants were seated comfortably in a sound-proof booth with their eyes closed and tongue flat and protruded. A piece of tubing connected to an air compressor was placed 5cm from the tongue surface. Participants sat with their back and head against a board to control distance from the tubing and listened to white noise over sound-isolating headphones to mask the sound of the air puffs. Puffs of air approximately 330ms long were delivered onto the tongue at two different strengths, a light puff of approximately 0.27 Pa and a strong puff of approximately 0.53 Pa, in random order. Participants were asked to correctly identify the strength of the puff and logged their responses on a keyboard. There were 80 trials in total, such that each puff strength appeared 40 times. The first 40 trials were treated as a practice section and were not analyzed further, leaving 40 experimental trials.

2.4 Statistical analyses

The contact acuity d-prime score (CA) and aerotactile acuity d-prime score (AA) were calculated for each subject in R (R Development Core Team, 2011) using the package psyphy (Knoblauch, 2012). Linear modeling and Pearson's product-moment correlation tests were performed using R (R Development Core Team, 2011).

3. RESULTS

Pearson's product-moment correlation between CD and AA was significant ($r=0.73$, $p<0.001$). In contrast, there was no significant correlation between the CD and CA ($r=-0.16$, $p=0.5516$), nor between the two acuity measures ($r=0.02$, $p=0.9428$).

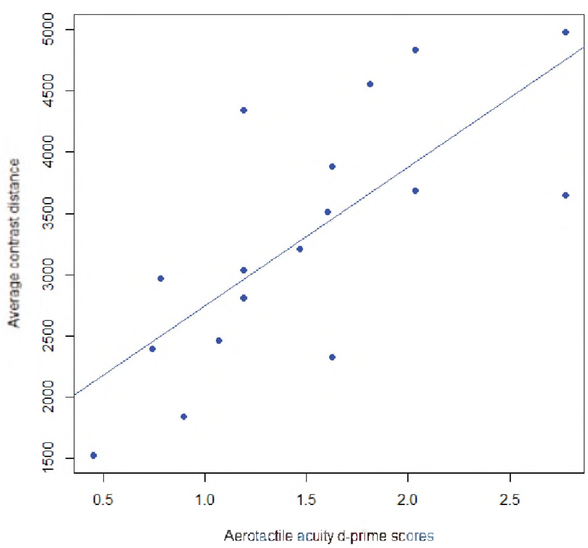


Figure. 1. Contrast distance vs. aerotactile acuity

Multiple linear regression shows that AA was a significant predictor of CD ($\beta=2015.4$, $t=3.009$, $p=0.0101$) while CA ($\beta=826.6$, $t=-1.421$, $p=0.1788$) and the interaction between the two acuities were not ($\beta=-914.2$, $t=-1.421$, $p=0.1788$). Simple linear regression of the relationship between AA and CD shows that aerotactile acuity was a significant predictor of contrast distance ($\beta=1134.6.4$, $t=4.181$, $p<0.001$, adjusted $r^2=0.51$), while linear regression of the relationship between CA and CD shows that contact acuity was not ($\beta=-248.2$, $t=-0.609$, $p=0.5516$, adjusted $r^2=0.51$).

We initially set out to replicate Ghosh et al's (2010) results for the contact modality. In the overall analysis, contact acuity results were not significant. Following Ghosh et al. (2010), we performed a median split based on participants' acuity scores and repeated the correlation tests. There was a small, though not statistically significant, tendency toward a positive correlation between CA and CD in the low acuity group ($r=0.53$, $p=0.1461$). With more data or a more rigorous replication of Ghosh et al. (2010)'s contact acuity protocol, this tendency would likely be significant.

4. DISCUSSION

Our results indicate that speakers' aerotactile acuity is an independent predictor of their produced contrast distance. This is perhaps unsurprising given the high volume of airflow that passes over the anterior portion of the tongue during the production of sibilants. Nevertheless it indicates that aerotactile feedback is important for producing sibilant contrasts.

The fact that multiple sensory factors play independent roles in speech allows for the possibility that any modality could reasonably play an independent role. For example, the work by Ménard et al. (2009) on the speech of blind and sighted individuals indicates that the visual modality is also an independently important factor in speech production. Thus, speech production can perhaps be best understood as an interplay between many independent modalities.

REFERENCES

Ghosh, S., Matthies, M., Maas, E., Hanson, A., Tiede, M., Ménard, L., … Perkell, J. (2010). An investigation of the relation between sibilant production and somatosensory and auditory acuity. *JASA, 128*(5), 3079-3087.

Gick, B., & Derrick, D. (2009). Aero-tactile integration in speech perception. *Nature, 462*, 502-504.

Hsiao, S., & Gomez-Ramirez, M. (2011). Touch. In J. A. Gottfried (Ed.), *Neurobiology of sensation and reward* (141-160). Boca Raton, FL: CRC Press.

Knoblauch, K. (2012). psyphy: Functions for analyzing psychophysical data in R (Version 0.1-7) [R package]. Available from http://CRAN.R-project.org/package=psyphy

Ménard, L., Leclerc, A., Brisebois, A., Aubin, J., & Brasseur, A. (2009). Production and perception of French vowels by congenitally blind adults and sighted adults. *JASA, 126*(3), 1406-1414.

Jakobson, R., & Halle, M. (1956). *Fundamentals of language*. The Hague: Mouton.

Iskarous, K., Shadle, C., & Proctor, M. (2011). Articulatory-acoustic kinematics: The production of American English /s/. *JASA, 129*(2), 944-954.

R Development Core Team (2011). R: A language and environment for computing. Vienna, Austria: R Foundation for Statistical Computing. Available from http://www.R-project.org/.

Tremblay, S., Schiller, D., & Ostry, D. (2003). Somatosensory basis of speech production. *Nature, 423*, 866-869.

ACKNOWLEDGEMENTS

Many thanks to Dr. D. W. Francis for his engineering help in the construction and repair of our apparatus, to Kate Komova for her work in analyzing our data, and to Michael McAuliffe for his guidance regarding the statistical tests used in this paper. Research funded by an NSERC Discovery grant to the third author.

FROM QUANTAL BIOMECHANICS TO WHOLE EVENTS: TOWARD A MULTIDIMENSIONAL MODEL FOR EMERGENT LANGUAGE

Bryan Gick[1, 2]

[1]Dept. of Linguistics, University of British Columbia, 2613 West Mall, Vancouver, BC, Canada, V6T1Z4 gick@mail.ubc.ca

[2]Haskins Laboratories, 300 George St., New Haven, CT, USA 06511

1. INTRODUCTION

Nature is replete with nonlinearities. In speech, the simple act of closing the lips corresponds with a flood of mechanical and sensory nonlinearities: visual, kinematic, somatosensory, aerodynamic, acoustic-auditory, and so on. The term “quantal” has been applied to a subset of nonlinear effects in speech that help in achieving some auditory-perceptual goal (Stevens 1972, 1989). Stevens & Keyser (2006) define quantal effects as discontinuities "between the displacement of an articulatory parameter and the acoustical attribute that results from this articulatory movement," arguing that these effects provide the basis for phonological categories. Nonlinear effects fitting this definition of quantal thus share the three properties of: (a) originating in articulation, (b) having demonstrable auditory-perceptual implications, and (c) being advantageous in the communicative process.

According to this view, some of the nonlinearities described above for lip closure qualify as "quantal" while others – specifically those with no direct articulatory-auditory link – do not. However, many studies in the last decade have identified effects that are important in speech but that are independent of auditory-perceptual goals. For example, Tremblay et al. (2002) show that jaw movements associated with speech-related constrictions maintain their own proprioceptive trajectories. Likewise, Ménard & al. (2009) found that sighted speakers use lip movements differently from blind speakers for speech. Similar work has shown independent effects in aerodynamic (Gick & al. 2012), tactile (Gick & al. 2008) and aerotactile (Gick & Derrick 2009, Francis & al. this volume) dimensions.

In this model, speakers do not aim to achieve “movement goals/targets” as defined in physical, auditory-perceptual or somatosensory space, but rather “whole events” in a fully multidimensional space. The present paper demonstrates how a biomechanical nonlinearity (as described in Gick & al. 2011; see below) can be translated into a multidimensional density map, creating a topology of relative densities. Within this space, those nonlinearities that facilitate the production of ecologically successful events function as attractors to behavior. These “facilitiative nonlinearities” become the common currency of human ecological space, and coupled with an iterative learning simulator, are the basis for an emergent model of phonology or other patterned behavior.

2. VOCAL TRACT SPHINCTERS AS QUANTAL BIOMECHANICAL DEVICES

Previous models of speech articulation have been based predominantly on spatial targets in a 2-dimensional (midsagittal) vocal tract space. The present paper describes an approach to speech articulation based on a view of vocal tract constrictors as physiological sphincter mechanisms. In this model, a particular constrictor (sphincter) can produce only one kind of constriction, and it does so using the inherent quantal biomechanical properties of sphincters. Sphincters in this model differ from articulators in previous models in that they: (a) generate constrictions rather than targeted motions; (b) constrict from multiple directions in 3 spatial dimensions; (c) are functionally idiosyncratic to fixed locations in the vocal tract; (d) are highly constrained in degrees of freedom; (e) are fully predictive of location, degree, articulator and shape; (f) provide no inherent hierarchical structure; (g) exhibit robust mechanical nonlinearities (quantal effects) that generate discrete output from continuous, noisy input; (h) enable entirely feed-forward control; and (i) function as attractors to behavior in a multidimensional behavior space.

Figure 1 shows an example of three biomechanical nonlinear regions (highlighted) corresponding to quantal-like regions in the actions of the three lip sphincters used to produce labial approximants, fricatives and stops (adapted from Gick & al. 2011, which gives a more detailed description of this graph). These regions represent regions of biomechanical efficiency.

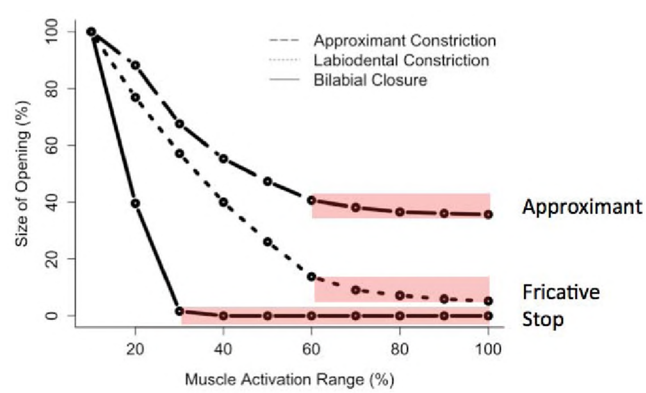


Figure. 1. Facilitative nonlinearities (highlighted) in lip opening size as a function of muscle activation.

3. GETTING FROM QUANTAL REGIONS TO MULTIDIMENSIONAL SPACE

Figure 2 shows how a potentially advantageous nonlinearity (shown in Figure 1) can be reduced in dimensionality in order to be incorporated into a multidimensional functional space. This process reduces the 2-dimensional size-by-activation space of Figure 1 to a single "biomechanical advantage" dimension, with three attractor regions based on the slope of the Figure 1 curves.

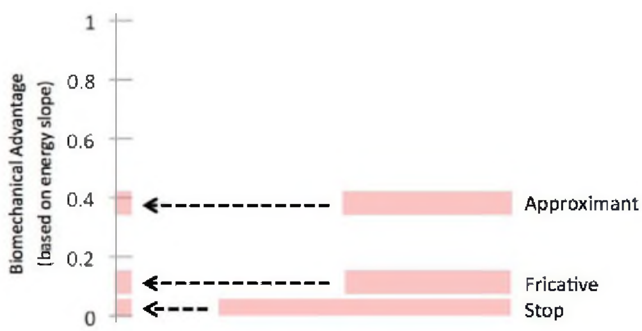


Figure. 2. Dimensionality reduction of lip data.

Figure 3 shows the lip data in the biomechanical advantage dimension now plotted in a multidimensional space, along with hypothetical "auditory advantage" and "visual advantage" dimensions. In this density- or heat-map plot, high-advantage areas are plotted as dense or hot regions.

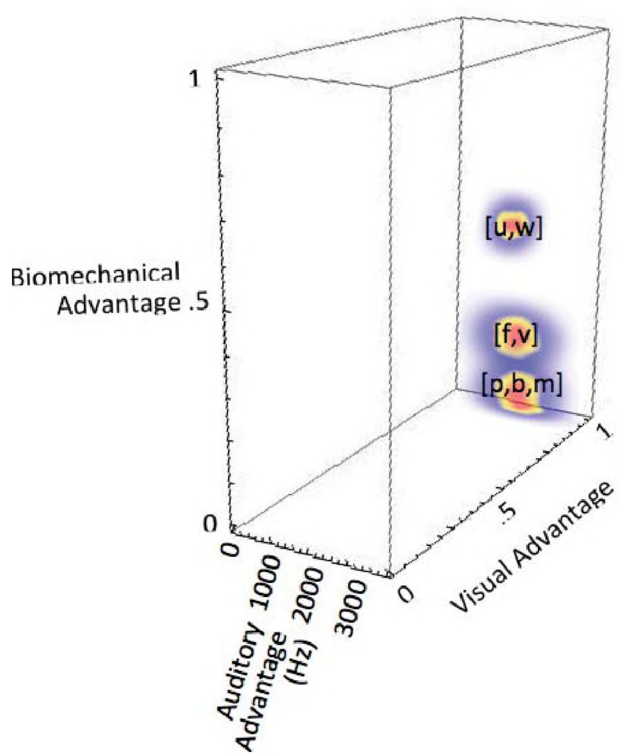


Figure. 3. Labial sound categories plotted in 4-D space.

4. DISCUSSION

The present paper is part of a research program whose aim is to identify a much wider range of potentially advantageous nonlinearities, and explore their role in an emergent model of phonetics and phonology. This work models the operational space in which spoken communication takes place as a fully multidimensional (physical, social, ecological) "whole event" space. Within this space, behavior is attracted to nonlinearities that afford users communicative advantage, allowing previously unrelated dimensions of human behavior to interact in a single functional space. The common currency of this space is communicative advantage, and it is only by offering demonstrative communicative advantage that a behavior may be adopted as exerting an effect. This approach interfaces with previous models that use nonlinear dynamical systems to simulate emergent language (e.g., Smith & Thelen 2003, Gafos 2006, etc.), models that have sought nonlinear relationships as important keys to understanding speech and language (Perkell 2010), and other empirically based, multimodal approaches to emergence of language (Mielke 2011).

REFERENCES

Gafos, A.I. and Benus, S. (2006). Dynamics of Phonological Cognition. *Cognitive Science 30*, 1-39.

Gick, B., Chiu, C., Flynn, C., Stavness, I., Francis, N. and Fels, S. (2012). Producing whole speech events: Differential facial stiffness across the labial stops. *Proc. of Acoust. 2012 HK*. Ppr 2pSC1, 6pp.

Gick, B. and Derrick, D. (2009). Aero-tactile integration in speech perception. *Nature*. *462*. 502-504.

Gick, B., Jóhannsdóttir, K., Gibraiel, D. and Muehlbauer, J. (2008). Tactile enhancement of auditory and visual speech perception in untrained perceivers. *JASA* 123(4), EL72-76.

Gick, B., I. Stavness, C. Chiu and S. S. Fels. (2011). Categorical variation in lip posture is determined by quantal biomechanical-articulatory relations. Proceedings of Acoustics Week in Canada 2011. *Canadian Acoustics* 39(3).

Mielke, J. (2011). A phonetically based metric of sound similarity. *Lingua*. doi:10.1016/j. lingua.2011.04.006.

Ménard, Lucie, Sophie Dupont, Shari R. Baum, & Jérôme Aubin. (2009). Production and perception of French vowels by congenitally blind and sighted adults. *JASA*, 126 (3) 1406-1414.

Perkell, J.S. (2012). Movement goals and feedback and feedforward control mechanisms in speech production. *J. Neurolinguistics*. *25*:382-407.

Smith, L.B. and E. Thelen. Development as a dynamic system. TiCS 7, 343-348.

Stevens, K. N. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. In E. David & P. Denes (Eds.), *Human communication: A unified view*. NY: McGraw-Hill. 51-66.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics* , 17, 3–4.

Stevens, K. N. and Keyser, S. J. (2010). Quantal theory, enhancement and overlap. *Journal of Phonetics*, 38, 10-19.

ACKNOWLEDGEMENTS

The author acknowledges the many collaborators who have contributed to this work, and funding from NSERC.

Familiar Talker Advantages in Formant-Based and Concatenative Synthetic Speech

Jacqueline Jones
Dept. of Linguistics, University of Calgary, 2500 University Drive NW, Alberta, Canada, T2N 1N4, jmjone@ucalgary.ca

1. INTRODUCTION

Access to synthetic speech technology has never been easier than it is today. Home computers come bundled with text-to-speech software, as do some eReaders and smart phones. The technology has come a long way since Stephen Hawking's recognizable DECTalk voice in the late 1980s. And yet despite synthetic speech's increased intelligibility, decreased cost, and lessened reliance on bulky equipment, there is still a recognizable peculiarity in synthetic speech. While extensive research has been done on both synthetic speech perception and familiarity effects, no study has yet examined whether synthetic voices are treated as individual voices, or whether they are perceived by the listener as a broad category of "synthetic speaker" that encompasses all voices produced in the same manner.

It has been established that listener perception of synthetic speech improves with training, but past research has also identified limits to this improvement (Nygaard et al, 1998; Greenspan et al, 1988). The perceptual benefits listeners gain from being familiar with a speaker in natural speech is known as the "Familiar Talker Advantage."

This research examines this familiar talker advantage and how it relates to synthetic speech by utilizing two kinds of synthetic speech. In this study, listeners were trained to identify words produced by four different synthetic voices, created using two different synthesizing processes.

2. METHOD

Participants were trained with a synthetic speaker produced by one of two types of synthesis (formant-based and concatenative) via a series of sentence transcription tasks. Participants were then tested on their ability to transcribe sentences produced by novel synthetic voices to determine if the training generalized across speakers and synthesis types. Finally, participants completed a post-test phase to examine the retention effects of any benefits from training.

2.1 Participants

24 young adult monolingual speakers of Canadian English participated in these experiments. Two participants were removed from the results for problematic answers or manipulation of testing equipment during testing. All participants were recruited from the University of Calgary's introduction to linguistics class in exchange for marks towards their research participation requirement for that class. Prior to starting the research, participants were asked to confirm that they were native English speakers and asked to provide a self-assessment of their familiarity with synthetic speech. None reported any speech or learning deficit, and all reported having normal hearing and minimal exposure to synthetic speech.

2.2 Stimuli

Four sets of stimuli were created for this research. The threshold stimuli consisted of 70 pre-recorded spondees – bisyllabic words with equal stress on both syllables – produced by a native speaker of American English. Training, Testing, and Post-Test stimuli consisted of pre-recorded sets of Harvard Sentences produced by synthetic speech. The formant voices were produced from presets in the eSpeak text-to-speech freeware program, version 1.45.05. The concatenative voices used were from the Ivona Voices commercial software suite. The names chosen from the Ivona Library were Joey and Eric.

2.3 Procedure

Participants were placed randomly into four groups.

During the threshold phase, all four groups listened to 70 pre-recorded spondees produced by a natural speaker of American English. The volume the tokens were played at was progressively reduced. The volume at which participants were achieving correct answers 50% of the time was used for all further phases of the experiment.

The training phase consisted of 60 sentences produced by a synthetic speaker. Groups 1 and 2 trained with the concatenative voice Eric, and groups 3 and 4 trained with formant voice Wheatley. The participants listened to the sentence a single time and were asked to transcribe what they heard. The sentence was then repeated, and feedback was given in the form of a transcription displayed on screen. The response and response time were recorded.

In the test phase, immediately after the training phase, participants were asked to transcribe 20 Harvard sentences. For two groups, 10 of the sentences were produced using the same synthetic speaker as presented in training, while 10 were produced by a voice which used a different type of synthesis. The other two groups were tested with 10 sentences presented by a voice that differed in synthesis type, and 10 sentences produced by a different synthetic speaker that shared a synthesis type with the voice used in training.

The post-test phase took place between three and five days after the initial training and testing. In the post-test phase, participants were presented with 20 novel Harvard sentences, of the same type and distribution as those used in the test phase.

3. RESULTS

Three ANOVAs were run to determine what effects trained voice, stimulus type, and test number had on the percentage of unique content and function words correctly identified, and response times. Post-hoc t-testing between individual pairs of variables was used to identify the directions of the significant interactions. Please contact the author for the complete results of the experiment.

Overall, participants identified individual words in testing correctly significantly ($p < 0.001$) more often when trained with the Formant-based synthesized voice Wheatley, regardless of stimulus type or the properties of the individual words (content or function).

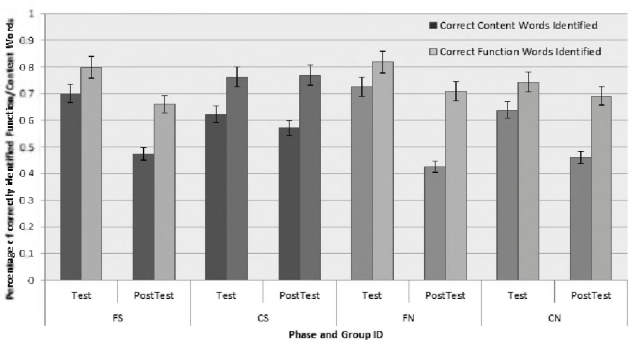


Figure 1. Word Identification scores across groups and phases.

Figure 1 shows word identification scores across all groups during the testing and post-testing phase. Darker colours indicate content word identification and lighter colours represent function word identification. Overall performance is worst during the post-test across all groups, and performance by those trained with formant-based voices is significantly better ($p < .001$). Those trained with formant voices do worse in the post-test than their counterparts trained with concatenative voices ($p < .0001$).

Function words were correctly identified more often regardless of any other factor. In examining individual results, when function words were incorrect, they were often transcribed as other function words – 'a' transcribed as 'the', for example.

Overwhelmingly, participants performed better in test one than in test two. Further refinement of the methodology is necessary to determine whether this drop in performance is a significant indication of failure to retain benefits gained in training over time.

4. DISCUSSION

The results of this research show support for the existence of a familiar talker advantage in synthetic speech. The results also show evidence of the existence of a familiar synthesis type advantage. The group trained with Eric and tested with Eric and Joey had better performance with the similar voice (Joey) than with the voice they were trained with (Eric). This could indicate a between-type perceptual benefit to concatenative synthesis. However, it is not possible to determine the extent to which this performance boost is due to a between-type perceptual benefit, a combination of Joey's novelty causing an increase in attention of the participant, or even whether Joey is perceived as a naturally more intelligible voice.

The results suggest that listeners gain a perceptual benefit from familiar synthetic talkers, regardless of synthesis type, in a similar way to the benefit gained from natural familiar talkers. Training with a concatenative synthetic speaker can potentially grant a “synthesis type familiarity” advantage, but the exact nature of this advantage cannot be determined from the current results, as there are too many variables between the acoustic properties of the two synthesis types to confidently determine which are causing the perceptual benefit.

Participants trained with the less intelligible formant-based synthesis voice had overall better performance with all synthesized voices presented to them. A perceptual benefit appears to emerge across synthesis types, but only in one direction and if the training is relatively intense. This may have an analogue in the research done on bilingual talkers in English and German, where listeners had more success with unfamiliar languages. (Winters et al, 2008). Effectively, this means that the more intelligible synthesized voices do not transfer any kind of familiar talker advantage to less intelligible voice, but that training with less intelligible voices may grant a perceptual benefit to synthetic speech in general.

REFERENCES

Greenspan, S. L. et al (1988). Perceptual learning of synthetic speech produced by rule. *Journal of Experimental Psychology*. 14, 421-433.

Nygaard, L. et al (1998). Talker-specific learning in speech perception. *Attention, Perception, & Psychophysics*. 60, 355-376.

Winters, S. J., et al (2008). Identification and Discrimination of Bilingual Talkers across Languages. *The Journal of the Acoustical Society of America*. 123, 4524.

AUTHOR NOTES

This research was conducted while the author was an undergraduate student at the University of Calgary. The author can be contacted at jmjone@ucalgary.ca.

LINGUOPALATAL CONTACT DIFFERENCES BETWEEN JAPANESE GEMINATE AND SINGLETON STOPS

Alexei Kochetov
Dept. of Linguistics, University of Toronto, 100 St. George St., Toronto, Ontario, Canada, M5S 3G3
e-mail: al.kochetov@utoronto.ca

1. INTRODUCTION

Japanese is known to exhibit a phonemic contrast between geminate (long, moraic) and singleton consonants (e.g. *itta* 'went' vs. *ita* 'was'; Vance, 1987). This contrast has received considerable attention in phonetic literature, with most studies focusing on its most salient, durational properties – differences between geminated and singletons in the duration of the consonant and the preceding or following vowel (see Kawahara, to appear, for an extensive review). A few recent studies have extended the analysis to non-durational acoustic correlates, such as pitch, amplitude, and creakiness of adjacent vowels (Kawahara, 2006; Idemaru & Guion, 2008). In one of these works, Idemaru & Guion (2008) found that Japanese geminate stops tend to be produced with creaky voice, suggesting a tighter, fortis-like constriction compared to the weaker constricted (lenis) singletons (cf. Sakuma, 1929/1963 cited in Vance, 1987). Previous research on similar contrasts in other languages has revealed that differences in duration tend to be accompanied by differences in the degree of linguopalatal contact (e.g. Payne, 2006 on Italian).

Until now, however, there has been little articulatory work on the contrast, and the differences in degree of contact have not been directly investigated. This paper presents some preliminary results of an electropalatographic (EPG) study of Japanese geminate and singleton stops /t: k:/ vs. /t k/ evaluating the hypothesis that geminates are characterised by both longer duration and tighter constriction.

2. METHOD

The participants were 3 females, native speakers of Japanese from Shizuoka (JF1), Shiga (JF2), and Ibaraki (JF3), at the time of the study residing in Toronto, Canada. The materials included 4 segmental minimal pairs with voiceless geminate and singleton coronal and velar stops shown in Table 1. The words were randomized and presented in a carrier phrase *kabeni wa ___ mo kaite aru* '___ is also written on the wall'. Nine repetitions of each utterance were elicited, resulting in the total of 216 tokens (8 words * 9 repetitions * 3 speakers).

Electropalatographic (EPG) data were collected using custom-made artificial palates with 62 built-in electrodes and a WinEPG system by *Articulate Instruments* (Wrench et al., 2002). The EPG method provides information about the contact between the tongue and the palate over time, at a sampling rate of 100 Hz. Analysis was performed using the *Articulate Assistant* software (Wrench et al., 2002) and involved measurements of the stop closure duration (in seconds) and the degree of the contact ('Whole Total': the number of activated electrodes divided by the total number of electrodes) taken at the point of maximum contact. Statistical analysis involved t-tests with minimal word pairs performed separately for duration and degree of contact, followed by correlations between the two measures (all done separately for each participant).

Table 1. Words used in the study.

Contrast	Geminate	Singleton
/t:/ vs. /t/	*matte* 'wait' (request)	*mate* 'wait' (imper.)
	hetta 'reduce' (past)	*heta* 'unskilled'
/k:/ vs. /k/	*sekki* 'stone tools'	*seki* 'seat'
	shikke 'stormy weather'	*shike* 'humidity'

3. RESULTS

The results showed that geminate stops were about twice as long as their singleton counterparts, as shown in Figure 1. On average, the duration of single /t/ and /k/ was 68 ms and 65 ms, while the duration of their geminate counterparts was 143 ms and 147 ms, respectively. The geminate/singleton differences were highly significant ($p < .001$) for all 3 speakers, and for both coronals and velars. More interestingly, geminates showed on average greater degree of linguopalatal contact compared to singletons, as illustrated in Figure 2. These differences were relatively small, being on average .64 vs. .60 for coronals and .63 vs. .57 for velars. Nevertheless, the differences were significant for most of the comparisons: /t:/ vs. /t/ for JF1 ($p < .01$) and JF3 ($p < .001$), and /k:/ vs. /k/ for JF1 ($p < .05$), JF2 ($p < .001$), and JF3 ($p < .01$). There were also some differences in the overall degree of contact among the speakers (JF1 > JF2 > JF3), possibly reflecting individual anatomical differences in the shape of the palate. Correlations between duration and degree of contact (see Figure 3) were significant for all 3 speakers (JF1: $r(72) = .33$, $p < .01$; JF2: $r(72) = .36$, $p < .01$; JF3: $r(72) = .58$, $p < .001$), indicating that longer duration implies greater degree of contact.

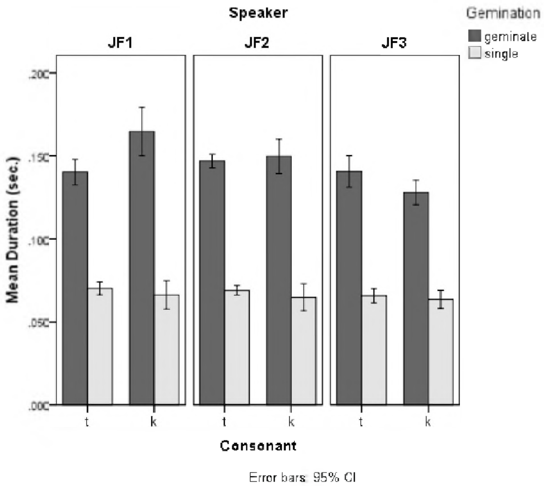


Figure 1. Mean duration (in seconds) of geminate and singleton coronal and velar stops separately by speaker.

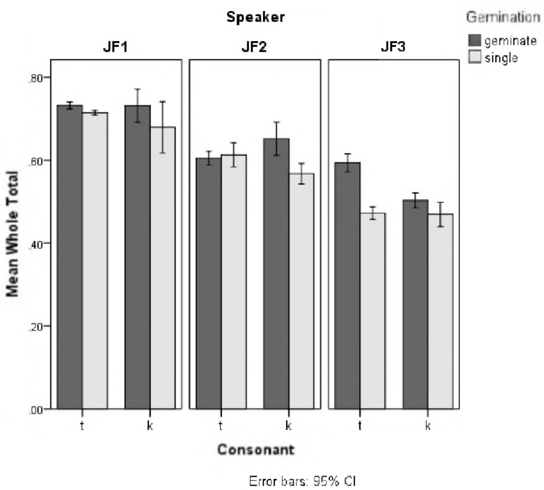


Figure 2. Mean degree of contact (Whole Total) of geminate and singleton coronal and velar stops separately by speaker.

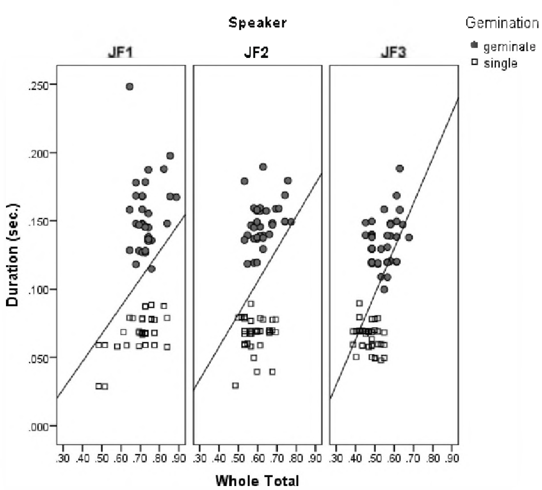


Figure 3. Correlations between duration (in seconds) and degree of contact (Whole Total) for individual tokens, separately by speaker.

4. DISCUSSION AND CONCLUSION

The results of this preliminary study confirm the hypothesis that durational differences in the production of Japanese geminate/singleton stops can be accompanied by differences in the degree of contact. The longer geminates tend to be articulated with a tighter constriction, compared to the shorter and weaker-articulated singletons. The results thus confirm impressionistic observations made by Sakuma (1929/1963) and predictions based on acoustic analysis (vowel creakiness next to geminates) made by Idemaru & Guion (2008). They also show the Japanese geminate/singleton contrast shares some key phonetic properties with similar contrasts in other languages, like Italian (Payne, 2006). More work is necessary to determine whether the results are representative of spoken Japanese in general and can be extended to other segmental/prosodic contexts and to other consonants (e.g. geminate/singleton affricates, fricatives, and nasals). Similarities and differences between languages in the phonetic realization of the geminate/singleton contrast also require further investigation, as these may clarify the nature of the relation (automatic or learned) between duration and degree of contact.

To conclude, the current study presented some evidence for the role of degree of contact in the realization of the Japanese geminate and singleton stops, contributing to the growing body of work on non-durational acoustic and articulatory correlates of the contrast.

REFERENCES

Idemaru, K., & Guion, S. (2008). Acoustic covariants of length contrast in Japanese stops. *Journal of International Phonetic Association*, 38, 167–186.

Kawahara, S. (2006). A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language*, 82, 536–574.

Kawahara, S. (to appear). The phonetics of obstruent geminates, *sokuon*. *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*. Mouton.

Payne, E.M. (2006). Non-durational indices in Italian geminate consonants. *Journal of the International Phonetic Association* 36, 83–95.

Sakuma, K. (1929). *Nihon Onseigaku*. [Reprinted 1963. Tokyo: Kazama Shobō]

Vance, T.J. (1987). *An introduction to Japanese phonology*. New York: SUNY Press.

Wrench, A.A., Gibbon, F.E., McNeill, A.M. & Wood, S.E. (2002). An EPG therapy protocol for remediation and assessment of articulation disorders. In J.H.L. Hansen & B. Pellom (eds.), *Proceedings of ICSLP-2002* (pp. 965–968). Denver, CO.

ACKNOWLEDGEMENTS

Thanks to Maho Kobayashi for assistance with the design of the materials and to the audiences at the National Institute for Japanese Language and Linguistics, University of Aizu, Hanyang University, and Hiroshima University for feedback on earlier versions of the paper. The work was supported by SSHRC.

ANALYSIS OF TONGUE SHAPES DURING THE PRODUCTION OF KANNADA CONSONANTS

Alexei Kochetov[1], Sreedevi N.[2], and Midula Kasim[2]
[1]Dept. of Linguistics, University of Toronto, 100 St. George St., Toronto, Ontario, Canada, M5S 3G3
e-mail: al.kochetov@utoronto.ca
[2]Dept. of Speech Sciences, All India Institute of Speech and Hearing, Manasagangothri, Mysore, Karnataka, India, 570006

1. INTRODUCTION

Like many languages of South Asia, Kannada (Dravidian) exhibits a relatively complex set of lingual consonants that differ in place of articulation – dental, retroflex, alveolopalatal, and velar. Among these, retroflex consonants are particularly interesting, being produced with the tongue tip curled behind the alveolar ridge towards the hard palate. While a number of articulatory studies have examined retroflexes in South Asian languages (e.g. Švarný & Zvelebil, 1955; Narayanan *et al.*, 1999), few have systematically compared them to other lingual articulations. Moreover, there has been hardly any articulatory work on Kannada, in contrast to studies of consonants in other Dravidian languages such as Tamil and Telugu. The goal of this paper is to examine how the tongue shape for Kannada retroflex stop /T/ (IPA /ʈ/) differs from the shapes for alveolopalatal affricate /c/ (/ʧ/), dental stop /t/, and velar stop /k/. The data come from our earlier ultrasound study of Kannada lingual consonants produced by 4 native speakers of the language (Kochetov, Sreedevi, & Kasim, to appear). The focus of that study was on the relative displacement of the tongue in the production of these consonants, calculated as the difference between the tongue shapes during the consonant closure and the rest position. It was found that the retroflex stop and the palatal affricate were characterized by greater displacement compared to the dental and velar stops, suggestive of the greater articulatory complexity of the former two.

What was not directly investigated in that study is how exactly the retroflex tongue shapes differ from those of other places of articulation or from the rest position. That is, do the differences involve the entire tongue contour or only part of it (e.g. the tongue front or the anterior tongue body)? This question is important in light of the controversy about the role of the tongue body in retroflex articulations. Some researchers have proposed that retroflexion is accompanied by the tongue body backing – velarization or pharyngealization, which is either optional (Bhat, 1974) or obligatory (Hamann, 2003). This was predicted largely on phonological grounds, as backing in retroflexes would account for their cross-linguistic patterning with back vowels. Other researchers, based on some articulatory data, have argued that retroflexes require the tongue body 'bracing' or stabilization – to facilitate the palatal constriction and the characteristic forward movement of the tongue tip (Narayanan *et al.*, 1999; Best *et al.*, 2010). In the current study we address this question by performing statistical analyses of pairs of consonant tongue shapes, with the goal to determine to whether they differ in specific regions of the tongue, and in which direction.

2. METHOD

As mentioned earlier, the study used ultrasound data from Kochetov *et al.* (to appear), where 2 female (F1, F2) and 2 male speakers (M1, M2) produced 6 repetitions of the following Kannada words: /aTTa/ 'garret', /acca/ 'pure', /atta/ 'that side', and /akka/ 'elder sister'. Geminate consonants were used to obtain the duration adequate for the analysis (given the 15 frames per second rate of the system). The data were recorded using a PI 7.5 MHz *SeeMore* ultrasound probe by *Interson*, connected through a USB port to a laptop computer and captured by a DVD recorder. Video frames were selected for a number of time points, and tongue contours were traced using *EdgeTrak* (Li *et al.*, 2005). Each traced contour was saved as a set of 100 X and Y points. Figure 1 plots the resulting contours for 6 tokens of each of the 4 consonants (taken during the closure) produced by one of the speakers (M1). The front of the tongue is on the right. Note that ultrasound does not usually capture the tongue tip proper, and particularly when it is curled back (as for /T/). It is clear from the figure that there is considerable variation within each consonant category, while there are also differences among the consonants.

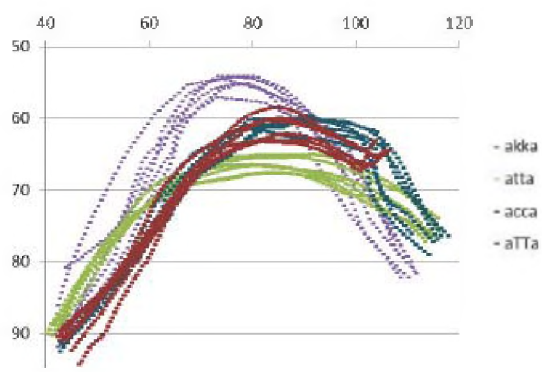


Figure 1. Traced tongue contours (in mm) for consonants produced by speaker M1 (6 tokens per consonant).

These differences were evaluated using a series of Smoothing Spline Analyses of Variance (SS-ANOVAs), the method that compares smoothing splines for two datasets and determines whether they are significantly different from each other (see Davidson, 2006 for details). The two curves are considered significantly different in a given region if their 95% confidence intervals do not overlap. For the purposes of the analysis, the tongue contour was divided into 3 regions – the posterior and anterior tongue body, and the tongue front (the blade and the tip, if visible). The input to the analysis were X and Y points for the most extreme tongue position during the consonant closure (point 0) for pairs of words with retroflex /T/ (/aTTa/) and the other consonants (/acca/, /akka/, /atta/). In addition, a comparison was made between point 0 and the point at 10 frames prior to it (-10; or 333 ms before) for /aTTa/, which corresponded to the neutral position of the tongue (the rest position). The analysis was performed using the *assist* package of the *R* programming language (version R 2.14.1; www.r-project.org/). Overall, four SS-ANOVAs were run for each participant, with six tokens for each consonant, and each token being based on 100 X and Y data points.

3. RESULTS

Figure 2 presents the output of SS-ANOVAs for one of the speakers, M1. It can be seen in the upper left image that the curves for retroflex /T/ and alveolopalatal /c/ are almost identical in the posterior and part of the anterior tongue body regions (the leftmost 2/3). The difference between the two is mostly in the tongue front, which is raised for /T/ (and partly invisible due to the sound wave refraction) and lowered for /c/. Compared to velar /k/ (upper right), the entire tongue for /T/ is considerably fronted, with the anterior tongue body being in a lower position. The posterior tongue body fronting for /T/ is also notable when compared to dental /t/ (lower left), which has an overall lower and more stretched tongue shape. Finally, compared to the rest position, /T/ shows considerable fronting of the tongue body, and raising of it anterior part, together with the tongue front (which is presumably curled back).

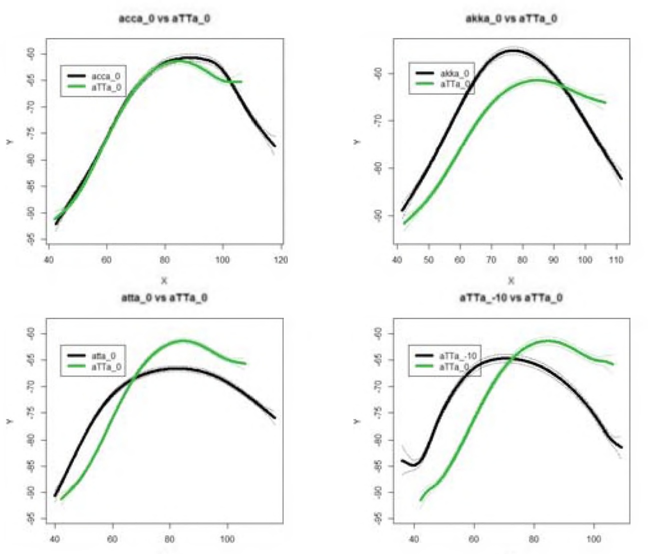


Figure 2. Results of SS-ANOVA for 4 pairs of tongue contours (in mm) based on the data in Figure 1 (speaker M1).

The results for the other speakers were overall similar. Specifically, the retroflex posterior tongue body was more front than for /k/ and the rest position, and similar to /c/. The anterior tongue body for /T/ was lower than for /k/, but higher than for /t/ and the rest position, and either slightly lower or higher than for /c/. The tongue front for /T/ was substantially higher than for the other articulations, except for /c/. In general, the retroflex and the alveolopalatal were most similar in terms of their tongue shapes, differing primarily in the direction of the tongue tip – up or down.

4. DISCUSSION AND CONCLUSION

The results of the statistical comparison of tongue shapes for Kannada lingual consonants are parallel to our earlier findings of displacement differences based on the same data. Posterior coronals, /T/ and /c/, have similar tongue body shapes and involve the greatest displacement from the rest position. The finding of the tongue body fronting for the retroflex stop, however, is unexpected given the prediction that the curling of the tongue tip should involve backing of the tongue body (Bhat, 1974; Hamann, 2003). More research is necessary to determine whether the results are representative of Kannada retroflex articulations in general. Yet, this finding appears to be compatible with earlier X-ray and MRI results for Tamil (Švarný & Zvelebil, 1955; Narayanan *et al.,* 1999), which found that the pharyngeal cavity was wider for retroflexes (and hence the more front tongue body) than for dentals or the rest position. At the very least, this suggests that the tongue body backing is not an obligatory property of retroflexion. In fact, the articulator can move in the opposite direction, likely as part of the stabilization phase facilitating the formation of the tip-up constriction and the subsequent tip forward movement, which is consistent with the observations made by Narayanan *et al.* (1999) and Best *et al.* (2010).

REFERENCES

Best, C.T., Bundgaard-Nielsen, R.L., Kroos, C., Harvey, M., Baker, B., Goldstein, L. & Tiede, M. (2010). How does a language contrast 4 distinct coronal stop places? Poster presented at LabPhon12, Albuquerque NM, 8–10 July.

Bhat, D.N.S. (1974). Retroflexion and retraction. *Journal of Phonetics,* 2, 233–237.

Davidson, L. (2006). Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *Journal of the Acoustical Society of America*, 120, 407–415.

Hamann, S. (2003). *The phonetics and phonology of retroflexes.* Utrecht: LOT.

Kochetov, A., Sreedevi, N., & Kasim, M. (to appear). A preliminary ultrasound study of Kannada lingual articulations. *Journal of Indian Speech and Hearing Associaiton.*

Li, M., Kambhamettu, C., & Stone, M. (2005). Automatic contour tracking in ultrasound images. *Clinical Linguistics and Phonetics*, 19, 545–554.

Narayanan, S., Byrd, D. & Kaun, A. (1999). Geometry, kinematics, and acoustics of Tamil liquid consonants. *Journal of the Acoustical Society of America*, 106, 1993–2007.

Švarný, O. & Zvelebil, K. (1955). Some remarks on the articulation of the "cerebral" consonants in Indian languages, especially in Tamil. *Archiv Orientální,* 23, 374–407.

ACKNOWLEDGEMENTS

The work was supported by grants from *All India Institute of Speech and Hearing* and *Social Sciences and Humanities Research Council of Canada.*

EMOTION CO-EXISTS WITH LEXICAL EFFECTS: A CASE STUDY

Tatiana V. Kryuchkova[1], Benjamin V. Tucker[2]
[1]Dept. of Linguistics, University of Alberta, 4-32 Assiniboia Hall, Alberta, Canada, T6G 2E7, kryuchko@ualberta.ca
[2]Dept. of Linguistics, University of Alberta, 4-32 Assiniboia Hall, Alberta, Canada, T6G 2E7, benjamin.tucker@ualberta.caa

1. INTRODUCTION

In speech production many lexical factors influence the way in which words are produced. Munson and Solomon (2004) find that in laboratory speech shorter vowel duration is associated with larger neighborhood density and lower lexical frequency. While lexical factors have been shown to affect the way in which speech is produced, emotion has also been shown to affect acoustic features of speech, such as vowel and word duration, fundamental frequency, and intensity (e.g. Zupan et al., 2009).

With emotion (joy, fear, anger, happiness, etc.) being an important part of daily communication, emotional and lexical acoustic variation may come into conflict such that during emotionally produced speech lexical effects found in laboratory speech may not be present. The present work develops the results of previous work (Kryuchkova & Tucker, 2011), which found a robust effect of neighborhood density on word duration in production of acted emotional speech. This effect interacted with lexical Frequency and Emotion. In the present study we further examine the possible interaction of emotion and the lexicon by investigating effects of neighborhood density and lexical frequency for words and also at the vowel level.

2. EXPERIMENT

2.1 Materials

The stimulus set comprises 260 real words of English taken from Wurm & Seaman (2008), for which counts of lexical frequency, neighborhood density, and morphological family size were available. Two professional male actors (T and D) recorded the stimuli in five emotional modalities: neutral, anger, joy, content, and fear. Participants read items from a list: fillers were added at the beginning and end to account for list intonation. Each actor recorded the stimuli over two sessions. The recordings of three emotional modalities: anger, joy and neutral - were annotated and acoustic measures extracted using PRAAT (Boersma & Weenink, 2010). Mean vowel fundamental frequency (F0) and vowel duration were extracted for each word. Measurements for the mean word F0 and word duration were taken from the previous study and mean intensity for words and vowels was extracted.

2.3 Methods

The statistical analysis was performed using linear mixed effects regression modeling (*lme4*, Bates, Maechler, Bolker, 2011) with words as random effects and duration, mean F0, and mean intensity as dependent variables. Markov chain Monte-Carlo sampling (10,000 simulations) was used to estimate reported p-values (*languageR*, Baayen 2010). First, the dependent variables were modeled individually as a function of Emotion and Speaker to check whether the acoustic characteristics of the recorded speech were indeed modified by the emotional modality. Second, the dependent variables were modeled as a function of the predictors lexical Frequency (logged) and Neighborhood Density; the model also accounted for possible interactions of the two lexical predictors with Emotion and Speaker. The count for Number of Syllables and inherent vowel length (tense vs. lax) were also included in the model as control variables.

3. RESULTS

3.1 Emotion related acoustic variation

All six dependent variables are affected by the emotional modalities portrayed by the actors. The pattern of the effects for mean Intensity and mean F0 are similar at the word and vowel level (Figure 1, top panels)

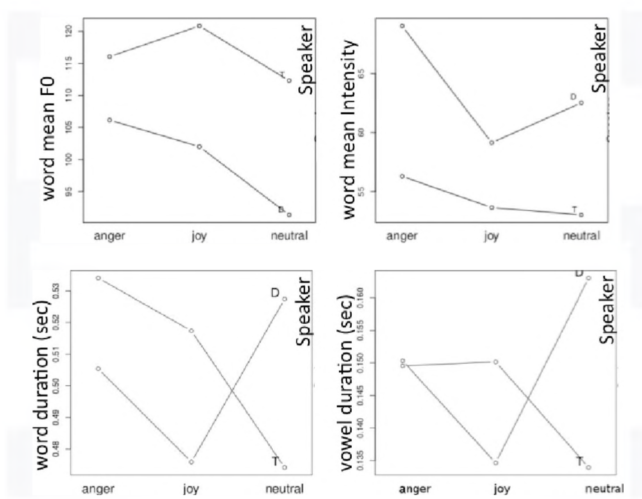


Figure. 1. Effects of Emotion and Speaker on F0, Intensity and duration.

The effect pattern for Duration is different for T at the word and vowel levels. For T anger has the longest word durations of all three emotional modalities, but at the vowel level T's anger and joy have similar vowel durations (Figure 1, bottom panels).

The results of this stage of analysis confirm that in the current dataset the target acoustic features (mean F0, mean Intensity and duration) are modified by the acted emotional modality both at the word and vowel levels. The patterns of variation are similar but not identical.

3.2 Lexical effects

No significant effects were found for intensity measures.

For vowel duration, Neighborhood Density significantly interacts with Speaker ($p<0.005$): higher Neighborhood Density is associated with longer duration; the slope of the effect is steeper for D. The effect of Neighborhood Density on vowel duration does not interact with lexical Frequency and lexical Frequency is not predictive of vowel duration. For word duration, the effect of Neighborhood Density interacts with Emotion: joy and neutral show a slightly steeper slope than anger, but for all three emotions larger Neighborhood Density is associated with shorter durations ($p<0.01$). For words, the effect of Neighborhood Density also significantly interacts with lexical Frequency ($p<0.005$), so that for words with lower lexical frequency larger Neighborhood Density is associated with shorter durations, but for words with high lexical frequency the effect disappears.

Neighborhood Density has a significant positive effect on the mean vowel F0 ($p<0.05$), with larger Neighborhood Density associated with higher F0. A similar effect of Neighborhood Density is observed at the word level ($p<0.054$). In both cases Neighborhood Density interacts with Speaker: T has overall higher mean F0 and a steeper slope for the effect. Lexical Frequency is also predictive of the mean word and vowel F0. At the word level, lexical Frequency interacts with Neighborhood Density ($p=0.0002$): for low frequency words the effect of Neighborhood Density is positive: larger Neighborhood Density is associated with higher mean word F0. For high frequency words the effect of Neighborhood Density on word duration reverses: larger Neighborhood Density in high frequency words is associated with lower F0. Neighborhood Density interacts with Vowel Type ($p=0.03$) for mean vowel F0: thus for lax vowels higher mean F0 is associated with larger Neighborhood Density, for tense vowels higher mean F0 is associated with smaller Neighborhood Density.

4. DISCUSSION

Our dataset was comprised of words recorded by two professional actors simulating anger, joy, and neutral emotion. To convey emotion, in this case study, both actors employed slightly different manipulations of duration, mean F0, and mean Intensity. Differences in vocal training and idiosyncratic differences in the way the actors convey emotion are likely the source of this variation.

Neighborhood Density is predictive of Duration and F0 at both word and vowel levels of analysis (Table 1); both longer duration and higher F0 are associated with higher Neighborhood Density. Like the first experiment in Munson and Solomon (2004) we find longer vowel duration for larger Neighborhood Density. We do not, however, find an effect of lexical Frequency on vowel duration. This may be due to differences in the word sets elicited (30% of our words were bisyllabic while Munson and Solomon (2004) only used monosyllabic words). Unlike Munson and Solomon (2004) we also investigated F0 and word duration. We find that lexical predictors, such as Neighborhood Density are also predictive of other acoustic characteristics at the word level. Most importantly, lexical effects, Neighborhood Density, are not removed by the acted emotional modality. The strength of the effect of Neighborhood Density is sometimes modified by the emotional aspect of the speech but is never the less present.

Table 1. Summary of statistically significant ($p<0.05$) lexical effects (ND=Neighborhood Density; Freq=lexical Frequency)

Dependent variable	Lexical predictors		
	ND (direction of the effect)	ND x Freq	ND x Vowel Type
Word Duration	Yes (x Emotion) (duration decreases)	yes	
Vowel Duration	Yes (x Emotion) (duration increases)		
Mean word F0	Yes (x Speaker) (F0 increases)	yes	
Mean vowel F0	Yes (x Speaker) (F0 increases)		yes

REFERENCES

Baayen, R.H. (2010). LanguageR: Data sets and functions with "Analyzing linguistic data: A practical introduction to statistics". R package 1.0 .http://CRAN.R-project.org/package=languageR.

Bates, D., Maechler, M., and B. Bolker (2011). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-42. http://CRAN.R-project.org/package=lme4

Boersma, P. and D. Weenik. (2010). Praat: doing phonetics by computer (version 5.1.29) [computer program]. Available from hhtp://www.praat.org/

Kryuchkova, T. and B. V. Tucker. (2011). Lexical effects on the production of emotional speech. Presented at the 162nd Meeting of the Acoustical Society of America (San Diego, CA).

Munson, B. and N. Solomon. (2004). The effects of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research*, vol47, 1048-1058.

Wurm, L. and S. Seaman. (2008). Semantic effects of naming and perceptual identification but not in delayed naming: implications for models and tasks. *Journal of Experimental Psychology*, vol34, 381-398.

Zupan, B., Neumann, D., Babbage, D. R., and B. Willer. (2009). The importance of vocal affect to bimodal processing of emotion: Implications for individuals with traumatic brain injury. *Journal of Communication Disorders*, vol42, 1-17.

PREDICTING ACCENTEDNESS: ACOUSTIC MEASUREMENTS OF CHINESE-ACCENTED ENGLISH

Vincent Porretta[1] and Benjamin V. Tucker[2]
[1]Dept. of Linguistics, University of Alberta, Alberta, Canada, T6G 2E7, porretta@ualberta.ca
[2]Dept. of Linguistics, University of Alberta, Alberta, Canada, T6G 2E7, bvtucker@ualberta.ca

1. INTRODUCTION

Listeners perceive foreign-accented speech as different from native speech because it deviates from native speaker acoustic targets. These deviations may occur across many acoustic dimensions such as word duration, vowel duration, vowel formant values, and voice onset time. Researchers have shown that non-native speakers produce longer and less variable word durations than native speakers and these measures are correlated with accentedness ratings [3]. Listeners are sensitive to these types of deviations and are capable of detecting accentedness in as little as a single 30 ms burst release [6]. Two previous studies employed distance measures to quantify these deviations in an attempt to model foreign-accented speech. For Arabic speakers of English, distance measurements for vowel duration, first and second formant values, and formant movement predict accentedness ratings of Arabic speakers' English vowels [9]. Similarly, for English speakers of Thai, distance measures of F0 valley and F2 values in diphthongs predict accentedness [12]. These findings suggest that native listeners compare non-native vowel tokens to the distributional properties of their learned native language.

The present study investigates the following questions: 1) How does Chinese-accented English differ from American English across gross acoustic measures such as word duration, vowel duration, and vowel formant values? 2) Which of these acoustic measures are most predictive of accentedness rating?

2. METHODS

The speech of nine male native Chinese speakers and one male native English speaker (low mean accentedness rating = 1.04) was analyzed for this study. Recordings and accentedness ratings were retrieved from the Wildcat Corpus of native- and foreign-accented English [11]. Forty-one monosyllabic words, each with three repetitions, were examined, totaling 123 tokens per talker. Measurements of word duration, vowel duration (of both monophthongs and diphthongs) and format values for F1-F3 (measured at the midpoint) were hand-measured in PRAAT [3] and extracted for analysis.

3. ANALYSIS AND RESULTS

Accentedness ratings were modeled using ordinary least squares linear regression performed in R using the *rms* package [8]. Prior to analysis, formant values were log transformed and plotted revealing considerable variation in vowel space. Additionally accented speakers produce word durations longer and shorter than the native speaker; however, they produce vowel durations longer than the native speaker. The vowel-to-word ratio was calculated by dividing the vowel duration of a word by the total duration of that word. If accentedness is a result of non-native productions approaching (to varying degrees) native-like acoustic targets, quantifying the distance from the native speaker norm allows examination of how variation along different variables affects perceived accentedness. For each numeric variable the token value of each non-native speaker was subtracted from that of the native speaker. The absolute value of that difference yielded a positive number representing the distance between the non-native production and a typical, native-like production. Collinearity between variables was verified and all numeric variables were centered.

These variables were used as input for the model. Restricted cubic splines (df = 5) were used for all three formant variables to account for nonlinearity of the variables. Interactions between variables were checked and included if they were significant. For this model two significant interactions resulted; the first between *Word Duration Distance* and *Vowel Duration Distance*, and the second between *log F1 Distance* and *log F2 Distance*. Outliers (less than 5% of the data) were identified and removed from the model following the procedures of Baayen [1] and the model was refitted. A significant effect resulted for Vowel-to-Word Duration Ratio (β = 1.9444, t (1127) = 2.75, p = .0061). As the ratio increases the predicted accentedness rating increases indicating that accentedness is in part predicted by the amount of a monosyllabic word that consists of a vowel. There is a similar result for the interaction between Word Duration Distance and Vowel Duration Distance (β = -38.3949, t (1127) = -2.38, p = .0174). As illustrated in Figure 1, high accentedness ratings result when speakers produce word durations far from the norm of a native speaker, but vowel durations that are similar to the norm. When both distances are high (i.e., lower ratio) accentedness ratings decrease, indicating that listeners may be sensitive to the ratio between vowel length and word length thus influencing their evaluation of accentedness. The interaction between log F1 Distance and log F2 Distance (Fig. 2) indicates that when the distances are equally high, accentedness rating increases. When F1 distance is low, but F2 distance is high, accentedness rating decreases. However, when F2 distance is low and F1 distance is high, accentedness rating increases dramatically. This indicates that there is a strong effect for F1 values that are far from native speaker norms which is perhaps a strong cue to listeners that the speaker has a foreign accent.

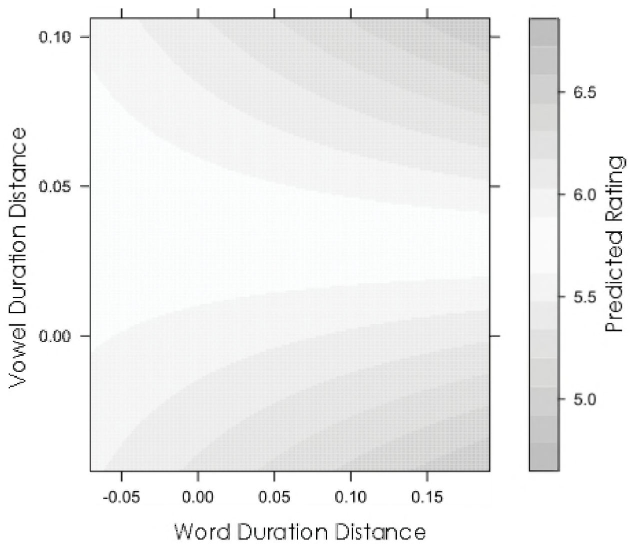


Figure 1: Contour plot of interaction between Vowel and Word Duration Distances for Predicted Rating.

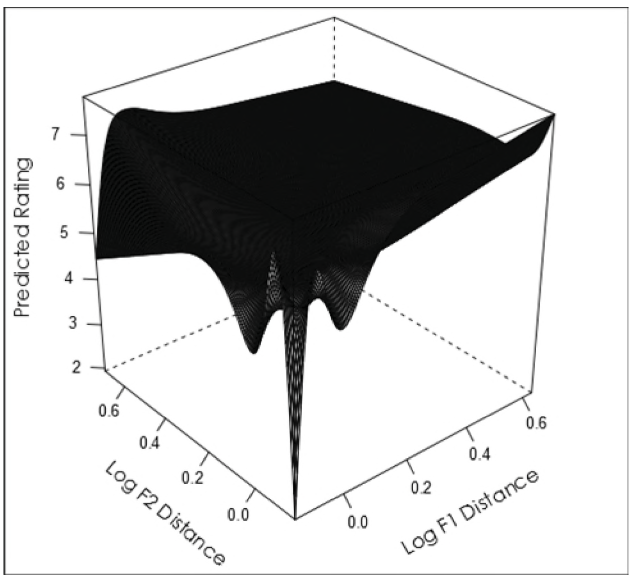


Figure 2: Perspective plot of interaction between log F1 and log F2 Distances for Predicted Rating

4. DISCUSSION AND CONCLUSIONS

This study demonstrates that foreign accentedness ratings of native Chinese speakers can be predicted using measures of acoustic distance. These findings replicate previous research [5, 6] that seek to model non-native vowel productions using measures of distance from typical native speaker values. The present results add to these findings by showing the relationship between word and vowel durations. For Chinese accented speech, vowel duration and word duration (and how they interact) play an important role in how listeners evaluate accentedness. Additionally, deviations of F1 from typical values have a strong effect on the degree of perceived accent. It appears that speakers who are more able to approximate typical native speaker values on these measures are perceived as less accented supporting the idea that, in evaluating degree of foreign accent, listeners compare gross acoustic features like word duration, vowel duration, and formant values to typical native speaker values. This suggests that listeners store distributional information about acceptable native-like productions with which they compare productions from new speakers [10]. By understanding which acoustic variables affect the perception of accentedness and how acoustic information is evaluated by listeners, it is possible to investigate their role in perceptual learning of accented speech. Current research on adaptation to foreign-accented speech [4, 5, 7] does not address the role of standard acoustic measurements, much less acoustic distance measures. Moreover, we plan to further investigate how these variables correlate with markers of cognitive processing when listening to accented speech of varying degrees.

REFERENCES

[1] Baayen, R.H. 2008. Analyzing linguistic data: A practical introduction to Statistics using R. Cambridge: Cambridge University Press.

[2] Baker, R.E., Baese-Berk, M., Bonnasse-Gahot, L., Kim, M., Van Engen, K.J., Bradlow, A.R. 2011. Word durations in non-native English. Journal of Phonetics, 39(1), 1-17.

[3] Boersma, P., Weenink, D. 2011. Praat: doing phonetics by computer. Version 5.3.04, retrieved 12 January 2012 from http://www.praat.org/

[4] Bradlow, A.R., Bent, T. 2008. Perceptual adaptation to non-native speech. Cognition. 106, 707-729.

[5] Clarke, C.M., Garrett, M.F. 2004. Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America, 116,* 3647–3658.

[6] Flege, J.E. 1984. The detection of French accent by American listeners. *The Journal of the Acoustical Society of America*, *76*(3), 692.

[7] Jongman, A., Wade, T., Sereno, J. 2003. On improving the perception of foreign-accented speech. In M. J. Sole, D Recasens, & J Romero (eds). *Proceedings of the XVth International Congress of Phonetic Sciences, Barcelona, Spain,* 1561-1564. Barcelona, Spain: Universitat Autonoma de Barcelona.

[8] Harrell, F.E., Jr. (2012). rms: Regression Modeling Strategies. R package version 3.5-0. http://CRAN.R-project.org/package=rms

[9] Munro, M.J. 1993. Productions of English vowels native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech, 36(1),* 39–66.

[10] Pierrehumbert, J. (2001) Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee and P. Hopper (eds.) *Frequency effects and the emergence of lexical structure*. John Benjamins, Amsterdam. 137-157.

[11] Van Engen, K.J., Baese-Berk, M., Baker, R. E., Choi, A., Kim, M., Bradlow, A.R. 2010. 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.

[12]Wayland, R. 1997. Non-native production of Thai: Acoustic measurements and accentedness ratings. *Applied Linguistics, 18(3),* 345-373.

MODELLING VOWEL INHERENT SPECTRAL CHANGE IN SPONTANEOUS SPEECH

Michelle Sims, Benjamin V. Tucker, and Terrance M. Nearey
Dept. of Linguistics, University of Alberta, AB, Canada, T6G 2E7

1. INTRODUCTION

Phonetic research strongly supports the notion that formant trajectories vary more or less continuously over time even for vowels usually classed as monopthongs (Strange et al. 1983, Nearey and Assmann 1986, Hillenbrand and Gayvert 1993). One possible source for such formant movement is vowel inherent spectral change (VISC). Several studies investigated the role of VISC in vowel perception, arguing that listeners' use F1 and F2 contours in vowel identification. However, this dynamic information has not been investigated in spontaneous speech. It is possible that the VISC- relaed dynamic spectral movement seen in carefuly produced vowels will not be evident in conversational speech due to coarticulatory effects of the surrounding consonants (Strange et al. 1983). The present study investigates this issue by looking at the vowel productions in a corpus of spontaneous speech. Several alternate accounts of the nature of VISC in speech production are also briefly discussed.

1.1 Theories of Vowel Inherent Spectral Change

In their work on the perception of VISC, Morrison and Nearey (2007) discuss three predominate approaches of analysing vowel movement: the onset+offset, onset+slope, and onset+direction hypotheses. The onset+offset, or ΔF, approach holds that a vowel's ending F1 and F2 values are important to the perception of vowels. The onset+slope, or $\Delta F/\Delta t$, approach models VISC as a function of time, stating that modelling vowel movement through time enhances the perception of vowels. The onset+direction model, on the other hand, states that it is the overall direction of movement, not necessarily the offset or slope, that provides a perceptual cue to vowel identity. VISC directions are measured in terms of the rise and fall in F1 and F2 (i.e. F1 rising and F2 falling, F1 falling and F2 rising, no movement, etc.).

Research on vowel identification has suggested that the onset+offset hypothesis best captures the perceptual cues listeners use (Morrison and Nearey 2007, Hillenbrand et al. 2001, to name a few). Studies on the production of vowels in careful citation speech have added to this theory with models that include vowel duration and pitch (Hillenbrand et al. 2001). The present paper tests these approaches to modelling VISC in spontaneous speech. As much of the research on VISC has focused on perceptual studies or citation speech, it is of interest to see if these theories also hold for vowels produced in everyday conversations.

2. METHOD

The present study makes use of a dataset of 54 monosyllabic irregular English verbs that differ between their past and present tense forms based on a single vowel alteration. For example, we included irregular verbs like sing/sang in our dataset, but excluded irregular verbs such as weep/wept and is/were. We extracted the vowels of these verbs from the Buckeye Corpus of Conversational English (Pitt et al., 2007), yielding 7,034 tokens of eleven different monophthongs from 40 adult speakers (20m/20f)

F1, F2, and F3 contours for each vowel were gathered using FormantMeasurer (Morrison and Nearey 2011) and hand-corrected. In accordance with Nearey and Assmann (1986), we marked the onset of each vowel at 24% of its entire duration and the offset at 64%. In doing so, we attempt to mitigate, to some extent, the influence of coarticulation from the surrounding consonantal context.

We performed a cross validation discriminant analysis to test each VISC hypothesis' ability to capture the spontaneous speech data. The analysis is based on a linear parametric technique trained on all various combinations of F1 and F2 onsets, offsets, slope ($\Delta F/\Delta t$), direction (all 9 combinations of F1 and F2 falling, rising and no movement), pitch and duration. We performed the discriminant analysis both on males and females (separately and combined). We then used t-tests to test the significance of the VISC movement and differences across males and females.

3. RESULTS

Figures 1 and 2 show VISC movement for males and females, respectively. On the plots, the arrow indicates the average vowel offset and the labelled end represents the average vowel onset. The linear VISC movement shown is gathered from the onset+offset model. Though the vowel spaces are significantly different between genders ($p<0.001$ or all vowels), the trends in the VISC movements for males are not significantly different from females ($p>0.05$ for all vowels). The slopes for each VISC movement across males and females, are significant ($p<0.005$ for all vowels). It is interesting to note the extreme fronting of /u/ that is characteristic of the Ohioan dialect.

Table 1 illustrates the outcomes of the discriminant analysis. For simplicity, we have only included those models that perform the best. The percentages indicate the amount of improvement each model contributes to vowel discrimination compared to a model consisting of a single F1 and F2 measurement. These improvements are seen in the separate models for males and females, as well in the combined gender model. Regardless of gender, a model consisting of formant onsets, offsets, pitch, and vowel duration performed the best at the discrimination task.

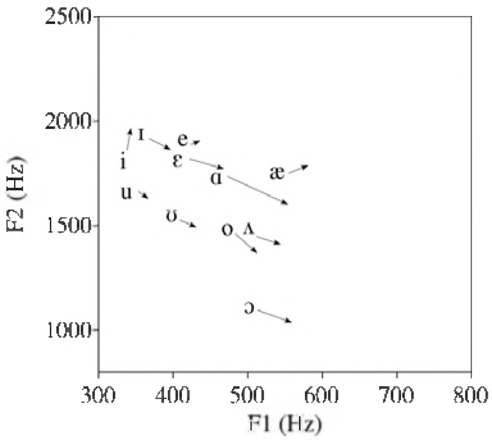


Figure. 1. Average vowel inherent spectral change for males.

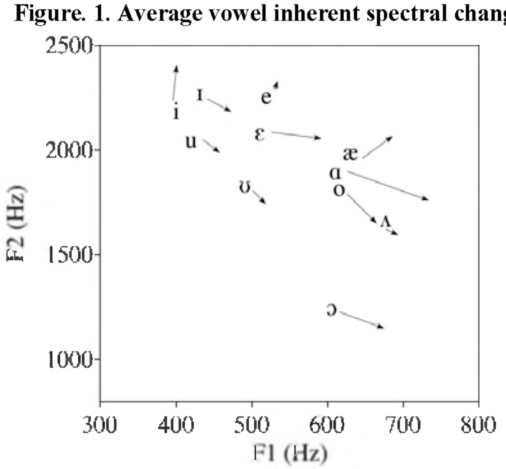


Figure. 2. Average vowel inherent spectral change for females.

Table 1. Results of the discriminant analysis: amount of improvement with the individual addition of Slope, Direction and Offet compared to a Onset+Pitch+Duration model.

	...+Slope	...+Direction	...+Offset
combined genders	6.57%	6.74%	12.35%
males only	4.80%	4.69%	10.30%
females only	7.80%	7.95%	15.77%

4. DISCUSSION

Contrary to theories of dynamic vowel specification, the present study finds evidence for vowel inherent spectral change in spontaneous speech productions. The trends in F1 and F2 frequency shifts in English monophthongs seen here compliments similar findings from carefully produced vowels in a limited consonantal environment. That is to say, we find support for VISC movement in spontaneous speech across a wide variety of phonetic contexts in line with previous research in more controlled conditions.

All in all, our models of VISC in spontaneous speech support the perception research and studies on citation speech. We find a slight superiority for a combined onset+offset+duration+pitch model in capturing the dynamic spectral properties of vowels in conversational English. These results are in line with both Hillenbrand et al.'s (2001) research on citation vowels and other studies comparing the different approaches to VISC analysis (Morrison and Nearey 2007). It is notable that even though the vowel spaces differ between males and females, the trends in VISC movement remain the same. This is evident in both the discriminant and difference tests. Moreover, we find support for the direction and types of formant movement first found by Nearey and Assmann (1986).

These findings have immediate implications for perceptual experimentation. Most studies on vowel identification make use of a onset+offset theory of VISC, with the data here point to the discriminative importance of this hypothesis in spontaneous speech productions. Theories of speech processing, too, could benefit from this more complex and integrated model of dynamic vowel specification.

REFERENCES

Hillenbrand, J. M., Clark, M. J., & Nearey, T. M. (2001). Effects of consonant environment on vowel formant patterns. Journal of the Acoustical Society of America, 109, 748–763.

Hillenbrand, J. M., & Gayvert, R. T. (1993). Vowel Classification Based on Fundamental Frequency and Formant Frequencies. Journal of the Acoustical Society of America, 36, 694-700.

Morrison, G. S., & Nearey, T. M. (2007). Testing theories of vowel inherent spectral change. Journal of the Acoustical Society of America, 122, EL15–EL22.

Morrison, G. S., & Nearey, T. M. (2011) FormantMeasurer: Software for efficient human-supervised measurement of formant trajectories. [Software release 2011-05-26].

Nearey, T. M., & Assmann, P. F. (1986). Modeling the role of inherent spectral change in vowel identification. Journal of the Acoustical Society of America, 80, 1297–1308.

Pitt, M.A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., and Fosler–Lussier, E. (2007). Buckeye Corpus of Conversational Speech (2nd release) [www. buckeyecorpus. osu. edu] Columbus, OH: Department of Psychology, Ohio State University (Distributor).

Strange, W., Jenkins, J. J., & Johnson, T. L. (1983). Dynamic specification of coarticulated vowels. Journal of the Acoustical Society of America, 74, 695–705.

ACKNOWLEDGEMENTS

Work supported by Social Sciences and Humanities Research Council of Canada Grant number 410-2011-0386 to B. V. Tucker.

CONSONANTAL DURATION SCALING IN ACCENTED WORDS

Richard Yanaky
Dept. of Linguistics, University of Alberta, 4-32 Assiniboia Hall, Edmonton, AB, T6G 2E7, Canada
yanaky@ualberta.ca

1. INTRODUCTION

There are many names for accented speech. These include *functional stress, focus, emphatic, stressed word, sentential stress, prosodic prominence*, etc. No matter the name, they will all be defined here as the act of accenting part of a sentence (be it a word or a phrase) to change the pragmatics of a sentence. That is, the context in which an otherwise identical sentence is spoken reflects a change in meaning of an utterance. For a clear example,

1a) I thought the idea was brilliant.
1b) I *THOUGHT* the idea was brilliant. (but I changed my mind later)

By accenting part of the sentence, we are using this to assign another semantic element to the sentence via prosody. The above example is assigning a contrastive element to the sentence. Similarly, we can stress a part of a sentence to indicate focus, emphasis, and many other syntactic/semantic elements to change the pragmatics of an utterance. This study sheds more light on the acoustic structure of accented speech.

2. METHOD

This experiment consisted of a simple speaking task which was recorded and analyzed acoustically via Praat. The participants (3 female native speakers of Canadian English) were given identical paired sentences, where one was an unaccented question, and the second was an accented version of the same sentence, where the accented element was the verb. For example:

2a) "Mary [verbed]$_1$ the ball?"
2b) "Mary *[VERBED]$_1$* the ball?"

This sentence pair was repeated twice per speaker, with 19 different verbs in frame. The verbs are common monosyllabic words. The goal was to capture the durational differences of the different onset consonants of accented words, as the hypothesis is that they will scale differently according to the type of consonant. That is, the onset consonant of an accented word will occupy a larger percentage of the total word duration than it would otherwise occupy in an unaccented position. This hypothesis is brought forth by the idea of domain-initial strengthening [1]. The accenting of a word will create a new prosodic boundary, which will be observable by different prosodic factors, including differences in duration.

3. RESULTS

Table 1 shows the given statistical significances of accented vs. unaccented durations. All statistical analyses were done in statistical analysis software, R, using Welch's Two Sample T-Test.

Variable	P-Value
Word Duration	**< 2.2e-16**
Vowel Duration	0.7335
Vowel Duration after obstruents	0.4851
Vowel Duration after Stops	**0.0007329**
Vowel Duration after Nasals	0.3859
Stop Duration (Closure + Aspiration)	**0.00023**
Stop & Affricate Closure Duration	**0.0002611**
Stop Closure Duration	**0.0007261**
Stop Release Duration	0.8091
Affricate Duration (Closure + Frication)	0.06522
Affricate Closure Duration	**0.02363**
Affricate Frication Duration	0.3212
Fricative Duration	0.109
Nasal Duration	**0.003771**
Liquid Duration	0.4751
Voiced Stop Duration (Closure + Aspiration)	**0.0442**
Voiceless Stop Duration (Closure + Aspiration)	**4.375e-05**
Voiced Affricate Duration (Closure + Frication)	0.1355
Voiceless Affricate Duration (Closure + Frication)	**0.01432**
Voiced Fricative Duration	0.09445
Voiceless Fricative Duration	0.08191

Table 1. Accented vs. Unaccented significance level (T-Test)

Figures 1-6 show the visualizations of the average segment's duration, as proportionate to word duration. All findings show an increased proportional duration of the onset consonant in accented words compared to their unaccented counterparts, as hypothesized. The strongest statistical findings however, were in favour of consonants with a full oral stricture, namely stops, affricates, and nasals, all with $p < 0.05$. If these onset consonants are segmented even further where possible, the period of full closure is more important than the release, as seen in the smaller p value of stop and affricate closures vs. their released aspiration and frication, respectively. The other consonants without full stricture (fricatives and liquids) still show an increased overall duration within accented words, but without significant p values. P values are further significant again in favour of voiceless consonant onsets as compared to their voiced counterparts.

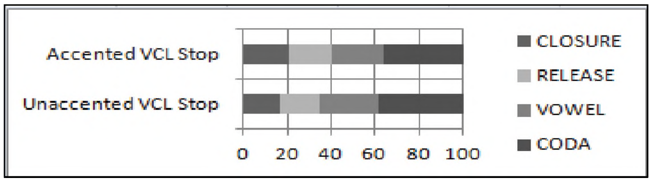


Fig. 1. Avg voiceless stop onset proportions. Accented words are on avg 36% longer than the unaccented.

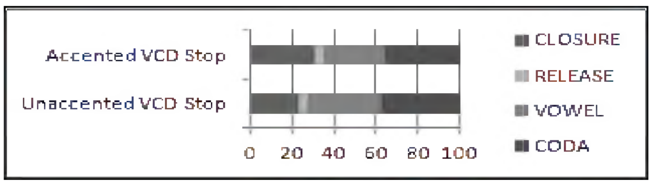


Fig. 2. Avg voiced stop onset proportions. Accented words are on avg 30% longer than the unaccented.

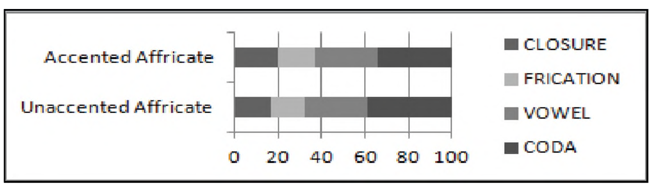


Fig. 3. Avg affricate onset proportions. Accented words are on avg 19% longer than the unaccented.

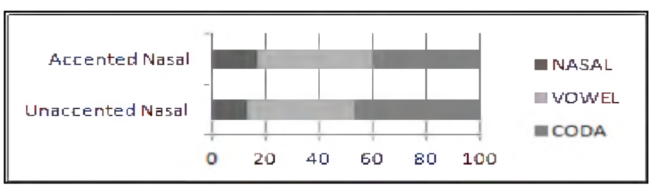


Fig. 4. Avg nasal onset proportions. Accented words are on avg 14% longer than the unaccented.

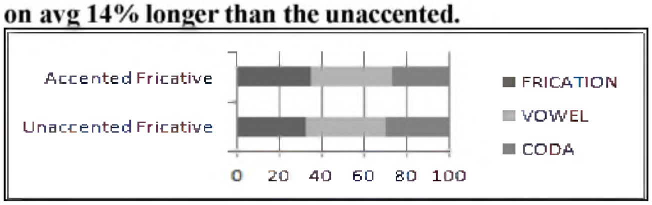


Fig. 5. Avg fricative onset proportions. Accented words are on avg 28% longer than the unaccented.

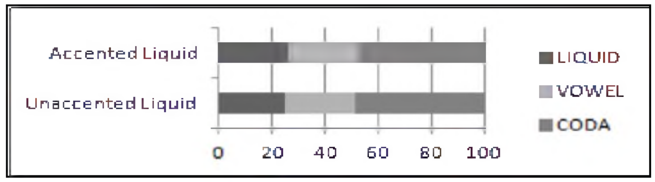


Fig. 6. Avg liquid onset proportions. Accented words are on avg 28% longer than the unaccented.

5. DISCUSSION

There are several ideas as to why consonants with full oral constrictions seemed to result in the most significant findings. Since this type of accenting performs a pragmatic function, functional theories of prosodic prominence and boundaries [2] seem to be a logical point to follow. One hypothesis is that it is the closure period which acts as the silent pause, which cues a different meaning from the norm in the syntax. Fuller constrictions would then likely be more susceptible to holding this pause longer since the closure gestures aren't being made in the fricative and liquid consonants. As such, for these other consonants, a different prosodic cue to accented words are likely more relevant-most likely intonation, intensity, or something else in the surrounding environment. On a similar note, van Santen and Shih [3] indicate that suprasegmental timing (duration) is highly dependent on the suprasegmental unit, and not the prosodic context. If we are to accept this, the analyzed accented words should be analyzed as their own suprasegmental unit, and as such, be in support of a prosodic boundary hypothesis, as boundaries are found outside suprasegmental units.

A second idea is that sonority and degree of stricture somehow plays a role with size of effect, since it was the consonants with high degrees of stricture that proportionally 'grew', while the vowel's duration 'shrank' in comparison to this. If further studies prove this to be true, it could very well weaken any prosodic boundary, or domain initial strengthening hypothesis. Both hypotheses will require further testing, but causation aside, this proves to be another potential cue for fine prosodic identification and voice synthesis.

REFERENCES

[1] Cho, T. & Keating, P. (2009). Effects of initial position versus prominence in English. Journal of Phonetics 37(4), 466-485.

[2] Wagner, Michael & Watson, Duane (2010). Experimental and theoretical advances in prosody: A review. Introduction to special issue of Language and Cognitive Processes 25(7). 905-945.

[3] van Santen, J.P.H & Shih, C. (2000). Suprasegmental timing models in Mandarin Chinese and American English. Journal of the Acoustical Society of America. 107(2), 1012-1026.

The Vocal Signature of the Male Northern Elephant Seal

Caroline Casey[1], Colleen Reichmuth[1], Brandon Southall[1],
Isabelle Charrier[2], and Nicolas Mathevon[3]
[1] Institute of Marine Sciences, University of California Santa Cruz, Santa Cruz, California,
USA, 95060 cbcasey@ucsc.edu
[2] Equipe Communications Acoustique, Université Paris Sud, Orsay, France
[3] Laboratoire de Biologie Animale, Université Jean Monnet, Saint-Etienne, France

1. INTRODUCTION

In the late 1800s, the northern elephant seal (*Mirounga angustirostris*) was extensively hunted for blubber. Consequently, the population was reduced to fewer than 100 individuals located on Guadalupe Island off the coast of Baja California. Since then, the northern elephant seal population has steadily recovered and grown to more than 175,000 individuals. This species, therefore, provides an opportunity for researchers to better understand the lasting effects of a severe population bottleneck on the communication and behavior of a recovering mammalian species.

Northern elephant seals breed annually during the winter months at islands and mainland rookeries along the western coast of the United States and Mexico. These animals maintain a highly polygynous breeding system in which adult males establish dominance hierarchies that determine access to estrous females. Males typically arrive in early December and establish their dominance position within the colony and stay until mid-March, a period extending over three months without access to food or water. Competition between males is intense, with only a small subset of individuals gaining access to breeding females. While dominance relationships may be established through physical contact, these hierarchies are maintained through the use of stereotypic threat displays that include distinctive vocalizations (Bartholomew and Collias, 1962; Le Boeuf and Peterson, 1969) and associated visual and seismic cues. These complex displays play an important role in settling otherwise costly interactions between competing males, as stereotyped acoustic signals often elicit appropriate behavioral responses from spatially separated individuals without physical contact.

The acoustic threat displays of these males have been described as 'an extremely loud resonant clapping sound with a metallic quality which suggests the exhaust noise made by a diesel engine' (Le Boeuf and Peterson, 1969). While many early accounts of the temporal features of these calls define them as containing short pulses with regular spacing (Shipley, 1981) and some variations of this call type have been noted, no study to date has attempted to describe in detail the various call types within breeding populations. Additionally, while several studies have investigated the function of these signals in maintaining social structure (Bartholomew and Collias, 1961; Sandegren, 1976; Shipley, 1981), the manner in which these vocalizations convey information to receivers remains unknown.

The aim of this study is to: (1) describe the variability in the stereotypic calls of male northern elephant seals within an established breeding rookery; (2) examine the acoustic characteristics of individual male threat vocalizations to evaluate whether they are stable within and across breeding seasons; and (3) evaluate the spectral and temporal features of these signals for their dependence on body size and dominance rank. These objectives, particularly (3), enable an assessment of whether honest signaling serves to communicate resource holding potential in this species, or alternatively, whether it is likely that these calls serve as an individual signature that males use to identify one another through learned association.

2. METHODS

More than 200 sub-adult and adult male elephant seals were identified and observed between the months of December and March over two consecutive breeding seasons at the Año Nuevo Rookery in San Mateo County, California, USA, located 60 miles south of San Francisco. To identify focal males within the breeding season, males were temporarily marked using black hair dye. A subset of individuals were opportunistically tagged with numeric plastic roto-tags inserted into the hind flipper to facilitate tracking between seasons. GPS coordinates were recorded daily for the location of each focal male at the main study site at Año Nuevo to assess movement patterns and rival familiarity during each of the two breeding seasons. Dyadic interactions between breeding males were opportunistically scored throughout the season, and the outcome of these interactions fed into a linear ranking system to quantitatively evaluate the dominance level of each individual. Body size parameters for focal individuals were derived through repeated photometric sampling. High-quality, complete calling bouts were digitally recorded and analyzed from sub-adult and adult males to evaluate spectral, temporal, and sound pressure characteristics across both seasons. These call qualities were evaluated for dependence on body size and rank in order to assess the possible role of honest signaling. Calls from a subset of individuals that were repeatedly sampled within and across seasons were investigated to determine whether the calls of male northern elephant seals are individually reliable and easily discernible from one another.

3. RESULTS

Several call types with varying structural and temporal patterns were identified and described at the Año Nuevo breeding colony. These distinctive vocal signatures were present across age-classes, and were not segregated by the size or dominance status of the caller. The lack of obvious correlations between the call type displayed by an individual and his resource holding ability was consistent with earlier observations that the outcome of a vocal challenge could not be predicted on the basis of call parameters alone.

The results of the individual acoustic analysis did confirm a significant divergence among sexually mature males, even among those males sharing the same call type. The calls of individuals displayed certain unique temporal, spectral, and phrasing patterns that were easily recognizable and statistically reliable. These individual differences were retained across multiple years.

4. DISCUSSION

By their very nature elephant seals are an ideal mammalian system for studies of vocal communication, as they congregate in large, easily accessible groups each year in predictable locations and are relatively undisturbed by human presence. We have found that the calls of male northern elephant seals are individually unique; with the structural and temporal patterns of the call being easily recognizable to even the human ear. In this species, where individuals spend a significant amount of time fasting during the breeding season and thus need to conserve energy, selective pressures for avoiding harm and conserving energy would favor the accurate assessment of dominance stature between competing males. Analysis of the calls indicates that there are no obvious honest indicators of resource holding potential within the construct of the vocalization, lending weight to the hypothesis that these signals serve as reliable vocal signatures that males use to identify one another during the breeding season. Males may be learning the individually unique vocal signatures of their rivals through experience and using this information in decision-making regarding when to engage in competitive interactions for access to females. The opportunity we have to study this species at this field site allows us to pursue questions about the role of associative learning in animal communication by combining both observational and experimental approaches.

REFERENCES

Le Boeuf, B., Peterson R.(1969) Dialects in elephant seals. Science, 166, 1654-1656.

Shipley, C., Hines, M., Buchwald, J. (1981) Individual differences in threat calls of northern elephant seal bulls. Animal behavior, 29, 12-19.

Bartholomew, G., Collias, N. (1961) The role of vocalization in the social behavior of the northern elephant seal. Animal behavior, 10, 7-14.

Sandegren, F. (1976) Agonistic behavior in the male northern elephant seal. Behavior, 57, 136-158.

UNDERSTANDING THE MASKING EFFECTS OF NOISE ON COMMUNICATION IN NATURAL ENVIRONMENTS

Robert Dooling[1], Sandra Blumenrath[1], and Ryan Simmons[1] and Kurt Fristrup[2]
[1]Dept. of Psychology, University of Maryland, College Park, MD, USA, 20742. rdooling@umd.edu
[2] Natural Sounds Program, National Park Service, Fort Collins, CO, USA.

1. INTRODUCTION

Anthropogenic noises can cause a variety of adverse effects on birds and other wildlife. These effects include stress and physiological changes, auditory system damage from acoustic overexposure, and masking of communication and other important biological sounds. A precise understanding of these effects is of interest to many groups including biologists, environmentalists, and government regulators, as well as city planners and roadway and construction engineers. However, for a number of reasons, it is difficult to reach a clear consensus on the causal relationships between noise levels and these adverse effects. One reason is that there are surprisingly few studies in animals that can definitively identify anthropogenic noise alone as the principal source of stress or physiological effects. A second reason is that, while all humans have similar auditory capabilities and sensitivities, the same is not true for all animals. Still another issue is separating the various effects of noise. There are well documented adverse consequences of elevated noise on humans including hearing loss, masking, stress, physiological and sleep disturbances, and changes in feelings of well-being, and it would not be too surprising to find a similar range of effects in animals.

There are four overlapping classes of anthropogenic noise effects on animals (i.e. PTS, TTS, Masking, and Other Effects) with particular spatial relationships. Using birds, I will suggest a framework for conceptualizing the separate and integrated effects of anthropogenic noise, particularly those of masking. This is useful because independent of other effects, masking of communication signals and other important biological sounds (e.g., sounds of an approaching predator) can potentially have significant adverse consequences for species' behavior and population viability. Most vocal species rely on acoustic communication for species and individual recognition, mate selection, territorial defense, parent-offspring communication and detection of predators/prey. Understanding how and to what extent masking can affect communication between individuals is an important first step toward determining the level of impact to them, and to the species.

1.1 Masking Effects on Different Aspects of Hearing.

Common sense and our own experience tell us that acoustic communication can be severely constrained if background noise is of a sufficient level. Such noise decreases signal-to-noise-ratios and therefore limits the acoustic space (the combination of sound frequencies and levels that are audible) of a sound. Noises can be continuous or intermittent, broadband or narrowband, and predictable or unpredictable in time or space. These noise characteristics determine the strategies that birds might employ to minimize the effect of noise on acoustic communication. Background noise makes it harder for an animal (or human) to detect sounds that may be biologically relevant, to discriminate among these sounds, to recognize these sounds, and to communicate easily. Studies on the effect of noise on hearing in birds and humans show that signal discrimination requires a higher signal-to-noise ratio than detection; recognition requires a higher signal-to-noise ratio than discrimination; and comfortable communication requires an even higher signal-to-noise ratio. We can use this information to estimate the effect of anthropogenic noise on acoustic communication in birds.

1.2. Masking and the Spectrum of Noise

The simplest kind of masking experiment is to measure the sound detection thresholds for pure tones in the presence of a broadband noise. These signal-to-noise ratios in masking (i.e., critical ratios) are now available for 14 different species of birds so we have a fairly good idea of how the average bird hears in noise. Most laboratory studies estimating the effects of noise on signal detection use continuous noises with precisely defined bandwidths, intensities, and spectral shapes. Traffic noise on heavily traveled roads can approximate these features (e.g., relatively continuous, relatively constant spectrum and intensity). This provides the opportunity to move from laboratory results based on continuous noises to predictions of behaviors in the field (e.g., communication distance) that might be affected by anthropogenic noises such as highway noise. From masking studies in birds, humans, and other animals, it is known that the noise in the frequency region of a signal is the most important acoustic feature in masking the signal--- not noise outside that frequency band. Highway noise, for instance, has more energy below 1 kHz than above, and bird vocalizations generally contain more energy above 1 kHz than below. Thus, the masking effects of highway noise on bird vocalizations are less than would be expected from noise of the same level in the same frequency range of bird vocalizations.

1.3. Modeling the Effect of Traffic Noise

To evaluate the effect of masking noise on bird communication, we developed a model that integrates the spectrum and level of the masking noise, the bird's hearing in quiet and noise, the spectrum and level of a signaling bird's vocalizations, and the acoustic characteristics of the environment. The model assumes that the spectrum and amplitude level of the noise and the signaler's vocalization are both known at the location of the receiver. These values can either be measured directly or they can be estimated by applying signal attenuation algorithms to both the noise source and the signals of the sender. The algorithms adjust the spectra and level of the noise and of the signal transmitted over distance and through different habitats (e.g. meadows, forests) between the communicating birds. The challenge for the receiver is to hear the signal in the presence of noise. This is dependent on the species-specific auditory capabilities of the receiver such as how well it hears in noise (i.e., its critical ratio) and the signal-to-noise ratio at the receiver's location. Using a human parallel, the model also incorporates the notion that different auditory behaviors (e.g., communicating comfortably versus just being able to detect that something was said) require different signal-to-noise ratios.

2. RESULTS

Imagine a specific case illustrated for a background noise level at the listening bird of 60 dB(A) – a level typical of traffic noise measured roughly 300 meters from a busy 6 lane highway. The example assumes the calling bird is vocalizing at a peak sound pressure level of 100 dB through an open area and the vocalization is affected by excess attenuation, beyond the loss due to spherical spreading, of 5dB/100 meters. In such a noise, a comfortable level of communication between two birds requires a distance between them of less than 60 meters. Recognition of a bird vocalization by the receiver can still occur at greater inter-bird distances up to about 220 meters. Discrimination between two vocalizations is possible at inter-bird distances up to 270 meters. And finally, simple detection of another bird's vocalization can occur at distances up to 345 meters in this noise.

The distance values above as computed for a 60 dB SPL level of traffic noise can be used to construct a receiver-centric map of distances corresponding to the four different types of auditory communication behaviors. In such a plot, communication distance between the sender (along the periphery of the circle) and receiver (at the center) is represented as the radius "r" for a set of concentric circles defining the boundaries of each of the four levels of communication described above. While any increase in ambient noise level from anthropogenic sources can potentially affect acoustic communication, which auditory behaviors are affected depend on the noise level. The inner (smallest) circle represents the case where the sender is close to the receiver. This represents a signal-to-noise that is sufficiently large that the sender and receiver can communicate comfortably (i.e. about 15 dB above the critical ratio). As the sender moves away from the receiver, the signal level and therefore signal-to-noise ratio, at the receiver drops. At some distance, the receiver can no longer communicate comfortably but can recognize a sender's different vocalizations. If the sender moves even further away, the receiver can still discriminate between two vocalizations but cannot reliably recognize them. Finally, at the outer perimeter, the signal level at the receiver results in such a low signal-to-noise ratio that the receiver can just detect that some kind of a sound has occurred. The distance over which masking from anthropogenic noise sources occurs can be quite large. This schematic provides a way of estimating and quantifying the risk to acoustic communication in birds at different distances from a noise source.

3. DISCUSSION

This approach of considering communication from the standpoint of the receiver may provide a useful metric for evaluating the actual noise impact on individuals, or collectively on populations, in areas subject to anthropogenic noise exceeding ambient levels. For instance, in determining risk to a species, the communication distances derived from this model might be considered in relation to other aspects of biology such as territory size.

REFERENCES

H. Brumm and H. Slabbekoorn. (2005). Acoustic communication in noise. Adv. Study Behav., 35: pp. 151-209

R.J. Dooling, B. Lohr, and M.L. Dent.(2000). Hearing in birds and reptiles. In R.J. Dooling, A.N. Popper, and R.R. Fay, eds., Comparative Hearing: Birds and Reptiles. pp. 308-359. Springer-Verlag, New York,

B. Lohr, T.F. Wright, and R.J. Dooling. (2003). Detection and discrimination of natural calls in masking noise by birds: estimating the active space signal. Anim. Behav., 65: pp. 763-777

UNDERWATER PASSIVE ACOUSTIC LOCALIZATION OF PACIFIC WALRUSES IN THE NORTHEASTERN CHUKCHI SEA

Brendan Rideout[1,2], Stan E. Dosso[1], and David Hannay[2]
[1]School of Earth and Ocean Sciences, University of Victoria, 3800 Finnerty Rd., Victoria, BC, Canada, V8P 5C2,
bprideou@hawaii.ca
[2]JASCO Research Ltd., 2305-4464 Markham St., Victoria, BC, Canada, V8Z 7X8

1. INTRODUCTION

This paper presents results from an MSc thesis [1] research project whose aim is the passive acoustic localization, and localization uncertainty analysis, of vocalizing, submerged Pacific walruses (*Odobenus rosmarus divergens*) when the acoustic propagation environment and receiver locations are not precisely known. In this case, passive acoustic localization involves using the natural vocalizations of the walrus to estimate the three-dimensional (3-D) position of the animal. The primary goal of this work is to provide a tool for collecting biological data (e.g., underwater swim speed) on wild walruses and other vocalizing underwater marine mammals which would not be possible with two-dimensional methods and which provides a rigorous uncertainty estimate.

To perform localization, arrival times for direct and interface-reflected (i.e., sound which reflects off the surface and/or bottom) acoustic paths at a set of underwater acoustic receivers are processed using an iterated linearized Bayesian inversion of the ray-tracing equations. Environmental parameters (water depth and sound speed correction) and receiver parameters (3-D hydrophone locations and synchronization times), in addition to walrus positions and vocalization times, are treated as unknowns and constrained with prior estimates and prior uncertainties. Prior estimates for walrus positions and times are given large uncertainties so as not to significantly influence the final solution. Maximum *a posteriori* (MAP) estimates for all unknowns (walrus positions and vocalization times, environmental parameters, and receiver parameters) are calculated simultaneously. To treat cases where both the data errors (i.e., arrival time) and prior uncertainties are only known in a relative sense, an inversion approach previously developed for array element localization [2] is extended here using the Akaike Bayesian information criterion (ABIC) [3] (a maximum entropy condition).

2. THEORY

Let $\mathbf{m}$ be the vector of unknown model parameters (i.e., 3-D locations and times for each vocalization, environmental parameters, and receiver parameters) and $\mathbf{t}$ be the vector of direct and interface-reflected arrival times for each of the vocalizations. The relationship between $\mathbf{m}$ and $\mathbf{t}$ is non-linear, but can be linearized about an arbitrary starting model $\mathbf{m}_0$ by retaining the first order term from the Taylor expansion:

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

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

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

This yields an inverse problem which is ill-conditioned due to both the source and receiver positions being unknown. Here, regularization is used to stabilize the problem and incorporate prior information (e.g., measurements of receiver positions and water depth) into the problem. As a result, in seeking to estimate the most probable walrus location, the following objective function is minimized over $\mathbf{m}$:

$$\Theta(\mathbf{m}) = [\mathbf{d} - \mathbf{d}(\mathbf{m})]^T \mathbf{C}_\mathbf{d}'^{-1} [\mathbf{d} - \mathbf{d}(\mathbf{m})] + \mu[\mathbf{m} - \hat{\mathbf{m}}]^T \mathbf{C}_{\hat{\mathbf{m}}}'^{-1} [\mathbf{m} - \hat{\mathbf{m}}] \qquad (3)$$

where $\mathbf{C}_\mathbf{d}'$ and $\mathbf{C}_{\hat{\mathbf{m}}}'$ are the relative data covariance and relative prior covariance matrices, respectively, $\hat{\mathbf{m}}$ is the vector of prior estimates, and $\mathbf{d}(\mathbf{m})$ is the model-predicted data vector. It can be shown that the MAP solution for $\mathbf{m}$ is:

$$\mathbf{m}_{\mathrm{MAP}} = \hat{\mathbf{m}} + \left[\mathbf{J}^T \mathbf{C}_\mathbf{d}'^{-1} \mathbf{J} + \mu \mathbf{C}_{\hat{\mathbf{m}}}'^{-1}\right]^{-1} \mathbf{J}^T \mathbf{C}_\mathbf{d}'^{-1} \left[\mathbf{d} - \mathbf{J}\hat{\mathbf{m}}\right]. \qquad (4)$$

The optimum value of the trade-off parameter μ is selected by minimizing the ABIC over μ. The derivation and expression for the ABIC can be found in Ref. [1]. After calculating $\mathbf{m}_{\mathrm{MAP}}$ for this iteration, $\mathbf{m}_{\mathrm{MAP}}$ is set as the starting model for the next iteration; due to the linearization step, the solution converges over several iterations. Following convergence, the posterior covariance matrix is calculated:

$$\mathbf{C}_\mathbf{m} = \sigma_0^{\,2}\left[\mathbf{J}^T \mathbf{C}_\mathbf{d}'^{-1} \mathbf{J} + \mu \mathbf{C}_{\hat{\mathbf{m}}}'^{-1}\right]^{-1} \qquad (5)$$

where $\sigma_0^{\,2} = \dfrac{\Theta(\mathbf{m}_{\mathrm{MAP}})}{N}$ and N is the number of data. This matrix includes parameter variances as diagonal elements and covariances as off-diagonal elements.

In some instances, knowing the relative localization uncertainty between estimated source positions (e.g., when estimating the uncertainty in a swim speed measurement) may be more informative than the absolute uncertainty values given by diagonal elements of $\mathbf{C_m}$. It can be shown that the relative uncertainty between two estimated model parameters is:

$$\sigma_{pq} = \left[C_{m,p,p} + C_{m,q,q} - 2C_{m,p,q}\right]^{1/2}$$

where $C_{m,p,p}$ is the posterior variance of m_p and $C_{m,p,q}$ is the covariance of m_p and m_q.

3. RESULTS AND DISCUSSION

From August to October 2009, a set of 40 underwater acoustic recorders were deployed throughout the northeastern Chukchi Sea, northwest of Alaska, to record marine mammal vocalizations. These recorders continuously collected 24-bit acoustic data with a sampling frequency of 16 kHz. Three of these recorders were deployed approximately 400 m apart near the Hanna Shoal, a ~30 m deep region of the Chukchi Sea where Pacific walruses are known to congregate during the summer months. Given that water depth was essentially uniform, a range independent propagation environment is assumed.

Pacific walruses are considered a gregarious animal, and make a variety of different vocalizations both in-air and underwater. One of the underwater vocalizations, called a 'knock', is particularly well suited for localization due to its impulsive-like characteristics (i.e., short duration and fast rise time). From the Hanna Shoal data, a sequence of 11 walrus knocks was localized using the approach outlined in the previous section. The results of this localization are shown in Figure 1. To perform this localization, arrival times for the direct and interface-reflected rays at each of the three hydrophones were identified. The error bars in Figure 1 represent single standard deviations. The slope of the least-squares linear regression fit to the estimated walrus position and vocalization time values yields an estimated walrus swim speed of (0.98 ± 0.19) m/s, with swim speed uncertainty estimated using a Monte Carlo approach.

Given the relative and absolute uncertainties, several observations can be made about the estimated track. The zig-zags in the track in the x-y plane are within relative uncertainties and, thus, are unlikely to be real. Also, there is limited evidence that the walrus is deeper at the end of the track than at the beginning.

The uses of this technique are not limited to locating Pacific walruses; other marine mammals (such as sperm whales) which produce impulsive-like vocalizations, and non-biologic sources, could potentially be located.

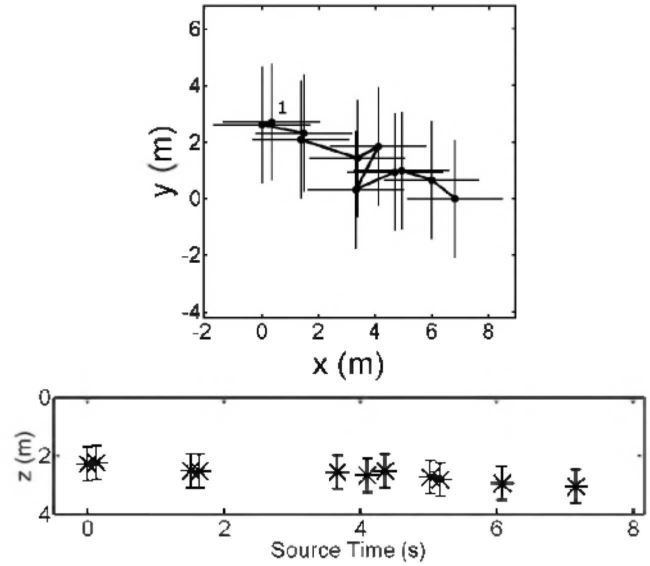


Figure 1 - Estimated horizontal position with relative x and y uncertainties (top) and estimated z position with absolute uncertainties for each of the eleven walrus knocks. The earliest knock in the top plot is labeled. Straight lines between knock locations are meant to illustrate the order in which the knocks were produced, and do not necessarily indicate the path the walrus took.

REFERENCES

[1] Rideout, B. P. (2011). Underwater Acoustic Localization and Tracking of Pacific Walruses in the Northeastern Chukchi Sea. MSc. Thesis, University of Victoria.

[2] Dosso, S. E. et al. (1998). Array element localization for horizontal arrays via Occams inversion. J. Acoust. Soc. Am., 104, 846-859.

[3] Akaike, H. (1974). A new look at the statistical model identification. IEEE Trans. Auto. Cont., 19, 716-723.

ACKNOWLEDGEMENTS

The authors would like to thank JASCO Research Ltd., ConocoPhillips Company, and Shell Exploration and Production Company for allowing access to underwater acoustic data collected on their behalf during the summer of 2009. This work was supported in part by the Natural Sciences and Engineering Research Council, the British Columbia Ministry for Advanced Education, and the National Science Foundation.

AUTHOR NOTES

The work was conducted while Brendan Rideout was a student at the University of Victoria in Victoria, British Columbia. He is currently a PhD student at the University of Hawaii.

NOISE AND BLOOD PRESSURE: A CROSS SECTIONAL AND LONGITUDINAL STUDY OF TH EFFECTS OF EXPOSURE TO LOUD NOISE ON RESIDENTS OF CALABAR, NIGERIA

Ubon Asuquo[1], Michael Onuu[2], Aniefiok Akpan[3] and Affiong Asuquo[4]

[1]Department of Mathematics, Trinity Western University, Langley, BC, Canada. (2011/2012 Session)

[2, 3] Department of Physics, University of Calabar, Calabar, Nigeria.

[4] Department of Anaesthesiology, University of Calabar Teaching Hospital, Calabar, Nigeria.

1. INTRODUCTION

Effects on the systemic circulation, such as the constriction of blood vessels, have been observed under laboratory and field conditions. Many studies have shown blood pressure to be higher in noise exposed workers and in populations living in noisy areas around airports and on noisy streets than in control populations [1, 2]. Other investigations indicate no blood pressure effects. The overall evidence suggests that a weak correlation exists between long-term noise exposure and blood pressure elevation, or hypertension. In real life, community noise interferes with a number of activities, for example recreation, sleep, communication, and concentration [3, 4]. The risk of adverse effects on health must be considered in the light that noise as a stressor may operate through physiological responses modified in complex ways by individual psychological processes. To our knowledge, this is one of the few longitudinal studies of the effect of chronic noise on blood pressure.

1.1 Hypotheses

The following null hypotheses were used for the research:

(i) there is no immediate relationship between exposure to loud noise and high blood pressure;
(ii) there is no long-term relationship between exposure to loud noise and high blood pressure.

2. MATERIALS AND METHODS

Measurements were taken in sixteen zones-eight high noise zones (study group) and eight low noise zones (control group). The high noise zones had average A-weighted noise levels of 80 dB or above, a level which may be hazardous to the hearing of most people. The low noise zones had average A-weighted noise levels of 50 dB or below, a level above which most people complain.

2.1 Objective measurements

Physical measurements were taken between 7 a.m. and 9 a.m., and 2 p.m. and 4 p.m. on working days (Monday to Friday) by an accredited Acoustician. About fifty random readings were taken at different locations within each zone and the sound level of each zone was calculated. The materials used for this measurements are the sound level meter, (Bruel & Kjaer) type 2203 with octave band filter (Bruel & Kjaer) type 1613, for measuring the sound level of the zones, a factory calibrated sphygmomanometer (SF 60502) for measuring blood pressure. During the noise level measurements, the sound level meter was held in such a way that the microphone was at least one meter from any reflecting surface and 1.2 m from the ground, corresponding to the ear level of an average person.
Blood pressure measurements and the physical examination of respondents were done by an accredited Medical Doctor. Ambient temperature, atmospheric pressure, weight and height of respondents were also measured.

2.2 Subjective measurements

A 39-item questionnaire was used for subjective assessment of the respondents in all the sixteen zones. The respondents were given questionnaires which contained standard questions tailored toward getting their reactions about the effects of noise on them. The questionnaires were structured, among other things, to elicit information on general sociodemographic characteristics, viz: age, sex, educational level, medical and occupational history of diseases and conditions that could cause hearing impairment.

3. RESULTS

The effect of noise on blood pressure was assessed by comparing the blood pressure of respondents who had lived or worked/schooled in the high noise zones with that of respondents in the low noise zones. The respondents taken were those that have lived in these areas for at least three years, a period believed to be enough for adaptation to the noise environment by the respondents. The results of some of the measurements are shown in tables 1, 2 and 3.

Table 1. Blood pressure of respondents in the high noise zone (HNZ)

S/N	CODE	100/50- 140/90	141/91-180/140
1	HNZ1	70	61
2	HNZ2	49	52
3	HNZ3	112	93
4	HNZ4	115	116
5	HNZ5	68	54
6	HNZ6	32	55
7	HNZ7	45	70
8	HNZ8	45	55

Table 2. Blood pressure of respondents in low noise zone (LNZ)

S/N	CODE	100/50- 140/90	141/91-180/140
1	LNZ1	22	26
2	LNZ2	12	13
3	LNZ3	26	11
4	LNZ4	93	74
5	LNZ5	111	154
6	LNZ6	13	24
7	LNZ7	160	107
8	LNZ8	209	80

Table 3. Blood pressures of Respondents in Calabar Timber Market in July 2002 and July, 2006

S/N	Systolic/diastolic blood pressure, July, 2002 (mmHg)	Systolic/diastolic blood pressure, July, 2006 (mmHg)
1	120/80	140/80
2	110/80	120/80
3	120/80	120/80
4	120/80	130/80
5	120/80	130/80
6	120/80	120/90
7	140/90	140/90
8	110/80	120/80
9	130/90	130/90
10	120/80	140/80
11	120/90	130/90
12	120/80	130/80
13	100/70	120/70
14	130/100	160/100
15	130/90	140/90

First, the coefficients of correlation were calculated for the noise measurements to determine how related the subjective responses, assessed by the use of questionnaires as the study instrument were to the objective responses measured with the sound level meter. Thereafter, using regression analysis, figure 1 was obtained when distribution of blood pressures within range 100/50 to 140/90 (High 1) in the high noise zones were compared with that within the same range in the low noise zones (Low 1).

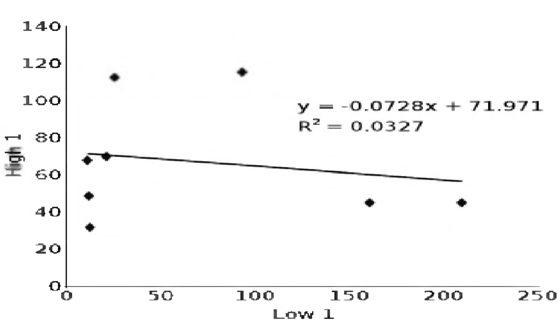


Figure 1. A graph of High 1 against Low 1.

1.

In figure 2, comparison is made between systolic pressures of the same respondents measured in 2002 (sys 02) and 2006 (sys 06). A similar result was seen when the diastolic blood pressures were also compared.

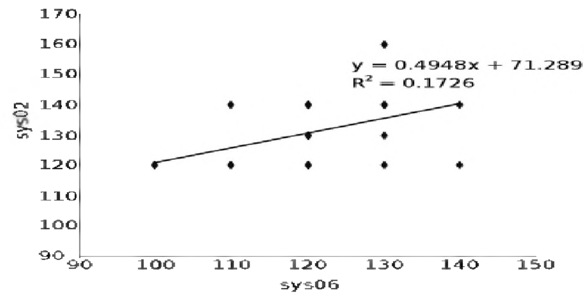


Figure 2. A graph of systolic blood pressure measured in 2002 against systolic

4. DISCUSSION AND CONCLUSIONS

The equation in figure 1 shows that a negative relationship exists between the blood pressures measured in the high noise zones and that measured in the low noise zones. Also, the level of association is low. This is shown by R2 which is only 3.3%. We therefore conclude that the null hypothesis that there is no immediate relationship between exposure to loud noise and high blood pressure should be accepted.

In figure 2, there is a positive relationship in the comparison of the systolic pressures measured in 2002 and 2006. Also, R2 = 17.3%. With the probability value (p-value) of 0.000 (< 0.05), we reject the null hypothesis that there is no long-term relationship between exposure to loud noise and high blood pressure. This finding replicates the positive results obtained in some retrospective longitudinal studies [5].

REFERENCES

1. Ising, M. (1979). *Study on the Quantification of Risk for the Heart and Circulating System Associated with Noise Workers*. Institute for Water and Air Hygiene of the BGA, Berlin.

2. Andr´en, L. (1982). Cardiovascular effects of noise. *Acta Medica Scandinavica,* 65, 1–45.

3. A.O. Akpan, E.O. Obisung and U.E. Asuquo (2009). Comparative study of impacts of industrial and aircraft noise on workers and community residents in parts of Rivers State, Nigeria. African journal of Environmental Pollution and Health, Vol 7, number 2. Pg 52- 58

4. Asuquo, U.E., Menkiti, A.E. Onuu, M.U. and Opaluwa, E.H.O. (2000). Environmental Noise Studies in some areas of Calabar and Uyo. *Global Journal of Pure and Applied Science*, 7(2), 339–344.

5. Pickering, T.G. and Gerin, W., (1988). Ambulatory blood pressure monitoring and cardiovascular reactivity testing for the evaluation of the role of psychosocial factors and prognosis in hypertensive patients, *American Heart Journal*, 116, 655–672.

EFFECTS OF EXPOSURE TO LOUD NOISE ON THE HEARING OF THE RESIDENTS OF CALABAR, NIGERIA

Ubon Asuquo[1], Michael Onuu[2] and Affiong Asuquo[3]
[1]Department of Mathematics, Trinity Western University, Langley, BC, Canada. (2011/2012 Session)
[2]Department of Physics, University of Calabar, Calabar, Nigeria.
[3]Department of Anaesthesiology, University of Calabar Teaching Hospital, Calabar, Nigeria

1. INTRODUCTION

As the society develops technologically, more and more people are exposed in one way or the other to noise exceeding 65 dB(A). In some countries, more than half of the population is exposed. When one realizes that at 65 dB(A) sleeping becomes seriously disturbed and most people become annoyed, it is clear that noise is a genuine environmental health problem.

Problems associated with noise – induced hearing loss are not new. In the middle ages, workers in certain profession such as blacksmithing, mining, and church bell ringing were known to become deaf, or partially deaf, after years of work [1]. However, with technological development, the number of workers exposed to excessive noise increased significantly as has the number of people exposed to other sources of noise, such as transportation noise and loud music [2]

The risk of much increased rates of occupationally acquired hearing loss must be met by strong preventive measures in engineering and medicine in both developed and developing countries

1.1 Hypotheses

The following null hypotheses were used for the research;

i. There is no relationship between objective measure of noise using sound level meter and the subjective measure of noise using questionnaires.
ii. There is no relationship between exposure to loud noise and noise – induced hearing loss.

2. MATERIALS AND METHODS

The materials used for this study are the sound level meter for measuring the sound level of the zones, an audiometer for testing hearing acuity and a 39-item questionnaire for subjective assessment of the respondents.

2.1 Site description

Guided by our preliminary measurements, sixteen sites were selected. Measurements were taken in these sixteen zones- eight high noise zones and eight low noise zones. The high noise zones had average A-weighted noise levels of 80 dB or above, a level which may be hazardous to the hearing of most people. The low noise zones had average A- weighted noise levels of 50 dB or below, a level above which most people complain

2.2 Measurements

Noise levels at schools and their surroundings were measured objectively using a precision sound level meter (Bruel and Kjaer) type 2203 with octave band filter (B&K) type 1613 by an accredited Acoustician from the Department of Physics, University of Calabar, Nigeria.
Measurements were taken between 7 am and 9 am, and 2 pm and 4 pm on working days (Monday to Fridays). About fifty random readings were taken at different locations within each zone and the sound level of each zone was calculated.

Hearing assessment of the respondents was done by a trained Technician using a pure tone audiometer in the Department of Otolaryngology, University of Calabar Teaching Hospital, Calabar using the audiometer. The frequency of the audiometer was varied from 64 Hz to over 8000 Hz. Amplitude was varied by 5 dB increments. The oscillator was used to change pitch so that a range of sounds can be tested. Hearing threshold measurements were carried out separately for both the right and left ears of the 2000 subjects using the Hughson-Westlake (ascending/descending) procedure. Each audiometric test was preceded by an auriscopic examination in order that any coexistence significant aural pathological conditions such as wax impaction, tympanic membrane perforation, etc., could be detected.

Subjective measurement was done using questionnaires. A total of 1300 questionnaires were distributed in the high noise zones while 1200 were distributed in the low noise zones. 1100 (85.6%) were collected back from the high noise zones while 1050 (87.5%) were collected back from the low noise zones. However, 1000 questionnaires were accepted for analysis from both the low and the high noise zones. This is because 100 and 50 questionnaires were rejected from the high noise zones and the low noise zones, respectively because the respondents did not stayed or worked/schooled in the area for up to three years.

3. RESULTS

The effects of exposure to loud noise on hearing of people in Calabar were assessed by comparing the hearing acuity and subjective responses of respondents in the high noise zones with that of respondents in the low noise zones. The respondents taken were those that have resided or done business in these areas for at least three years, a period believed to be enough for the respondents to be adapted to the noise environment of the zone. Results of physical measurements are presented in Tables 1, 2, 3 and 4.

Table 1. Noise levels in the high noise zones.

Code	Background Noise ± 0.5 dB(A)	A-Weighted SPL ± 0.5 dB(A)	Lmax ± 0.5 dB(A)
HNZ 1	56.0	118.0	121.5
HNZ 2	51.0	116.5	120.0
HNZ 3	54.0	109.5	112.0
HNZ 4	57.0	102.0	108.0
HNZ 5	54.5	110.0	115.0
HNZ 6	50.5	116.0	119.0
HNZ 7	60.5	129.0	131.0
HNZ 8	60.0	112.0	118.0

Table 2. Noise levels in the low noise zones.

Code	Background Noise Level ± 0.5 dB(A)
LNZ 1	47.5
LNZ 2	40.0
LNZ 3	44.5
LNZ 4	40.0
LNZ 5	43.5
LNZ 6	38.0
LNZ 7	47.0
LNZ 8	44.5

Table 3. Hearing assessment in the high noise zones

Threshold levels (dB)	Number of Respondents	
	Right ear	Left Ear
< 25	540	367
26 – 40	392	560
41 – 60	60	40
61 - 80	5	25
> 80	3	8

Table 4. Hearing assessment in the low noise zones

Threshold levels (dB)	Number of Respondents	
	Right ear	Left Ear
< 25	620	720
26 – 40	375	251
41 – 60	3	20
61 - 80	2	7
> 80	1	2

3.1 Correlation between the subjective and objective responses

To determine how related the subjective responses, assessed by the use of questionnaires as the study instrument were to the objective responses measured with the sound level meter, the coefficients of correlation[3] were calculated for the noise measurements.

The objective responses measured with the sound level meter represent x-variates and the subjection responses represented by corresponding average value per zone, as y-variates. Substituting these data into the correlation equation, we have the correlation coefficient to be 0.66 and 0.56 for the high noise zones and the low noise zones respectively. The results show that there are good correlations between the two measurements in both zones.

3.2. Effects of noise on hearings

The hearing assessment of respondents in the high and low noise zones were shown in Tables 3 and 4 respectively. Analyses of these tables were done using the Statistical Analysis Software (SAS). The results in Tables 3 and 4 show that majority of respondents had hearing loss. In the high noise zone, 1.45% had hearing loss in different degrees in the right ear while 15.82% had hearing loss in the left ear. In the low noise zones, 9.43% had hearing loss in the right ear while 6.98% had hearing loss in the left ear. The probability value is 0.0001. This is statistically significant and so we reject the null hypothesis that there is no relationship between exposure to loud noise and noise-induced hearing loss.

To further confirm our decision, we consider the chi-squares value. Here, the chi-square ($\chi 2$) value is 82.2509. The tabulated ($\chi 2$) value is 5.22. Therefore, since the calculated ($\chi 2$) is greater than the tabulated ($\chi 2$), this confirms the null hypothesis should be rejected. Therefore, we can conclude from the analysis that exposure to loud noise do cause noise-induced hearing loss, as seen in this study. This is in line with previous findings [4]

4. DISCUSSION AND CONCLUSIONS

Noise is a disturbance to the human environment that is escalating at such a high rate that it will become a major threat to the quality of human lives if nothing is done to reduce it. Noise has been a constant threat since the industrial revolution. Too much noise exposure may cause a temporary change in hearing or a temporary ringing in your ears (tinnitus). These short-term problems usually go away within a few minutes or hours after leaving the noise. However, repeated exposures to loud noise can lead to permanent, incurable hearing loss or tinnitus.

REFERENCES

1. B. Berglund and T. Lindvall, (Eds.) (1995). *Community Noise*, Centre for Sensory Research, Stockholm, Sweden.
2. U. E. Asuquo, M. U. Onuu, A. O. Akpan and A. U. Asuquo (2009). Noise and blood pressure: A cross sectional and longitudinal study of the effects of exposure to loud noise on residents in Calabar, Cross River State, Nigeria. *International Journal of Acoustics and Vibration*, 14(2), 1–15.
3. A. O. Adedayo (1998), *Understanding Statistics* JAS Publishers, Lagos.
4. O. Bethesda (1999). Noise-induced heaving loss. National Institute on deafness and other communication disorder. *Noise Induced Hearing Publication*, 97, 1433–1441.

ACKNOWLEDGEMENTS

PROF. A. I. MENKITI – PHYSICS DEPT., UNICAL., NIGERIA.

EFFECTS OF EXPOSURE TO LOUD NOISE ON THE HEARING OF THE RESIDENTS OF CALABAR, NIGERIA

Ubon Asuquo[1], Michael Onuu[2] and Affiong Asuquo[3]
[1]Department of Mathematics, Trinity Western University, Langley, BC, Canada. (2011/2012 Session)
[2]Department of Physics, University of Calabar, Calabar, Nigeria.
[3] Department of Anaesthesiology, University of Calabar Teaching Hospital, Calabar, Nigeria

1. INTRODUCTION

As the society develops technologically, more and more people are exposed in one way or the other to noise exceeding 65 dB(A). In some countries, more than half of the population is exposed. When one realizes that at 65 dB(A) sleeping becomes seriously disturbed and most people become annoyed, it is clear that noise is a genuine environmental health problem.

Problems associated with noise – induced hearing loss are not new. In the middle ages, workers in certain profession such as blacksmithing, mining, and church bell ringing were known to become deaf, or partially deaf, after years of work [1]. However, with technological development, the number of workers exposed to excessive noise increased significantly as has the number of people exposed to other sources of noise, such as transportation noise and loud music [2]

The risk of much increased rates of occupationally acquired hearing loss must be met by strong preventive measures in engineering and medicine in both developed and developing countries

1.1 Hypotheses

The following null hypotheses were used for the research;

i. There is no relationship between objective measure of noise using sound level meter and the subjective measure of noise using questionnaires.

ii. There is no relationship between exposure to loud noise and noise – induced hearing loss.

2. MATERIALS AND METHODS

The materials used for this study are the sound level meter for measuring the sound level of the zones, an audiometer for testing hearing acuity and a 39-item questionnaire for subjective assessment of the respondents.

2.1 Site description

Guided by our preliminary measurements, sixteen sites were selected. Measurements were taken in these sixteen zones-eight high noise zones and eight low noise zones. The high noise zones had average A-weighted noise levels of 80 dB or above, a level which may be hazardous to the hearing of most people. The low noise zones had average A- weighted noise levels of 50 dB or below, a level above which most people complain

2.2 Measurements

Noise levels at schools and their surroundings were measured objectively using a precision sound level meter (Bruel and Kjaer) type 2203 with octave band filter (B&K) type 1613 by an accredited Acoustician from the Department of Physics, University of Calabar, Nigeria.
Measurements were taken between 7 am and 9 am, and 2 pm and 4 pm on working days (Monday to Fridays). About fifty random readings were taken at different locations within each zone and the sound level of each zone was calculated.

Hearing assessment of the respondents was done by a trained Technician using a pure tone audiometer in the Department of Otolaryngology, University of Calabar Teaching Hospital, Calabar using the audiometer. The frequency of the audiometer was varied from 64 Hz to over 8000 Hz. Amplitude was varied by 5 dB increments. The oscillator was used to change pitch so that a range of sounds can be tested. Hearing threshold measurements were carried out separately for both the right and left ears of the 2000 subjects using the Hughson-Westlake (ascending/descending) procedure. Each audiometric test was preceded by an auriscopic examination in order that any coexistence significant aural pathological conditions such as wax impaction, tympanic membrane perforation, etc., could be detected.

Subjective measurement was done using questionnaires. A total of 1300 questionnaires were distributed in the high noise zones while 1200 were distributed in the low noise zones. 1100 (85.6%) were collected back from the high noise zones while 1050 (87.5%) were collected back from the low noise zones. However, 1000 questionnaires were accepted for analysis from both the low and the high noise zones. This is because 100 and 50 questionnaires were rejected from the high noise zones and the low noise zones, respectively because the respondents did not stayed or worked/schooled in the area for up to three years.

3. RESULTS

The effects of exposure to loud noise on hearing of people in Calabar were assessed by comparing the hearing acuity and subjective responses of respondents in the high noise zones with that of respondents in the low noise zones. The respondents taken were those that have resided or done business in these areas for at least three years, a period believed to be enough for the respondents to be adapted to the noise environment of the zone. Results of physical measurements are presented in Tables 1, 2, 3 and 4.

Table 1. Noise levels in the high noise zones.

Code	Background Noise ± 0.5 dB(A)	A-Weighted SPL ± 0.5 dB(A)	Lmax ± 0.5 dB(A)
HNZ 1	56.0	118.0	121.5
HNZ 2	51.0	116.5	120.0
HNZ 3	54.0	109.5	112.0
HNZ 4	57.0	102.0	108.0
HNZ 5	54.5	110.0	115.0
HNZ 6	50.5	116.0	119.0
HNZ 7	60.5	129.0	131.0
HNZ 8	60.0	112.0	118.0

Table 2. Noise levels in the low noise zones.

Code	Background Noise Level ± 0.5 dB(A)
LNZ 1	47.5
LNZ 2	40.0
LNZ 3	44.5
LNZ 4	40.0
LNZ 5	43.5
LNZ 6	38.0
LNZ 7	47.0
LNZ 8	44.5

Table 3. Hearing assessment in the high noise zones

Threshold levels (dB)	Number of Respondents	
	Right ear	Left Ear
< 25	540	367
26 – 40	392	560
41 – 60	60	40
61 - 80	5	25
> 80	3	8

Table 4. Hearing assessment in the low noise zones

Threshold levels (dB)	Number of Respondents	
	Right ear	Left Ear
< 25	620	720
26 – 40	375	251
41 – 60	3	20
61 - 80	2	7
> 80	1	2

3.1 Correlation between the subjective and objective responses

To determine how related the subjective responses, assessed by the use of questionnaires as the study instrument were to the objective responses measured with the sound level meter, the coefficients of correlation[3] were calculated for the noise measurements.

The objective responses measured with the sound level meter represent x-variates and the subjection responses represented by corresponding average value per zone, as y-variates. Substituting these data into the correlation equation, we have the correlation coefficient to be 0.66 and 0.56 for the high noise zones and the low noise zones respectively. The results show that there are good correlations between the two measurements in both zones.

3.2. Effects of noise on hearings

The hearing assessment of respondents in the high and low noise zones were shown in Tables 3 and 4 respectively. Analyses of these tables were done using the Statistical Analysis Software (SAS). The results in Tables 3 and 4 show that majority of respondents had hearing loss. In the high noise zone, 1.45% had hearing loss in different degrees in the right ear while 15.82% had hearing loss in the left ear. In the low noise zones, 9.43% had hearing loss in the right ear while 6.98% had hearing loss in the left ear. The probability value is 0.0001. This is statistically significant and so we reject the null hypothesis that there is no relationship between exposure to loud noise and noise-induced hearing loss.

To further confirm our decision, we consider the chi-squares value. Here, the chi-square ($\chi 2$) value is 82.2509. The tabulated ($\chi 2$) value is 5.22. Therefore, since the calculated ($\chi 2$) is greater than the tabulated ($\chi 2$), this confirms the null hypothesis should be rejected. Therefore, we can conclude from the analysis that exposure to loud noise do cause noise-induced hearing loss, as seen in this study. This is in line with previous findings [4]

4. DISCUSSION AND CONCLUSIONS

Noise is a disturbance to the human environment that is escalating at such a high rate that it will become a major threat to the quality of human lives if nothing is done to reduce it. Noise has been a constant threat since the industrial revolution. Too much noise exposure may cause a temporary change in hearing or a temporary ringing in your ears (tinnitus). These short-term problems usually go away within a few minutes or hours after leaving the noise. However, repeated exposures to loud noise can lead to permanent, incurable hearing loss or tinnitus.

REFERENCES

1. B. Berglund and T. Lindvall, (Eds.) (1995). *Community Noise*, Centre for Sensory Research, Stockholm, Sweden.
2. U. E. Asuquo, M. U. Onuu, A. O. Akpan and A. U. Asuquo (2009). Noise and blood pressure: A cross sectional and longitudinal study of the effects of exposure to loud noise on residents in Calabar, Cross River State, Nigeria. *International Journal of Acoustics and Vibration*, 14(2), 1–15.
3. A. O. Adedayo (1998), *Understanding Statistics* JAS Publishers, Lagos.
4. O. Bethesda (1999). Noise-induced heaving loss. National Institute on deafness and other communication disorder. *Noise Induced Hearing Publication*, 97, 1433–1441.

ACKNOWLEDGEMENTS

PROF. A. I. MENKITI – PHYSICS DEPT., UNICAL., NIGERIA.

On The Use Of Smartphones For Occupational Noise Monitoring: Instrumentation

Romain Dumoulin, Jérémie Voix
École de technologie supérieure, Montreal, Qc., Canada
jeremie.voix@etsmtl.ca

ABSTRACT

This paper presents an on-going research effort that focuses on a smartphone-based occupational noise monitoring platform. A laboratory calibration method for smartphones and embedded devices, together with an innovative "field" calibration are detailed. The evaluation of uncertainties associated with the use of smartphones for noise level measurement is discussed with regards to the main individual uncertainty components.

1. INTRODUCTION

Although noise induced hearing loss represents the number one occupational disease, individual workers' noise exposure levels are still rarely precisely known and infrequently tracked. Indeed, standardized noise exposure campaigns have as their principle disadvantages the cost of instrumentation and the practical difficulties associated with real world implementation. In opposition to these procedures, informal noise surveys carried out with basic and inexpensive sound level meters are not necessarily precise or accurate enough.

The *WikiLeq* project [1] proposes the use of smartphones as an alternate solution. The first step of the project focuses on the evaluation of the measurement quality through the assessment of uncertainties associated with the instrumentation. In a second step, the uncertainty associated with time and spatial sampling strategies for noise exposure assessment will be carefully evaluated.

2. ANDROID IMPLEMENTATION

WikiLeq is an open source framework for monitoring occupational noise based on the NoiseTube project [2]. It combines two simultaneous approaches: the first one features a personal full work shift dosimetric assessment, while the second features a participative sound pressure levels mapping.

An Android™ (v2.3.5) application has been developed for the personal full work shift dosimetric assessment: A-weighting and real-time octave band filters (63 to 8,000 Hz) have been implemented for a 16-bits audio stream acquired at 22,050 Hz. This application requires the user to define what type of microphone device is used and where it will be located on the user's body; an associated *calibration adjustment* is determined and applied to the measured L_{eq}. One-second octave-band equivalent level, $L_{eq,1s}$, are then computed and tagged with GPS data. A cumulative long-term L_{eq} is displayed within the application and will be in near future presented with the associated expanded measurement uncertainty (detailed in Section 4). Finally, the locally stored data is sent to the *WikiLeq* server for data aggregation required for the participative sound pressure levels mapping.

3. CALIBRATION PROTOCOLS

3.1. Laboratory calibration

This procedure, based on IEC 1183 standard [3], aims at separating the diffuse field sensitivity levels associated with the directional characteristics of a device from the one associated with the effects of its mounting position. The resulting *random-incidence calibration value*, regroups these two effects, and will be determined for a wide range of noise levels for one specific microphone device among supported by the WikiLeq application such as the phone embedded microphone, the "in-line" microphone on headphone cord and the microphone of a Bluetooth® earpiece.

- <u>Random Incidence Microphone Placement Error</u>

In a reverberation room, the reference microphone measurements, conducted on a head and torso simulator (HATS) at different mounting positions are compared to measurements done with the same microphone at the center-of-head position without the mannequin. Acoustical reflections and mannequin shielding effects lead to a random-incidence microphone placement error, ε_{mp} , that can be measured for every octave band at typical microphone positions: on the belt holster, in the breast pocket, on an "in-line" headphone cord and on top of the shoulder, as per ISO 9612 standard [4].

- <u>Random-incidence sensitivity levels</u>

The random-incidence sensitivity levels, G_{RI}, are calculated for each octave band from Eq. 1 below and measured in a semi-anechoic room. The free-field sensitivity level, G_F, is measured for a reference direction of sound incidence and directivity factor, γ, assessed from an IEC 1183 procedure.

$$G_{RI} = G_F - 10 \log (\gamma) \quad (1)$$

- <u>Calibration adjustment</u>

The obtained G_{RI} values are stored in the *WikiLeq* application calibration module and will be added to ε_{mp} to obtain a so-called octave-band *calibration adjustment* specific to the microphone device used and its mounting position.

3.2. Field calibration proposed approach

A laboratory calibrated smartphone will now be used to calibrate an uncalibrated microphone device in the field by assessing its free-field sensitivity levels with the following proposed approach: the calibrated and uncalibrated phones are physically brought close together to be immersed in the same sound-field and a real-time audio recording is performed. A dual channel FFT analysis, implemented in the Android "app", will estimate the transfer function between the two microphone devices and compute a free-field sensitivity levels, G_F, for the device under field calibration. The normalized random error for the frequency response magnitude is calculated from the coherence function in order to evaluate the quality of that field calibration. The values of G_F, are then stored in the newly calibrated phone app with an associated "field calibration uncertainty" that quantifies the quality of the proposed field calibration. An average directivity factor, $\bar{\gamma}$, for the three types of microphone device is calculated. The figure illustrates the calculation of the *calibration adjustment*.

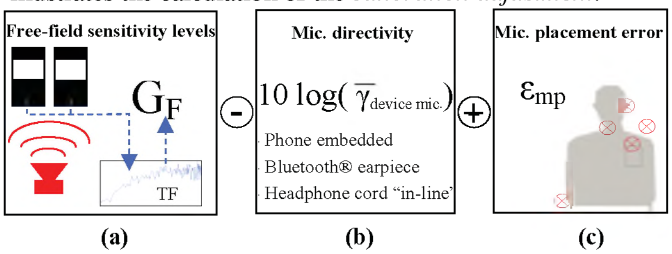


Fig.1. Calculation of the *calibration adjustment* with the field calibration values; assessing free-field mic. sensitivity relative to calibrated device (a), then accounting for mic. directivity (b) and mic. placement position (c).

4. INSTRUMENTATION UNCERTAINTIES

4.1. Main uncertainties considered

Prior published analysis on noise dosimeter and sound level meter measurements uncertainties has been revisited for the envisioned use of a smartphone. For a large population of IEC-compliant noise dosimeters, 4 main uncertainties [5] are given: microphone placement errors, frequency response errors, linearity and sensitivity errors. On top of these, the uncertainties associated with the *WikiLeq* lab and field calibrations as well as the long term environmental and aging drifts need to be considered.

- Microphone placement error is partially addressed in the *calibration adjustment* calculations. In a free-field situation (with a source at a particular angle of incidence) the microphone placement error will be underestimated [5]. It is reduced when the worker is mobile with respect to the noise source [5].
- Frequency response uncertainty is principally influenced by the data acquisition hardware (microphone, pre-amplifier,…) since A-weighting and octave-bands filters are digitally implemented. The quality of the laboratory calibration process impacts significantly this uncertainty.
- Linearity and sensitivity uncertainties are due to the discrepancies between the measured sound pressure level and the reference sound pressure level for a 30 dB range over 90 dB sound pressure level (at which sensitivity is assessed). Again, the quality of the laboratory calibration process impacts significantly this uncertainty while it rises naturally at higher noise levels because of the stalling "slope" of the calibration response curve.
- Laboratory calibration error, ε_{mp} , may lead to an error since it is measured without considering the directivity of the microphone under calibration. Uncertainties associated with the "laboratory" measurements of the sensitivity levels and directivity factors must also be taken into account.
- Field calibration error includes the above-mentioned laboratory calibration error and microphone placement error, as well as the error due to the use of an averaged directivity factor and the uncertainties associated with the dual channel FFT analysis applied.
- Long term environmental and aging drifts error may be taken at first from published work: Electret and MEMS microphones showed a negative correlation between sound level and temperature with a global error around 1 dB [6], while the variability from different phones of the same model was determined as negligible [7].

4.2. Discussion about the evaluation of uncertainties

The determination of these main individual uncertainty components aims to evaluate a practical value of the instrumentation error for the overall measured noise level. In ISO 9612 [4], the standard uncertainties associated with the instrumentation and the measurement position are defined for standardized instrument and are based on empirical data. These empirical data must be used, since the tolerance limits given in IEC 1252 [8] would lead to an overestimation of the instrumentation uncertainties [4]. The practical approaches for the assessment of each of these uncertainty components are still to be developed and preferably using "hands-on" approaches as in [5] and [9] for the specific constrains of the proposed use of smartphone-based instruments. For example, based on the evaluation tests defined in IEC 1252 as reference, the sinusoidal signals source should be modified for more "real world" industry noises.

REFERENCES

[1] http://wikileq.org/en

[2] Niessen, M. E. (2009) "NoiseTube: Measuring and mapping noise pollution with mobile phones", Environmental Engineering.

[3] IEC standard 1183 (1994) "Electroacoustics— Random-incidence and diffuse-field calibration of sound level meters."

[4] ISO, ISO standard 9612 (2009) "Acoustics – Determination of occupational noise exposure –Engineering method".

[5] Giardino, D. A., & Seiler, J. P. (1996). "Uncertainties associated with noise dosimeters in mining", *100*, J.A.S.A. 100

[6] Renterghem, T. V, et al. (2011) "On the ability of consumer electronics microphones for environmental noise monitoring", Journal of Environmental Monitoring, 544-552.

[7] Hondt, E. D., et al. (2011) "Participatory noise mapping works !" Pervasive and Mobile Computing, 1-27.

[8] IEC standard 1252 (1993) "Electroacoustics — Specifications for personal sound exposure meters."

[9] Payne, R. (2004) "Uncertainties associated with the use of a sound level meter." NPL Report DQL-AC 002

A COMPARISON OF ISO 9613-2 AND ADVANCED CALCULATION METHODS: PREDICTIONS VERSUS EXPERIMENTAL RESULTS

Panos Economou[1], Richard J. Peppin[2]
[1]P.E. Mediterranean Acoustics Research & Development Ltd, Cyprus, panos@pemard.com
[2]Richard J. Peppin, Scantek, Inc., USA, PeppinR@ASME.org

1 INTRODUCTION

Standardization provides methodologies by which independent investigations of the same situation are able to derive the same conclusions. However, standardization is sometimes also perceived as absolute and accurate, beyond which one should not investigate matters deeper. Moreover, often enough the engineering community tends to neglect the science (or lack of it) underlying standardized methods and just follows prescriptions. Standardization provides algorithms that can be turned into software code. Software developers are always looking for ready-made algorithms with great market potential. The responsibility of the accuracy of these methods does not lie with the developers but with the standards organizations. This is not the case with algorithms based on pure scientific research where the full responsibility lies with those who turn it into software applications. Faced with a choice between prescribed methods of standardization with simple mathematical code vs. accurate scientific findings, PEMARD applied the latter developing complicated mathematical computation, albeit slower, yet delivering more accurate results in its commercially available software application[1].

2 COMMERCIALLY AVAILABLE THEORETICALLY BASED SOFTWARE APPLICATION THEORETICAL BACKGROUND

The Simulation and prediction of outdoor sound propagation using advanced calculation methods are based on principles of physics with an effort to try to avoid empirical or approximate methods, often found in published outdoor propagation standards.

The commercially available software application, utilizes sound ray modeling which solves Helmholtz's sound wave equation, accounting for sound diffraction to any order, sound wave reflection from finite size surfaces of finite impedance using Fresnel Zones and spherical wave reflection coefficient concepts, respectively. The approach uses flow resistivity instead of sound absorption coefficient. It also takes into account geometrical spreading, atmospheric absorption, and atmospheric turbulence. The software application also has an in-house developed algorithm to detect valid diffraction and reflection sound paths from source to receiver in a proper 3D environment. It is based on the image source method and the Geometrical Theory of Diffraction according to Keller[2].

3 ISO 9613-2[3] BACKGROUND

This standard is an empirical standard[4], which at the time of its preparation and publication, only few dedicated acoustical software applications existed for which most of the potential users of such software were "computer-phobic" since program user interface was not as convenient as it is today. It can therefore be said that there were good grounds to apply simpler empirical methods at the time.

There are however, many limitations in this method. The standard is vague, allowing the user to interpret the methodology according to their understanding, thus the user decides on whether vertical diffraction paths are important or not, as well as other significant parameters affecting final outcome.

4 PRESENTATION OF COMPARISON OF RESULTS AMONG COMMERCIALLY AVAILABLE SOFTWARE APPLICATION, ISO 9613-2 AND PUBLISHED MEASURED DATA.

4.1 Published measured data used as comparison reference

Two cases are presented based on sound measurements taken and presented in the Delta Report of 2006, "Nord2000: Validation of the Propagation Model for The Danish Road Directorate"[5].

Distance S - R		
	Thick barrier	Multiple barrier
4.5 m	Case 33	Case 36

Table 1: Cases used for the validation of NORD 2000 and implemented with ISO 9613-2 and the commercially available software application.

4.2 Comparison results

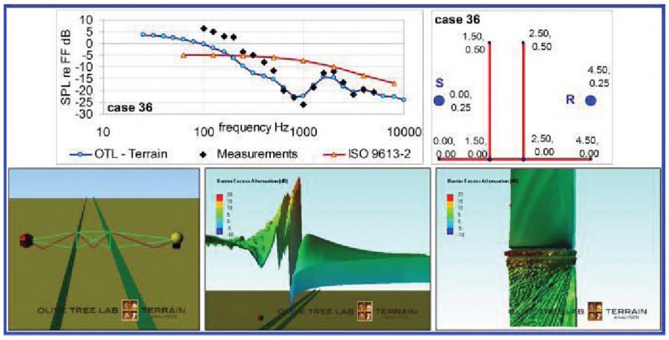


Figure.1. Results for Case 33.

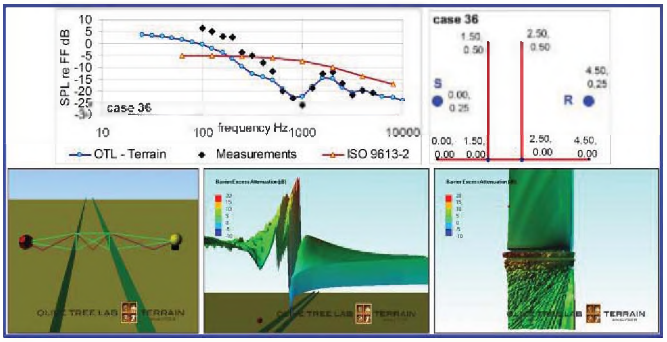


Figure.2. Results for Case 36.

4.3 Measurements Data

Measurement data used were the sound measurements results used to validate the Nord 2000 model. However, there is very little information available on the measurement methodology used. Some of the cases were able to be tracked down, which are not included in this paper (but are included in the report by Delta) and are described by K.B. Rasmussen in his work "On the effect of terrain profile on sound propagation outdoors."[6]

4.4 ISO 9613-2 results

Other than the obvious deviations of ISO 9613-2 calculation results from sound measurements results, ISO calculations lacks detail depriving the interpretation of the outdoor sound propagation mechanisms which come into play. On the other hand high-resolution results, allow for the interpretation of the sound propagation mechanisms that take place over ground and obstacles.

4.5 Theoretical Approach- e.g. Commercially available software application results

Even though the results from the software application matched fairly well with the results from measured data, they were expected to have a better agreement. This is because the software application was validated against other measured data[7], which as previously mentioned, there is limited information on the details of how the validation data were obtained. Furthermore, slight lateral shifts of source or receiver with respect to the barrier produce significant changes in results.

Another factor usually ignored in measurements studying sound interference phenomena, is the diffraction effects from loudspeaker cabinet units which contaminates sound measurements results by effectively turning one sound source into many sources.

5 CONCLUSIONS

ISO 9613-2 is an empirical method which is simple to understand and implement, widely used ever since its publication in 1996. It has served the acoustical community well, but this paper shows that it yields results that are inaccurate and imprecise.

Advanced calculation methods provide a unified approach in acoustics with one calculations engine to deal with most topics in acoustics. They offer the ability to simulate complicated environments by using simple rules and they apply accurate general solutions without vagueness to all scenarios, including special cases. Perhaps new scientific methods can now be used that allow the study of outdoor sound propagation mechanisms. The software application simulates sound propagation in a three dimensional environment with a possibility of eventually including all phenomena deemed important in acoustics.

The main disadvantage of advanced calculation methods is that they are still computationally expensive, and a better understanding of the science behind them is needed by end users for the proper interpretation of the results.

6 REFERENCES

1. Olive Tree Lab – Terrain, An Outdoor Sound Propagation Calculation Software. http:// www.otlterrain.com
2. Keller. J. B. (1962). "Geometrical Theory of Diffraction," JOSA 52(2).
3. International Standards Organization, "Acoustics - Attenuation of sound during propagation outdoors - Part 2: General method of calculation," International Organization for Standardization, (1996).
4. Attenborough, K., Li, K.M. & Kirill, H. (2007) "Predicting Outdoor Sound, Taylor & Francis. (New York).
5. DELTA, Danish Electronics Light & Acoustics: www.delta.dk. Nord2000. "Validation of the Propagation Model for The Danish Road Directorate." (2006).
6. Rasmussen. K.B. (1985). "On the effect of terrain profile on sound propagation outdoors," J. Sound Vibrat. 98(1), 40.
7. http://www.otlterrain.com/ValidationProjects.aspx.

INDUSTRIAL APPLICATION OF SIMPLESILENCE TECHNOLOGY: EVAPORATOR FAN NOISE CONTROL

Anthony Gérard[1], Michel Pearson[2]

[1]Soft dB, 1240 rue Beaumont, suite 206, Montréal (QC), Canada, H3P 3E5 a.gerard@softdb.com
[2]Soft dB, 1040 Avenue Belvédère, suite 215, Québec (QC), Canada, G1S 3G3 m.pearson@softdb.com

1. INTRODUCTION

Tonal noise control from evaporator fans used for cooling a cold storage room is described. The low frequency tone at blade passage frequency (BPF=90Hz) can result in environmental noise penalty (added to the global noise level) in the residential area near the cold storage room. For low frequency, passive techniques are bulky, inefficient and cannot be applied to the food industry due to hygienic constraints, but the SimpleSilence technology is better adapted to those frequencies and have a great potential for hygienic and "at the source" control.

Tonal noise originates from non-uniform flow that causes circumferential varying blade forces and gives rise to considerably larger radiated dipolar sound at the BPF and its harmonics[1]. For the evaporator studied in this paper, the primary non uniform flow mainly came from the motor struts and the ice accretion in the upstream flow field of the fan (in the casing or on the struts supporting the motor). For low speed axial fans, the circumferential flow pattern having the same number of lobes as the number of blades (B) will emit intensive sound at the BPF.

2. METHOD

The concept of SimpleSilence technology is to produce destructive interference between the primary source and a B-periodic obstruction[2,3]. The flow control obstruction (Fig. 1) is located such that the secondary radiated noise is of equal magnitude but opposite in phase compared to the primary noise.

Figure. 1. Left: example of flow control obstructions installation (upstream view). Right: rotor and its environment (strut, ice accretion, radiator…)

The obstruction can be located upstream or downstream. When located upstream, the secondary noise is caused by obstruction/rotor interaction and when it is located downstream, the secondary noise is caused by rotor/obstruction interaction. The magnitude of the secondary noise can be controlled by the axial distance between the rotor and the obstruction, or by the radial extent of the lobes of the obstructions. The phase of the secondary noise can be controlled by the angular location of the obstruction. On the downstream side of the fan, a positioning device allowed for the obstruction to be moved in the angular position. The fan (right of Fig. 1) had B=4 regularly spaced "basic blades" and its rotational velocity was 1370 R.P.M., corresponding to a BPF at 91 Hz. The fan diameter was 91 cm. The B=4 period trapezoidal obstructions used in this study (shown in Fig. 1) were made of stainless steel and had an inner diameter of 67 cm and an outer diameter of 80 cm. Acoustic pressure measurements were performed at 3 microphone locations (2 downstream and 1 upstream), using a multifunction 4 channel acoustic measuring system.

3. RESULTS

3.1. Control near the fan

The average of sound pressure spectra measured at 3 microphone locations around fan 1 and fan 2 are shown in Fig. 2. The attenuations of the BPF were about 6 dBA and 3 dBA for fan 1 and fan 2 respectively. The global attenuations were about 0.8 dBA and 2 dBA for fan 1 and fan 2 respectively.

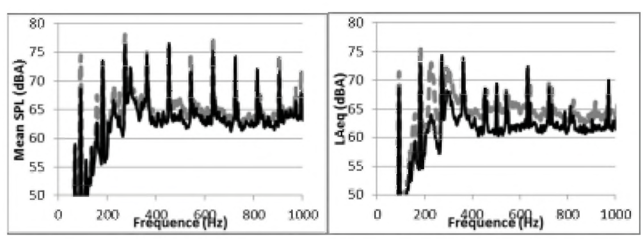


Figure. 2. Averaged sound pressure level spectra from Fan 1 (left) and Fan 2 (right). Dashed grey line: fan noise without the flow control obstruction and solid line: fan noise with flow control obstruction

Close to the evaporator, better attenuations of the BPF tone were obtained in the upstream flow field (up to 10 dB for the two upstream microphones).

3.2. Control in the residential area

The A-weighted sound pressure spectra measured in the residential area near the cold storage room are shown in Fig. 3 (without the obstructions and with the obstruction installed for 5 fans). The BPF tone was decreased by 2.4 dB (several harmonics was also attenuated). The equivalent continuous A-weighted sound pressure level was 55.8 dBA without the flow control obstructions and 51.9 dBA with the flow control obstructions, leading to an attenuation 3.9 dB.

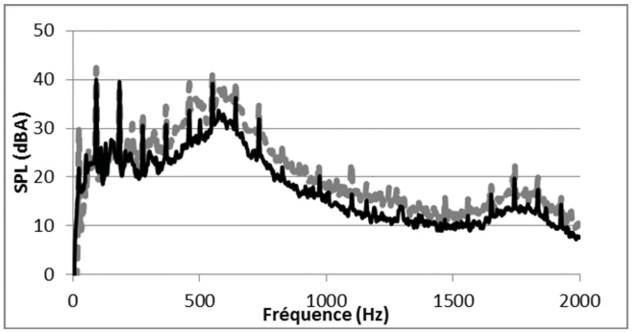


Figure. 3. Sound pressure level spectra measured in the residential area near the cold storage room. Dashed grey line: without flow obstructions and solid line: with flow control obstructions.

4. DISCUSSION

The reduction of the broadband noise from fan 2, and other evaporator fans in the cold storage room (not shown in this paper), was unexpected. For example, the spectrum from fan 2 without flow control obstruction exhibits subharmonic narrowband peaks around the first and the second harmonic of the BPF, which is probably caused by rotating blade flow instabilities at the blade tip. The obstruction probably controlled these instabilities so that the narrowband peak and low frequency broadband noise were decreased. CFD simulations could give more insight in the flow topology without and with the obstruction.

No aerodynamic performance measurements were performed in the cold storage room. However, experiments on several other fans with and without flow control obstruction located in the upstream flow field showed that the control obstruction has low or almost no effect on the fan efficiency. For controlling the BPF tone from an automotive engine cooling axial fan (Siemens-VDO)[3], the effect of a 121 cm2 obstruction (the area of the rotor was 707 cm2) was negligible on the static pressure, the flow and the efficiency of the fan. Same trends were observed for Valeo cooling fans. For large centrifugal fans, the flow decreased by 1% to 2% when the obstruction was in place[4]. The worst flow loss (3.5%) was measured for a tractor axial fan using a large sinusoidal obstruction in the upstream flow field[4].

The proposed approach is well adapted to acoustically compact fan (large wavelength, compared to the source dimension), for which one unsteady lift mode has a major contribution to the radiated noise at BPF. Increasing the rotation Mach number or the radius of the fan leads to an increase of the number of the radiating circumferential unsteady lift modes contributing to the noise at a single frequency (especially the modes B-1, B and B+1). In[5], a method has been proposed to combine a B-1 and a B lobed obstruction to enhance spatially the control of the BPF for a small radiator cooling fan in anechoic conditions.

The control approach is adapted to the control of PC fans, air conditioning fans, residential heat pump, automotive fans, engine cooling fans…

REFERENCES

1. Blake, W.K. (1986). Mechanics of flow-induced sound and vibration. Vol. 2: Complex flow structure interaction, Academic Press Inc., London
2. Gérard, A., Berry, A., Masson, P. and Gervais, Y (2009). Modeling of tonal noise control from subsonic axial fans using flow control obstructions, Journal of Sound Vibration., 321 (1-2), 26-44.
3. Gérard, A., Berry, A., Masson, P. and Gervais, Y (2009). Experimental validation of tonal noise control from subsonic axial fans using flow control obstructions, Journal of Sound Vibration., 321 (1-2), 8-25.
4. Goth, Y., Besombes, M., Chassaignon, C. and Gérard, A. (2012) Fan tonal noise reduction using calibrated obstructions in the flow - An experimental approach, Fan 2012, Senlis, France, 18-20 Apr. 2012.
5. Gérard, A., Moreau, S., Berry, A. and Masson, P (2012). Multimodal obstruction to control tonal fan noise: theory and experiments, 14th International Symposium on Transport Phenomena and Dynamics of Rotating Machinery, 27 feb-2 mar, Honolulu, HI, USA.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Dave Nadeau and Dave Leclerc from Soft dB, who helped in the installation of the obstructions in the cold storage room at -30°C.

Measurements of Propagation of Sound over Water at Low to Mid Frequencies Applicable to Wind Turbine Noise Calculation

Tim Kelsall

Director of Noise and Vibration, Hatch, 2800 Speakman Drive, Mississauga, ON L5K 2R7 - tkelsall@hatch.ca

The ISO 9613(2) standard for calculation of noise propagation over land was written excluding propagation over water. A series of measurements of sound at various distances from a large highway bridge have been taken and are compared to ISO 9613(2) calculations for the same scenario. They show that for lower frequencies, which are the most important for offshore wind turbine calculations, ISO 9613(2) provide a good approximation to the measured results and may prove suitable for offshore wind turbine noise predictions.

1. INTRODUCTION

Ontario has a moratorium on offshore wind turbines, at least in part due to the absence of a proven method of predicting their noise impacts. ISO 9613(2)[1] is used for this purpose on land; however the standard specifically excludes propagation over water.

A series of measurements were taken at night from a sailboat on Lake Ontario at various distances from the Burlington Skyway, a large highway bridge at the West end of the lake. This bridge is approximately 3km in length and 75m high, which is not dissimilar to a wind farm.

2. MEASUREMENTS

Measurements were taken at 0.54km, 1km and at 1km intervals out to 12 km from the bridge centreline. Temperatures, measured by an automatic meteorological station close to the bridge, dropped from 26 to 24°C over the measurement period from 10pm to 1am while relative humidity rose from 64 to 75%. The winds were very light (6km/h) and off shore, i.e. out of the harbour towards the Lake, and it was sufficiently calm that the boat had to use its engine between measurement points and the waves died away as the boat approached the shore. At each measurement point the boat probably coasted about 50m during the nominal 1 minute measurement. The microphone was tied to a stay (wire) on the stern (back) of the boat, approximately 3m above the water and the boat was turned so that the microphone had a clear view of the bridge. As a quick check of repeatability the measurement at 2km was repeated two hours later at a distance of 2.13 km. The results differed by 1.5 dB.

3. TRAFFIC MODELLING

A model of the bridge and its surroundings was created using Cadna/A software, which implements ISO 9613(2), and it took into account terrain, residential and built-up areas, nearby highways, road barriers and the difference between ground absorption over land and over water. The bridge itself was modeled using line sources with a sound power of 130 dBA re10^-12W, which was calculated using the RLS90 traffic noise prediction, based on typical tabulated traffic counts, traffic mix and the speed limit for the bridge traffic.

4. RESULTS

Figures 1-5 compare the measured results with the ISO 9613(2) model. The model and the measurements were compared at each octave band. The model agreed well with measurements at the lower octave bands.

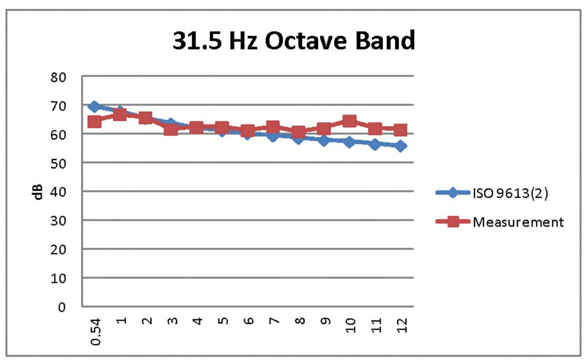


Figure 1 Comparing Model & Measurements at 31.5 Hz

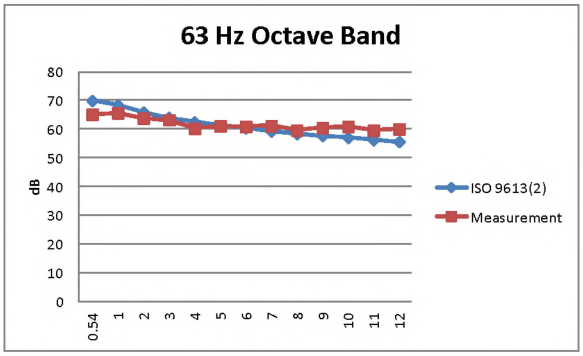


Figure 2 Comparing Model & Measurements at 63 Hz

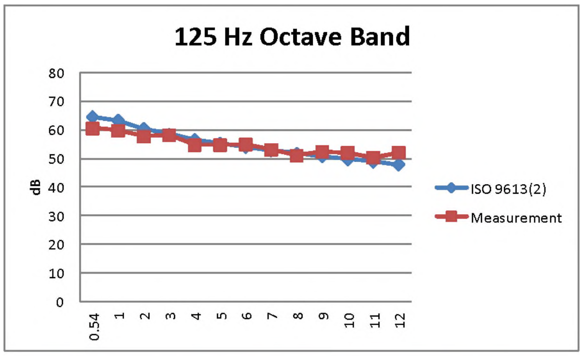


Figure 3 Comparing Model & Measurements at 125 Hz

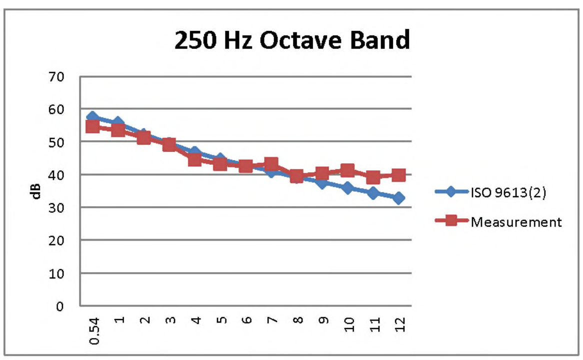


Figure 4 Comparing Model & Measurements at 250 Hz

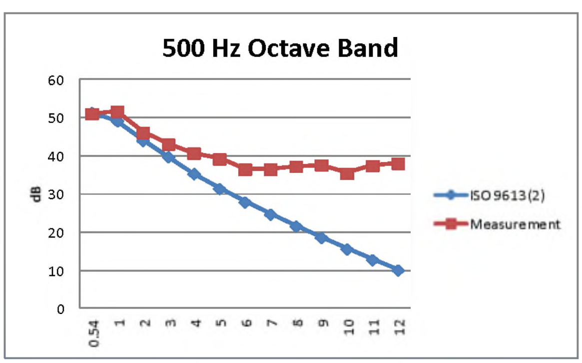


Figure 5 Comparing Model & Measurements at 500 Hz

There was good agreement at 31.5, 63, 125 and 250 Hz (Figures 1-4). The deviation was less than 4.1 dB for the 31 through 250 Hz octave bands at all measurements from 1-9 km from the bridge and less than 3.2 dB 2-8 km from the bridge. The close-in measurements were more affected by the barrier effect of the bridge deck.

At 500 Hz and above the measurements did not agree with the model at larger distances, primarily because higher frequencies drop off significantly at larger distances as air absorption becomes important, while the measured sound became increasingly dominated by the background noise of the waves hitting the hull of the boat and of wire halyards rattling inside the mast. However there was still adequate agreement out to 3-4km at 500 Hz. At 5 km air absorption at 500 Hz is already 13.5 dB and it roughly doubles for each octave above 500 Hz.

5. CONCLUSIONS

At long distances air absorption means lower frequencies dominate in any case. Even at 3 km wind turbine noise is dominated by frequencies of 500 Hz or below and at 4km by those below 250Hz. Although these are only preliminary results from a single measurement set, it appears that ISO 9613(2) may be suitable for predicting noise propagation from wind turbines over water at the distances of concern.

REFERENCES

1. ISO 9613(2) Attenuation of sound during propagation outdoors -- Part 2: General method of calculation

ACKNOWLEDGEMENTS

The Cadna modeling was done by Chloe Rourke, a Shad Valley student Intern. Initial data reduction was done by David Oliver, a student participating in Take Your Kid to Work day.

PERFORMANCE EVALUATION OF LINED DUCT BENDS –EXPERIMENT VS. THEORY

Ramani Ramakrishnan[1] and Romain Dumoulin[2]
[1]Department of Architectural Science, Ryerson University, Toronto, ON, rramakri@ryerson.ca
[2]ETS, Montreal, Quebec, dumoulin.acoustics@gmail.com

1. INTRODUCTION

Acoustical performances of simple elbow (round and rectangular) silencers, used in building HVAC systems, have been conventionally evaluated using empirical relations based on laboratory and/or field measurements[1,2]. Preliminary results of a simulation model had been presented already[3]. Experiments were conducted using a standing-wave-tube set-up at Concordia University. The experimental results from the liner set-up are applied to calibrate and validate simulation results of COMSOL Multiphysics application software[4]. The results of the validation will be presented in this paper.

2. BACKGROUND

The schematic details of a lined elbow fitting are shown in Figure 1. The liners are symmetric in Figure 1. The liner details are: liner depth is 'd'; the open air-way width is 'h'; and the liners are used for a minimum of two-duct width on either side. Ξ is the flow resistivity of the liner per unit thickness and ρc is the characteristic impedance of air.

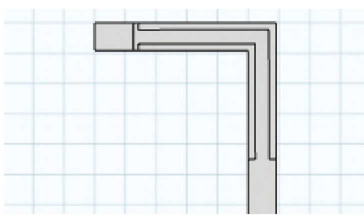


Figure 1. Details of a Lined Elbow.

The sound propagates along the centre axis from left to right. The baffle materials are bulk reacting and hence appropriate complex wave speed and material density (complex in this case) can be obtained from Bies and Hanson[5]. The mathematical modelling details were presented in Ramakrishnan and Watson[6].

3. EXPERIMENTAL SETUP

An existing standing wave tube, built in the 1980s, was modified to fit a vertical duct to simulate a lined elbow. The duct cross-section was 10" X 10". One inch duct foam liner (24" long) was used to line the end of the standing wave tube on all four sides. One four foot long 10" square tube was placed vertically over the end of the standing wave tube to simulate the lined-elbow configuration. The first 24" of the vertical pipe was also lined on all four sides with the one inch liner material.

The loud speaker at the far-end of the standing wave tube was used to generate, pink noise, white noise and band-filtered random noise. The sound pressure levels upstream and downstream of the liner section were measured to evaluate the noise reduction of the lined elbow.

4. COMSOL MODEL

The current investigation has made use of COMSOL, a powerful multiphysics numerical analysis tool and has attempted to provide results based on multi-dimensional analysis.

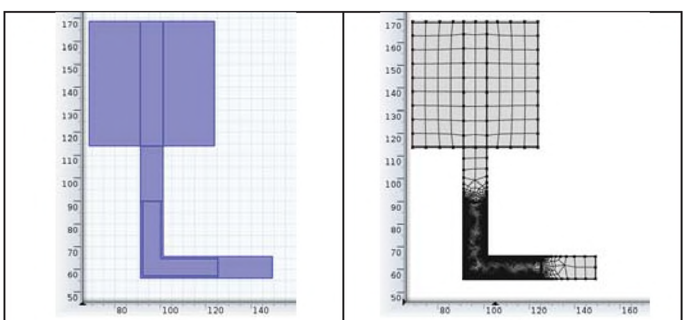
The elbow geometry can be easily modelled as 3-D. In this investigation, however, the elbow is modelled in a 2-D configuration as shown in Figure 1. The liner material is assumed to be isotropic and homogeneous fibrous material of given flow resistivity, 'Ξ'. The acoustic propagation in the liner material uses the complex propagation constant and complex density of bulk reacting material. A given acoustic field was assumed at the inlet of the elbow and the outlet is connected to a long anechoic termination. To accommodate high frequency calculation, COMSOL suggests using a length of pipe in front of the elbow within which scattered acoustic field is calculated to provide the required acoustic field at the inlet of the lined (or unlined) elbow. The application of COMSOL for simple rectangular ducts with baffles was validated in Ramakrishnan[7].

The elbow attenuation is given in Equation (1) below.

$$IL = \frac{W_{in}}{W_{out}}, dB \tag{1}$$

where, W_{in} is the sound power at elbow inlet and W_{out} is the sound power at the elbow outlet. The acoustic propagation from COMSOL model and the experimental results are presented in the next section.

A simple schematics of the lined elbow modelled in COMSOL and the FEM Mesh are shown Figure 2.



(a) (b)

Figure 2. COMSOL Model- a) Geometry; and b) FEM Mesh

The small inlet extension is to allow high frequency solutions and the large volume at the outlet is to simulate anechoic termination.

5. RESULTS AND DISCUSSION

The simulation software allows various results to be generated. The SPL (sound pressure level) variation inside the elbow at 500 Hz is shown in Figure 3 and it can be seen that the termination conditions are functioning properly.

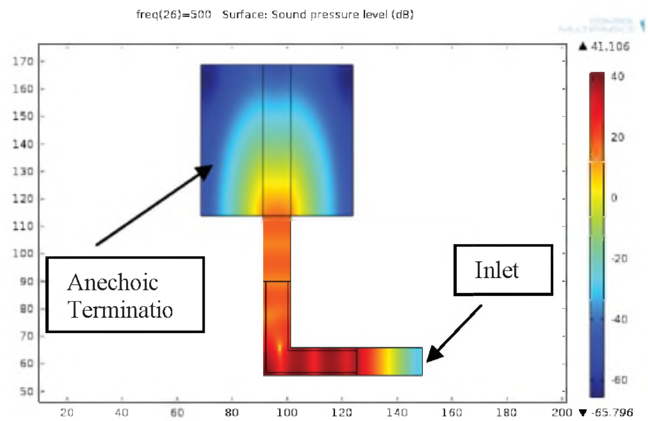


Figure 3. SPL Distribution.

The insertion loss of the lined elbow is calculated using Equation 1. The IL is calculated from 200 Hz till 5500 Hz in 5 Hz steps and the corresponding values in 1/3 octave bands are evaluated using conventional procedures.

The IL results are shown in Figure 4 below. The IL values from two simulations of the Ξ, the flow resistivity of the liner are presented in Figure 4. Figure 4 also shows the test results from the standing-wave-tube experimental setup.

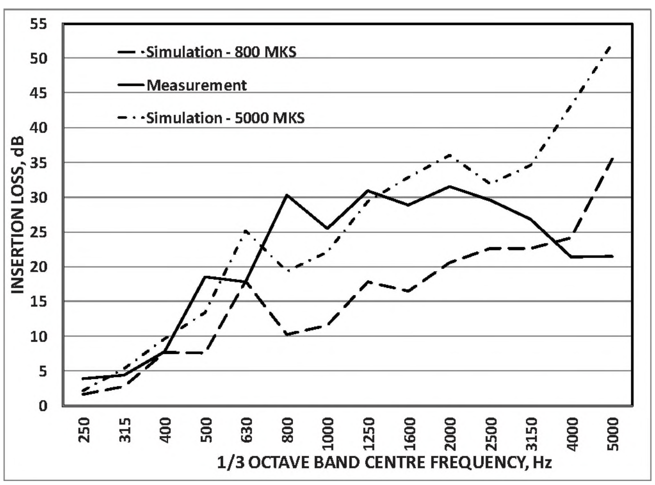


Figure 4. Attenuation of Lined Elbows.

It can be seen from the figure comparisons are reasonable in the low frequency, but differences are quite large in the high frequency range. The peak at 630 Hz band is mainly due to an anomaly in the model set-up. The 630 Hz is a duct mode of the large anechoic box and the larger the box, the anomaly will be diminished. The reasons for the large differences are highlighted below:

a) The noise floor of the test set-up both in low frequencies and in high frequencies was high. The ambient noise levels of the room were quite high. Further, the sound generator had limitations above 2500 Hz bands;

b) The true value of Ξ of the liner is unknown;

c) Only a 2-D model of the application software was applied. The duct sizes were quite small and 3-D model may provide better results; and finally,

d) The complex wave speed and density were obtained by using the elementary model of Delaney and Bazely [8]. It is seen that updated representations given by Bies and Hansen [5] for foams may provide more realistic results.

6. CONCLUSIONS

Attenuation results for duct elbow fittings were evaluated using two-D representation in commercial application software, COMSOL Multiphysiscs[3]. COMSOL model results were seen to be closer to the test data in low frequency, but diverge in the high frequency band. The simulation model needs further refining.

ACKNOWLEDGEMENTS

Jaques Payer, Josesp Hrib and Luc Demers of the BCEE Department of Concordia provided assistance to set up the lined elbow configuration and their contributions are duly acknowledged.

REFERENCES

1. Istvan Ver, "A Study to Determine the Noise Generation and Noise Attenuation of Lined and Unlined Duct Fittings," Report #5092, ASHRAE RP-265, June 1983.
2. "Sound and Vibration Control, Chapter 47" ASHRAE Handbook – HVAC Applications, 2007.
3. R. Ramakrishnan and R. Dumoulin, "Performance Evaluation of Duct Bends," Proceedings of Acoustics Week in Canada, *Canadian Acoustics*, Vol. 39, No.3 pp 144-145 (2011).
4. COMSOL Multi-Physics Application Software, Boston, MASS, Version 4.3. (2012).
5. D. A. Bies and C. H. Hansen, *Engineering Noise Control-Theory and Practice*, 3rd Edition, Spon Press, London. (2003).
6. R. Ramakrishnan and W. R. Watson, "Design Curves for Rectangular Splitter Silencers," *Applied Acoustics*, Vol 35, No.1, pp. 1-24 (1992).
7. Ramani Ramakrishnan, "Validation of COMSOL Multiphysics and Acoustical Performance of Splitter-Silencers," Proceedings of Acoustics Week in Canada, *Canadian Acoustics*, Vol. 38, No.3 pp 178-179 (2010).
8. M. E. Delany and E. N. Bazley, "Acoustical Properties of Fibrous Absorbent Materials," Applied Acoustics, Vol. 3, pp. 105-116 (1970).

NEAR FIELD SOUND OF A HRSG

Werner Richarz
Echologics Engineering Ltd. 6295 Northam Dr., Unit 1, Mississauga, ON, L4V 1W8, Canada wricharz@echologics.com

1. INTRODUCTION

The prediction of near field sound for large heat recovery steam generators (HRSG) is largely based on empirical data based on historical data. One of the reasons is that the underlying source mechanisms are rather complex. The unsteady pressure field inside the HRSG is a combination of turbulent and acoustic pressures. The former are linked to the unsteady flows and the former are sound emitted from the gas turbine exhaust as well as flow generated sound inside the HRSG. Acoustic data for gas turbine units are often available in terms of octaves, rendering more detailed studies virtually impossible as the input data must by synthesized from the sparse data. This note provides some more detail that may be used to determine the sound pressure of the acoustic near field.

2. NEAR FIELD SURVEY

The methodology is based on a series of field measurements comprising both single microphone surveys and cross-spectral density/transfer functions measurements. The locations of the measurements (field points) on which form basis of the analysis are shown in figure 1. Accelerometers were placed off-diagonal on the panels to 'catch' a large number of panel modes. The surface pressure was measure with a ½" condenser microphone held in place by a compact holder secured on a magnet. The near field microphone was held 1m from the surface.

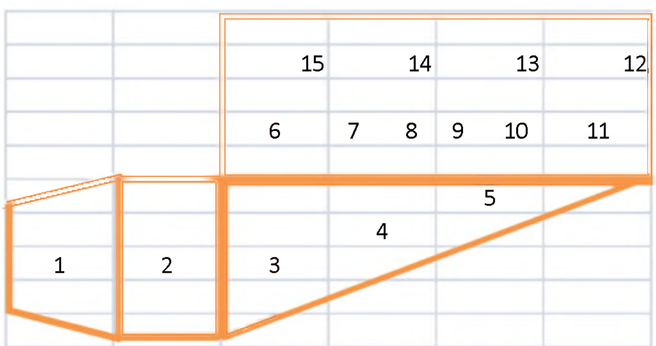


Figure 1. Measurement locations (not to scale)

2.1 Single microphone surveys

Overall levels are plotted in Figure 2. The levels diminish with distance from the gas-turbine connection. The local cross-sectional area could have been used; however, the odd-shaped elbow provides scope for 'creativity'.

From the perspective of occupational health, only A weighted levels are required, while for off-site assessments octave band level usually suffice. The octave spectra scale approximately as $\Phi(N)=A\log(N)+B$, where N in the octave band index (16 Hz→ N=1).

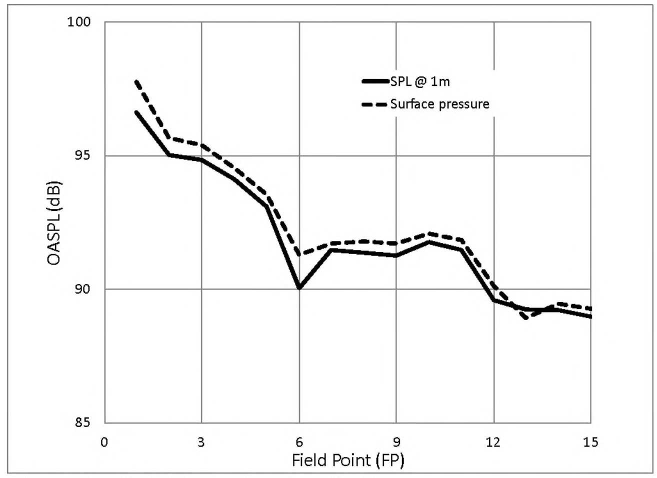


Figure 2. OASPL in the near field.

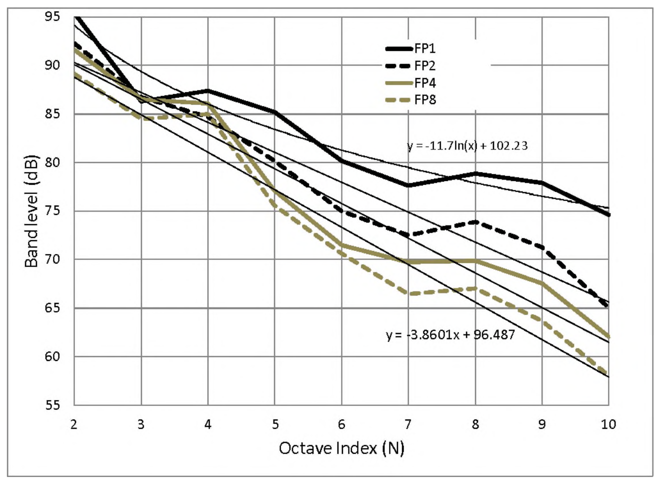


Figure 3. Typical octave band spectra

3. MULTI-TRANSDUCER SURVEYS

The exterior sound field is generated by the motion of the HRSG walls. If the walls were perfectly rigid, then the surface velocity (**u**) is zero and the local acoustic flux (**u**p) is zero irrespective of the unsteady pressure (p).

It is common practice to assume that the near field sound pressure is determined by the local normal displacement of the panels. Most practitioners have developed empirical schemes that relate the interior pressure to the panel response. Modern CFD codes can be configured to model the low frequency regime (f<30Hz), but the detailed predictions are difficult to validate.

One approach, that has yielded good results, consists of subdividing the HRSG into a series of virtual ducts of varied cross-sections. The panel excitation consists of the local sound pressure, deduced from one dimensional duct acoustics, and a correction factor for the local boundary layer. The latter is only important in the diffuser. Self-noise from the heat exchanger tube bundles can be added, but does not figure prominently in the near field.

Over the frequency range of interest (~16 Hz - 1000 Hz) it is generally assumed that near the radiation field near panels is dominated by plane waves and the radiated power by a panel of area A is |**u**|pA. This can be checked by simultaneous measuring the cross-spectral density of the panel acceleration and the near-field sound.

3.1. Acceleration-sound transfer function

The cross-spectral density of the panel vibration and the radiated sound is given by $\Phi_{ap}(f)$ and is a complex valued function. As only sound levels are of interest here, the magnitude suffices (note: $\Phi_{ap}(f) = |\Phi_{ap}|(f)e^{i\Theta(f)}$). The transfer function is obtained by dividing by the 'source spectrum'. The latter is always real, so there are no further manipulations. The transfer functions are obtained by dividing the cross –spectral density with the power spectral density of the acceleration. The curves are similar in shape and amplitude, and their ensemble average is shown in figure 4.

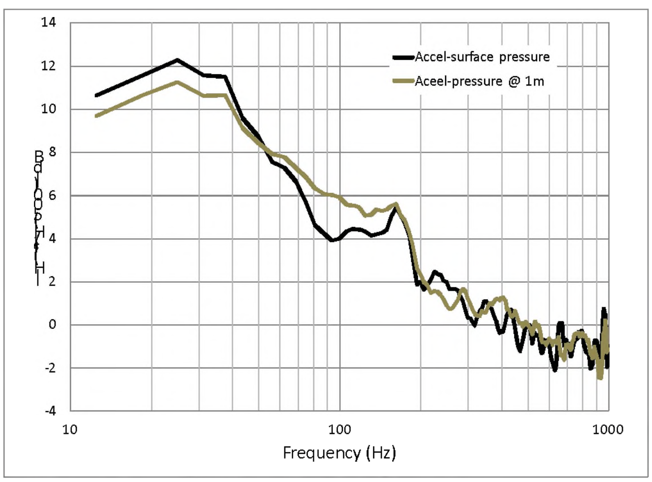


Figure 4. Normalized transfer functions

Even though the near field pressure ought to be dominated by the local panel motion, there are other, uncorrelated contributions from the entire side of the HSRG. In principle it is possible to identify the contribution of multiple sources [1]. However, this requires a considerable amount of instrumentation and even with modern data acquisition systems is time-consuming.

3.2. Coherence function for acceleration and sound

The coherence function provides some of the indication of assumed linear input/output model. The coherence function is $C_{ap}(f)= |\Phi_{ap}(f)|^2 \{|\Phi_{aa}(f)||\Phi_{pp}(f)|\}^{-1}$. For a noise free linear process $C_{ap}(f)= 1$; otherwise $0<C_{ap}(f)<1$. Some typical coherence functions are shown in figure 6.

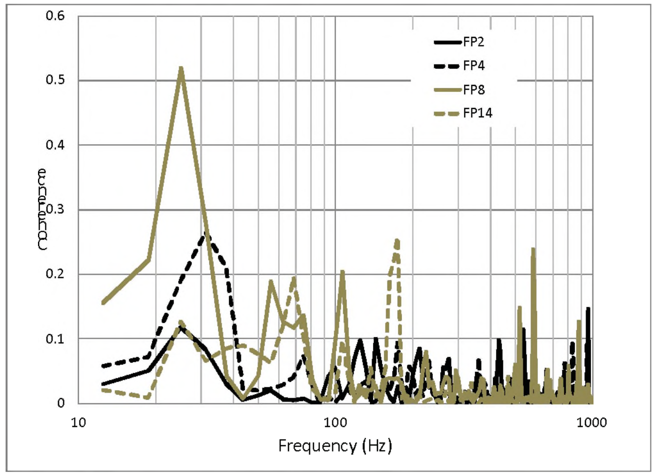


Figure 5.

Significant coherence is only evident at low frequencies. The running integrals in Figure 6 provide a different perspective. Frequency bands with low coherence contribute only a little to the integral. Substantial contributions are found in regions

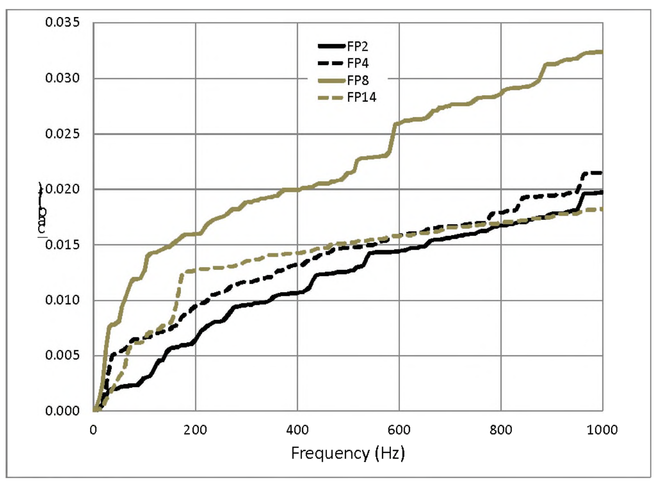


Figure 6.

4. CONCLUDING REMARKS

Coherence measurements of the panel acceleration and near field pressures do not support the assumption of local source dominance.

REFERENCES

Bendat, J.S., Piersol, A.G. Random Data, 2nd ed. Wiley Interscience, 1986

NOSE CONE DEVICE AND BUILT-IN CALIBRATION CHECK AS ESSENTIAL FEATURES OF A STAND-ALONE INSTRUMENT FOR UNATTENDED MID- AND LONG-TERM NOISE MEASUREMENTS

Daniel Vaucher de la Croix[1] and Erik Aflalo[1]
[1]ACOEM / 01dB, 200 chemin des Ormeaux, 69 576, Limonest, France, daniel.vaucherdelacroix@acoemgroup.com

1. INTRODUCTION

Environmental noise generated by ground transportation, construction sites or recreational activities, is coming from all directions, and such multiple sources are usually located at random positions with respect to the measurement point. With the instrument placed vertically (while sources mainly coming from a reference direction of 90° from its axis), the goal is to meet the requirements of the IEC 61672-1 standard [1] on sound level meters taking into account noise incidence from the horizontal direction. Emphasis will be given on the main technical difficulty and the adequate solution meeting the standard requirement.

In addition, mid- to long-term remote operation of the device imposes a constant control of the measurement quality. At first an acoustic calibration of the instrument is made during installation on site. Then during unattended operation of the instrument, the integrity of the complete measurement chain (including the microphone) is automatically tested, relying on a built-in electrical check. The test consists in injecting a sinusoidal charge into the microphone membrane, at selected frequencies. The (partly user-defined) multiple-frequency check allows assessing of a possible degradation of the microphone as well as the electronic components.

2. 0° and 90° incidence directions

In unattended noise measurement situations, the direction from the source is generally unknown. Apart from aircraft noise, the sources are located on the ground. The optimal position of the sound level meter for unattended noise measurements is thus vertical, with an instrument design so as to fulfill the IEC 61672-1 standard for ground sources (90° incidence). Our goal was to achieve a sound level meter design able to fulfill both 0° (aircraft noise [2] and measurement pointing at the source) and 90° reference directions (ground source noise measurements).

Additionally the risk of wrong measurements in environmental noise assessment must be reduced as low as possible. A built-in electrical check (multi-frequencies charge injection) allows for testing the whole measurement chain, including the microphone and offers the advantage of a better assessment of a possible degradation of the microphone membrane or of the electronic components. This will be discussed on the second part of this paper.

2.1 Design constraints

IEC 61672-1 gives directional response requirements for the configuration of a sound level meter as stated in the instruction manual for the normal mode of operation or for those components of a sound level meter that are intended to be located in a sound field. The specifications apply for plane progressive sound waves at any sound-incidence angle within the indicated range (focus on +/- 30°), including the reference direction. At any frequency, the design goal is to get equal response to sounds from all directions of sound incidence (§ 5.3 of [1]).

High frequency response depends on the diameter of the diaphragm: they are improved when the diameter of the diaphragm decreases. Nevertheless the positive effect is altered by a higher background noise. As a consequence 1/2' microphones are usually selected, as being the best compromise between costs, frequency response and background noise for environmental noise assessment.

2.2 Solution: Acoustic Cone

Various shapes of the mechanical design of DUO at the early stage, and several distances between the microphone and the body were made and tested using 3D prints:

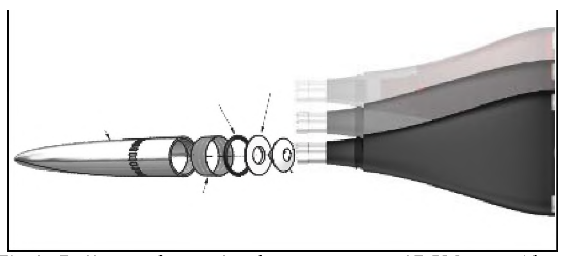


Fig.1: Different shapes for the upper part of DUO sound level meter body tested for optimum frequency/directional response

Fig. 2: Electro acoustic tests performed on 3D prints to validate the mechanical design of the body

2.3 Directional response results

θ = 30° directional response is displayed below, as an example of requirement to achieve. The benefit of the cone corresponds to the difference between the green and the blue curves.

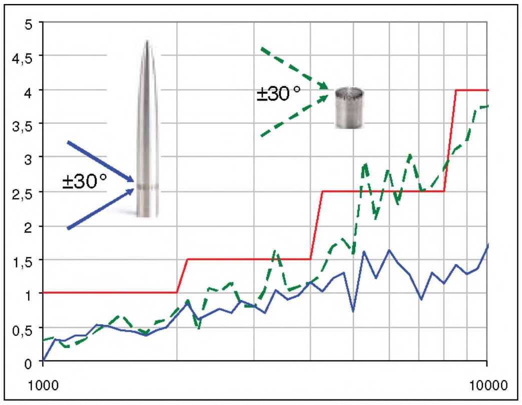


Fig. 3: In blue, directional response of DUO [3] with acoustic cone for θ = 30° (reference direction 90°); in red, tolerance values of IEC 61672-1 without expanded uncertainty

3. BUILT-IN ELECTRICAL TESTS

3.1 Multi-frequencies charge injection principle

The test consists in injecting a stable reference signal though a reference capacitor that "simulates" an acoustic signal at the microphone output (see the dash lines in Fig. 4 below):

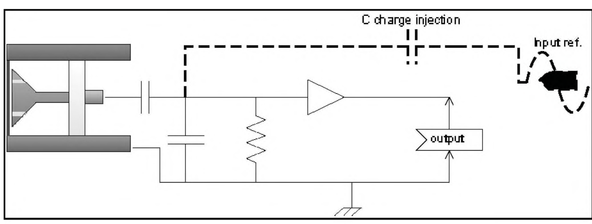


Fig 4: charge injection principle

The reference signal is a frequency user defined sine at a selectable level between 0 and 5 V peak. The capacity for the reference C charge injection is typically around 0.2 pF.

The test consists in creating reference values on a valid system and measure through the sound level meter measurement chain itself the difference in dB with the current situation. The value of the deviation will be representative of a variation of the system.

Charge injection will behave as an impedance comparison between the condenser microphone and the charge injection reference capacitor. If the impedance of the microphone is changed (typically a mechanical damage of the active part of the membrane that will change its capacity) the charge injection method will detect it.

3.2 Test conditions

This test consists in injecting a stable reference Tests were performed in an anechoic room at LNE (French Laboratoire National d'Essais), with DUO used in vertical position.

3.3 Checking description

For each test, a microphone is damaged on purpose and the following measurements are performed before/after damage:

- A-weighted background noise level
- Sensitivity @ 1 kHz
- Multi-frequencies charge injection response at 250 Hz, 500 Hz, 1 kHz, 2 kHz and 4 kHz

Relying on this procedure we could successfully detect the following list of defaults:

- Punching of the microphone membrane
- Water drop on the microphone membrane
- Light & heavy dust on the microphone membrane
- Shock on the edge of the microphone
- Small & large cuts on the edge of the microphone
- Bad contact at the inner pin of the microphone

4. DISCUSSION AND CONCLUSIONS

The requirement for DUO [3] as a sound level meter used for unattended measurements made mandatory the possibility to setup 0° and 90° reference directions for capturing all types of sources. A key to success has been the close cooperation between the 01dB and G.R.A.S. R&D teams. Taking into account the acoustic front end design at the early stage of the development allowed determining and optimizing the constraints: position of the microphone, shape of the body of the sound level meter, cone and wind screens characteristics.

In addition, multi-frequencies CIC is identified as a very useful tool for ensuring reliable measurements between two calibrations on unattended measurement systems. When performed periodically automatically and remotely (one to four times a day) this test will allow controlling the stability of the noise measurement equipment and securing the validity of the measurement between two checks.

DUO [3] is the first and single sound level meter approved in France [4], in Germany [5] and in Switzerland [6] with both reference directions 0° and 90°.

REFERENCES

[1] IEC 61672-1 2002-05 "Electroacoustics-sound level meters – part 1: Specifications"

[2] ISO 20906:2009 Acoustics "Unattended monitoring of aircraft sound in the vicinity of airports"

[3] 01dB, a brand of ACOEM: "DUO Product data sheet" – December 2011

[4] Certificat d'examen de type LNE-21674 (21 juillet 2011) delivered by LNE (Laboratoire National d'Essais), France

[5] PTB-1.63-4052726 delivered the 6th February 2012 by PTB (Physikalisch-Technische Bundesanstalt), Germany

[6] Zulassungszertifikat CH-A3-12096-00 (20th February 2012) by Bundesamt für Metrologie METAS, Switzerland.

[7] NF S31-010 December. 1996 « Acoustique -- Caractérisation et mesurage des bruits de l'environnement - Méthodes particulières de mesurage »

[8] IEC 61672-3 2006-10 "Electroacoustics-sound level meters – part 3: Periodic Tests"

DOWNTOWN MONTREAL NOISE CONTROL – CHALLENGES ON THE RISE WITH MIXITY AND URBAN DENSITY

Frederic Bouchard[1], Daniel Vaucher de la Croix[2], and Charles Gagné[3]
[1]Dept. of Noise Control, Ville Marie Burough of Montreal, 800 Maisonneuve blvd 17th Floor, Montréal, Québec, Canada, H2L 4L8, frederic.bouchard@ville.montreal.qc.ca
[2]ACOEM / 01dB, 200 chemin des Ormeaux, 69 576, Limonest,,France, daniel.vaucherdelacroix@acoemgroup.com
[3]Alliance Predictive, 1100 Thomas-Tremblay, Sherbrooke, J1G 5G5, Canada, charles@alliancepredictive.com

1. INTRODUCTION

Cities consist of a large urban mix where different human activities coexist. Housing can be found next to stores and transports, schools and hospitals, industrial facilities and service firms, leisure and tourism activities. During the day and at night, part of the city is sleeping while the other works or relaxes. Keeping harmonious local noise situations is beneficial to all inhabitants. Today, the accumulation of different ways of life leads local authorities to elaborate new policies for the management of noise pollution. Also, a new concept relative to the measurement of sound levels in several synchronized points has come out. Adapting metrology to this concept leads to innovative quantifying of urban noise events, for a better knowledge of the noise situation. For a very concrete illustration of the various operating steps addressed in this approach, this paper presents typical situations in the Ville Marie borough of the city of Montreal (all dB levels expressed are in Leq fast, i.e: RMS averages).

2. INNER CITY NOISECONTROL

2.1 Caseload overview

Bedrooms only 40m away from massive air conditioning units blaring 90 dBA each, resulting in 65 dBA exposure versus an authorized 40 dBA limit; sub bass from an outdoor party traversing brick and stone walls into a condo 200m away; business owners buying overly powerful and noisy ventilation systems from sellers unaware of noise limits; another bedroom 50m away from a 100 000 Watt sound system inside an after-hours nightclub; an entire touristic neighbourhood haunted by a fluctuating whistle tone coming from a dust filled muffler half a kilometre away, etc. All these problems are solvable once the noisy party has been shown clear evidence they're above municipally authorized limits.

2.1 Helped by technology

With the evolution of integrating sound level meters and spectrum analyzers we can pinpoint which frequency range needs to be reduced by how many decibels instead of simply divulging global levels. Firms like Acoustilog in New York City have inspired us to equip ourselves with high quality calibrated wave recorders that we can use for proof in court.

All-in-one devices have started to appear, such as DUO smart noise monitoring which has the calibrated recording feature, as well as a GPS and modem for remote control. One of the big differences to other systems is that 01dB designed DUO for unattended noise measurement outdoor, including a number of very unique features like waterproof microphone, automated system check (based on the CIC principle) and full compliance with the IEC 61672-1 standard for class 1 sound level meter, on both 0° and 90° of source incidence [1].

3. EXAMPLES OF PRACTICE

3.1 More fear than harm

A plaintiff was certain that his upstairs neighbours were using a rotating press machine all night because a bass tone was audible and structurally transmitted in the wall. Finally it was a dryer on four legs with a tool case over it that was backed to the wall, transmitting the frequency of 63 Hz all the way down to the plaintiff's apartment.

Another such case involved a tabletop ventilator placed on a wobbly wood table about 1' by 2' with two foldable U-shaped legs. Two tones at 31.5 and 100 Hz were transmitted into the apartment downstairs and although the plaintiffs urged the Ville Marie noise control department to call the police and emit warnings by bailiff, a technician went upstairs and sat the shaky table on a 1 inch thick Styrofoam baseplate and the problem was instantly solved. The reduction was 15 dBZ at 31.5 Hz and 23 dBZ at 100 Hz.

3.2 Discotheque soundproof diagnostics

The borough was asked by the police department to confirm soundproofing of a new 3000 person capacity nightclub before the liquor board could deliver its alcohol permit. It is located 200 m away from existing and proposed condo buildings. Before the club completed soundproofing we measured 15 dBZ emergences at 63 Hz (i.e. kick drum or boom frequency), even after attenuating the 18 cardoïd subwoofers which had rendered the sound quality mediocre.

Even though there was only 1 dBA emergence at 180 m away, the musical thump was still audible and we measured an 8 dBC emergence (74.7 dBC with music and 66.8 without). Every time a bus or train passed the level at 50Hz shot up to 86 dBZ from 56 dBZ so care was taken to not integrate these.

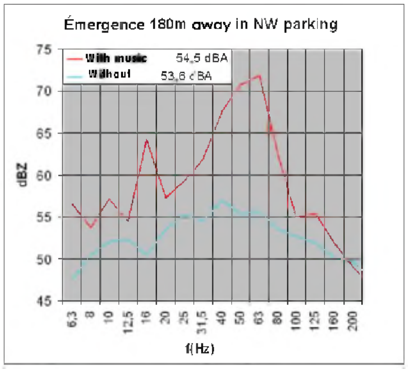


Figure 1. Bass emergence levels 180m away from club (95 dBA global and 106 dBZ at 63 Hz inside)

We monitored inner and outer noise levels with two DUO systems on six channels: dBA, lin and third octaves 40, 50, 63 and 80 Hz, while repeating the same house music loop. Then further soundproofing was done and emergence levels dropped to 0 dBZ at 60 Hz with a remaining 10 dBZ at 50 Hz that was inaudible. No variations in dBA were observed with music on or off, and only 3 dBC varied:

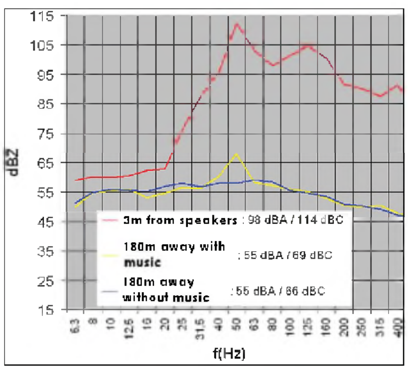


Figure 2. Emergence levels after final soundproofing

The club's sound system has been set to Montreal's 98 dBA noise limit at minimum 3m from speakers or musical instruments. No noise complaints have ever been registered towards the nightclub after several months of operation.

3.3 Outdoor music festival monitoring

We have put the DUO technology to good use by placing several instruments in strategic positions between dwellings and festival sites involving amplified music and pyrotechnics. Standalone remote controlled waterproof units with modems, GPS, manual and triggered wave recording, direct audio and SMS warnings thus greatly helped efficiently collecting & processing all relevant information.

The department received a complaint about bass tones from an outdoor dance party that were infiltrating a condo unit. The meters used by the sound and lighting company had no spectrum analyzers and from what their sound technicians and engineers could measure, levels were within the 80 dBA / 100 dBC authorized limits at 35m from speakers. However when a noise control technician did a spectrum analysis between the speakers and the plaintiff's residence west of them, he noticed as much energy at 50Hz (100 dBZ) as directly in front of the speakers, although the global level was 9 dBA lower (79 vs 88). He reported this to the sound engineer who realized that the vertical subwoofers had been laid to the side, thus sending off the bass tones laterally.

During a summer festival, one of the neighbouring residents complained about bass pulses reaching his bedroom, where the energy level in the 63 Hz octave was 5 dBZ over NR curve 55 that covered the rest of the spectrum. The Ville Marie noise control department had that reduced by the sound engineer without any audible difference at the site of the party. The department also saw much variation in the hi-mid frequencies, due to insufficient compression or limiting of a percussive instrument. The sound engineer quickly remedied this problem as well. The department's actions increased the quality of the sound instead of simply lowering the master volume and frustrating the organisers and clients. Everyone was happy, especially the plaintiff.

4. DISCUSSION AND CONCLUSIONS

Although the limits are expressed in dBA, pure tones between 22 and 11 314 Hz warrant a 5 dBA penalty as do impulsivity and speech transmission. In some cases emergences are calculated with L95% levels in order to eliminate parasite noises like traffic and animal noises (especially dogs and birds). And even though technology helps (definitely DUO smart noise monitors bring much more flexibility than any other classical instruments previously used in similar situations), a lot of what the Ville Marie noise control department does involves mediation and dealing with human tolerance, perception and expectations.

REFERENCES & ACKNOWLEDGEMENTS

[1] IEC 61672-1 2002-05 "Electroacoustics-sound level meters – part 1: Specifications"

[2] 01dB, a brand of ACOEM: "DUO Brochure - COM1002 Rev B" – May 2012

[3] Certificat d'examen de type LNE-21674 (21 juillet 2011) delivered by LNE (Laboratoire National d'Essais), France

[4] PTB-1.63-4052726 delivered the 6th February 2012 by PTB (Physikalisch-Technische Bundesanstalt), Germany

[5] Zulassungszertifikat CH-A3-12096-00 (20th February 2012) by Bundesamt für Metrologie METAS, Switzerland.

[6] Ordonnance nr.2 – Règlement 4996 sur le Bruit « Ordonnance sur le Bruit dans les lieux habités », 8 Juin 1977, Montréal

DUO is the first and single sound level meter approved in France [3], in Germany [4] and in Switzerland [5] with both reference directions 0° and 90°.

VARIATION OF ASTC RATINGS DUE TO FLANKING TRANSMISSION WITHIN A RESIDENTIAL CONVERSION

Clair W. Wakefield[1] and Andrew. P. Williamson[1]
[1]Wakefield Acoustics Ltd., 301 - 2250 Oak Bay Avenue, Victoria, B.C., Canada, V8R 1G5

1. INTRODUCTION

Swift House is an older building in Victoria's downtown core that originally served warehouse and/or commercial purposes but many years ago was converted to a homeless shelter having many small rooms. In 2011, the building was transformed again, this time into larger, self-contained affordable housing units. The units were also brought up to BC Building Code 2006 (BCBC) requirements including provision of adequate airborne sound insulation.

Swift House is a two-storey building with poured concrete slab floors supported by cast-in-place concrete columns (of various shapes – often octagonal) spaced quite uniformly throughout the building. Under the new floor plan, unit demising (party) walls frequently terminate at these concrete columns. In some situations, however, demising walls terminate at lightweight steel stud and gypsum board corridor walls, some of which extend continuously from one living unit to the next. In other situations, one end of a demising wall will terminate at a concrete column while the other terminates at a gypsum board wall. Wakefield Acoustics Ltd. conducted a series of Apparent Sound Transmission Class (ASTC) tests between units and found the variability of measured ASTC's to be much wider than typically observed in new condominium buildings. This paper describes the ASTC tests conducted and discusses the suspected reason for the high variability in results.

2. TYPICAL FLANKING TRANSMISSION EFFECTS

The BCBC recognizes that residential unit demising walls rarely provide the same degree of airborne sound insulation in field (real world) situations as they do in laboratory test situations. Reduced ASTC ratings in field tests are sometimes due to sound leakage via cracks and/or gaps, or to other flaws in wall construction. More commonly, ASTC's are limited by structure-borne sound which "flanks" around the demising wall via continuous sidewalls and/or floor and ceiling assemblies. The BCBC therefore recommends that architects and builders chose wall assemblies with a lab-rated STC at least 5 points higher than the Code's minimum airborne sound insulation requirement of ASTC 50. Party walls having lab ratings of about STC 55, or slightly more, are then commonly selected. In multi-family residential buildings, it is not generally cost-effective to construct wall assemblies with a lab ratings greater than STC 60, since flanking will usually prevent such superior performance from being realized in the field.

3. ASTC TESTS OF TWO NOMINALLY IDENTICAL DEMISING WALLS

ASTC tests were conducted between two pairs of units having identical demising walls construction as follows:

- 2 layers of 16 mm Type X Gypsum Board,
- 152 mm, 25 ga. steel studs,
- 150 mm batt insulation,
- 2 layers of 16 mm Type X Gypsum Board,
- 152 mm 25 ga. steel studs,
- 1 layer of 16 mm Type X Gypsum Board.

Note that this demising wall includes a second 152 mm, uninsulated cavity which serves as a pipe chase for the half of the wall that separates back-to-back kitchens. Without this second cavity, this wall would be identical to BCBC Wall No. S9a, which received a lab-rated STC 59. With the second cavity present, the wall's lab-rated STC would be somewhat greater – perhaps STC 62 to 65.

The first test of this wall assembly (between Units 208 and 210) yielded ASTC 49. Figure 1 is a partial floor plan including the test wall. It is seen that one end of the demising wall terminates in a rectangular concrete column while the other terminates in a gypsum board corridor wall which was believed to run continuously (including the gypsum board sheeting) from one unit to the next.

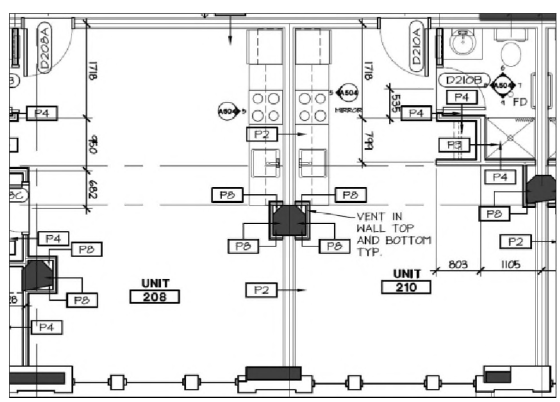


Figure 1; Floor Plan of Units 208 & 210 Showing Test Wall

Since the wall between Units 208 and 210 did not meet the BCBC's minimum ASTC 50 requirement, a second test was conducted of an identical demising wall between Units 213 and 215. Figure 2 is a plan of Units 213 and 215 showing the second test wall. Note, in this case, both ends of the demising wall terminate at octagonal concrete columns. This test yielded ASTC 59 - one of the highest field test results that Wakefield Acoustics Ltd. has ever measured. Initially it was not clear how such as large (10 STC point) variation could occur between identical wall assemblies.

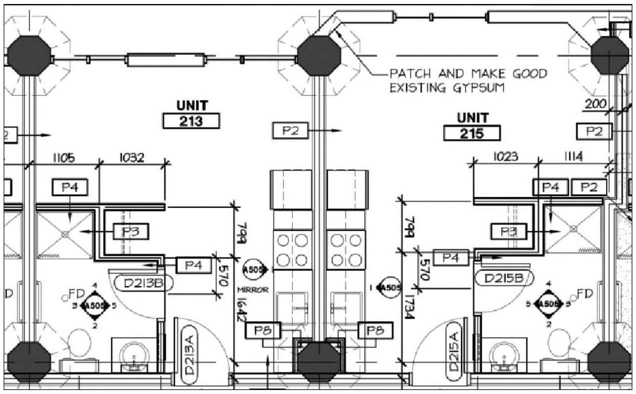


Figure 2; Floor Plan Units 213 and 215 Showing Test Wall

Figure 3 shows the Apparent Transmission Losses (ATL's) in 1/3rd octave bands from 125 to 4,000 Hertz as required in the standard ASTC test. The ATLs' obtained between Units 213 and 215 are consistently greater (by 4 to 20 dB) than those obtained between Units 208 and 210.

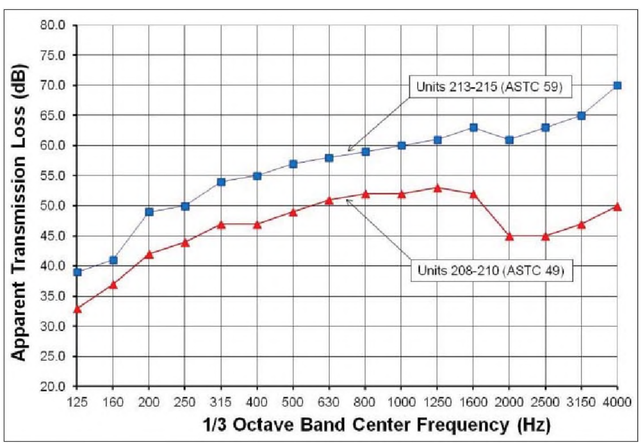


Figure 3; ATL's for Two Identical Demising Walls.

The low ATL's obtained between Units 208 and 210 were not due to sound leakage via gaps/cracks as such leakage tends to only degrade wall performance at high-frequencies.

4. ASTC TEST OF A WALL WITH NO CONCRETE COLUMN TERMINATIONS

A third ATSC test was conducted between the Bedroom of Unit 212 and the Living Room of Unit 214. Since no kitchen was involved, this demising did not include a second cavity and was constructed as follows:

- 2 layers of 16 mm Type X Gypsum Board,
- 152 mm, 25 ga. steel studs,
- 150 mm batt insulation,
- 2 layers of 16 mm Type X Gypsum Board.

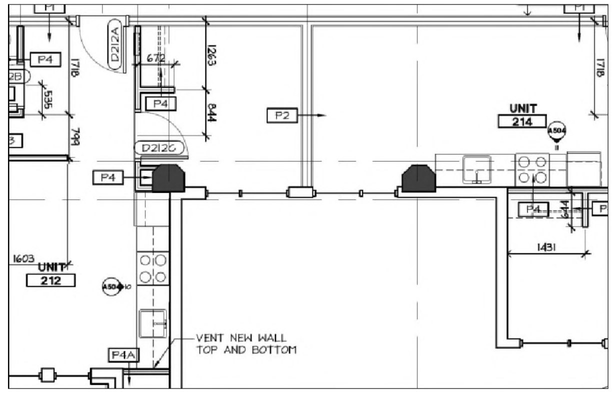


Figure 4 – Floor Plan Units 212 and 214 Showing Test Wall

This assembly is then BCBC Wall No. S9a and has a lab-rated STC of 59. The field test yielded ASTC 46. Figure 5 compares the ATL's obtained from the field test with the TL's predicted for this wall assembly using Marshall Day Acoustic's sound insulation prediction software, INSUL (Version 6.3). INSUL predicted STC 55. Again, there is no evidence of sound leakage as the difference between measured and predicted TL's diminishes in the highest frequency bands.

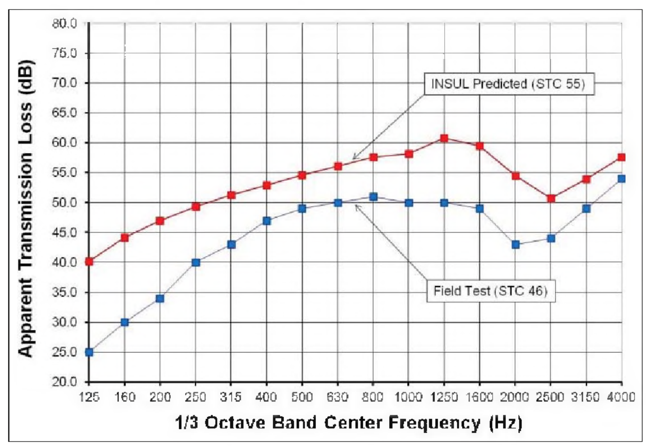


Figure 5 –Measured ATL's versus INSUL-Modeled TL's

5. CONCLUSIONS

The very substantial differences observed between the ASTC's measured between two nominally identical demising walls, and those between the measured ASTC and modeled STC of another wall assembly, are largely due to variations in structure-borne sound which, in the field, "flanks" around the primary noise barrier, principally via paths provided by continuous sidewalls. The floors and ceilings of the living units tested were reinforced concrete slabs. The rigidity of these slabs minimized their excitation by the intense sound fields created in the "source rooms" and therefore minimized the transmission of structure-borne sound into adjacent units. Flanking transmission was strongest (and therefore ASTC's lowest) where one or both ends of the demising wall was coupled directly to a gypsum board sidewall, particularly where the sidewall continued on, without a structural break, into the adjacent unit. ASTC's were maximized, and approached lab test values, where both ends of the demising wall terminated in one of

NONLINEAR GEOACOUSTIC INVERSION VIA PARALLEL TEMPERING

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

Charles W. Holland
Applied Research Laboratory, State College, Pennsylvania 16804-0030 USA, cwh10@psu.edu

1. INTRODUCTION

Knowledge of seabed geoacoustic and scattering properties is important for sonar, geophysical, and geotechnical applications in shallow-water environments. Since direct measurements can be time consuming and expensive, inferring *in-situ* information about seabed model parameters from the inversion of ocean acoustic data has received a great deal of attention. A Bayesian approach to geoacoustic inversion [1,2] provides quantitative uncertainty analysis and has been applied for a variety of acoustic data types, including the inversion of full-field, reflection, dispersion, and reverberation measurements.

In Bayesian inversion, unknown model parameters are considered random variables constrained by measured data and prior information, with the goal of estimating integral properties of the multi-dimensional posterior probability density (PPD), such as marginal probability distributions. For nonlinear problems, such as geoacoustic inversion, numerical methods must be applied to estimate these integrals. In particular, the Markov-chain Monte Carlo method of Metropolis-Hastings sampling (MHS) has been applied in virtually all Bayesian geoacoustic inversions to date [1]. However, MHS can be inefficient for strongly nonlinear inverse problems involving PPDs with multiple modes (i.e., multiple isolated regions of high probability in the parameter space). For such problems MHS has the potential to miss important regions of the parameter space and to significantly under-estimate parameter uncertainties. Multi-modal PPDs have been observed for all of the geoacoustic data types mentioned above.

This paper applies the method of parallel tempering [3,4] to achieve efficient and effective sampling of a particularly challenging multi-modal problem involving the inversion of acoustic reverberation data for geoacoustic and scattering parameters. Parallel tempering has the ability to transition freely between multiple PPD modes by running parallel Markov chains at a series of increasing sampling temperatures T, with probabilistic interchanges between chains. High-T chains provide wider sampling of the parameter space and the possibility of bridging isolated modes, while low-T chains provide more precise local sampling but are prone to become trapped in localized regions of the space. Parallel tempering improves sampling by providing interchange between chains at different temperatures. Including higher-T chains ensures that the lower-T chains can access all regions of the space while still providing efficient local sampling, resulting in a robust ensemble sampler.

2. EXAMPLE

This section compares MHS and parallel tempering for Bayesian geoacoustic inversion of simulated (noisy) reverberation data. Simulation provides a number of advantages for such comparisons, in that an appropriate model parameterization is known and the error statistics are also known and controlled. Hence, characteristics of the inversion, such as PPD multi-modality, arise solely from the physics of the forward problem, and are not an artifact of a poor choice or parameterization or unaccounted-for sources of error.

The (range-independent) seabed model assumed for the reverberation inversion problem is illustrated Fig. 1. The seabed is represented by an upper sediment layer of thickness h=5 m, sound velocity v_1=1470 m/s, density ρ_1=1.4 g/cm^3, and attenuation α_1=0.5 dB/wavelength, overlying a semi-infinite basement with corresponding parameters v_2=1660 m/s, ρ_2=1.8 g/cm^3, and α_2=0.1 dB/wavelength. Acoustic backscatter occurs at rough interfaces at the top and bottom of the sediment layer and at heterogeneities within the volume of the sediments. The spatial roughness of the upper and lower interfaces is assumed to be isotropic and characterized by a two-dimensional power-law spectrum $R_i(k) = w_i k^{-\gamma_i}$, where k is the magnitude of the horizontal wave vector, w_i is the spectral strength, and γ_i is the spectral exponent, with i=1,2 corresponding to the upper and lower interfaces, respectively (values used here are $w_1 = w_2$= 0.02 and $\gamma_1 = \gamma_2$ = 3). The volume-scattering intensity cross-section for the sediments is given by $S_V = 10^{-6}$ m^{-3}. Finally, the standard deviation of the data errors, σ =1 dB, is also considered an unknown parameter in the inversion. The reverberation data are shown in Fig. 2.

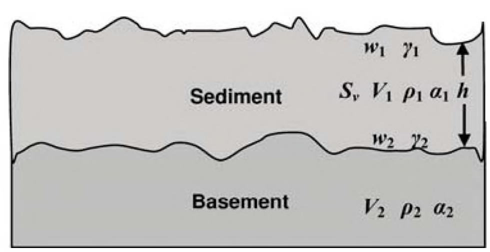


Figure 1. Schematic diagram of the two-layer seabed model indicating unknown parameters. The error standard deviation is also included as an unknown in the inversion.

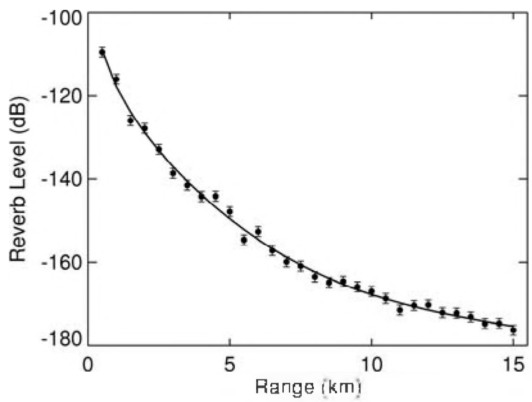


Figure 2. Simulated noisy reverberation data (circles) and modelled data (solid line) computed for the highest-probability model. Error bars indicate the maximum-likelihood standard-deviation Estimate evaluated at the most-probable model.

Both MHS and parallel tempering were applied to the reverberation inversion problem outlined above. The parallel tempering approach made use of a total 30 parallel chains, with 16 chains at $T = 1$, 8 chains at $T = 2$, 4 chains at $T = 4$, and 2 chains at $T = 8$. Results are based on the T=1 and 2 chains, with the samples collected at $T = 2$ re-weighted so as to remove the sampling bias which otherwise occurs for sampling at non-unity temperature [5].

All 13 model parameters described above were included in the two inversion approaches. However, given space constraints and to highlight sampling of multi-modal PPD structure, results are considered here only in terms of joint marginal probability distributions for sediment thickness and sound velocity which is highly multi-modal, as shown in Fig. 3. This figure compares inversion results for MHS and parallel tempering after various numbers of samples, as indicated on each panel (for parallel tempering the total number of samples is indicated, including samples at higher temperatures which were not included in the marginal distribution estimate).

Considering first the MHS results (left column of Fig. 3), over the first 10^5 samples (top panel), the method has sampled only a single mode of the PPD, and if sampling was terminated here the multi-modality would go undetected and parameter uncertainties would be substantially under-estimated. A second PPD mode is detected by $2\text{x}10^5$ samples, and third mode by $5\text{x}10^5$ samples. The marginal distribution does not change significantly in going from 10^5 to 10^6 samples, but by $2\text{x}10^6$ samples additional modes are detected. From the behavior in Fig. 3, it appears unlikely that MHS has visited all modes even with $2\text{x}10^6$ samples, and it is clear the sampling is not even close to convergence (typically requiring many transitions between all modes).

Considering the parallel-tempering results (right column of Fig. 3), it is clear that the multi-modality of the joint marginals for h and v_1 are mapped out far better with 10^5 total samples using parallel tempering (top panel on right) than with $2\text{x}10^6$ samples using MHS (bottom panel on left). In particular, the parallel-tempering marginal distribution includes multi-modal structure which is not apparent with MHS for sample sizes up to 20 times larger. Further, the fact that there is little practical difference in results for different parallel-tempering sample sizes in Fig. 3 indicates convergence by 10^5 samples, and hence large sample sizes are not required to map the complicated multi-modal PPD structure using this approach.

REFERENCES

[1] S.E. Dosso and J. Dettmer, 2011. ``Bayesian matched-field geoacoustic inversion, Inverse Problems, **27**, 1-23.

(2011). [2] S.E. Dosso, C. W. Holland, and M. Sambridge, 2012. Parallel tempering for strongly nonlinear geoacoustic inversion, In Press: J. Acoust. Soc. Am.

[3] C. J. Geyer, Markov chain Monte Carlo maximum likelihood, In: *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface* (New York, 1991), pp. 156-163.

[4] D. J. Earl and M. W. Deem, 2005. Parallel tempering: Theory, applications, and new perspectives, Phys. Chem. Chem.Phys.,**7**, 3910-3916.

[5] B. F. Brooks and L. N. Frazer, 2005. Importance reweighting reduces dependence on temperature in Gibbs samplers: An application to the inverse coseismic geodetic problem, Geophys. J. Int., **161,** 12-20.

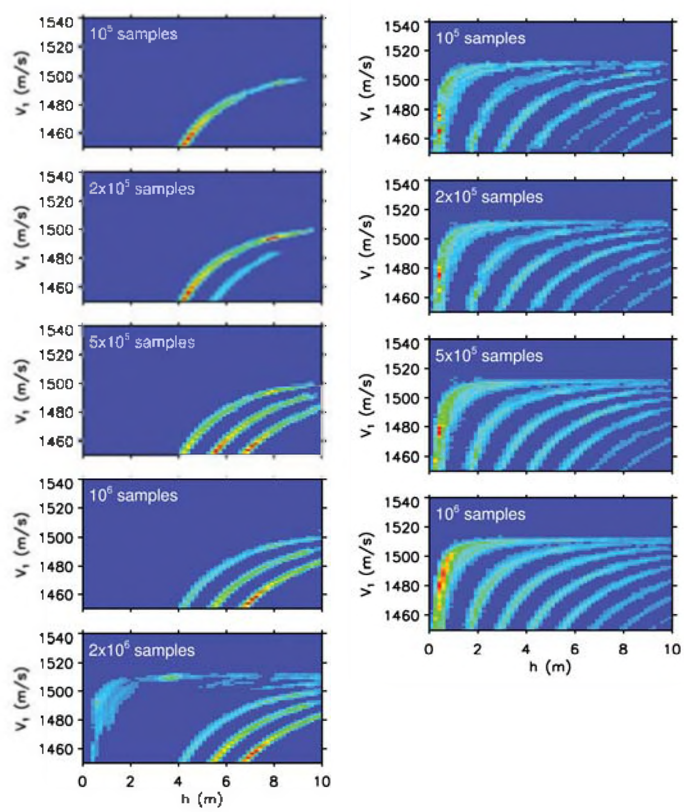


Figure 3. Comparison of joint marginal probability distributions for h and v_1 as computed using MHS (left column) and parallel tempering (right column) using the number of samples indicated on each panel.

EFFICIENT BAYESIAN MULTI-SOURCE LOCALIZATION

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

1. INTRODUCTION

This paper develops and illustrates an efficient approach to the simultaneous localization of an unknown number of ocean acoustic sources, based on minimizing the Bayesian information criterion (BIC) over source parameters [1]. A Bayesian formulation is developed in which the number, locations, and complex strengths (representing amplitudes and phases) of an unknown number of sources are considered random variables constrained by acoustic data and prior information. The BIC, which balances data misfit with a penalty for extraneous parameters, is minimized using simulated annealing, with Gibbs sampling applied over source locations. Closed-form maximum-likelihood (ML) expressions for source strength and noise variance at each frequency allow these parameters to be sampled implicitly, substantially reducing the dimensionality and improving the efficiency of the inversion. Gibbs sampling and the implicit formulation provide an efficient scheme for adding and deleting sources with a reasonable acceptance rate during the optimization. A simulated example is presented which considers localizing several quiet submerged sources in the presence of multiple loud near-surface interfers.

2. THEORY

This section develops the Bayesian approach to multiple-source localization. Consider data $\mathbf{d} = \{\mathbf{d}_f,\ f = 1, N_F\}$ consisting of complex (frequency-domain) acoustic fields at an array of N_H hydrophones for N_F frequencies. The field at each frequency is assumed to be due to $s = 1, N_S$ acoustic sources at locations (ranges and depths) $\mathbf{x} = \{\mathbf{x}_s,\ s = 1, N_S\} = \{(r_s, z_z)\ s = 1, N_S\}$ with complex strengths $\mathbf{a} = \{[a_f]_s\}$. Errors on $\mathbf{d}_f$ are assumed to be complex Gaussian distributed with unknown variance v_f. Let $\mathbf{m} = \{\mathbf{x},\mathbf{a},\boldsymbol{v}\}$ represent the set of unknown model parameters. Data and parameters are considered to be random variables related by Bayes' rule

$$P(\mathbf{m} \mid \mathbf{d}, N_S) = \frac{P(\mathbf{d} \mid \mathbf{m}, N_S) P(\mathbf{m} \mid N_S)}{P(\mathbf{d} \mid N_S)}. \tag{1}$$

In Eq. (1), the posterior probability density (PPD), $P(\mathbf{m}|\mathbf{d})$, represents the state of information for the parameters incorporating both data information, $P(\mathbf{d}|\mathbf{m})$, and prior information, $P(\mathbf{m})$. Interpreting the conditional probability $P(\mathbf{d}|\mathbf{m})$ as a function of $\mathbf{m}$ for the (fixed) observed data $\mathbf{d}$ defines the likelihood function $L(\mathbf{m}) \propto \exp[-E(\mathbf{m})]$, where E is the data misfit (log likelihood) function. Given the assumptions stated above, the likelihood is given by

$$\begin{aligned} L(\mathbf{x},\mathbf{a},\mathbf{v}) &= \prod_{f=1}^{N_F} \frac{1}{(\pi v_f)^{N_f}} \exp\left\{ -\left| \mathbf{d}_f - \sum_{s=1}^{N_S} a_{fs} \mathbf{d}_f(\mathbf{x}_s) \right|^2 / v_f \right\} \\ &= \frac{1}{\prod_{f=1}^{N_f} (\pi v_f)^{N_f}} \exp\left\{ -\sum_{f=1}^{N_F} \left| \mathbf{d}_f - \mathbf{D}_f \mathbf{a}_f \right|^2 / v_f \right\} \end{aligned} \tag{2}$$

where $\mathbf{d}_f(\mathbf{x}_s)$ represents the modelled acoustic fields for a unit-amplitude, zero-phase source at location $\mathbf{x}_s$ and $\mathbf{D}_f$ is an N_H by N_S complex matrix of acoustic fields defined

$$[\mathbf{D}_f]_{hs} = [\mathbf{d}_f(\mathbf{x}_s)]_h. \tag{3}$$

Equation (2) can be written $L \propto \exp[-E]$ where the misfit function is given by

$$E(\mathbf{x},\mathbf{a},\mathbf{v}) = \sum_{f=1}^{N_F} \left(| \mathbf{d}_f - \mathbf{D}_f \mathbf{a}_f |^2 / v_f + N_H \log_e v_f \right). \tag{4}$$

Implicit sampling over source strengths and noise variances is derived by setting $\partial E / \partial \mathbf{a}_f = \partial E / \partial v_f = 0$ to obtain ML estimates

$$\begin{aligned} \hat{\mathbf{a}}_f &= \left(\mathbf{D}_f^H \mathbf{D}_f \right)^{-1} \mathbf{D}_f^H \mathbf{d}_f \\ \hat{\mathbf{v}}_f &= \frac{1}{N_H} \left| \left[\mathbf{I} - \left(\mathbf{D}_f^H \mathbf{D}_f \right)^{-1} \mathbf{D}_f^H \right] \mathbf{d}_f \right|^2 \end{aligned} \tag{5}$$

where ^{H} indicates Hermitean (conjugate transpose) and $\mathbf{I}$ is the identity matrix. Applying these in Eq. (4), the misfit can be written

$$E(\mathbf{x}) = N_H \sum_{f=1}^{N_F} \log_e \left| \left[\mathbf{I} - \mathbf{D}_f (\mathbf{D}_f^H \mathbf{D}_f)^{-1} \mathbf{D}_f^H \right] \mathbf{d}_f \right|^2. \tag{6}$$

Evaluating Eq. (6) for specific $\mathbf{x}$ automatically applies the ML solution for $\mathbf{a}$ and $\mathbf{v}$, and hence accounts for the corresponding variability in source strengths and variances implicitly. This greatly reduces the dimensionality and improves the efficiency of multiple-source localization.

Determining the number of sources that contribute significantly to the acoustic field may be considered an application of model selection; i.e., seeking the most appropriate N_S given the measured data **d**. In Baye's rule, Eq. (1), the conditional probability $P(\mathbf{d}|N_S)$ may be considered the likelihood of N_S, and is referred to as the Bayesian evidence for N_S. Since the evidence serves as a normalizing factor in Bayes' rule it can be written

$$P(\mathbf{d} \mid N_S) = \int P(\mathbf{d} \mid \mathbf{m}, N_S) P(\mathbf{m} \mid N_S) d\mathbf{m}. \quad (7)$$

Unfortunately, numerical solution of this integral is not practical for all models sampled in the localization algorithm. Rather, an asymptotic point estimate, the BIC, is applied here:

$$-2\log_e P(\mathbf{d} \mid N_S) \approx BIC = -2\log_e L(\hat{\mathbf{m}}, N_S) + N_S \log_e N_d \quad (8)$$

where $\hat{\mathbf{m}}$ is the ML source location obtained by minimizing Eq. (6) and N_d is the number of data. As the BIC is based on the negative log likelihood, low BIC values are preferred. The first term on the right of Eq. (8) favours models with low misfits; however, this is balanced by the second term which penalizes unjustified free parameters. Minimizing the BIC provides the smallest number of acoustic sources which fits the data to within uncertainties, or, conversely, the largest number of sources resolved by the data.

The multiple-source localization algorithm developed here optimizes over the number and locations of acoustic sources, as well as complex sources strengths and noise variance at each frequency, by minimizing the BIC. This minimization is carried out by applying heat-bath (Gibbs sampling) simulated annealing with fast cooling. Source locations are treated as explicit parameters, and source strengths and variances as implicit parameters. Each iteration of the simulated annealing process consists of Gibbs sampling each location parameter as well and an attempt to either add or remove a source. Sources are added by Gibbs sampling from the conditional probability distribution defined by the existing sources, and when sources are removed the remaining sources are Gibbs sampled to compensate for the change in acoustic fields. Implementation of the implicit formulation, Eq. (6), requires a large number of complex matrix inversions which are handled efficiently using a parallel implementation of Gauss-Jordan elimination that is stable without pivoting since the matrices are diagonally dominant.

3. EXAMPLE

This section presents a (simulated) example of the multiple-source localization algorithm involving 2 submerged sources and 3 louder near-surface interfering sources, with acoustic fields recorded at $N_F = 3$ frequencies of 200, 300, and 400 Hz at a 24-hydrophone vertical array spanning a 100-m water column. The ranges, depths, and signal-to-noise ratios (SNR, taken to be constant over frequency) of the sources are as follows: source 1 (8 km, 4 m, 10 dB), source 2 (3 km, 2 m, 8 dB), source 3 (5.5 km, 2 m, 6 dB),

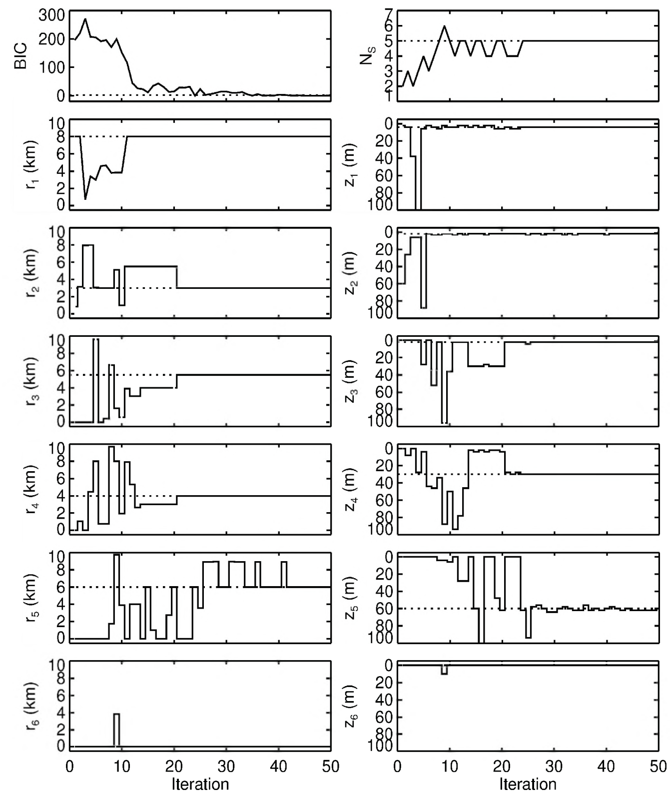


Figure 1. Inversion results as a function of simulated annealing iteration for BIC, number of sources, and ranges and depths of up to 6 sources (a maximum of 7 sources was allowed, but never accepted in the inversion). An absent source is assigned zero range and depth. The BIC is arbitrarily shifted so the minimum value corresponds to zero. Sources are ordered according to SNR. Dotted lines indicate true values.

source 4 (4 km, 30 m, 4 dB), and source 5 (6 km, 60 m, 0 dB). The source search region is 0-10 km in range and 0-100 m in depth, with from 1 to a maximum of 7 sources allowed in the search. This formulation includes a total of $2N_S(1+N_F)+N_F$ (real) unknowns (e.g., up to 51 for 6 sources), of which $2N_S$ are treated as explicit parameters and the remaining as implicit parameters.

The results of the localization are shown in Fig. 1. The BIC drops quickly (although not monotonically) and the number of sources settles into the correct value of $N_S = 5$ by about iteration 30 of the simulated annealing process. All source ranges and depths are correctly determined by about iteration 40, with the order in which the sources converge approximately following that of decreasing source SNR (i.e., the highest SNR source converges first, followed by the second highest, etc.).

REFERENCES

[1] S.E. Dosso, 2012. Acoustic localization of an unknown number of sources in an unknown environment. Canadian Acoustics, **40**, 3-12.

THE EVOLUTION OF AN ACOUSTIC HOMING SYSTEM FOR UNDERWATER VEHICLES

Garry J. Heard[1], Nicos Pelavas[1], Carmen E. Lucas[1], R. Fleming[1], D. Watson[2]
[1] DRDC Atlantic, PO Box 1012, Dartmouth, NS CANADA B2Y 3Z7
[2] Omnitech Electronics Inc., 10 Akerley Blvd. Suite 1, Dartmouth, NS CANADA B3B 1J4

1. INTRODUCTION

Project Cornerstone was a joint Natural Resources Canada and Defence R&D Canada project using deep diving Autonomous Underwater Vehicles (AUV) to assist in the mapping of the Arctic seafloor in support of the Canadian submission to the United Nations Convention on the Law Of the Sea (UNCLOS). The project made use of a pair of modified International Submarine Engineering Explorer AUVs [1]. One of the modifications added to the AUVs was the Long-Range Acoustic Bearing (LRAB) homing system developed at DRDC (patent applied for). This brief paper describes the initial testing of the homing concept, the Explorer version of the LRAB, and a new first-stage of system miniaturization that will allow the homing device to be used in medium-sized AUVs.

2. CONCEPT TESTING

Figure 1 shows the Concept Test Array (CTA) being tested in a tent on the shore-fast ice at CFS Alert during spring 2009. The CTA was built with PVC tubes and plate. Seven hydrophones of an available analogue array, each on a separate cable, were used to provide the CTA receivers.

The CTA was lowered to mid-water depth, while a portable source was moved to a number of different locations around the central location of the CTA. Data were collected for various frequencies, source levels, and source depths.

Following the trial, the data were analyzed and a robust bearing estimation algorithm was developed. Results from the initial tests were good and proved the concept to be viable.

Figure 1. An initial Concept Test Array being readied for testing under the ice at CFS Alert.

3. FIRST GENERATION LRAB

DRDC and Omnitech Electronics have developed a low-cost, low-power, digital array technology that has been used to develop a number of different underwater sensing platforms such as the Northern Watch Underwater Sensing System [2]. This array technology is eminently suitable for the necessarily power conscious application of acoustic homing in the limited energy environment of an AUV.

A contract was established with Omnitech Electronics Inc. who used the DRDC digital array technology to construct a seven-element receiving array and a small, low-power processor, which interacts with the AUV control computer to provide relative bearings to the acoustic signal source. Figure 2 shows the seven-element, first generation, LRAB mounted in the nose cone of an Explorer. Figure 3 shows the processor unit that is mounted within the AUV pressure hull. This processor handles all of the AUV acoustic telemetry and when desired provides bearing and altitude estimates of the acoustic source relative to the vehicle's longitudinal axis.

Figure 4 illustrates the results of acoustic homing under the ice in the Arctic Ocean near the Sever Spur during spring 2010. The AUV activated the homing mode when it was expected to be within 50 km of an ice camp that was drifting freely. The 300-km trek required almost three days of travel time for the AUV during which the ice camp moved a significant distance. The abrupt course change at ‘Homing Start’ is a consequence of the unknown camp drift.

The inset in Figure 4 shows the tight maneuvers of the AUV when it arrived at the camp location. The AUV was forced to

Figure 2. The LRAB array in the ISE Explorer nose cone.

appropriate N_S given the measured data **d**. In Baye's rule, Eq. (1), the conditional probability $P(\mathbf{d}|N_S)$ may be considered the likelihood of N_S, and is referred to as the Bayesian evidence for N_S. Since the evidence serves as a normalizing factor in Bayes' rule it can be written

$$P(\mathbf{d} \mid N_S) = \int P(\mathbf{d} \mid \mathbf{m}, N_S) P(\mathbf{m} \mid N_S) d\mathbf{m}. \quad (7)$$

Unfortunately, numerical solution of this integral is not practical for all models sampled in the localization algorithm. Rather, an asymptotic point estimate, the BIC, is applied here:

$$-2\log_e P(\mathbf{d} \mid N_S) \approx BIC = -2\log_e L(\hat{\mathbf{m}}, N_S) + N_S \log_e N_d \quad (8)$$

where $\hat{\mathbf{m}}$ is the ML source location obtained by minimizing Eq. (6) and N_d is the number of data. As the BIC is based on the negative log likelihood, low BIC values are preferred. The first term on the right of Eq. (8) favours models with low misfits; however, this is balanced by the second term which penalizes unjustified free parameters. Minimizing the BIC provides the smallest number of acoustic sources which fits the data to within uncertainties, or, conversely, the largest number of sources resolved by the data.

The multiple-source localization algorithm developed here optimizes over the number and locations of acoustic sources, as well as complex sources strengths and noise variance at each frequency, by minimizing the BIC. This minimization is carried out by applying heat-bath (Gibbs sampling) simulated annealing with fast cooling. Source locations are treated as explicit parameters, and source strengths and variances as implicit parameters. Each iteration of the simulated annealing process consists of Gibbs sampling each location parameter as well and an attempt to either add or remove a source. Sources are added by Gibbs sampling from the conditional probability distribution defined by the existing sources, and when sources are removed the remaining sources are Gibbs sampled to compensate for the change in acoustic fields. Implementation of the implicit formulation, Eq. (6), requires a large number of complex matrix inversions which are handled efficiently using a parallel implementation of Gauss-Jordan elimination that is stable without pivoting since the matrices are diagonally dominant.

3. EXAMPLE

This section presents a (simulated) example of the multiple-source localization algorithm involving 2 submerged sources and 3 louder near-surface interfering sources, with acoustic fields recorded at $N_F = 3$ frequencies of 200, 300, and 400 Hz at a 24-hydrophone vertical array spanning a 100-m water column. The ranges, depths, and signal-to-noise ratios (SNR, taken to be constant over frequency) of the sources are as follows: source 1 (8 km, 4 m, 10 dB), source 2 (3 km, 2 m, 8 dB), source 3 (5.5 km, 2 m, 6 dB), source 4 (4 km, 30 m, 4 dB), and source 5 (6 km, 60 m, 0 dB). The source search region is 0-10 km in range and 0-100 m in depth, with from 1 to a maximum of 7 sources allowed in the search. This formulation includes a total of $2N_S\,(1+N_F)+N_F$ (real) unknowns (e.g., up to 51 for 6 sources), of which $2N_S$ are treated as explicit parameters and the remaining as implicit parameters.

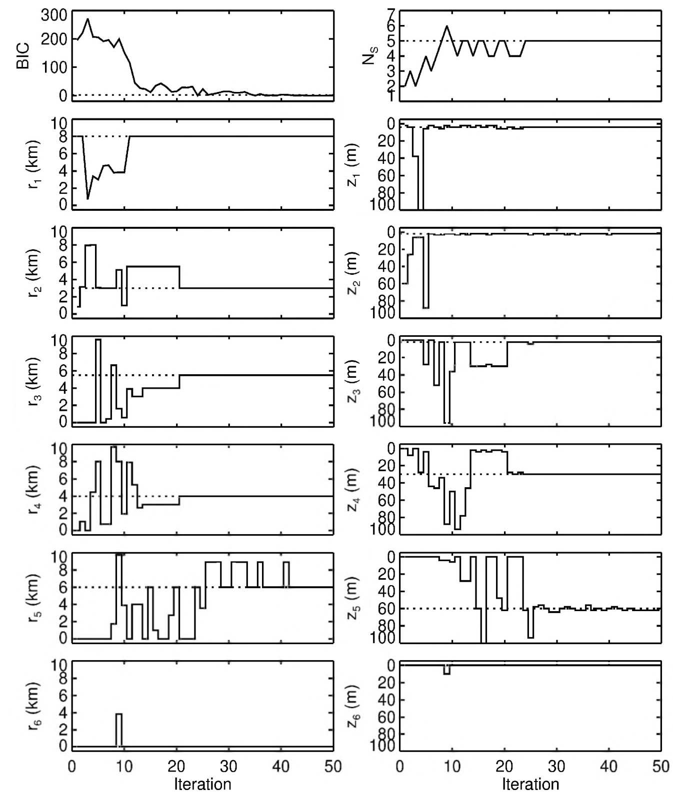


Figure 1. Inversion results as a function of simulated annealing iteration for BIC, number of sources, and ranges and depths of up to 6 sources (a maximum of 7 sources was allowed, but never accepted in the inversion). An absent source is assigned zero range and depth. The BIC is arbitrarily shifted so the minimum value corresponds to zero. Sources are ordered according to SNR. Dotted lines indicate true values.

The results of the localization are shown in Fig. 1. The BIC drops quickly (although not monotonically) and the number of sources settles into the correct value of $N_S = 5$ by about iteration 30 of the simulated annealing process. All source ranges and depths are correctly determined by about iteration 40, with the order in which the sources converge approximately following that of decreasing source SNR (i.e., the highest SNR source converges first, followed by the second highest, etc.).

REFERENCES

[1] S.E. Dosso, 2012. Acoustic localization of an unknown number of sources in an unknown environment. Canadian Acoustics, **40**, 3-12.

STRUCTURAL ANALYSIS OF MULTI-FLUID SHELL SYSTEMS SUBJECTED TO AN EXTERNAL ACOUSTIC PULSE

Serguei Iakovlev [1], Garrett Dooley[2], Kyle Williston[1], and Jonathan Gaudet [3]
[1] Dept. of Engineering Mathematics and Internetworking, Dalhousie University, Halifax, Nova Scotia, Canada
[2] Lunenburg Industrial Foundry & Engineering, Lunenburg, Nova Scotia, Canada
[3] O'Neil Scriven and Associates Ltd, Halifax, Nova Scotia, Canada

1. INTRODUCTION

In our recent work [1], we have demonstrated that when an elastic shell filled with and submerged into different fluids is subjected to an external acoustic pulse, the diversity of the internal reflection and focusing phenomena is such that it leads, for certain combination of the parameters of the fluids, to a very considerable increase of the pressure in the fluid which in some cases can be as high as 110% of the peak incident pressure. It appears that such a high pressure should, at least in some instances, result in a high structural stress, an effect of obvious practical significance, and addressing such a possibility and the corresponding structural dynamics is the subject of the present study.

2. MATHEMATICAL FORMULATION AND SOLUTION METHODOLOGY

We consider a thin elastic circular cylindrical shell filled with and submerged into different fluids. We assume that the shell is thin enough, and that its deflections are small in comparison to its thickness, so that the linear shell theory can be employed; we further assume that the Love-Kirchhoff hypothesis holds true. We note that although using the Reissner-Mindlin model was shown to provide more accurate results, employing the Love-Kirchhoff model is still acceptable for the purposes of the present study.

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

As was in established [2], a single most important parameter determines the appearance of the hydrodynamic fields observed in the system, namely ζ, which is defined as the ratio of the sound speed in the internal fluid to that in the external one. Changing ζ implies varying the acoustic properties of the fluids.

The problem is approached with the methodology developed in our earlier work [3-6], i.e. we apply the Laplace transform time-wise to the wave equations and then separate the spatial variables in order to obtain the pressure components as Fourier series with time-dependant coefficients. Then, the hydrodynamic and structural parts are coupled using a ID finite-difference approach.

We note that although the present approach is used, in this case, to model a structurally simple system (a single shell), it can also be successfully employed to address more complex structures [7,8]. Using a more complex model for the shell is also possible [9].

3. RESULTS AND CONCLUSIONS

A steel shell is considered with the thickness of 0.01 m and radius of 1 m, submerged into and filled with fluids of the same density but the acoustic speeds varying according to the changes of ζ. The interaction with a cylindrical pulse is analyzed, and the rate of the exponential decay is assumed to be 0.0001314 s while the initial pressure in the front is 250 kPa.

It seems suitable to first address how the high internal pressure affects the stress state of the shell. To that end, Fig. 1 shows the snapshots of the acoustic field inside the shell and the corresponding stress state of the shell at the same instant for ζ=0.50. As is clearly seen from the acoustic field image, this is a scenario where the high pressure occurs in a region immediately adjacent to the shell surface, a case that is of highest practical interest.

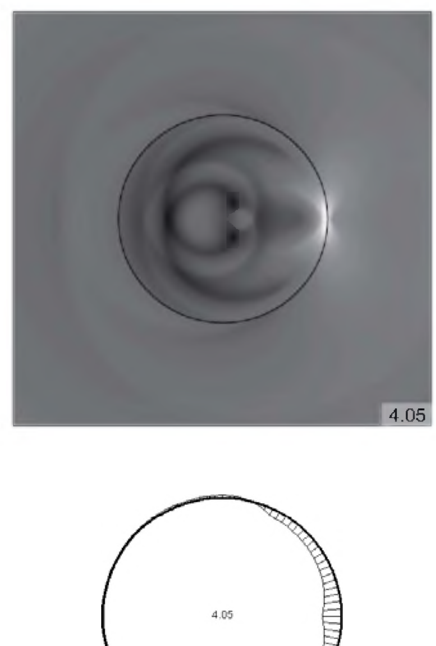


Figure 1. The internal acoustic field and the corresponding stress state of the shell for ζ=0.50.

From the practitioner's point of view, however, the estimates of the maximum stress attained in the system for each combination of the parameters are a primary concern, not the detailed structure of the stress field.

In order to assess such extremities of the stress state, we analyze the highest compressive and tensile stress induced in the shell for the values of ζ that were shown to result in the most interesting internal acoustic patterns, Fig. 2 and 3. We underscore that we consider the tensile and compressive stress separately due to the fact that they have been seen to be affected very differently by the changes in the systems similar to the one at hand [7], and we will show that this is indeed a trend that is seen in the present case as well.

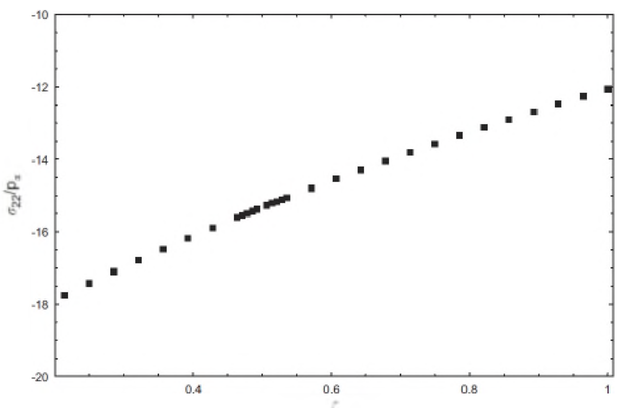


Figure 2. Variation of the maximum compressive stress depending on ζ (normalized to the peak incident pressure).

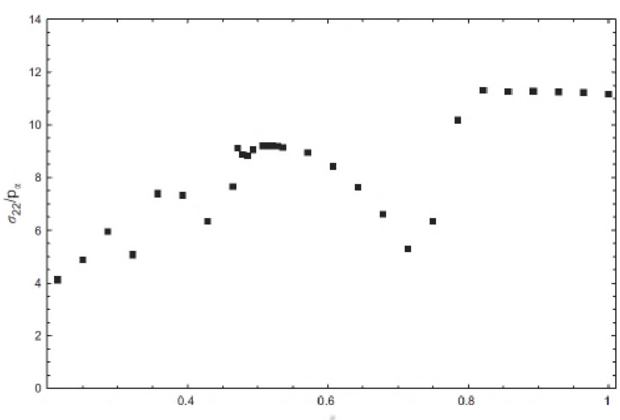


Figure 3. Variation of the maximum tensile stress depending on ζ (normalized to the peak incident pressure).

As one can see, the compressive stress exhibits a very regular pattern of change over the entire range of the values of ζ considered, and thus is not affected by the multitude of the phenomena in the internal fluid. At the same time, the tensile stress is profoundly affected by the changes of the acoustic properties of the fluids, so much so that it prompts a separate discussion of the matter.

Namely, despite the existence of the regions where the change of the peak tensile stress is rather regular, there also are regions where its changes are quite dramatic. More specifically, there are four intervals that can be identified: the intervals [0.5,0.7] and [0.82,1] which are characterized by a very regular change, the interval [0.2,0.5] where the behavior is somewhat oscillatory and no particular trend can be identified, and the interval [0.7,0.82] which is characterized by a sudden and dramatic change of the maximum tensile stress.

Within this last interval, the tensile peak is extremely sensitive to the changes of ζ: increasing ζ by only 13% (from 0.714 to 0.821) results in more than a two-fold increase of the maximum tensile stress. This sensitivity is something that the practitioner definitely needs to be aware of at the pre-design stage: even though it is very rarely a case that the properties of the fluids used in a system can be arbitrary chosen, the structural effects of varying the acoustic speeds we outlined here are too important to be ignored.

REFERENCES

[1] Iakovlev, S., Williston, K., Dooley, G., Gaudet, J. (2011) Analysis of the extreme acoustic pressure in multi-fluid shell systems subjected to an external pulse, *Canadian Acoustics*, 204-205.

[2] Iakovlev, S. (2009) Interaction between an external shock wave and a cylindrical shell filled with and submerged into different fluids, *Journal of Sound and Vibration* 322 (1-2), 401-437.

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

[4] Iakovlev, S. (2007) Submerged fluid-filled cylindrical shell subjected to a shock wave: Fluid-structure interaction effects, *Journal of Fluids and Structures* 23 (1), 117-142.

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

[6] Iakovlev, S. (2008) Interaction between a submerged evacuated cylindrical shell and a shock wave. Part II: Numerical aspects of the solution, *Journal of Fluids and Structures* 24 (7), 1098-1119.

[7] Iakovlev, S. (2004) Influence of a rigid coaxial core on the stress-strain state of a submerged fluid-filled cylindrical shell subjected to a shock wave, *Journal of Fluids and Structures* 19 (7), 957-984.

[8] Iakovlev, S., Gaudet, J., Dooley, G., MacDonald, B. (2010) Hydrodynamic fields induced by the shock response of a fluid-filled submerged cylindrical shell containing a rigid co-axial core, *Journal of Sound and Vibration* 329 (16), 3359-3381.

[9] Leblond, C., Iakovlev, S., Sigrist, J.-F. (2009) A fully elastic model for studying submerged circular cylindrical shells subjected to a weak shock wave, *Mecanique & Industries* 10 (3-4), 275-284.

ACKNOWLEDGEMENTS

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

ACCURATE MODELING OF THE STRUCTURE OF THE ACOUSTIC FIELD RADIATED BY A SUBMERGED CYLINDRICAL SHELL RESPONDING TO AN EXTERNAL PULSE

Serguei Iakovlev [1], Hugo A. F. A. Santos[2], Kyle Williston[1], Robynne Murray[3], and Martin Mitchell [1]
[1]Dept. of Engineering Mathematics and Internetworking, Dalhousie University, Halifax, Nova Scotia, Canada
[2]Dept. of Civil Engineering and Architecture, Technical University of Lisbon, IST, Lisbon, Portugal
[3]Dept. of Mechanical Engineering, Dalhousie University, Halifax, Nova Scotia, Canada

1. INTRODUCTION

In recent years, a number of models have been introduced for modeling the interaction between cylindrical shells and non-stationary acoustic loads based on the Kirchhoff-Love theory of thin shells [1-5]. These models allowed for a rather comprehensive analysis of the interaction in some of the more practically interesting cases, and certain new insights into the phenomenology of the acoustic response of fluid-contacting shells followed.

At the same time, all of the models mentioned had certain shortcomings that were not essential for the general understanding of the interaction, but that presented some rather serious inconveniences when a more detailed analysis was intended. Specifically, the structure of the acoustic field re-radiated by the shell back into the fluid as it was responding to the load was found to be different than that seen in experiments [e.g. 6, 7]: the model did not reproduce the two distinctly different waves, A_0 and S_0, that have been experimentally observed.

It was therefore necessary to overcome this limitation, and there were two ways to do that. The first one was to consider the shell as a 3D solid body, thus abandoning all the simplifying assumptions of shell theories, and this has been done in [8] where a fully elasto-dynamic methodology of modeling the system in question has been developed and successfully validated.

The second way to deal with the shortcomings of the model based on the Kirchhoff-Love theory was to consider a more advanced shell theory that would still take advantage of the unique geometry of a shell, but, at the same time, would be more accurate in terms of reproducing the structural behavior of the shell (and, therefore, the dynamics of the radiated field as well), perhaps at the expense of its computational attractiveness as it compares to its Kirchhoff-Love counterpart. Such an undertaking is the subject of the present study.

2. MATHEMATICAL FORMULATION AND SOLUTION METHODOLOGY

We consider a thin evacuated elastic circular cylindrical shell submerged into fluid, and model it using the Reissner-Mindlin theory of shells [9]. The fluid is assumed to be irrotational, inviscid, and linearly compressible, thus the wave equations are used to model the fluid dynamics. The fluids and the shell are coupled through the dynamic boundary condition on the interface.

The problem is approached with the methodology developed in our earlier work [1-5], i.e. we apply the Laplace transform time-wise to the wave equations and then separate the spatial variables in order to arrive at the expressions for the transforms of the internal and external pressure in a form of a series of modified Bessel functions of the first (internal fluid) and second (external fluid) kind. The pressure is then obtained as a Fourier series with time-dependant coefficients which, for the radiation pressure, depend on the unknown normal displacements of the shell.

Then, the same series form is used for the shell displacements and, substituting them into the shell equations, we arrive at the systems of the ordinary differential equations for each of the displacement harmonics. The systems are then approached numerically, using a finite-difference scheme, and the resulting normal displacement is used to compute the radiation pressure.

We note that although the present approach is used here to model a single shell, it can also be successfully employed to address systems consisting of several structures, for example a shell with a rigid core [10,11].

3. RESULTS AND CONCLUSIONS

A steel shell is considered with the thickness of 0.03 m and the radius of 1 m, submerged into water. The interaction with a cylindrical external acoustic pulse is analyzed, and the rate of the exponential decay is assumed to be 0.0001314 s while the initial pressure in the front is 250 kPa.

Fig. 1 shows three images of the radiated field for the chosen shell and pulse using three different models. The first image is that produced by the experiments [7], and it clearly shows the existence of two radiated waves, the labeled S_0 wave, or symmetric Lamb wave, and the A_0 wave, or pseudo-Rayleigh wave. The second image is numerically simulated and it is produced by a model based on the Kirchhoff-Love theory (KL model). The third image is numerically simulated using the Reissner-Mindlin shell theory (RM model).

As is apparent from the images, the KL model, although correctly reproducing the overall shape of the radiated field, fails to accurately reproduce its details, the drawback we mentioned earlier. The RM model, however, correctly reproduces both the S_0 and A_0 waves, thus eliminating the issues inherent to the KL model. At the same time, still being a shell model, it relies on certain simplifying assumptions regarding the structural behavior of the shell, and thus is very attractive from the computational point of view. More specifically, simulations of the stress state of the shell based on this model take only 13% longer than those based on the KL model, the latter has been demonstrated to be highly computationally efficient in a number of contexts.

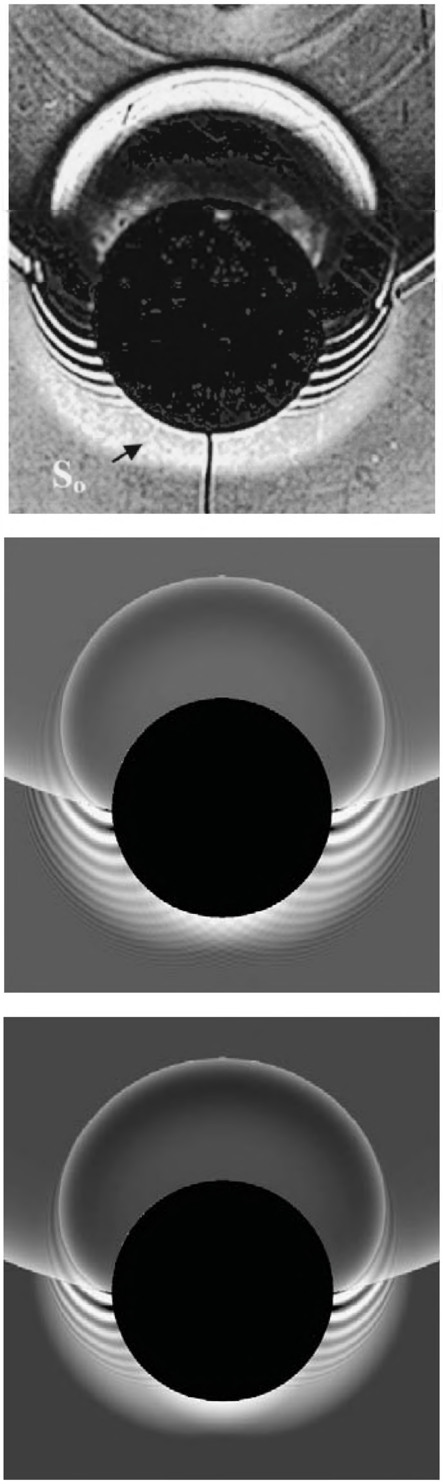


Figure 1. The radiated field for a steel shell with the thickness-to-radius ratio of 0.03 submerged into water: experiments [7], top image; KL model, middle image; RM model, bottom image.

Along with the qualitative validity of the simulated results, their quantitative validity had to be established as well. To that end, we compared the numerically simulated time-histories of the acoustic pressure at several points of the fluid to the available experimental time-histories for similar systems, and observed a very good match for the S_0 wave. At the same time, due to the lack of the available experimental data, we have not so far been able to assess the quantitative accuracy of the modeling of the A_0 wave.

REFERENCES

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

[2] Iakovlev, S. (2007) Submerged fluid-filled cylindrical shell subjected to a shock wave: Fluid-structure interaction effects, *Journal of Fluids and Structures* 23 (1), 117-142.

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

[4] Iakovlev, S. (2008) Interaction between a submerged evacuated cylindrical shell and a shock wave. Part II: Numerical aspects of the solution, *Journal of Fluids and Structures* 24 (7), 1098-1119.

[5] Iakovlev, S. (2009) Interaction between an external shock wave and a cylindrical shell filled with and submerged into different fluids, *Journal of Sound and Vibration* 322 (1-2), 401-437.

[6] Ahyi, A. C., Pernod, P., Gatti, O., Latard, V., Merlen, A., Uberall, H. (1998) Experimental demonstration of the pseudo-Rayleigh (A_0) wave, *Journal of the Acoustical Society of America* 104, 2727-2732.

[7] Derbesse, L., Pernod, P., Latard, V., Merlen, A., Decultot, D., Touraine, N., Maze, G. (2000) Acoustic scattering from complex elastic shells: visualization of S_0, A_0, and A waves, *Ultrasonics* 38, 860-863.

[8] Leblond, C., Iakovlev, S., Sigrist, J.-F. (2009) A fully elastic model for studying submerged circular cylindrical shells subjected to a weak shock wave, *Mécanique & Industries* 10 (3-4), 275-284.

[9] Naghdi, P. M., Cooper, R. M. (1956) Propagation of elastic waves in cylindrical shells including the effects of transverse shear and rotational inertia, *Journal of the Acoustical Society of America* 28, 56-63.

[10] Iakovlev, S., Mitchell, M., Dooley, G. (2011) Modeling two-fluid response of thin elastic shells, *International Journal for Numerical Methods in Fluids* 65 (11-12), 1389-1406.

[11] Iakovlev, S., Gaudet, J., Dooley, G., MacDonald, B. (2010) Hydrodynamic fields induced by the shock response of a fluid-filled submerged cylindrical shell containing a rigid co-axial core, *Journal of Sound and Vibration* 329 (16), 3359-3381.

ACKNOWLEDGEMENTS

S.I., K.W., R.M. and M.M. acknowledge the financial support of the Natural Sciences and Engineering Research Council (NSERC) of Canada. H.S. acknowledges the financial support of the Fundação para a Ciência e a Tecnologia (FCT) of Portugal. S.I. also acknowledges the financial support of the Killam Trusts and Dalhousie University.

A BAYESIAN FRAMEWORK FOR GEOACOUSTIC INVERSION OF WIND-DRIVEN AMBIENT NOISE IN SHALLOW WATER

Jorge E. Quijano, Stan E. Dosso, and Jan Dettmer
School of Earth and Ocean Sciences, University of Victoria, 3800 Finnerty Road, Victoria, B.C., Canada, V8P 5C2
jorgeq@uvic.ca

1. INTRODUCTION

Knowledge of seabed geoacoustic parameters in shallow water is of great importance for the operation of active-sonar systems, to properly identify and classify acoustic returns, and to predict levels of reverberation. Estimation of parameters such as sediment sound speed, density, attenuation, and layering structure remains a challenging task due to constraints on hardware, data collection and analysis, and cost of maritime surveys. This work presents a Bayesian framework for estimating seabed parameters based on inversion of wind-driven ambient-noise data collected at a vertical linear array (VLA) of hydrophones[1].

Remote sensing for acoustic characterization of large areas of the seabed has relied on active-source methods, in which broadband pulses are transmitted in the water column and echoes reflected from the seabed are analyzed to determine geoacoustic parameters. Although the reflection data obtained this way is rich in information on the seabed structure, there are concerns due to increasing levels of man-made noise in the oceans, which are likely to have a negative impact on marine mammals. In addition, the high power demand of active-source systems makes them unsuitable for implementation in autonomous underwater vehicles (for which battery life is limited).

The use of the ambient-noise field produced by wind-driven breaking waves to probe the seabed has been suggested[1] as an alternative (passive) method of remote sensing. This field is modelled as an extended source consisting of surface acoustic monopoles arranged in a thin layer near the air-water interface (Fig.1).

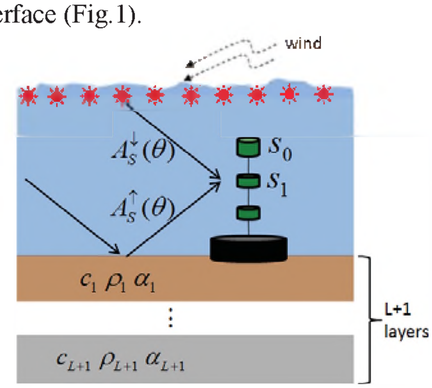


Figure. 1. Estimation of the bottom loss by measurement and beamforming of the wind-driven ambient-noise field.

The resulting field is recorded at an N-element VLA and separated into angle-dependent power flux components. The input data $\mathbf{d}$ to the inversion algorithm in this work consist of the angle- and frequency-dependent bottom loss (BL), computed as $10\log_{10}\left(A_s^{\uparrow}(\theta)/A_s^{\downarrow}(\theta)\right)$ since the upward power flux differs from the downward component by a single interaction with the sediment.

In this work, Bayesian inversion is applied to estimate the joint posterior probability density (PPD) of geoacoustic parameters[1-2]. Properties of the PPD such as the maximum *a posteriori* model, mean model, parameter marginal probability distributions, and inter-parameter correlations can be estimated from sampling the PPD. A fundamental step in the inversion is the selection of a model parametrization (e.g., number of seabed layers) consistent with the data information content. To this end, a trans-dimensional (trans-D) inversion approach[2] is applied.

2. METHOD

The estimated BL data carry information of the seabed geoacoustic parameters[1]. Geoacoustic inversion of these data poses a high-dimensional non-linear problem that can be efficiently handled by the Bayesian framework, which requires a model to represent the physical system (i.e. the seabed) that gives rise to the data. With the trans-D method[2], models I_k from a set of K candidates are included in the estimation of the joint PPD

$$P(\mathbf{m}_k, I_k \mid \mathbf{d}) = \frac{P(\mathbf{d} \mid \mathbf{m}_k, I_k)P(\mathbf{m}_k \mid I_k)P(I_k)}{P(\mathbf{d})}, \quad (1)$$

where $P(\mathbf{d}|\mathbf{m}_k, I_k)$ is the likelihood function (defined here based on the assumption of Gaussian-distributed residuals[1]), and $P(I_k)$ is the prior distribution for the parametrization, assumed here as a discrete uniform distribution. The distribution $P(\mathbf{m}_k|I_k)$ is the prior for the geoacoustic parameters $\mathbf{m}_k$ corresponding to a layered seabed with k interfaces. The vector $\mathbf{m}_k$ is defined as

$$\mathbf{m}_k = [c_1\ \rho_1\ \alpha_1\ h_1 \ldots c_{k+1}\ \rho_{k+1}\ \alpha_{k+1}\ SNR_1 \ldots SNR_F]^T \quad (2)$$

where c_i, ρ_i, α_i, and h_i are the sound speed, density, attenuation and thickness of the i^{th} layer, respectively, and the SNRs account for the unknown strength of the wind-driven ambient-noise data (i.e. the useful signal) versus other unwanted sources of noise. Details on the interpretation of the SNR parameters and the computation of $\mathbf{d}$ from the ambient-noise field are given elsewhere[1].

The PPD is sampled by a reversible-jump Markov chain Monte Carlo (rjMCMC) algorithm[2], which uses an extended Metropolis-Hasting (MH) criterion that allows trans-D jumps between parameterizations I_k, quantifying the uncertainty due to the lack of knowledge of the model parameterization.

For data collected with a drifting array, the PPD evolves with time as the array moves over sediments in which the number of layers, the depth of interfaces of high acoustic contrast, or the geoacoustic parameters change as a function of range. Sequential datsets can then be obtained by discretizing continuous-time recordings of ambient noise. For this application, a particle filter[3] is used to update the estimated geoacoustic parameters from one array position to the next as new data become available.

3. RESULTS

Geoacoustic inversion of experimental ambient-noise data from moored VLAs has been shown to produce results in agreement with active-source methods[1]. Results with simulated sequential data are shown in this summary. To generate the sequential data, the environment shown in Fig.2 (a) was input to the numerical propagation model OASES[4] for computation of the range-dependent ambient-noise field in a 32-element VLA with 0.18 m inter-element spacing. Conventional beamforming was used to estimate the BL at 8 frequencies in the range 550 Hz to 1400 Hz.

The BL data at 20 uniformly-spaced grazing angles from 14° to 90° was provided to the sequential Bayesian trans-D Monte Carlo algorithm for estimation of the PPD (details found in ref. 3). Figure 2 (b) shows the mean value of the sediment sound speed and density. Note that the estimated geoacoustic parameters and the depth (and number) of acoustic interfaces closely resemble the true profiles.

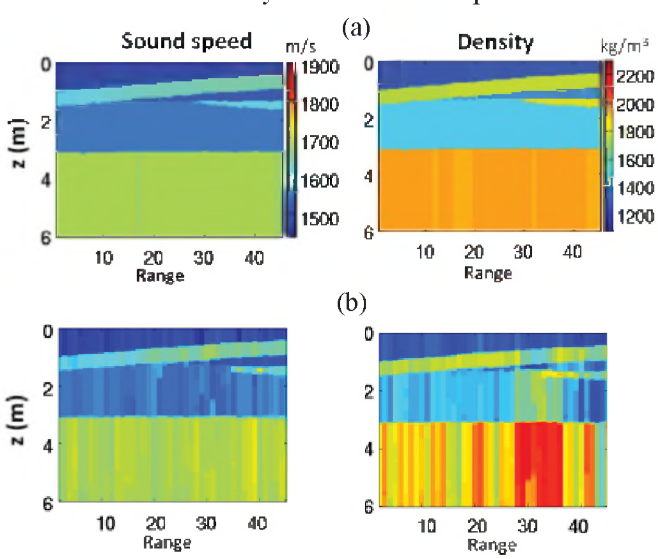


Figure. 2. (a) True seabed environment input to OASES to generate simulated data. (b) Mean sediment sound speed and density from the estimated PPD via inversion.

Figure 3 shows the marginal PPDs for sound speed and density at range=25 (arbitrary units) in Fig.2 (b), from which parameter uncertainties can be quantified as the spread of the PPD support around the true values (shown as dashed lines).

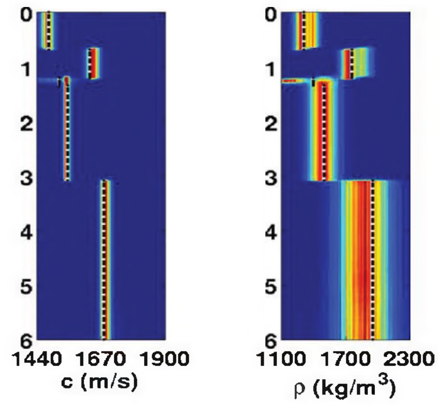


Figure. 3. Marginal probability profiles for the sediment sound speed and density at range=25 in Fig. 2. True profiles are indicated by dashed lines.

4. CONCLUSION

Geoacoustic inversion of ambient-noise data has great potential as a remote sensing technique. It has advantages over traditional active-source methods in terms of reduced power and ship time requirements, simplified and unobtrusive surveys, and zero environmental impact. Ongoing research with sequential datasets suggests the possibility of obtaining true-depth (rather than travel-time based) images of the seabed layering structure, along with estimates of parameter uncertainties, required for a meaningful interpretation of range-dependent variability of geoacoustic parameters.

REFERENCES

1. J. E. Quijano, S. E. Dosso, J. Dettmer, L. M. Zurk, M. Siderius, and C. H. Harrison, "Bayesian geoacoustic inversion using wind-driven ambient noise", J. Acoust. Soc. Am. 131, 2658-2667 (2012).
2. J. Dettmer, S. E. Dosso, and C. W. Holland, "Trans-dimensional geoacoustic inversion", J. Acoust. Soc. Am. 128, 3393–3405 (2010).
3. J. Dettmer, S. E. Dosso, and C. W. Holland, "Sequential trans-dimensional Monte carlo for range-dependent geoacoustic inversion", J. Acoust. Soc. Am. 129, 1794–1806 (2011).
4. H. Schmidt, OASES version 3.1 User guide and reference manual; http://acoustics.mit.edu/faculty/henrik/oases.html. Last accessed 07/30/2012.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of the Office of Naval Research postdoctoral fellowship and the Ocean Acoustics Program (ONR-OA Code 3211).

SIMULATION STUDY OF JOINT TRANS-DIMENSTIONAL BAYESIAN INVERSION OF SCATTERING AND REFLECTION DATA

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

1. INTRODUCTION

Ocean acoustic reverberation modelling and sonar performance prediction in shallow waters require good estimates of seabed geoacoustic parameters and scattering parameters defining seafloor roughness. Direct measurements of these parameters are time consuming and expensive, and it may be advantageous to estimate *in-situ* seabed parameters based on indirect measurements, e.g., as the solution to an inverse problem. This paper develops a joint trans-dimensional (trans-D) Bayesian inversion approach which is applied to synthetic seabed scattering and reflection data with the goal of determining the ability of such data to resolve geoacoustic and scattering parameters.

2. METHOD

Bayesian inversion requires specifying the posterior probability density (PPD, product of the prior distribution and likelihood function), and a method to sample PPD. The remainder of this section gives an overview of the creation of the synthetic data and of the Bayesian inversion approach as it is applied here; more general and complete descriptions of Bayesian inversion are given elsewhere[1,2].

2.1 Synthetic Data

The scattering data represent monotonic back-scatter strengths generated from Jackson's perturbation-theory scattering model[3] assuming a two-dimensional seabed roughness power spectrum (the von Karman spectrum) [4]

$$W(\mathbf{K}) = \frac{w_2}{\left(|\mathbf{K}|^2 + K_0^2\right)^{\gamma/2}}, \tag{1}$$

where w_2 is the spectral strength, K_0 is the spectral cut off, and γ is the spectral exponent. The vector $\mathbf{K}$ is the transverse component of the incident wave vector (i.e. $|\mathbf{K}|=k_0\cos(\theta)$, where θ is the incident grazing angle and k_0 is the wave number in the ocean). Data were created over an angular range of 6–24° at six frequencies (600 Hz, 900 Hz, 1200 Hz, 1800 Hz, 2400 Hz, and 3600 Hz). Gaussian-disturbed errors are added to the data which are correlated over angle but independent between frequencies.

The reflection data correspond to spherical reflection coefficients calculated recurvsively over a layered seabed[5]. Reflection data were generated over an angular range of 20–85° and averaged over 1/3-octave bands for six centre frequencies (630 Hz, 800 Hz, 1000 Hz, 1500 Hz, 2500 Hz, and 3600 Hz). The angular spacing between data points is non-uniform and increases with angle from approximately 0.3–10°, which is typical of experiment measurement techniques[6]. Data errors are positively correlated (over angle) with correlation between to data points decreasing exponentially with angular distance.

2.2 Posterior Probability Density

The PPD contains all information considered in Bayesian inversion, and can be expressed using Bayes rule as

$$\mathrm{PPD}(\mathbf{m}_j \mid \mathbf{d}) = \frac{\mathcal{L}(\mathbf{d} \mid \mathbf{m}_j)\pi(\mathbf{m}_j)}{\mathcal{Z}}, \tag{2}$$

where $\mathcal{L}$ is the likelihood function, π is the prior distribution, and $\mathcal{Z}$ the evidence. The vectors $\mathbf{d}$ and $\mathbf{m}_j$ represent the data and the model parameters, where subscript j indicates the model dimension which is variable in trans-D inversion. Here the number of sediment layers is treated as unknown and sampled in the inversion.

The prior distribution represents information known about the model before the introduction of the data. The prior distribution used here is

$$\pi(\mathbf{m}_j) \propto \frac{(j-1)!}{z_B^{(j-1)}} \frac{H^j}{(\Delta c \Delta\rho \Delta\alpha)^j}, \tag{3}$$

where Δc, $\Delta\rho$, and $\Delta\alpha$ are the range of possible sound velocity, density, and attenuation values for a given sediment layer; H is a normalizing constant that accounts for the assumed correlation between geoacoustic parameters; z_B is depth of the basement. The prior distribution can be formulated proportionally since Bayesian inversion considers only the ratio of prior distributions. The remaining parameters, such as w_2, K_0, and γ have uniform priors and are absorbed by the proportionality.

The likelihood function describes both the physics (forward model) and the data error statistics assumed in the inversion. As this work describes a simulation study, the same forward model and data error statistics used to create simulated data are assumed in the inversion.

2.3 Sampling Scheme

It is not in general possible to interpret the PPD in an analytic manner. Thus in Bayesian inversion it is common to sample the PPD using Markov-chain Monte Carlo algorithms; here, the reversible-jump Markov-chain Monte Carlo (rjMCMC) algorithim[7] with parallel tempering[8] is used. The sampled PPD is then interpreted in terms of its

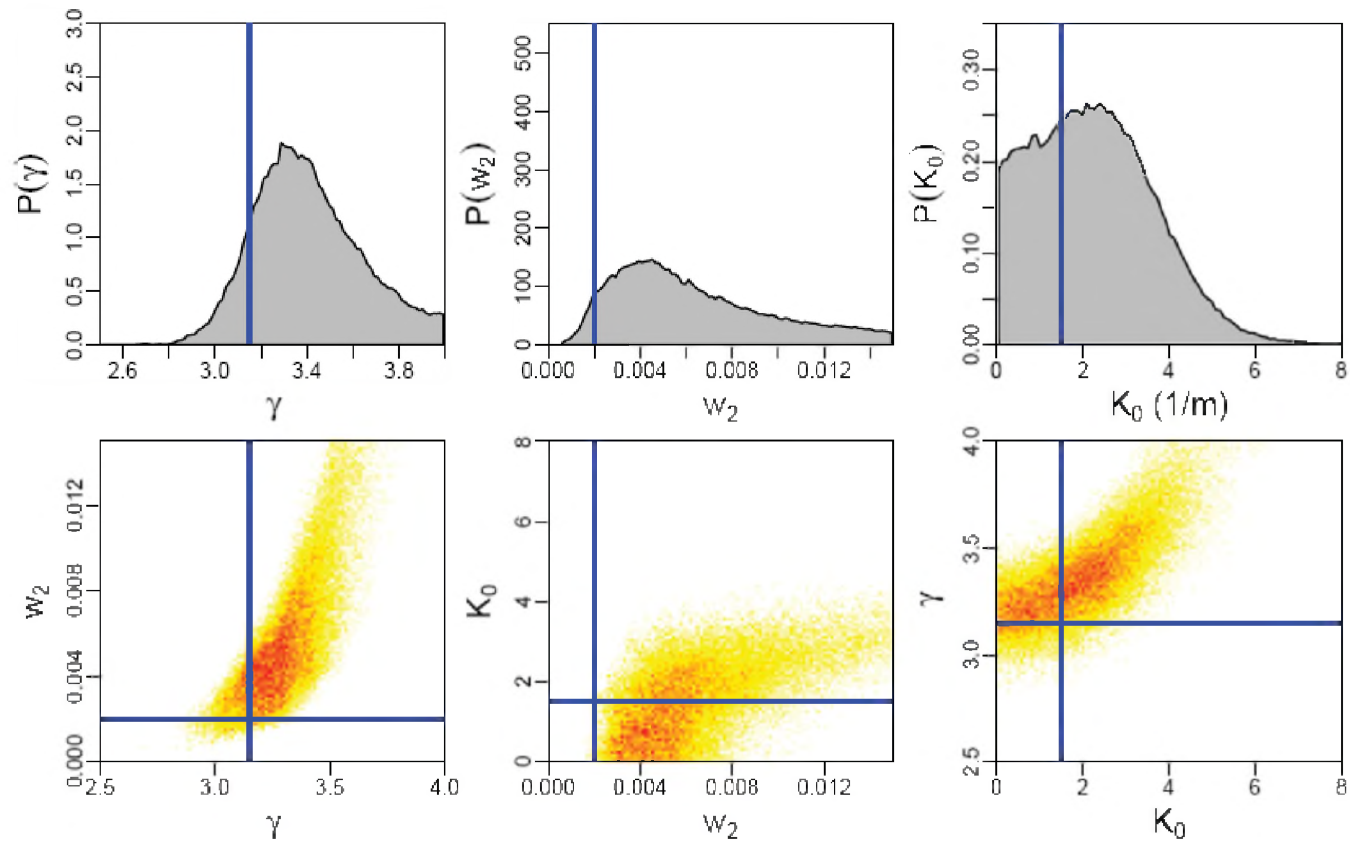


FIG. 1. Top: one-dimensional marginal posterior distributions of the scattering parameters. Bottom: two-dimensional marginal posterior distributions of the scattering parameters. The horizontal and vertical lines indicate the true value.

moments, parameter uncertainties (variances, marginal distributions, credibility intervals), and parameter inter-relationships (correlations and joint marginal distributions).

To adequately approximate the PPD 600,000 models were sampled from it using the rjMCMC algorithm. These are thinned by one third to reduce sample correlation; only these remaining samples are considered here.

3. RESULTS

The one- and two-dimensional marginal distributions of the scattering parameters (w_2, K_0, and γ) for the inversion are shown in Fig. 1. The parameter distributions are centred near the true values, and the uncertainties indicate a useful level of resolution of the roughness spectrum. Geoacoustic parameters are also well resolved, but are not shown here due to space constraints. This work indicates that Bayesian inversion is an adequate way of determining the acoustic scattering properties of marine sediment.

ACKNOWLEDGEMENTS

This work is supported by the Office of Naval Research.

REFERENCES

[1] D. G. T. Denison, C. C. Holmes, B. K. Mallick, and A. F. M. Smith (2002). *Bayesian Methods for Nonlinear Classication and Regression*, 1-277 (Wiley, New York).
[2] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin (2004). *Bayesian Data Analysis*, 73-100 (Chapman, New York).
[3] J. E. Moe and D. R. Jackson (1994). "First-order perterbation solution for rough surface scattering cross section including the effects of gradient", J. Acoust. Soc. Am. 96, 1748-1754.
[4] D. R. Jackson, and M. D. Richardson (2007). *High-frequancy seafloor acoustics*, 178-200, 1st edition (New York).
[5] L. M. Brekhovskikh and O. A. Godin (1992). *Acoustics of Layered Media II: Point source and bounded beams*, Wave Phenomena, 1st edition (Springer, New York).
[6] C. W. Holland and J. Osler (2000). "High-resolution geoacoustic inversion in shallow water: A joint time- and frequency-domain technique", J. Acoust. Soc. Am.. 107, 1263-1279.
[7] P. J. Green (1995). "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination", Biometrika, 82, 711-732.
[8] A. Jasra, D. A. Stephens, and C. Holmes (2007). "Population-based reversible jump Markov chain Monte Carlo", Biometrika, 94, 787-807.

AUV LOCALIZATION IN AN UNDERWATER ACOUSTIC POSITIONING SYSTEM

Dugald Thomson and Stan E. Dosso
School of Earth & Ocean Science, University of Victoria, Victoria, BC, Canada
dugalthomson@gmail.com

1. INTRODUCTION

Precise positioning of autonomous underwater vehicles (AUVs) is an important problem for the ocean science community as it attempts to extend its reach to ever-greater depths. Terrestrial Global Positioning Systems are of little use for an underwater target as the high-frequency/low-power signals they employ are unable to penetrate beyond the surface layers of the ocean due to reflection and absorption by seawater. The Integrated Acoustic System (IAS) being designed at the University of Victoria aims to overcome this obstacle by developing a high-precision underwater acoustic positioning system, similar in operation to a commercial long baseline positioning unit, that receives power and transmits data via the Victoria Experimental Network Under the Sea (VENUS) infrastructure, a cutting-edge underwater cabled observatory. The IAS will be capable of localizing a target within the IAS range to a sufficient accuracy for use as ground truth for testing onboard navigation systems.

The IAS, located in a range which covers an area of approximately 1.5 km by 1.5 km, is comprised of five 3-m high hydrophone towers mounted on the seabed at depths of 60 m to 130 m, located in the four corners of the range plus one in the centre, as depicted in Fig 3. AUVs operating in the range are outfitted with a transducer that periodically emits an acoustic pulse. Pulse time difference of arrivals (TDOA) are used to localize the AUV using a method known as multilateration, and the IAS employs the IEEE 1588 precision timing protocol (PTP), allowing a precision of +/- 10 μs in clock timing (Lentz & Lecroart, 2009), a substantial improvement over the milliseconds-order accuracy offered by less precise network timing protocols typically employed in a data communication network.

2. METHOD

The analysis summarized here consists of two distinct studies, involving a ray-based Bayesian inversion algorithm developed to estimate AUV position and uncertainties. The first study estimates the non-linear localization accuracy for a target located within the range, based on a Monte Carlo method of estimating uncertainties of the source-location in x, y, and z. The second study maps the target positional uncertainty as a function of position within the range by estimating the linearized posterior uncertainties of the source-location in x, y, and z, as well as the lateral uncertainty in $r = [x^2+y^2]^{1/2}$.

2.1 Linearized Uncertainty Model

Each simulation is executed for three distinct test cases which were developed to simulate target positions representing a favourable source-receiver geometry (test case 1), a poor source-receiver geometry (test case 2), and an average over geometries in terms of a series of random source positions drawn from within the range (test case 3). The simulation scenarios were developed to investigate: (1) modelling transmission paths accounting for refraction due to a depth-varying SSP instead of using straight rays through a constant sound-speed approximation, (2) inverting for a potential sound-speed bias in the measured profile, (3) accounting for errors in hydrophone position by including these positions as unknowns in the inversion, and (4) applying path correction factors to account for lateral variability in the sound-speed profile. Each scenario is studied using a Monte Carlo method in which a large number of noisy data sets are inverted to derive statistical measures to quantify the various effects. In addition, inversions for scenarios 2–4 are carried out for a single source transmission, as well as for 20 source transmissions, to determine the degree to which the over-determined inverse problem improves localization accuracy. Linearization error is computed by comparing the results of the non-linear Monte Carlo analysis to the linear uncertainty estimates of the model covariance matrix.

2.2 Linearized Uncertainty Model

Once linearization errors are verified as described above, the posterior uncertainties of the source-location in x, y, z and r are calculated. Since the source-location uncertainty varies with source location, uncertainties are calculated for the source at each point within a grid of positions over the area of the test bed. At each grid point, the source-location uncertainties are estimated using a linearized Bayesian approach that includes the effects of arrival-time errors as well as uncertainties in hydrophone locations and sound speed. A complete description of these methods can be found in Dosso & Ebbeson (2006).

3. RESULTS

Monte Carlo analysis of the scenarios described in Sec. 2.1 were carried out, and the results from scenario 1 are shown in Fig. 1, while the results from a test that combines the factors described in scenarios 2–4 into a single inversion are shown in Fig. 2.

Comparing results when inversions are based on straight rays versus refracted (curving) rays, Fig. 1 shows systematic

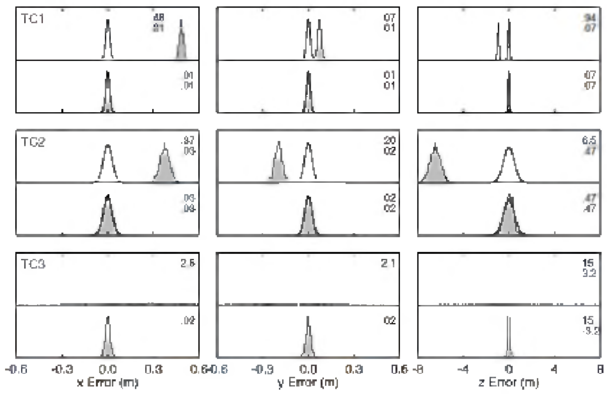


Figure 1. Histogram of errors for x, y, and z (left, centre, and right panels, respectively) for test cases 1, 2, and 3 when inversions are based on straight rays (top distribution in each panel) and refracted rays (bottom distribution). RMS errors and standard deviations (except for TC3) in metres are given in each panel.

error in the straight-ray model that biases the target position away from the true position. With the curving-ray model, however, results show excellent agreement with the true locations, and linearization error, shown in the difference between histogram results and the linearized uncertainty estimate from the continuous line in TCs 1 and 2, is seen to be small.

Figure 2 shows that localization results are substantially improved when the sources of error described in scenarios 2–4 are inverted for as parameters in the model, and linearization errors are much reduced. Over-determined inverse problems, where data from 20 source transmissions are inverted rather than from a single transmission, show improved positional accuracy, particularly in *z*.

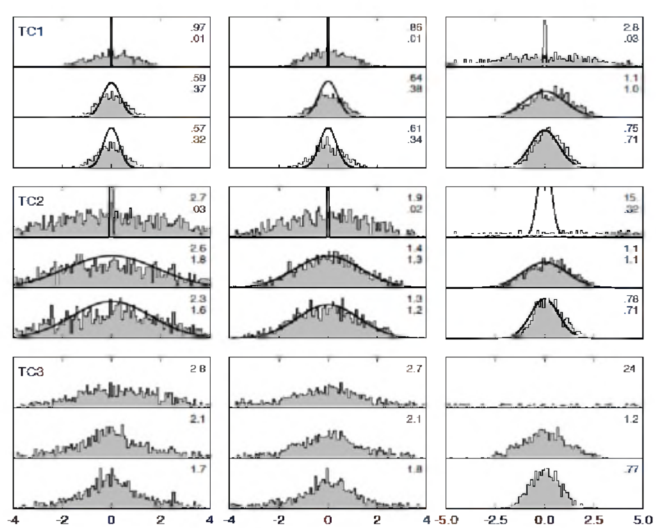


Figure 2. Histogram of errors for *x*, *y*, and *z* for test cases 1, 2, and 3 when sound speed bias, hydrophone positions, and path correction factors are not included as inversion parameters (top distributions), when these factors are included as inversion parameters for 1 source transmission (middle distributions), and for 20 source transmissions (bottom distributions). RMS errors and standard deviations (except for TC3) are given in metres.

Since linearization errors are shown to be small for the test cases investigated above, linearized uncertainty estimates can now be used to estimate target positional uncertainty for locations throughout the range (Fig. 3) with a high level of confidence. Figure 3 shows that the lateral component of uncertainty is lowest for target locations towards the centre of the range, while the vertical uncertainty is lowest when the target is located above a hydrophone.

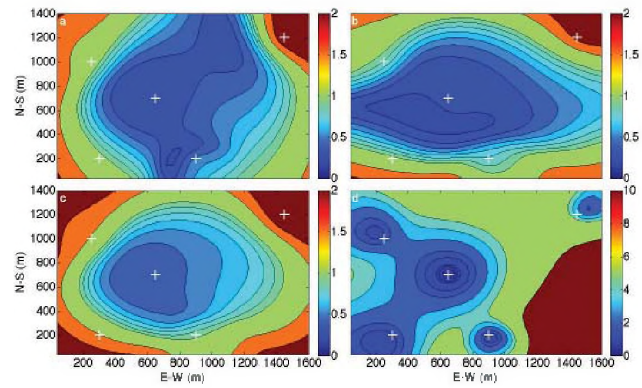


Figure 3: Linearized localization u over the range. Panels (a)-(d) show absolute errors in *x*, *y*, *r*, and *z*, respectively, for a source at 10-m depth. Contours represent uncertainty in metres. Hydrophone locations are depicted as white crosses.

4. DISCUSSION

The modeling studies described in this paper served as simulation tests for a ray-based Bayesian inversion algorithm developed for an acoustic positioning system for AUV localization in the IAS test range, which should become operational some time in 2012-13 within the VENUS infrastructure, and will serve as an functional 'ground truth' test bed for AUV operations.

REFERENCES

Lentz, S. & Lecroart, A. (2009). Precision timing in the NEPTUNE Canada network. *OCEANS 2009-EUROPE*, 1-5.

Dosso, S.E. & Ebbeson, G. (2006). Array element localization accuracy and survey design. *Canadian Acoustics*, 34(4), 1-11.

Thomson, D., Dosso, S. E., Gamroth, E., & Bradley, C. (2012). Modelling Uncertainty in an Underwater Acoustic Positioning System. *Canadian Acoustics*, *40*(1), 30–37.

AUTHOR NOTE

This work was conducted while Thomson was a student at the University of Victoria (2012).

UNDERWATER SOUND MEASUREMENTS OF HIGH FREQUENCY SONARS USING A SEABED-MOUNTED RECORDER

Graham Warner[1], Andrew McCrodan[2], Jeff MacDonnell[3], and Eric Lumsden[4]
[1,2]JASCO Applied Sciences, 2305-4464 Markham St., BC, Canada, V8Z 7X8 [1]graham.warner@jasco.com
[3,4]JASCO Applied Sciences, 202-32 Troop Ave., NS, Canada, B3B 1Z1

1. INTRODUCTION

There is concern among regulatory agencies that sonars operating below 180 kHz (upper frequency limit for high frequency marine mammal listeners) could produce sounds causing auditory injury or induce behavioural effects, and that sonars operating above 180 kHz could leak energy into the audible range. The regulatory permit for Statoil USA E&P Inc.'s 2011 marine survey program in the Alaskan Chukchi Sea required underwater sound measurements of high and low frequency sound sources used in the program. We used a seabed-mounted Autonomous Multichannel Acoustic Recorder (AMAR, JASCO Applied Sciences) to accurately measure high frequency sonar sound levels (Warner and McCrodan, 2011).

2. METHODS

2.1 Measurement Equipment

Underwater sound level measurements were made using a seabed-mounted Autonomous Multichannel Acoustic Recorder (AMAR, JASCO Applied Sciences) with a Reson TC4014 hydrophone (nominal sensitivity -186 dB re 1 V/uPa). The AMAR recorded 16 bit samples at 687.5 kHz for 27 hours. The end-to-end sensitivity of the recording system was calculated by adding the factory TC4014 frequency-dependent calibration sensitivity to the calibrated digitization gain of the AMAR.

The AMAR mooring consisted of a 120 lbs single chain link for ballast, an acoustic release, and a float collar surrounding the AMAR. The AMAR was deployed twice, each time in 37 m of water. On deployment, the ballast sank the AMAR with the hydrophone approximately 2 m above the seafloor. On recovery, the acoustic release was triggered to release the mooring from the ballast and the equipment was retrieved once it floated to the surface. The ballast was then retrieved using a grapple.

2.2 Measurement Procedure

Underwater sound measurements were made at Statoil's exploration lease area in the Chukchi Sea (Figure 1, right). The AMAR was deployed at approximately 71.7N, 164.3W in 37 m water. Three sonars were measured: a towed GeoAcoustics 159D side-scan sonar (112.5 kHz), a hull-mounted Kongsberg EM2040 multibeam sonar (205 kHz), and a hull-mounted Kongsberg EA600 single-beam sonar (200 kHz). The sonars were taken past the AMAR along several 1 km long parallel track lines. The lines were offset at horizontal distances between 0 and 400 m from the AMAR (Figure 2).

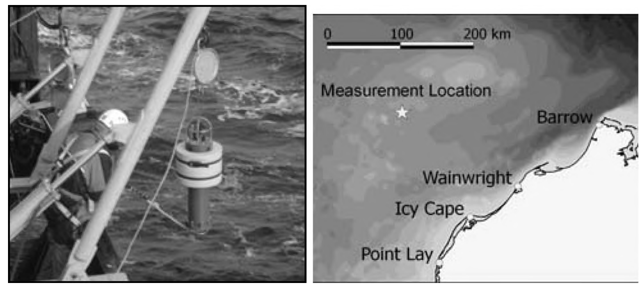


Figure. 1. Photograph of an AMAR with float collar being deployed (left) and map of the measurement location off northwest Alaska (right).

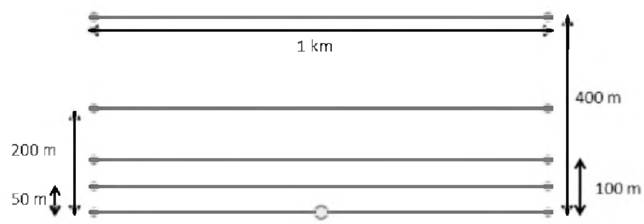


Figure.2. Sonar track lines relative to the AMAR (hollow circle).

2.3 Data Analysis

For each sonar, the acoustic data were band-pass filtered around the operating frequency. Sonar pulse sound levels were calculated using the filtered data and in-beam pulse levels were plotted versus slant range. An empirical equation of the form RL=SL-ALog(r) was fit to the 90% rms SPL data and distances to sound level thresholds were estimated. Sound levels from in- and out-of-beam pulses were then back-propagated using spherical spreading and absorption loss at the centre frequency (Francois and Garrison, 1982). Source levels were plotted versus horizontal or vertical angle.

3. RESULTS

Sonar pulses were detected on all five track lines for the side-scan sonar, lines offset 0, 50, and 100 m for the multibeam sonar, and lines offset 0 and 50 m for the single-beam sonar. The empirical fit equation to the 90% rms SPL in-beam pulse levels was RL=229.3-29.5Log(r) for the side-scan sonar, RL=189.0-27.3Log(r) for the multibeam sonar, and RL=287.6-89.3Log(r) for the single-beam sonar. Table 1 lists the distances to threshold levels for the three sonars.

Table. 1. Threshold distances as determined from the empirical fit equations of the 90% rms SPL in-beam pulse levels.

90% rms SPL threshold (dB re 1 uPa)	Side-scan distance (m)	Multibeam distance (m)	Single-beam distance (m)
190	22		12*
180	47		16*
170	100		21*
160	230		27*
150	490*	27	35*
140	1100*	62	45
130	2400*	140*	58
120	5100*	330*	75

*Extrapolated beyond measurement range.

Received levels were back-propagated to estimate source level versus angle off broadside for the side-scan (Figure 3) and multibeam (Figure 4) sonars and versus angle off vertical for the single-beam sonar (Figure 5).

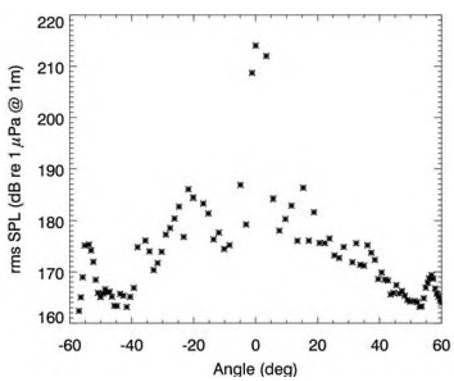


Figure. 3. Source levels as a function of azimuthal angle off broadside for the side-scan sonar.

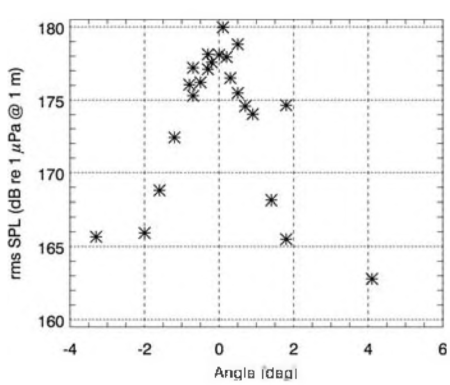


Figure. 4. Source levels as a function of azimuthal angle off broadside for the multibeam sonar.

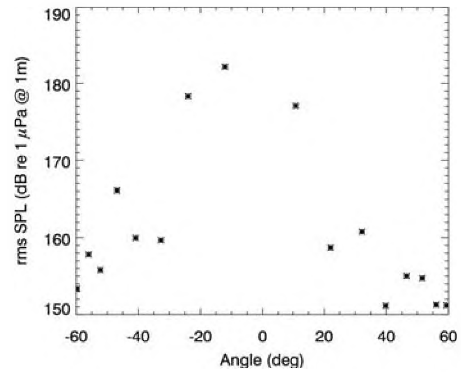


Figure. 5. Source levels as a function of angle off vertical for the single-beam sonar.

4. DISCUSSION

Statoil's regulatory permit for their 2011 Alaskan Chukchi Sea survey program required acoustic measurements of high frequency sonars. We measured underwater sound levels from side-scan, multibeam, and single-beam sonars using a high sample rate seabed-mounted AMAR. The measurement system allowed good control of source-receiver geometry compared to previous measurements using ship-based recorders (e.g. Chorney et al., 2011) and is suitable for both high and low frequency sound source measurements.

REFERENCES

Chorney, N.E., G. Warner, J. MacDonnell, A. McCrodan, T. Deveau, C. McPherson, C. O'Neill, D. Hannay, and B. Rideout. 2011. Underwater Sound Measurements. (Chapter 3) In: Reiser, C.M, D.W. Funk, R. Rodrigues, and D. Hannay. (eds.) 2011. Marine mammal monitoring and mitigation during marine geophysical surveys by Shell Offshore, Inc. in the Alaskan Chukchi and Beaufort seas, July–October 2010: 90-day report. LGL Rep. P1171E–1. Rep. from LGL Alaska Research Associates Inc., Anchorage, AK, and JASCO Applied Sciences, Victoria, BC for Shell Offshore Inc, Houston, TX, Nat. Mar. Fish. Serv., Silver Spring, MD, and U.S. Fish and Wild. Serv., Anchorage, AK. 240 pp, plus appendices.

Francois R.E. and G.R. Garrison. 1982a. Sound absorption based on ocean measurements: Part I: Pure water and magnesium sulfate contributions. JASA 72:896-907.

Francois R.E. and G.R. Garrison. 1982b. Sound absorption based on ocean measurements: Part II: Boric acid contribution and equation for total absorption. JASA 72:1879-1890.

Miller, J.H., A.E. Bowles, B.L. Southall, R.L. Gentry, W.T. Ellison, J.J. Finneran, C.R. Greene Jr., D. Kastak, D.R. Ketten, P.L. Tyack, P.E. Nachtigall, W.J.Richardson, and J.A. Thomas. 2005. Strategies for weighting exposure in the development of acoustic criteria for marine mammals. J. Acoust. Soc. Am. 118:2019 (Abstract).

Warner, G., and A. McCrodan. (2011). Underwater Sound Measurements. (Chapter 3) In: Hartin, K.G., L.N. Bisson, S.A. Case, D.S. Ireland, and D. Hannay. (eds.) 2011. Marine mammal monitoring and mitigation during site clearance and geotechnical surveys by Statoil USA E&P Inc. in the Chukchi Sea, August-October 2011: 90-day report. LGL Rep. P1193. Rep. from LGL Alaska Research Associates Inc., LGL Ltd., and JASCO Research Ltd. For Statoil USA E&P Inc., Nat. Mar. Fish. Serv., and U.S. Fish and Wild. Serv. 202 pp. plus appendices.

ALTERNATIVE SOURCE FOR MARINE GEOPHYSICAL EXPLORATION

Paul Yeatman[1], Sara Stout-Grandy[2], and Bruce Armstrong[2]
[1]GeoSpectrum technologies Inc., 10 Akerley Boulevard, Dartmouth, NS, Canada, B3B 1J4, paul.yeatman@geospectrum.ca
[2]GeoSpectrum technologies Inc., 10 Akerley Boulevard, Dartmouth, NS, Canada, B3B 1J4

1. INTRODUCTION

High power acoustic sources, such as the air guns used by the marine geophysical industry, are coming under increasing scrutiny by environmentalists for their potentially harmful impact on marine life, and in particular, mammals[1]. As a result, the marine geophysical industry is attempting to develop alternatives to replace traditional air guns and other impulsive broadband sources[2]. This paper discusses the fundamental physics issues associated with low frequency sources, particularly when they are operated near the water surface, and why it is so difficult to achieve even a small fraction of the power that an air gun produces with electro-acoustic or other marine vibroseis technologies. Our investigations show that GeoSpectrum's MPS technology is a strong, environmentally friendly, candidate for geophysical site survey exploration and we discuss what we have learned, as supported by Encana through the Deep Panuke Fund.

2. PHYSICAL CONSTRAINTS

When operating in shallow water, or near the water surface, it is very difficult to produce acoustic energy that is sustained. Near the water surface can be defined as within a fraction of a wave length. As can be seen in Figure 1 below, the power level achieved at 5m depth is about the same as in the free field at 25Hz, but at lower frequency it is attenuated and at higher frequencies it is in fact boosted by the surface reflection. This is because, fundamentally, the air/water boundary cannot sustain pressure so the reflected acoustic wave is 180° out of phase with the incoming acoustic wave[3]. The reflected wave interferes with the incoming wave, which, at shallow depths and low frequencies, results in destructive interference.

With an air gun operating at 5m depth, it would appear to be attenuated by well over 10dB, but this is not the case. The chart reflects steady state signaling. An air gun is impulsive and the detrimental reflection does not occur immediately. The 5m depth gives a reflected path of 10m, or 6.7ms. With a primary pulse width of 20ms, only 2/3 of the pulse is attenuated. For a coherent source, the time domain advantages are largely lost due to the greater pulse length required.

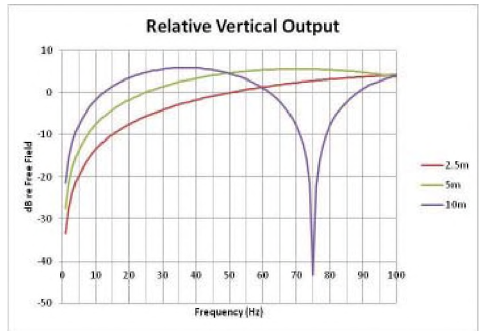


Figure 1 Surface Reflection Attenuation

A further constraint is the size and weight consideration for power at low frequency. The power generated by an acoustic source is proportional to the square of the volume velocity it achieves[4].

To put this in perspective, a piston displacement that generates 1 W at 1 kHz would produce only 10^{-4} W at 10 Hz. Thus, low frequency projectors just either produce large displacements or have large radiating areas, or both. No matter how you look at it, a powerful low frequency projector is large.

3. MPS

The MPS[5] takes advantage of the interactions of closely spaced transducers to increase the radiation impedance and thus increase the relative power transmission at low frequencies. The MPS uses a very large active surface area to spread the energy and stresses seen by the source and also minimize the magnitude of motion required to achieve the desired sound pressure level. This results in a relatively robust system, compared to devices with moving parts that require seals and other regular service parts.

Figure 2 MPS

As can be seen in Figure 2 above, the MPS is built like a set of baffles, and so for a cylindrical form, the radiating area is in fact the area of the end of the cylinder times double the number of elements in the array. For a 5" diameter cylinder with 20 elements spaced at 1" we have 785 in^2 of active area. If this were a moving coil with a 5" face, the active area would be less than 20 in^2 and if this were a radiating cylinder, the active area would be in the order of 345 in^2.

The operating frequency of a transducer greatly affects the transducer size and effectiveness. We have already noted that as we go down in frequency we have to go up in displacement to maintain the same volume velocity of water. To reduce the operating frequency, we need to make the transducer less stiff as well. With air backed transducers, such as flex-tensionals, we can make the operating frequency lower by reducing the mechanical stiffness of the flexing member. This will impact the robustness of the transducer relative to static pressure, and if we go low enough in frequency, we will need to add internal pressure (pressure compensation) to avoid damaging the transducer. As noted in Section 2, the depth of operation will impact performance, with the general note that the lower frequency you wish to operate, the deeper you need to operate. This is in conflict with the fact that the lower frequency transducers will are less able to withstand static pressure. As we head toward 10Hz, the need for pressure compensation becomes unavoidable. To achieve air gun equivalent power cannot be achieved (with existing technology) without an unmanageably large source.

An MPS, using flex-tensional benders, will be more compact and more robust than a transducer producing the same power, using ring shells, hydraulic transducers, moving coils, etc. The MPS provides a very large operating area in a relatively small space and does not require moving seals. The simplicity of the bender element makes it a relatively low cost solution with minimal parts to go wrong. Because the MPS is made up of a large number of identical elements, there is also a clear means of graceful degradation, as no one part will lead to catastrophic system failure.

4. COHERENT ADVANTAGES

When operating with a coherent source, you know the frequency, amplitude and phase content of your source signal and therefore can correlate the receive signals to the source signals in more than one dimension as it were. The received data can also be processed in the time domain, as it can be continuous or longer in duration than an impulsive source. Source signal control allows the operator to avoid destructive or interfering frequencies and also generate sequenced or coded signals for even greater ability to increase S/N ratios and thus reduce the needed power levels. It is also possible to avoid frequencies that pose environmental concerns or even to tailor the operational frequency content to each specific area.

For oil and gas exploration, no operational alternative to air guns exist, and the goal is clearly a lofty one. However, for other purposes, where higher frequencies of operation are possible, because depth of penetration requirements is less, the MPS provides an existing, viable, environmentally friendly solution. This could be for shallow hazard surveys, well monitoring, etc. The operation can be higher in frequency than marine seismic, but still below or above the frequencies of concern for marine mammals or those marine mammals of concern in the area of operation.

REFERENCES

1 Weilgart, Dr. L., et al, (2010). Report of the Workshop on Alternative Technologies to Seismic Airgun Surveys for Oil and Gas Exploration and their Potential for Reducing Impacts on Marine Mammals.
2 Johnson, G., Nichols, S. (2009). Marine Vibrator Acoustic Source Specification.
3 Chapman, D. M. F. (2004). You Can't Get There From Here: Shallow Water Sound Propagation and Whale Localisation.
4 Kinsler, L.E., Frey, A. R., et al, (1976). Fundamentals of Acoustics.
5 Armstrong, Dr. B.A., Yeatman, P.Y., (2005). Constructive Use of Acoustic interactions in a Multi Element Projector Array.

Management of Amplified Sound Generated by Band-Shell Performances in a Dynamic Urban Setting

Ellen Buchan
Ellen Buchan, P.Eng. Alberta Infrastructure, 6950-113 st, Edmonton, AB T6H 5V7, ellen.buchan@gov.ab.ca

1. INTRODUCTION

Over the years, residential areas surrounding the Edmonton Legislature Grounds' Band Shell have grown and residents have raised concerns about the sound levels generated from Band-shell performances. In response to the community's concerns, a project was started by Alberta Infrastructure to develop a system and procedure for monitoring and controlling these amplified performances. The intent was to provide Sound System Operators with a clear indicator of their sound levels so they could voluntarily take action to limit it. In addition, the system would record and maintain accurate records of event noise for future reference by the Legislature Grounds Staff.

The "first generation" Legislature Sound Monitoring System was introduced in 1997 at the Alberta Legislature Grounds Band Shell Site. It has continued to evolve over the last 15 years, expanding, incorporating new technology and functionality; shaping today's current system, which is the subject of this paper.

2. METHODS AND RESULTS

2.1 Primary System Functionality

The monitoring system relays real-time feedback to the event's Sound System Operator and Legislature Grounds Staff, indicating when the levels are approaching or exceeding the noise limits. To accomplish this, a warning light is triggered to indicate that the levels are approaching the limit and a horn is triggered to indicate that the limit has been exceeded.

The monitoring system also records the sound levels during an event. Continuous levels and corresponding statistical data are displayed and saved to hard drive and USB stick. The system user can view this data for future reference.

2.2 Site Plan

Figure 1 displays the Legislatures Grounds in relation to the surrounding community, Band Shell Site and Monitoring System Components. A microphone at measurement station #1 is encased in a custom designed weather protective housing positioned on top of a high pole about 30m north of the stage. A second microphone at measurement station #2 is enclosed in a pre-manufactured weather protective housing on the roof of the Legislature Building, about 200m northwest of the Band Shell. The locations of these measurement stations were chosen to be unobtrusive and intercept the line of site from the stage to the affected condominiums/apartments, northeast and northwest of the stage. The warning light is clearly visible from the band shell stage. The horn is located on the Band Shell, in very close proximity to the performers. The core data acquisition electronics and accessories are housed in an equipment cabinet mounted on the wall of the nearby washrooms.

Figure.1. Alberta Legislature Grounds Band Shell Sound Monitoring System Site Plan

2.3 Primary System Components

A variety of hardware pieces and software platforms have been evaluated and often implemented over the life of the Monitoring System. The current system's primary components are described as follows:

Both measurement station encasements' house a microphone that is protected from moisture, birds, and insects. Pre-amplifiers condition both microphone level input signals for line level data acquisition. An encrypted wireless transmitter/receiver pair facilitates the connectivity between measurement station #2 and the core data acquisition system. An equipment cabinet located near the Band Shell houses a desktop computer complete with a National Instruments Data Acquisition Card and output relays. This desktop computer runs National Instruments Sound and Vibration Toolkit Software and Microsoft Remote Desktop. Finally, a Wireless Air-Card connected to the desktop computer allows the Legislature Grounds staff to access the system remotely.

2.4 Maximum Permissible Daytime Sound Levels

Measurement Station #1

Edmonton Bylaw 7255 states that 65 dB(A) is the residential property line limits of allowable noise during the day (7am to 10pm) [1]. However, in development of the Monitoring System, a noise study, limited to Station #1 (30m from the Band Shell) and the Northeast Community (200 to 400m from the Band Shell), revealed that Band Shell music falling below the daytime noise bylaw limit was still definitely audible at the residences (and very likely annoying to the occupants). Therefore, basic bylaw compliance did not seem like a sufficient restriction.

Further investigation detected existing background noise levels in the northeast community ranging from 47 dB(A) to 56 dB(A). The maximum permissible level allowed at the site was selected so the resulting maximum level heard in the residences was more restrictive then the bylaws and equivalent to the ambient noise range maximum. The maximum permissible daytime sound limit was set to 75 dB(A) at measurement station #1, as a compromise between city bylaws and ambient levels.

Measurement Station #2

In 2006 noise complaints were issued by residents of the growing community located to the northwest of the Band Shell site. At this time, only Measurement Station #1 was in use. It was concluded that the system was not calibrated to keep events from disturbing new local residents to the northwest due primarily to the sound system directivity. It was reasonable to predict the SPL correlation between the northwest community (440m from the stage) and Measurement Station #2 (about 200m from the stage). It was also reasonable to assume that the range of background noise levels in the community northwest of the band shell was comparable to those measured in the noise study (they are similar urban residential communities). Measurement Station #2 was installed and intergraded into the Monitoring System and its maximum permissible daytime sound limit was calculated, then set to be 59 dB(A).

2.5 Sound Source Localization to Prevent False Alarm Triggers

Due to site limitations, Station #2 had to be installed 200m from the Band Shell. At this distance, Station #2's maximum permissible daytime sound level is only about 10dB higher than the ambient level atop the Legislature Building roof. As such, this Measurement Station is susceptible to false triggers (events exceeding Station #2's maximum permissible sound level that do not originate from the Band Shell Stage).

Sound source localization can be determined by comparing receiver level-differences. This logic was incorporated into the Monitoring System's software program to minimize false triggers. The program executes a level-difference calculation between Station #1 and #2 SPL values at the instant when the Monitoring System detects an exceedance. A warning signal is only triggered if this level-difference is within reasonable tolerances of the theoretically predicted level-difference. Further, this method of verifying source location was possible because the source signal arrives at Station #2 0.5 sec after Station #1, allowing ongoing 5sec and 60 sec Leq SPL data acquired by the Monitoring System to capture meaningful level-difference criteria.

The implementation of sound source localization logic into the software program resulted in a 98% reduction in false triggers at measuring station #2 and an unanticipated, but desirable, reduction in false triggers of 100% at station #1. These results were calculated with data acquired by the Monitoring System for two 9 hour daytime periods while no performances or events were taking place at the band shell.

3. DISCUSSION

The success of the Monitoring System has been measured by a decrease in the frequency of community complaints received by Legislature Grounds Staff. This trend is encouraging; however, it would be interesting to conduct a more in depth study on the communities' satisfaction with the noise levels originating from the site. Additional feedback could benefit system development in the future.

The Legislature Staff that operate the Monitoring System have provided mostly positive feedback about its ease of use. They can independently operate the "windows based" software interface and are able to input event schedules and view the records of past event SPL data, graphs statistics and more. Sound System Operators are educated on the function of the Monitoring System prior to performances/events. Numerous operators have been observed lowering their sound levels in response to feedback from the warning signals.

Introducing an automated Sound Monitoring System to the Legislature Grounds Band Shell performance site has reduced the level of community complaints reported to the Legislature Grounds Staff and has supported the continued programming and operation of this public venue (April through October, yearly).

REFERENCES

[1]City of Edmonton, "Bylaw No.7255"

ACKNOWLEDGEMENTS

The Author would like to acknowledge Kelly Kruger, Steven Bilawchuk, Jason McCrank for their contributions to the development of this system over the last 15 years.

AN EMPIRICAL COMPARISON OF THREE AUDIO FINGERPRINTING METHODS IN MUSIC AND FEATURE-LENGTH FILM

Thanh Pham, Matthew Giamou, and Gerald Penn
Dept. of Computer Science, University of Toronto; Toronto, Ontario, Canada, L4J 7P7
{mpham,mattgiamou,gpenn}@cs.utoronto.ca

1. INTRODUCTION

A classic problem in content-based audio retrieval is to identify the title of a song in a song database given a short audio query from one of the songs. Many techniques have been proposed to solve this problem (*inter alia*, Li et al., 2008; Mapelli and Lancini, 2003). A more recent trend has been to construct audio fingerprints by applying computer vision techniques to spectrogram images from short-term Fourier transforms of the audio data, rather than from the audio data directly (Ke et al., 2005). This work, however, generally takes the form of a proof of concept that the proposed method works at all, rather than a comparison with an appropriate non-vision-based baseline.

We have chosen one commercially very successful algorithm, Shazam (Wang, 2003), and compared it against two vision-based algorithms, the original CMU algorithm (Ke et al., 2005), and also Google's Waveprint algorithm (Baluja and Covell, 2008), an improvement upon the CMU algorithm which introduced the use of wavelet features as well as a few innovative hashing techniques to improve the CMU algorithm's time efficiency. We evaluate these three on two different datasets: one of 6000 proprietary CD-quality songs with an average length of 228 seconds, ranging from 4.4 seconds to 2051 seconds, with 12000 queries of 30 seconds each, the other of the mixed audio layers (talking, music, sound effects, pauses etc.) of 223 feature-length Hollywood films, together with 6690 queries of 30 seconds each. The task is to identify the song or film, respectively, of each of the queries corresponding to the respective dataset in turn. There are no negative answers --- every query occurs in exactly one song/film.

2. DATASETS

The feature-length film dataset was transcoded and down-sampled from 48KHz to 16KHz mono-channel PCM with 256 kbps bitrates. The dataset contains all of the popular movie genres. The video data were discarded and never used. On average, each film lasts about 6960 seconds, with lengths ranging from 2118 seconds to 12444 seconds.

From each transcoded film, we sampled 30 different 30-second queries at random positions. To emulate the audio of pirated films recorded in theatres, we re-recorded the queries using an inexpensive omni-directional microphone that has a frequency response from 50Hz to 13KHz, an impedance of 650 ohms and a sensitivity of -58dB +/- 3dB at 1 KHz. All the recordings took place in the same room, which had a baseline noise level of -42db. In total, we gathered 30*223=6690 queries.

The music dataset was sampled from 450 licensed audio CDs, with variety of genres and singers. Each piece was down-sampled to 11025Hz mono-channel with 352kb/s bitrates. From each down-sampled piece, we extracted two 30-second queries from random positions, for a total of 12000 queries. Each audio query was passed through an MP3 encoder, then an MP3 decoder, to introduce noise.

3. EXPERIMENTS AND RESULTS

For the music data experiment, we used 1000 songs and their corresponding 2000 audio queries for the threshold tuning. The remaining 5000 songs and their corresponding 10000 queries were used for evaluation testing only. For the film experiment, we used 100 films and their corresponding 3000 audio queries for threshold tuning, and the remaining 123 films and 3690 queries for testing. The experiments were conducted on a machine with a single 3.0GHz Intel Xeon CPU with a 4MB cache and 16GB RAM.

A comparison of per-query execution speed is shown in Table 1. The F-measures on the two datasets are shown in Tables 2 and 3. In brief, Waveprint's optimizations for speed pay very high dividends on the music data, but not on film audio, and its optimizations for quality pay very high dividends on film audio, but not on music data. Nevertheless, Shazam handily outperforms both vision-based algorithms, in both time and quality.

Dataset	Shazam	Waveprint	CMU
Music	Is	6s	21s
Movie	4s	89s	11s

Table 1. Speed Comparison per query on testing datasets.

3.1 Discussion of Speed Performance

Waveprint spends most of its time in the full-comparison and wavelet decomposition steps. To fingerprint a short segment of an audio snippet, the algorithm needs to perform multi-level wavelet decompositions on its spectral image. For silent or short pause moments, the corresponding spectral images usually have very low energy, and thus are highly similar to each other across all films. For these moments, the algorithm must spend extra time performing full comparisons. To achieve the top speed, we did not include dynamic temporal warping (DTW) in our implementation. The authors indicate that non-DTW system accuracy was only 0.76% faster.

Algorithm	Precision	Recall	F-measure
Shazam	0.9836	0.9990	0.9913
CMU	0.9631	0.9940	0.9786
Waveprint	0.9020	0.9952	0.9463

Table 2. Comparison on 5000 songs and 10000 queries.

Unlike Waveprint, instead of performing a full comparison between 2 spectral images, the original CMU algorithm only compares two descriptors extracted from the two spectral images. This step is implemented by applying direct indexing on descriptors extracted from spectral images. To cope with noise, the CMU algorithm sacrifices speed by including an EM temporal model and extra descriptors within a hamming distance of 2 of the original descriptors.

The Shazam algorithm is the fastest among the three because of its efficient hash signature design. Each hash is associated with an identified instance, so for each query call, matched hashes of every instance in the database can be quickly extracted by looking only once in the hash table. By making the strong assumption that distorted noise is linear, the algorithm forsakes DTW and simply computes the time difference of 2 anchor points coupled from each matched hash to accumulate scores across time. This score accumulation process takes O(n log n) in our implementation with n the number of matched hashes.

3.2 Discussion of F-measure performance

To account for the relative quality of Shazam's performance, we examined some of our audio queries in which there were: (1) lots of speech, (2) low-noise backgrounds, (3) pauses between speaker turns, or (4) silent moments. We also analyzed spectral images in these queries and the corresponding images reconstructed from their top 200 wavelets. In these reconstructed images, we can notice that a lot of important information from speech is neglected at high frequency bands by the CMU and Waveprint algorithms. This is due to a low-pass filtering stage with cut-off frequencies at 2kHz. The fundamental frequencies of musical instruments are usually below 2khz, and it is reasonable (or, at least, more reasonable than with speech) to assume that the important information above 2kHz derives primarily from harmonics of the fundamental frequency. Speech is different, i.e., it is more bursty, and has many different formants. By applying low-pass filters at 2khz to speech signals, the CMU algorithm and Waveprint neglect distinctive information of these higher formants. Increasing the cut-off frequency to an appropriately higher value such as 8khz would address this in principle, but was found to be computationally so inefficient as not to be testable. A literary analysis of the choice of 2kHz traces back to the original work of Haitsma and Kalker (2002), in which it is apparently selected because the ``the range from 300Hz to 2000Hz [is] the most relevant spectral range for the human auditory system." In terms of classifier performance, focussing on this range does not work well.

Also, the film dataset contains high energy at low frequencies due to low frequency background noises and our microphone. The Waveprint algorithm extracts the top 200 wavelets by magnitude, thus, if the top wavelets happen to lie at low frequency bands its fingerprints are not likely to have enough distinctive power. Low noise background, pauses and silent moments do not improve the F-measure, and in fact, they could worsen the performance because they occur in most films, leading to many false positive hits.

Algorithm	Precision	Recall	F-measure
Shazam	0.9862	0.9871	0.9867
CMU	0.2586	0.4650	0.3324
Waveprint	0.6279	0.9716	0.7628

Table 3. Comparison on 123 movies and 3690 queries.

The CMU algorithm's low precision could be due to using only descriptors within a hamming distance of 2 of the original descriptors.

4. CONCLUSION AND FUTURE WORK

Our comparison shows that the provenience of the audio data (music vs. voice and background noise from films) is enough to foil even simple design decisions that some algorithms (notably the two vision-based algorithms) use to improve speed and/or F-measure performance in other domains. Also, while these vision-inspired approaches are indeed competitive with Shazam on music data, the same cannot be said about their performance on film data. The F-measures are not even close. The spectral characteristics of the data seem to be responsible for this in great part.

It is also interesting that Shazam can perform so well on film identification --- with only the audio layer --- even though it was designed for music identification.

REFERENCES

Baluja, S. and M. Covell (2008). Waveprint: efficient wavelet-based audio fingerprinting. *Pattern Recognition* 41(11), 3467-3480.

Haitsma, J. and Kalker, T. (2002). A highly robust audio fingerprinting system. Proc. Intl. Sym. Music Info. Retrieval.

Ke, Y., D. Hoiem and R. Sukthankar (2005). Computer vision and music identification. Proc. Comp. Vision and Pat. Rec., 597-604.

Mapelli, F. and R. Lancini (2003). Audio hashing technique for automatic song identification. Proc. Intl. Conf. on Info. Tech., Research and Education, 84-88.

Li, Q. and Wu, J. and He, X. (2008). Content-based audio retrieval using perceptual hash. Proc. Intelligent Info. Hiding and Multimedia Sig. Proc., 791-794.

Wang, A. L.-C. (2003). An industrial-strength audio search algorithm. Proc. Intl. Sym. Music Info. Retrieval.

ACKNOWLEDGEMENTS

This research was funded by the GRAND NCE as part of the NEWS project.

HOW TO MODIFY A TESTED FIRE-RATED WALL TO IMPROVE ITS STC SOUND RATING, WHILE MAINTAINING ITS OFFICIAL FIRE-RATED QUALIFICATION

Harold Forester, Eng.
Contrôles Forester / Forester Controls
1434 Franklin, Laval, QC H7W 1K6, forester@videotron.ca

Codes and regulations governing building design and construction are generally issued by municipal and provincial governments. The regulations covering the designs for fire safety usually are based on standards issued by the Underwriters Laboratories Canada (ULC). Wall must be constructed in accordance with designs that have been tested in approved facilities to meet different ratings, such as 1 hour or 2 hour resistance to fire.

Walls are also designed to meet certain sound transmission loss values and are rated on their Sound Transmission Class or STC. There are hundreds of examples of wall partition designs that have been tested for their STC, but there is no general obligation for a designer to actually have to select a tested design to be able to meet sound criteria.

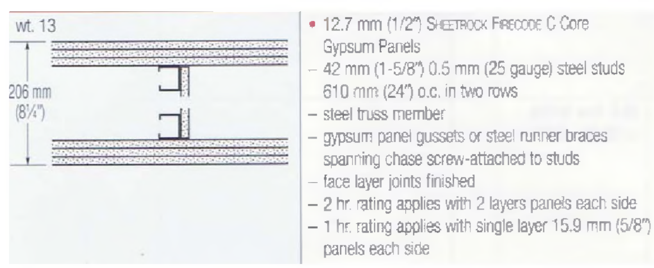
Several publications list both the fire and sound test results of many types of walls with huge numbers of permutations and combinations of materials, thicknesses, etc. The two most prominent lists are in Table A-9.10.3.1.A, Fire and Sound Resistance of Walls, in the National Building Code of Canada and in brochures issued by CGC Inc. and USG Corp. Unfortunately not all wall designs that have been tested for fire have also been tested for sound, and vice versa. This often creates a dilemma for the designer who is searching for a single design that will satisfy two independent criteria.

Fortunately, there is a legitimate way to work around this limitation. Underwriters Laboratories Canada has issued design guidelines for fire ratings (ULC document No. BXUVC). These guidelines outline modifications that are permitted to a tested design that do not affect the tested fire rating. Excerpts from ULC BXUVC are:

1. The dimensions and gauge of steel studs are minimums. The hourly ratings apply when the steel studs are of a heavier gauge and/or larger dimensions than specified in a design.
2. Spacing between parallel rows of studs are minimums unless otherwise stated in a design.
3. Additional layers of gypsum board are permitted to be added to any design.
4. Listed and labelled mineral fibre thermal building insulation processed from rock, slag and glass only may be used in ULC non load-bearing wall assembly designs consisting of wallboard and steel or wood studs with a fire-resistance rating not exceeding two hours when illustrated without insulation, without detracting from the rating assigned to the assembly.

Every one of the allowable modifications can be used to improve a wall's STC. So a designer starts by selecting a wall design that is close to the desired configuration that has been tested to the required fire rating. He can then proceed to modify that wall design step by step, in accordance with the ULC guidelines, until the wall conforms to a tested design with the desired STC.

Here is a specific example of modifying a wall. A mechanical shaft within a condo needed a 2 hour fire rating as well as a high STC rating of 62. The starting point was to select an approved chase wall (double stud) design #A36 from CGC brochure #SA-100, that has a 2 hour fire rating, per UL test No. Des U436, when constructed with two layers of ½" gypsum board on each side. It is shown below.



The following modifications were applied to that design:

1. Substitute 63 mm (2-1/2") CGC C-H 25 ga. steel studs @600 mm (24") o.c. on the rear side of the wall, replacing the 42 mm (1-5/8") 25 ga. studs.
2. Substitute 1" Sheetrock on the rear wall, replacing 2 layers of ½" GB. (The C-H studs and 1" Sheetrock are special designs that permit the studs and GB to be inserted from the front face where the shaft side of the wall is inaccessible.)
3. Substitute 64 mm (2-1/2") 25 ga, studs @ 600 mm (24") on the front side of the wall, replacing the 42 mm (1-5/8") 25 ga. studs.
4. Change the total external width to 9.0", to maintain the 2.0" spacing between the studs. The original width was 7-1/4".
5. Add 65 mm (2-1/-2") rock or glass fibre insulation batts to both sets of studs (original had no insulation).

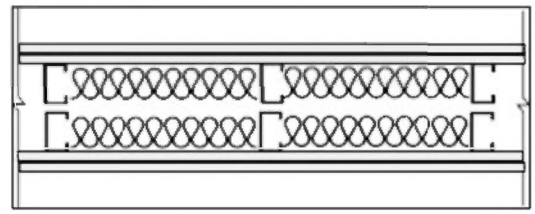
The resulting design was a shaft type of wall, 17 ft. high, with a 2 hour fire rating per UL test U436, and an STC rating of 62, per NRC test No. TL-93-305.

NRC-CNRC National Research Council Canada

TestID | TL-93-305

Element	Description:
1	single layer of 13 mm type X gypsum board
2	single layer of 13 mm type X gypsum board
3	65 mm steel studs at 610 mm on centre
4	65 mm of glass fibre insulation in cavity
5	16 mm gap filled with cross brace
6	65 mm steel studs at 610 mm on centre
7	65 mm of glass fibre insulation in cavity
8	single layer of 13 mm type X gypsum board
9	single layer of 13 mm type X gypsum board

2G13_SS65(610)_GFB65_AIR20_SS65(610)_GFB65_2G13



Finally, the following table lists a compendium of several different wall types with their corresponding fire and sound ratings. The ratings are based either on direct test results or by extensions of the original test results after modifications that comply with the ULC design guidelines, to improve their STC ratings. The list certainly does not cover all possible combinations of wall materials, and the fire ratings based on the extension of test results of similar walls may be questionable in some cases. Nevertheless this table can be very helpful for those trying to achieve both specified fire ratings and sound transmission ratings on wall types that have not been tested in the configuration of interest.

Compendium of Gypsum Board Wall Assemblies

Stud				GB		Absorptive	Max.	Resilent	STC		Fire resistance		Total
Size	Gauge	Spacing	Stud gap[1]	Qty	Size	filler	stud ht.	channel	Value	Test No.	Time	Test No.	width
CHASE WALLS													
2.5 "	25	24 " cc	20 mm	4	0.625 "	2.5 "	16'-0"	-	64	TL-93-302	2 hr*	UL Des U436	8.29 "
2.5 "	25	24 " cc	20 mm	4	0.500 "	2.5 "	16'-0"	-	62	TL-93-305	2 hr*	UL Des U436	7.79 "
2.5 "	25	24 " cc	65 mm	4	0.625 "	2.5 "	16'-0"	-	65	TL-93-321	2 hr*	UL Des W454/U493	10.06 "
2.5 "	25	24 " cc	65 mm	4	0.500 "	2.5 "	16'-0"	-	63	TL-93-308	2 hr*	UL Des U436	9.56 "
STANDARD WALLS													
3.625 "	25	24 " cc		4	0.625 "	3.0 "	13'-6"	-	56	USG 840818	2 hr	UL Des 411/419	6.13 "
3.625 "	25	24 " cc		4	0.625 "	3.5 "	13'-6"	-	56	NBC S6a	2 hr	NBC S6a	6.13 "
3.625 "	25	24 " cc		4	0.625 "	3.5 "	13'-6"	-	57	TL-92-369	2 hr*	UL Des W453/440	6.13 "
3.625 "	20	16 " cc		4	0.625 "	3.5 "	16'-9"	RC-1	61	NBC S14a	2 hr*	UL Des U404	6.63 "
3.625 "	20	16 " cc		4	0.625 "	3.5 "	16'-9"	-	55	NBC S6b	2 hr	NBC S6b	6.13 "
6.0 "	25	24 " cc		4	0.625 "	6.0 "	15'-0"	-	59	NBC S9a	2 hr	NBC S9a	8.50 "
6.0 "	25	16 " cc		4	0.625 "	6.0 "	20'-0"	-	59	NBC S9a	2 hr	NBC S9a	8.50 "
6.0 "	20	24 " cc		4	0.500 "	5.0 "	24-11"	RC-1	63	RAL-TL-87-141	2 hr	UL Des W453/U454	8.50 "
6.0 "	20	24 " cc		4	0.625 "	5.0 "	24-11"	RC-1	62	RAL-TL-84-139	2 hr	UL Des W453/U454	9.00 "
6.0 "	20	24 " cc		5	0.625 "	5.0 "	24-11"	RC-1	65	RAL-TL-84-150	3 hr	UL Des W453/U455	9.63 "
SHAFT WALLS, USING C-H STUDS													
4.0 "	25 C-H	24 " cc		1 2	1.000 " 0.500 "	3.0 "		RC-1	53	USG-040909	2 hr	UL Des U415	6.50 "
4.0 "	25 C-H	24 " cc		1 3	1.000 " 0.500 "	3.0 "		RC-1	58	USG-040910	2 hr	UL Des U415	7.00 "
2.5 " 2.5 "	25 C-H 25	24 " cc 24 " cc	51 mm	1 2	1.000 " 0.500 "	2.5 " 2.5 "	 13'-6"	- -	62	TL-93-305	2 hr*	UL Des U436	9.00 "

Notes:

All gypsum board (GB) panels to be fire rated Type X

1 - Space between studs in double stud (chase) walls

* 2 hour fire rating will qualify by extension

NBC - National Building Code

TL - National Research Council

RAL - Riverbank Acoustical Laboratories

UL - Underwriters Laboratories

USG - U.S. Gypsum Corp.

Acoustical Challenges in Long Term Care Facilities

Zohreh Razavi, P.Eng., M.A.Sc., LEED® AP BD+C

Acoustical Engineer, Stantec Consulting Ltd., zohreh.razavi@stantec.com
1100 - 111 Dunsmuir Street, Vancouver, BC, V6B 6A3, Canada, Ph: (604) 696-8472, Fx: (604) 696-8100

1. Abstract

Providing satisfactory acoustical environments in healthcare facilities can be ensured by applying recommended minimum design requirements provided in Sound and Vibration Design Guidelines for Hospital and Healthcare Facilities [1]. However, there are still acoustical challenges within long term care facilities that should be addressed, such as:

- Maintaining speech privacy between rooms and corridors with large undercut door openings for air flow;
- Maintaining STC ratings of the demising walls where the ceiling plenum is utilized for ducting and plumbing systems;
- Maintaining speech privacy between rooms and corridors while good speech intelligibility through corridors for caregivers to hear calls from residence inside the room is required.

I will discuss the aforementioned issues experienced in one Long Term Care (LTC) facilities, including all challenges for improving acoustical separation between a room, holding a person with dementia who was screaming during days and nights, and the public area, TV room / eating room. Proving the steps taken to improve this acoustical separation is discussed including all challenges on how to not affect fire separation of demising walls and door.

2. INTRODUCTION

Following a call from a manager of a long term care facility explaining noise issue within their facility, where one hundred and thirty elderly people reside, a review of their acoustical environment was undertaken.

The facility was a new building built in two levels with private rooms located on the perimeter of the building and the communal areas (e.g. TV room, dining room, etc.) positioned in the middle.

One elderly person who was suffering from dementia was relocated from another facility to one of the private rooms in this new facility. Her screams all day and night long resulted in discomfort environment for all other residence within the facility.

Steps taken to solve this issue are discussed in this paper.

3. METHOD

To evaluate level of background noise within the facility and acoustical separation between private rooms and between private rooms and the communal area, a site visit was arranged.

During our site visit, no scream was heard and the elderly person was quiet. We were told the subject quietness might be because of having companionship (e.g. our presence in the room).

In our site visit we reviewed architectural and mechanical drawings to evaluate STC rating of the demising walls and any openings through HVAC systems within the facility. Through this review we understood:

- Interior walls were all 152 mm steel studs with 15.9 mm Type X GWB on both sides filled with acoustical batt insulations, achieving STC 51.

- Doors to all private rooms were solid core with gaskets at the opening side which was not large enough to seal the gaps.

- The corridors were pressurized to have a positive pressure and undercut doors were used for air-intake and air exhausts and placed in the bathrooms.

- No electrical outlets were located back to back to affect the STC rating of the demising walls.

No acoustical treatments were considered anywhere within the building which is typical within most Long Term Care facilities [2, 3]. A white noise generator device was located inside the room to be played inside the room when the subject was screaming. The staff in the facility was using the device to mask the person's screams!

4. FIELD MEASUREMENTS

The white noise generator was used to investigate the effectiveness of the gaskets on the door.

The instrument used was a Bruel and Kjaer sound level meter Type 2250 with one octave band frequency analyzer and a free field ½" microphone.

The measured noise level was 70.4 dBA and 59.2 dBA, inside and outside the room, respectively. The measured noise level outside the room was conducted with door closed. To evaluate the effectiveness of tight seals around the door perimeter, figure tape was used to cover all gaps around the perimeter of the door. The measured noise level with the sealed door was 4.6 dBA lower than the one with existing conditions and it was equal to 54.6 dBA. The background noise within the facility during the measurements with all normal activities was equal to 50.5 dBA.

Measured noise level within the adjacent room was 40.78 dBA. No noise from the adjacent room could be heard through the demising wall or doors.

5. RECOMENDATIONS

It is found that the large opening under the door equal to 5/8" with no proper seals around the perimeter of the door was the main noise leaking points. Since the door's undercut was used for air intake, alternative should have been considered to seal this opening. The best option was to close the undercut door and add a quiet vent silencer. However this option due to door's fire rated properties wasn't the best solution. We suggested Pemko seals at the bottom of both sides of the inactive door and acoustical bottom door for the active door. It is also suggested to seal the perimeter of the door with draftseal self-adhesive bulb or Pemko products. It is decided to close the opening under the door and add a custom fabricated metal duct complete with louvre grilles and internal lining to provide fresh air into the room.

We also recommended to boxing in the duct with 5/8" thick Type X Gypsum Wall Board built on metal stud framing.

Since the air transfer duct was penetrated through a one hour fire rated wall assembly, a fire damper was required.

Addition of sound absorptive materials to the ceiling of the room to lower RT within the room was also suggested.

6. DISSCUSIONS

Installation of the duct requires cutting holes in the drywall of the wall beside the narrow door, so appropriate precautions must have been taken to contain dust while the work was in progress. We assumed that facility operator would have their own safety and environmental protocols to follow when doing maintenance or renovation work inside the facility. We also offered to pass on recommendations from the CSA standards that describe special measures for this type of situation if they were unsure what to do.

We also reminded them about the fire-rated labels attached to the ends of the doors. These labels indicated that both doors were fire rated and form part of the smoke separation that prevents smoke from migrating from the corridor into the resident room and vice versa. It was important not to compromise the integrity of these doors, so we were taking a careful approach to adding accessories to the door and frame.

7. CONCLUSIONS

We submitted recommendations in order to improve acoustical environment within LTC after receiving a phone call from the facility manager. This was to increase acoustical separations between a private room and the facility communal area. Phase 2 of acoustical upgrades to the private room was suggested once we realized the Phase 1 recommendation on improving gasketing around the perimeter of the door wasn't satisfactory. The owner of the facility planned our recommendations on their own, including all activities related to retaining a contractor and getting the work done. Thus, we didn't get involved with tendering or monitoring of construction activities.

Phase 2 of acoustical upgrade consisted of sealing the undercut at the entrance doors to the private room and providing an alternate means of allowing fresh air into the room via a special transfer duct specially designed to allow air movement but not noise transmission.

Not jeopardizing the fire and smoke separations provided by the door was one of our priorities.

At the time of writing this paper, the owner of the facility was implementing the recommendations. Thus, no outcome from this implementation was available at this time.

AKNOWLEGEMENT

I would like to thank Garth Balint timeless efforts. Without his inputs, this project wouldn't be initiated.

REFRENCES

1. Sound and Vibration Design Guidelines for Health Care Facilities, January 1, 2012;
2. Cohen-Mansfield J, Werner P, Hammerschmidt K, Newman JD. "Acoustic properties of vocally disruptive behaviors in the nursing home";
3. Dugan, 2003; Healthcare Design, 2002; Healthcare Design, 2003 "Research Institute on Aging of the Hebrew Home of Greater Washington, Rockville, MD 20852, USA.

SOUND ISOLATION UPGRADE OF EXISTING PERFORMING ARTS CLASSROOMS – DESIGN CHALLENGES AND THE PUSUIT OF SOUND FLANKING PATHS

Erwin Rebke
Alberta Infrastructure, Capital Projects Division, Technical Services Branch
3rd Floor, 6950-113 Street NW, Edmonton, AB, Canada T6H 5V7

1. INTRODUCTION

A new performing arts program was established within an existing urban high school to promote continuing education for dance, music and drama. To satisfy the immediate demand for space, an old industrial arts section of the school was converted by dividing the total floor area into three separate, but adjacent rooms using 190mm concrete block walls. Simultaneous activities within this new Dance/Music/Drama facility resulted in a high degree of sound interference. This author did an initial investigation in the Fall of 2005, providing recommendations with limitations due to budget constraints. The project did not proceed.

With a new and expanded budget, the project was recalled in January, 2011. A review was done of existing conditions and recommendations were revised to provide all the necessary elements and details to ensure optimum sound isolation.

2. EXISTING ROOM CONDITIONS

The old industrial arts section of the school was a large, open space with a total, approximate floor area of 43 meters by 18 meters. The floor consists of a monolithic concrete slab on-grade. The exposed structural roof is a concrete T-beam system. The height to the underside of the roof slab is approximately 7.5 meters. Two demising walls were constructed using 190mm concrete blocks, painted and sealed at all junctions and penetrations. The resulting rooms, excluding the built-in offices and storage have the following physical characteristics:

Dance Studio V= 2060 cubic meters; A = 275 square meters
Music Room V= 1610 cubic meters; A = 229 square meters
Drama Room V= 1630 cubic meters; A = 197 square meters

Each common block wall area is approximately 137 square meters (18.3m X 7.5m).

The existing sound isolation was measured and classified in terms of Noise Isolation Class (NIC) as follows:

Dance Studio to Music Room: NIC = 40
Drama Room to Music Room: NIC = 39

The corresponding level differences and lab-measured transmission loss values for a 190mm concrete block wall are compared in Figure 1.

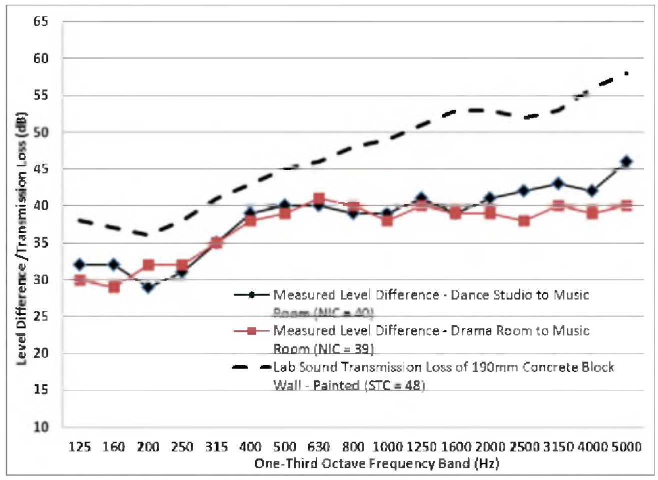


Figure 1. Original Sound Isolation

3. ACOUSTIC UPGRADE

One side of each concrete block wall was built-up using steel stud and gypsum board construction. Ideally, this additional wall system should be structurally separate from the existing block wall. However, the wall height of 7.5 meters did not allow construction of an unsupported wall. Steel hat channels were used to tie the new wall system back to the existing block wall. The amount of ties was limited to provide as much structural isolation as possible. Also, resilient channels were used to attach the gypsum board to the metal studding. The estimated Sound Transmission Class rating for the new composite wall system is STC 65.

A primary concern was the common floor slab and its' potential to flank sound underneath the new wall systems. With the Music Room located between both the Dance Studio and the Drama Room, an isolated floor system was placed within the Music Room. The installed floor system consists of the roll-out floor isolation system; Model RIM by Kinetics Noise Control Inc.

Specific concerns were addressed during construction that related to mechanical/electrical attachments and penetrations, as well as construction within and around stairs, bulkheads and storage areas.

4. POST CONSTRUCTION RESULTS

An initial, subjective evaluation of the sound isolation improvement indicated that the construction between the Dance Studio and Music Room is now optimized. However, the sound isolation between the Drama Room and Music Room still had significant sound flanking. At the top of the common wall and on the on the side of the new construction, the wall face is about 50mm from the T-beam web structure. The acoustic seal at the top of the new wall to the structure underside was questionable due to the limited access. Additional details were provided to seal this construction junction. Sound isolation tests at this point confirmed that significant sound flanking still existed. Results of the measured sound isolation for each room pair is as follows:

Dance Studio to Music Room: NIC = 61
Drama Room to Music Room: NIC = 47
The sound isolation improvements are shown in Figures 2 and 3 for the two room pairs.

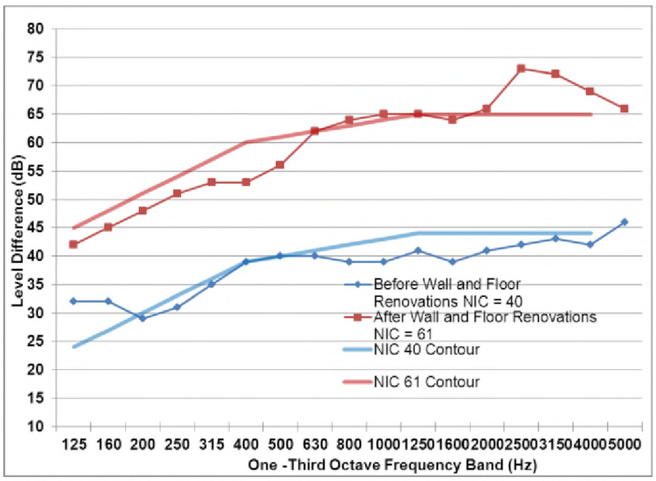


Figure 2. Sound Isolation Improvement – Dance Studio to Music Room.

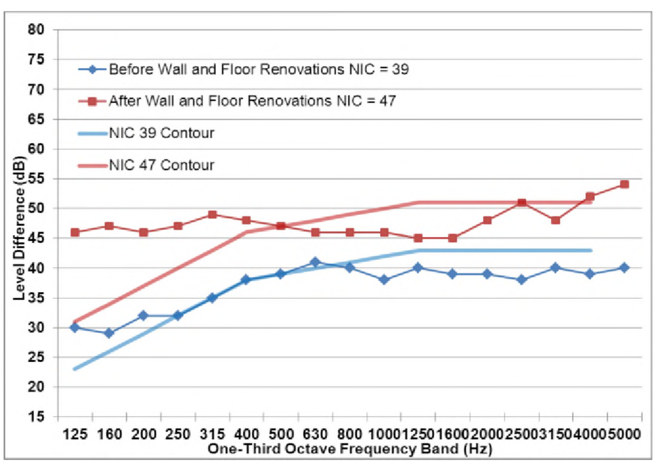


Figure 3. Sound Isolation Improvement – Drama Room to Music Room.

5. STEEL BEAM INVESTIGATION

The sound flanking investigation now focused on an array of steel hoist support beams that remained from the days when the area was used as an industrial arts shop. This steel beam array is common to both the Music Room and Drama Room, and penetrates the wall at five separate locations. All penetrations were sealed well and did not exhibit any sound leakage. However, it was surmised that the airborne sound was creating sufficient vibrations within the beams and radiating into the adjacent room from the web surfaces.

To validate the correlation of beam vibration to the receiving room sound pressure, pink noise was generated within the Drama Room at an overall, average level of 97 dBA. Using a two channel analyzer, simultaneous measurements were taken within the Music Room (receiving room). One channel measured the vibration velocities on the steel beam web, and the other channel measured the sound pressure levels within the room. Figure 4 shows a comparison of the two measurements. The correlation coefficient for the beam vibration velocities and the room sound pressure levels is 0.815 within the one-third octave frequency bands from 250Hz to 5000Hz.

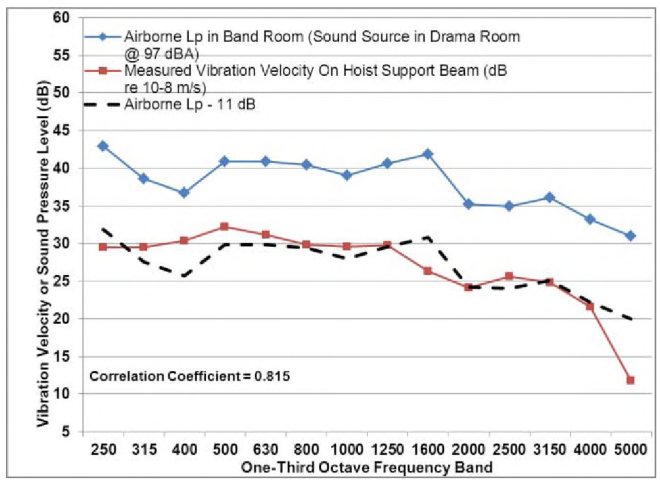


Figure 4. Simultaneous Beam Velocity and Room Sound Pressure Level Measurement

6. CONCLUDING REMARKS

The high sound isolation required for performing arts classrooms demands careful consideration in the acoustic design and the elimination of all structural sound flanking paths. All elements common with and penetrating through the sound-rated construction must be evaluated for their sound flanking potential.

ACKNOWLEDGEMENTS

The author acknowledges Russ Lewis for his technical assistance in the design and commissioning of this project.

ACOUSTIC METRICS FOR CLASSROOM PERFORMANCE – A LITERATURE REVIEW

Setareh Shams and Ramani Ramakrishnan
Department of Architectural Science, Ryerson University, Toronto, ON, rramakri@ryerson.ca

1. INTRODUCTION

Speech intelligibility, reverberation time and ambient sound are three commonly used acoustic metrics to assess the performance of classrooms. This literature review was undertaken by an undergraduate research project to enumerate all possible metrics that can be used to assess completely the classroom performance from an acoustic perspective. The review has attempted to evaluate the sufficiency of the three conventional metrics. The validity of applying other metrics was also assessed. Finally, the effectiveness of such metrics and their impact on subjective experiences were also studied. The results of the literature review are presented in this paper.

2. CONVENTIONAL METRICS

2.1 Speech Intelligibility and Reverberation Time

Majority of the papers studied often state reverberation times or speech intelligibility as the most accurate means of measuring how well students listen in classroom environments. There were always debates about which of these two commonly studied factors actually represented the acoustical conditions.

Barron and Lee, at least for concert halls, concluded that reverberation time determines early decay time, early-to-late sound index, and total sound-pressure level with volume [1]. Hence, reverberation time is one of the most important descriptors of perceived quality of a space and acoustic metrics.

Bradley indicates that speech intelligibility is the best choice for basing ideal acoustical conditions. When speech intelligibility is nearly perfect in classroom conditions, there is little impact on noise to academic achievement [2].

Despite his preference towards speech intelligibility, Bradley also uses both parameters in conjunction as a means of obtaining more accurate results for measuring classroom acoustic [3]. Reverberation time is not the sole descriptor of room acoustic conditions. There exists an important relationship between reverberation time and speech intelligibly that is also governed through various parameters, and both affect the outcomes of one another. Increased reverberation times can negatively impact speech intelligibility, and they can also increase speech intelligibility through increasing levels of early arriving speech sounds. Consequently, reverberation time has to be at an optimum so as to make the teachers voice as intelligible as possible. However, increased reverberation, will affect the speech-to-noise ratios as well. Therefore, a metric that combines predicators of both speech intelligibility and reverberation times such as the useful-to-detrimental sound ratios can be used.

Bradley states that there are numerous predicators of speech intelligibly that must be considered as well [4]. Some of these include signal-to-noise ratio, speech transmission index and useful/detrimental sound ratios. The most accurate was determined to be the 0.08-s useful/detrimental ratio (U_{80}). When using speech intelligibility as a metric, there are also individual factors that must be considered that can affect its validity. The most important of these factors include speech-to-noise ratios, reflected sounds, and the age of the listener [5]. Furthermore, as previously mentioned, reverberation time also has an important affect on speech intelligibility. As a result, it is critical to have reverberation times that are not very low while having some reflected sound in order to increase speech intelligibility. Bradley's results concluded that decreasing reverberation times influenced those of a younger age in their speech intelligibility, and therefore it is important to consider age in further studies.

2.2 Metrics and Age

Many papers also assert the importance of focusing data on age-specific subjects. Some data studied did not necessarily consider the significance of age and its relationship to acoustics. For instance, Neuman and Hochberg state that studies involving classrooms have focused on determining optimal reverberation through testing adults as opposed to children [6]. Because children adhere to different acoustical environments, the data that was thus determined in a way becomes rather inaccurate. Their study showed that in the same reverberant conditions, children scored less than adults, proving that children perceive sound in different and shorter reverberation times than adults do.

According to Bradley, different acoustical criteria are required, depending on the grade and age of students. For instance, in classes with Grade 1 students, because they are more easily distracted, teachers would adjust their voice levels based on the ambient noise level. Consequently, ambient noise is something that needs to be highly controlled in such cases [7].

Therefore, it is absolutely vital for researchers to be aware of their target age when studying conditions in a room and carrying out tests in order to obtain the most useful and effective results.

3. RESEARCH ISSUES

Though finding effective parameters in research can be difficult, the way in which tests are executed is also of great importance, where lack of attention to detail can hinder test results. One of these issues in research methods is monaural vs. binaural hearing. According to Bradley, test results can be accurate if children are tested in realistic classroom conditions with both ears [8]. Often, tests that are conducted by researchers neglect the importance of this. Furthermore, under experimental conditions, there is a lack of criteria that children experience in realistic classrooms, such as distractions, their own noises, and so on. The tests are conducted in unoccupied classrooms and fail to incorporate the subjective experiences of students. There are considerable differences between a room that is occupied and one that is not, and researchers must keep such issues in mind when conducting studies.

Bradley thus agrees that there is a greater need for acoustical conditions when classrooms were occupied and in operation [3]. His studies are completed in more realistic conditions during actual teaching conditions, and as a function of signal-to-noise ratios.

Hodgson and Nosal argue that often times, when calculating optimal reverberation times, experimenters forget to include the effects of noise [9]. Bradley argues that noise is much bigger overriding problem than poor room acoustics, especially when there is an absence of reverberation. Consequently, reverberation times have to be designed in such a way that they minimize the effects of background noise on the listener. In their paper, in order to determine optimal reverberation times, Hodgon and Nosal use the speech intelligibility metric of U_{50}. Experimental procedures often represent background noise in a more unrealistic manner. Therefore, Hodgson and Nosal state that "a physically realistic treatment of noise incorporates both the nearby noise source and the effect of reverberation on noise." Research conducted by Houtgast concludes the same issues. In his paper, Houtgast considers noise (from exterior) and how it affects intelligibility for students under the effects of reverberation [10].

Finally, Palovic points out that it is often difficult to make direct comparisons between results, because different researchers use different parameters when dealing with speech intelligibility [11].

4. SUBJECTIVE ANALYSIS

Subjective experiences are vital aspects that students undergo on a daily basis in classrooms. Studying them often exposes researchers to different and more detailed perspectives on classroom acoustics. However, the majority of papers studied neglect the true subjective experience of the listeners in the classrooms. Tests that were conducted were based on the technical and scientific approach to results. Many simply performed straightforward tests such as the Word Identification by Picture Identification, Fair-banks rhyme test, recordings of nonsense syllables, and so on. These tests focused solely on obtaining results that dealt with various acoustical parameters through the use of acoustical equipment, as opposed to how students truly perceived sound in a space. The results were then taken and compared to other quantitative results and analyzed through means such as graphs and charts. The only studies that mention subjective experiences were by Lochner and Burger, in which they tie subjective judgments to reflection patterns [12].

4.1 Subjective parameters

According to their study, Lochner and Burger believe that the execution of subjective judgments in a space is vital to how effective acoustical qualities in a room are carried out [12]. However, they also state that determining the physical units that can accomplish this in the most effective way possible is very difficult. In other words, finding the proper metric to evaluate subjective experiences is a big challenge. Nonetheless, they believe that the most accurate way is through reflection patterns in a room, and that "the quality of speech and music in a room is a function of the reflection pattern." They carried out tests with different observers using subjective articulation tests in varying reflection conditions, and used the observers' results to conclude speech intelligibility in relation to reflections. However, in the end, these speech intelligibility scores serve as a quantitative measure, and route right back to the issue of the difficulty of obtaining qualitative characteristics from such quantitative measurements.

A few other studies only briefly mention how some parameters can indicate subjective experiences. Steeneken and Houtgast used subjective intelligibility measurements in order to obtain the relation between speech transmission index values (an objective physical measure of speech transmission) and intelligibility scores [13], which again rely on quantitative measures to determine subjective measures. Bradley states that though reverberation time has been a strong indication of acoustical conditions in a room, early decay times have become more useful in terms of subjective evaluations [7]. They indicate the level of clarity as well as speech intelligibility in a room. Finally, Bradley also mentions that according to a study by Reichardt and Lehman that the "early/late-arriving sound-energy ratio for a 0.08-s early sound limit (C_{80}) has gained considerable acceptance as a correlate of subjective judgments of musical clarity" [14]. All these studies simply mention a brief measure that potentially has the ability to assess subjective analysis, yet they fail to actually present concrete results using qualitative analysis.

Though all of the research refer to how various metrics can direct us into the subjective experiences, but no quantitative basis had been presented. Thus, it is important that subjective experiences become a critical segment of experimental procedures. It becomes valuable to include how students truly perceive acoustics in a space subjectively, and finding a parameter to effectively do so becomes the next challenge.

5. REFERENCES

1. M. Barron and L.-Lee, "Energy relations in concert auditoriums. I, " J. Acoust. Soc. Am. 84, 618-628 (1988).
2. J. S. Bradley, "Speech Intelligibility studies in classrooms," J. Acoust. Soc. Am. 80, 846-854 (1986).

3. H. Sato and J.S. Bradley, "Evaluation of Acoustical Conditions for Speech Communication on Working Elementary school classrooms," J. Acoust. Soc. Am, 2064-2077 (2008).
4. J. S. Bradley, "Predictors of Speech Intelligibility in Rooms," J. Acoust. Soc. Am. 80, 837-845 (1986).
5. J.S. Bradley, "Effects of room acoustics on intelligibly of speech in classrooms for young children," J. Acoust Soc. Am, 922-933 (2008).
6. A.C. Neuman and I. Hochberg, "Children's Perception of Speech in Reverberation," J. Acoust. Soc. Am. 73, 2145-2149 (1983).
7. J.S. Bradley and H. Sato, "The intelligibility of speech in elementary school classrooms," J. Acoust. Soc. Am. 123, 2078-2086 (2008).
8. J. S. Bradley, "Auditorium Acoustics Measures from Pistol Shots," J. Acoust. Soc. Am. 80, 837-845 (1986).
9. M. R. Hodgson and E. M. Nosal, "Effect of noise and occupancy on optimal reverberation times for speech intelligibility in classrooms," J. Acoust. Soc. Am. 111, 931-939 (1991).
10. T. Houtgast, "The Effect of Ambient Noise on Speech Intelligibility in Classrooms," Appl. Acoust. 14, 15-25 (1981).
11. C.V. Pavlovic, "Derivations of primary parameters and procedures for use in speech intelligibility predications," J. Acoust. Soc. Amer. 82, 413-422 (1987).
12. J. P. A. Lochner and J. F. Burger, "The Influence of Reflections on Auditorium Acoustics" J. Sound Vib. 1, 426-454 (1964).
13. H. J. M. Steeneken and T. Houtgast, "Mutual Dependence of the octave-band weights in predicting speech intelligibility" Speech Commun. 28, 109-123 (1999).
14. J. S. Bradley, "Predictors of Speech Intelligibility in Rooms," J. Acoust. Soc. Am. 80, 846-854 (1986).

LOST IN TRANSLATION: A STUDY OF MUSICAL LANGUAGE AND ENGINEERING DESIGN

Colin Spohr[1]
[1]4th Year Electrical Engineering, University of Calgary, 2500 University Drive NW, Alberta, Canada, T2N-1N4
cspohr@gmail.com

1. INTRODUCTION

The success of any project truly depends on how the unique challenges associated with the project are managed and dealt with. When dealing with a client with a musical background, the acoustic engineer is then faced with the challenge of translating the subjective metrics as described to them by the client to the objective metrics needed to proceed with the design of a treatment regime.

1.1 Musical Language

The term "Musical language" is used in this paper not refer to a language based on musical sounds such as Solresol[1], but refers to the language used by individuals to describe various traits or characteristics of music or sound. This language has also been described as "The Language of Musical Acoustics"[2].The variance in terms used is partly due to our own unique perception of sound and our own experience. Our interpretation of what we hear is described in words that relate to the sound, and since these words are based on our own experience, they are therefore subjective by nature; pertaining to the experience of the individual.

Terms like "dry" and "dead" may essentially describe the same characteristic, but is that characteristic what the client intends to communicate? There are other terms that are commonly used but much less descriptive, but if one was without experience with a sound containing these characteristics, there would be an inherent difficulty to understand descriptions that make use of these terms. Moreover, even if one does have experience, it may not be how that individual would describe the characteristic of interest. It has been said that a picture is worth a thousand words, but it could it not be said that a sound can be worth the same amount of words or more?

1.2. Engineering Design

Before we actually perceive a sound, it is just pressure fluctuations about a mean atmospheric pressure, which oscillate in frequencies that change over time. We may describe a sound differently, but in the end, it is still the same sound. If a specific sound is measured by a device such as a sound level meter, we would see that the same sound measured would always produce the same results, the same amplitude, and would have the same frequency response.

During the course of the design process, it is essential for an engineer to have quantifiable values, targets or objective metrics to use to guide them on the path towards a design solution. For the case of the client with a musical background, the subjective metrics can be somewhat abstract, but one needs concrete objective metrics to facilitate problem analysis and solution design. The language used in engineering is much more concrete in nature. Words like "Parameters", "dimensions" and "requirements" are used in the design process and more often than not, these words are followed by quantities, with units, making them measureable. The key to overcoming this language gap, and arriving at a solution, is the translation of the subjective metrics to objective metrics.

1.2. Example Project

This paper makes use of a project in which a room that has been re-purposed for musical rehearsal is assessed and a treatment regime is designed. The room is comprised of 4 parallel walls and one angular wall with a total volume of 507 m^3. The walls are treated with high frequency absorption (HF) panels; the ceiling is treated with angular diffusers, while the floor is linoleum on concrete. The band consisting of 15 to 27 musicians is comprised of brass, wind and percussion instruments are situated in a tiered fashion.

2. REQUIREMENTS

The general reports from the client and the band members was that the room was described as too "dry", too "live", "Muddy and too loud. The general consensus from the band and the conductor was that they all had trouble hearing instruments located further away from them and subsequently had difficulty playing in tune with each other. As design metrics the solution would have to decrease the muddiness of the room, increase blend and tonality, decrease loudness and increase ensemble. These metrics are completely un-achievable without more information.

3. METHOD

The initial assessment was conducted using accepted equipment and methods. The reverberation measurements were taken from 5 source positions per each of 6 receiver positions at 3 decays each. Erroneous measurements were completely discarded and the remaining results were averaged to obtain the reverberation time RT60 of the room.

Terms used by the client and the users of the room to describe issues with the room were researched to gain some understanding of their meaning. The existing treatment regime was examined for its effect on the room characteristic and the shape of the curve generated from the RT60 times at each full octave band was compared to the subjective metrics given by the client and the users of the room.

4. RESULTS AND DISCUSSION

As even the interpretation of both the subjective metrics given by the client and the interpretation of the results of the assessment is the opinion of this paper's author, they are themselves subjective, but are based on research and experience of the author. The reverb times for the room are shown in Figure 1.

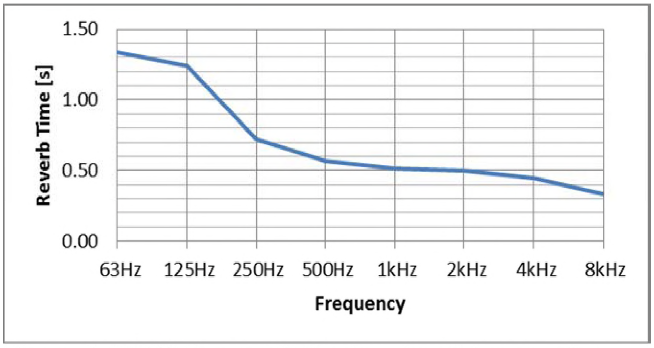


Figure 1. RT60 times at full octave band frequencies

4.1. Metric: Decrease "Muddiness"

Muddiness can be interpreted as the excess reverberation of the frequencies below 500 Hz when compared to those above 500 Hz, and therefore by this interpretation, it is clear from figure 1 that the room is indeed muddy as the low frequencies are over emphasized by the character of the room. To remedy this, HF panels would need to be removed and replaced with low frequency absorption (LF) panels as there are none currently in the room.

4.2. Metric: Increase "Ensemble"

Ensemble can be described as the ability of the performer's to play together as a cohesive unit[2]. This of course would depend on the ability of the musician to hear each other play. This interpretation is applicable as again the over treatment of the room with HF panels has removed the early sound, essential for the musician to hear each other as well as themselves. A proposed solution would be to remove more of the HF panels to increase the early sound so that the musicians could hear themselves more easily.

4.3. Metric: Increase "Blend"

Blend can be described as the mixing of sound produced by the different instruments so that they sound harmonious[2]. It could be said the lack of blend in this room is primarily a result of the lack of balance in the reverb decay times of the low frequencies vs. those of the high frequencies as shown in Figure 1. This lack of balance would be remedied by the solution proposed in section 4.2, allowing a more even reverb response in the room, increasing balance.

4.4. Metric: Increase "Tonality"

Tonality or tonal quality has been described as the accurate transmission of the sounds produced by the instruments[2]. The lack of tonality in this room can be interpreted as a direct result of the imbalance shown in Figure 1, and again would be remedied by the solution proposed in section 4.2.

4.5. Metric: Decrease "Loudness"

One interpretation is that the loudness or strength of sound in the room depends on the strength of the early sound, and the strength of the reverberant sound. As the walls were heavily treated with HF panels, it can be said that by the previous interpretation of loudness, the excess loudness is not a result of early sound, nor is it a result of reverberant sound when taking into account the results displayed in Figure 1. It would seem that the loudness was unfortunately a result of a large band, in a small room, but the solution proposed in 4.1 should also assist in lowering the overall level in the room.

4.6. Metrics: "Dryness and Liveliness"

To shed some light on the use of the clearly contradicting comments of the room being too live and too dead, the seating positions of the specific members were explored. The comment of the room being "dry" was given by a member seated next to a wall heavily treated with HF panels, and the comment of the room being too "live" was given by a member seated near the middle of the room. These reports were then seen to be a function of the seating position and the existing treatment regime, not a function of the room's overall characteristic.

5. CONCLUSION

This project presented unique challenges; the client was of a musical background and presented subjective metrics as requirements. These subjective metrics were compared to results of an RT60 assessment of the room and by using an understanding of the effect of different acoustic treatments, were converted into objective metrics to be used in the design of a treatment regime. This process allowed the client's requirements to be supported by experimentally determined objective metrics and ensured that the client's requirements were not lost in translation.

REFERENCES

1. François Sudre, 1827
2. Beranek L. (2004). "Concert Halls and Opera Houses, Music Acoustics and Architecture" 2nd Ed. Springer.

ACKNOWLEDGEMENTS

I would like to acknowledge Dr. Daryl Caswell giving the author the opportunity to undertake this project as well as Jesse Roy for her guidance along the way.

CAN AGE RELATED DECLINES IN SENSORY PROCESSING ALTER THE NATURE OF MULTISENSORY INTEGRATION IN THE PRESENCE OF ENERGETIC AND INFORMATIONAL MASKING?

Avivi-Reich, M., Puka, K. and Schneider, B.A.
Department of Psychology, University of Toronto Mississauga, Canada
Centre for Research on Biological Communication Systems, University of Toronto Mississauga
E-mail: meital.avivi@utoronto.ca

1. INTRODUCTION

Age-related deterioration in one's ability to comprehend speech in noise plays a primary role in the difficulties many older adults experience daily when communicating (CHABA, 1988). Listeners may intuitively look at the talker when experiencing difficulties in order to receive additional visual information when hearing is difficult. Visual cues have been found to facilitate speech perception when the acoustic information is incomplete or ambiguous (Grant and Braida, 1991; Helfer and Freyman, 2005). Previous studies have shown that as the auditory speech information becomes increasingly weak or distorted, the greater the gain exhibited from additional visual cues. According to this phenomenon, which is also referred to as "inverse effectiveness" (Macleod and Summerfield, 1990), it is reasonable to expect older adults to utilize visual cues accompanying the acoustic input when listening is difficult and the talker is visible, and possibly benefit more than young adults from these visual cues. However, the additional visual information provided by the face of the talker requires efficient integration of audio-visual information if this visual information is to enhance speech understanding. Previous studies provide evidence indicating that older adults experience both auditory and visual deterioration in perception of speech information with age (Campbell et al., 2007; CHABA, 1998). Age related declines in sensory processing, both auditory and visual, could alter the nature of multisensory integration. While some studies found no age difference in the ability to benefit from combining auditory and visual speech information (Sommers et al., 2005), others did find significant age differences (Campbell et al., 2007).It is possible that the degree of benefit obtained from visual cues might depend on the type of competing sounds that are present. All sounds, both noise and speech, activates regions along the basilar membrane which energetically interferes with the encoding of the target speech signal. This interference, which takes place whenever there is overlapping energy from several sound sources, is commonly referred to as peripheral or energetic masking. In addition, competing speech from one or more talkers is likely to interfere with the linguistic and semantic processing of the target speech. This interference which is believed to be taking place at more central levels in the auditory pathway (Schneider et al., 2007), is known as informational masking. The benefit provided by visual cues may depend on the degree of energetic and informational masking produced by competing sound sources, which, in turn, may be affected by aging. This study attempts to address the possible effects of age and type of masker on auditory visual speech perception.

2. METHOD

Twenty-four younger and 24 older adults with normal hearing and normal or corrected vision participated in these experiments. All participants were native English speakers in general good health and with no known neurological or cognitive deficits. Prior to the experimental session, participants' vision and hearing was tested as well as their English proficiency using a vocabulary test (Mill-Hill) and a reading comprehension assignment.

The participants were asked to repeat semantically anomalous target sentences with 3 keywords (e.g., "A *house* should *dash* to the *bowl*.") (Helfer, 1997, key words italicized) spoken by a female talker. These sentences are presented against either two-talker anomalous speech (two females), 12-talker babble, or speech-spectrum noise. The target sentences and the background noise are played over loudspeakers which are placed symmetrically in the frontal azimuthal plane at 45 degree angles to the left and right of the listener in a sound-treated booth. Masker signal started 1 second before the onset of the target sentence and continued up to the target sentence offset. Half of the sentences are played when the target sentences are accompanied by videos presenting the talker (Audiovisual); the other half are played over loudspeakers only (Audio). The entire face of the talker is presented to the participant at a natural size. The talker used for this study is a professional actress which was instructed to minimize facial expression to avoid possible bias. Each masker type and condition combination is played at four Signal to Noise Ratios (SNR) levels; and for each individual, the dB SNR corresponding to his or her 50%-correct performance threshold is computed from their individual psychometric functions along with the functions' slopes.

3. RESULTS

The results presented here are preliminary results based on 12 younger and 12 older adults who have completed the study.

Figure 1 presents the thresholds in dB SNR corresponding to 50% correct word-repetition in the audio only condition. This figure suggests that younger and older adults have comparable thresholds in noise but that older adults have higher thresholds than younger adults when the masker is either babble or speech. An ANOVA found significant main effects of masker and age along with a significant interaction between the two.

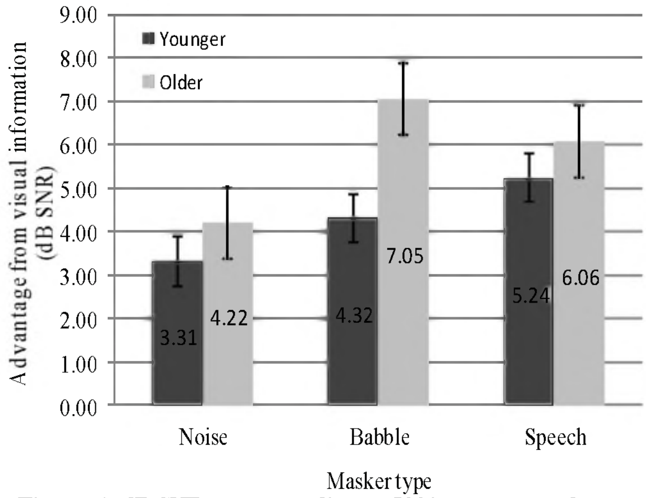


Figure. 1. dB SNR corresponding to 50% correct word-repetition as a function of masker type for younger and older adults in the auditory only condition.

Figure 2 presents the difference in dB SNR thresholds between the audio-visual condition and the audio only condition. This figure indicates that older adults receive greater benefit from available visual information than do younger adults. Indeed the performance in the audio-visual conditions of older adults is virtually equivalent to that of younger adults. While younger adults seem to increasingly benefit from visual cues as the informational masking increases, older adults seem to benefit the most when listening with in babble background noise.

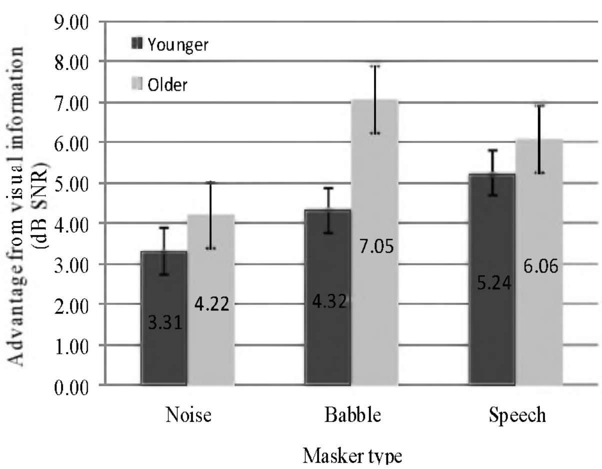


Figure. 2. The advantage in dB SNR (audio only minus audiovisual 50% correct word- repetition threshold) as a function of masker type for younger and older adults.

4. DISCUSSION

A previous study by Helfer and Freyman (2005) investigated the effect of additional visual speech information on the ability to correctly repeat sentences in the presence of either noise or competing speech in young adults. The results of this study showed substantial benefit from visual cues when repeating key word from semantically-meaningful sentences. We wanted to see whether the degree of benefit would be similar for sentences with no contextual support and whether there were age differences in the extent to which visual information facilitated release from masking in the presence of three different levels of informational and energetic interference. Comparing the present results to those of Helfer and Freyman (2005) imply that while the presence of contextual support does not affect the advantage from additional visual information in the presence of background noise, context does seem to increase the advantage in the presence of competing speech. In addition, the present results demonstrate age related differences in the use of visual cues based on masker type. The implications of these age-related differences will be discussed at the meeting.

REFERENCES

Campbell, M., Preminger, J.E., & Ziegler, C.H. (2007). The effect of age on visual enhancement in adults with hearing loss. *Journal of the Academy of Rehabilitative Audiology* , 40, 11 – 32.

CHABA. (1988). Speech understanding and aging. *The Journal of the Acoustical Society of America*, 83, 859 – 895.

Grant, K.W., & Braida, L.D. (1991). Evaluating the articulation index for auditory-visual input. *The Journal of the Acoustical Society of America*, 103, 2677 – 2690.

Helfer, K. S. (1997). Auditory and auditory-visual perception of clear and conversational speech.*Journal of Speech Language and Hearing Research*, *40(2),* 432-443.

Helfer, K.S., & Freyman, R.L. (2005). The role of visual speech cues in reducing energetic and informational masking. *Journal of the Acoustical Society of America*. 117, 842–849.

Helfer, K.S., & Freyman, R.L.(2008). Aging and speech-on-speech masking. *Ear and Hearing*. 29, 87–98.

Macleod, A., & Summerfield, Q. (1990). A procedure for measuring auditory and audiovisual speech-reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use. *British Journal of Audiology*, 24(1), 29 – 43.

Schneider, B.A., Li, L.,& Daneman, M. (2007). How competing speech interferes with speech comprehension in everyday listening situations. *Journal of the American Academy of Audiology*. 18, 559–572.

Sommers, M. S., Tye-Murray, N., & Spehar, B. (2005). Auditory-visual speech perception and auditory-visual enhancement in normal-hearing younger and older adults. *Ear and Hearing* , 26, 263 – 275.

ACKNOWLEDGEMENTS

This research was supported by a Canadian Institutes of Health Research Grant (MOP-15359). We thank Selasie Ametorwo for her assistance.

HEARING LOSS IN CLASSICAL ORCHESTRA MUSICIANS

Alberto Behar[1], Frank Russo[2], Marshall Chasin[3] and Stephen Mosher[4]
[1]SMART Lab., 350 Victoria Street, Toronto, ON. Canada, M5B 2K3, behar@sympatico.ca
[2]SMART Lab., [3]Musicians' Clinics of Canada, [4]National Ballet of Canada

1. INTRODUCTION

Quian et al. performed a noise exposure survey on members of the National Ballet orchestra, using noise dosimeters. They concluded that the exposure normalized for the 360 hr./year of musicians' activities with the orchestra is below the hazard level of 85 dBA. As a follow-up to the survey it was decided to perform hearing tests to the members of the orchestra. A questionnaire was also conducted to gather information on particulars of the surveyed members.

2. QUESTIONNAIRE

Musicians were assured that the questionnaire will be anonymous: no names were gathered nor included, something essential to the participants, due to the nature of their occupation. In addition to basic questions concerning demographics, questions were asked about the musician's principal instrument (some participants play more than one), the duration of exposure through practice and performance and exposure to other sources of noise.

Because of the small number of participants involved, musicians were divided in five groups according to their respective locations on the orchestra floor (See Figure 1) and similarity of the spectral profile of the sound they generate. Groups were as follows: Group 1 (Violins); Group 2 (Violas/Cellos); Group 3 (Woodwinds); Group 4 (Brasses); Group 5 (Percussion & Double Basses).

3. AUDIOMETRIC ASSESSMENT

After completing an audiological history and otoscopic examination a full audiometric battery including middle ear assessment was performed on all participants. Air conducted and bone conducted audiograms were obtained using a clinical audiometer with 5 dB steps. All measurements were conducted in a sound treated audiometric booth Speech testing (word recognition scores and speech reception thresholds), and admittance measures (tympanometry and acoustic reflexes) were also part of the evaluation. The results of the test were explained to each musician and a range of hearing loss prevention strategies, including the use of uniform attenuation earplugs was discussed.

4. RESULTS

4.1 Noise Exposure

Normalized $L_{ex,8\,hs}$ noise exposures levels were calculated and shown in Table 1 below.

Group	Average	St. Error
1	86.5	0.8
2	86.8	0.3
3	89.8	0.8
4	92.7	0.5
5	89.0	0.5

Table 1: Normalized Average and Standard Error noise exposures.

4.2 Questionnaire

Forty-four of the 52 musicians completed the questionnaire (85%). Twenty-one of the 44 respondents were female (48%). Average age of males was 51.7 years (SD = 11.1), while average age of females was 48.7 (SD = 9.9). Average age of participants was reasonably matched across the five groups.

Table 2 shows the numbers of years participants were playing in general and professionally. An analysis-of-variance (ANOVA) determined that the groups did not differ with respect to the length of their exposure.

Groups	1	2	3	4	5
Average playing	41.8	35.6	40.4	38.6	32.5
Average playing professionally	26.0	23.9	31.5	29.9	22.8

Table 2: Years of Playing

The majority of respondents reported listening to music through a speaker system (40). Many will also use earbuds with portable listening systems (21). Very few listen through circumaural headphones (5). Participants do not normally limit themselves to only one music listening device.

Finally, relatively few respondents reported involvement in noisy activities (9). However, it should be noted that a proper definition of "noisy activity" was not provided.

4.3 Audiometric Results

Figure 2 shows the average audiometric pure tone test results for the musicians. On average the data suggest only a slight to mild mid to high frequency sensory-neural hearing loss in the 4000-6000 Hz region. This is consistent with hearing losses observed in the earlier stages of other forms of noise exposure.

4.4 Measured and predicted hearing losses

The ISO 1999 Standard predicts hearing loss at different frequencies for males and females, according to age and the number of years of exposure at a given noise level. The algorithm presents the percentages of the population in 5% intervals.

Figure 3 shows the measured and calculated hearing losses for orchestra musicians. It may be observed that there are practically no differences between measured and calculated hearing losses at 3000, 4000, and 8000 Hz. Although the difference at 500 Hz is likely due to background noise in the audiometric booth, there is no obvious explanation for differences at 1000 and 2000 Hz. In any case, the differences between measured and predicted losses do not exceed the limits of measurement accuracy.

5. DISCUSSION

Pure tone audiometry showed that threshold varied as a function of instrument group and frequency region. Brasses and percussion/basses had the highest thresholds, bordering on clinically significant losses in the 4000-6000 Hz region. These differences across groups could not be explained by age, years of playing, or years of playing professionally, and are thus most likely due to differences in occupational noise exposure. Brass players had the highest noise exposure level (10 dB or greater than strings and woodwinds between 4000 and 8000 Hz). This finding is consistent with other noise-exposure surveys and audiometric investigations which raises some concern about long-term hearing health of brass players. Nonetheless, it is important to acknowledge that at the time of testing, none of the groups had hearing loss that would be considered outside the limits of normal hearing.

Noise exposure levels in the orchestra, normalized to 360 hr/year were below the hazard limit of 85 dBA with the exception of the brasses. Longer playing times will increase the risk. It seems reasonable to recommend that orchestras comparable to the National Ballet Orchestra adopt a Hearing Conservation Program (e.g., NIOSH, 1998), and that linear ear-plugs be considered for those orchestra members that are exposed to higher noise levels. On the basis of the current study, it appears that such interventions may be most necessary among brass players.

REFERENCES

Quian, C., Behar, A., Wong, W.: Noise exposure of musicians of a ballet orchestra. Noise & Health, Jan-Feb 2011, 13:50, 59-63.

ISO 1999 (International Organization for Standardization) (1990). Acoustics—Determination of occupational noise exposure and estimation of noise-induced hearing impairment. ISO TC/43 N1139. Revision of ISO 1999:1990. Second CD March 2010.

NIOSH (National Institute for Occupational Safety and Health) (1998). NIOSH criteria for a recommended standard: occupational noise exposure, revised Criteria 1998. DHSS(NIOSH) Publication No. 98–126. Cincinnati, OH: U.S. Department of Health and Human Services, Public Health Service, Center for Disease Control, National Institute for Occupational Safety and Health.

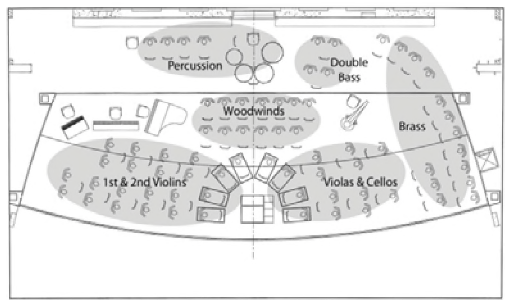


Figure 1: Location of the instruments groups on the orchestra floor

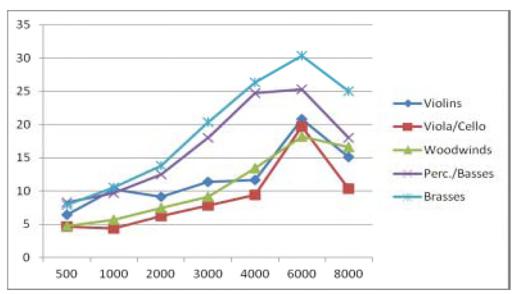


Figure 2: Average hearing loss of the different groups.

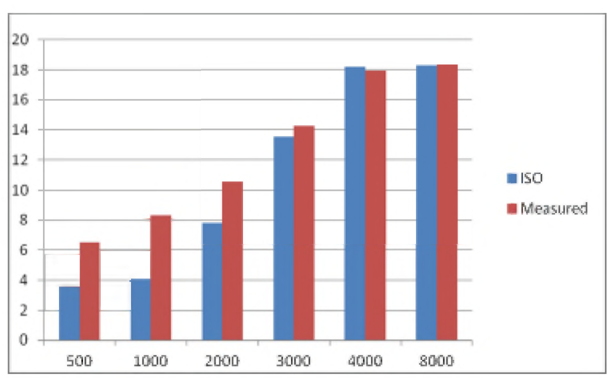


Figure 3. Measured and predicted hearing losses (as per ISO 1999) in dB.

ON THE INFLUENCE OF THE MATERIAL PROPERTIES OF THE EXTERNAL EAR ON OCCLUSION EFFECT SIMULATIONS

Martin K. Brummund[1], Franck Sgard[1,2], Yvan Petit[1], and Frédéric Laville[1]
[1]Dept. of Mechanical Engineering, École de technologie supérieure, 1100 Rue Notre Dame O., Montréal, Québec, Canada, H3C1K3 martin.brummund.1@ens.etsmtl.ca
[2]Service de la recherche, IRSST, 505 Boulevard de Maisonneuve O., Montréal, Québec, Canada, H3A3C2

1. INTRODUCTION

About 120 million workers worldwide are at risk of developing professional hearing loss (WHO, 2001). Hearing protection devices (HPD) such as earplugs (EP) represent the most frequently used short term solution to protect employees that are exposed to harmful noise levels. Nevertheless, research has also shown that workers often only tend to wear provided EPs for limited amounts of time (Berger, 2000) due to physical and auditory discomfort. One important source of auditory discomfort is the occlusion effect (OE). The OE occurs upon EP insertion and causes an uncomfortable distortion of the perception of the wearer's own voice and an amplification of physiological noises. Numerical modeling can contribute to further our understanding of the sound propagation in the external ear and ultimately the OE's underlying mechanisms. The present study aims at implementing a two-level fractional factorial design to examine how the material properties of the external ear tissues influence numerical predictions of the OE.

2. METHODS

The present study uses a 2D coupled linear elasto-acoustic finite element model for simulation of the OE. This simplified 2D model was developed in a previous study (Brummund et al., 2011) and successfully compared to an equivalent 3D model whose complex external ear geometry was reconstructed using 135 anatomical images of a female cadaver head (images were retrieved from the Visible Human Project® database of the US National Library of Medicine). Furthermore, the 2D model was validated against literature findings (Stenfelt & Reinfeldt, 2007).

2.1 Axis-symmetric model geometry

The axis-symmetric model comprises the domains: ear canal, a skin layer that covers the ear canal walls, a second skin layer that covers the lateral ear canal entrance region as well as soft tissue and bony tissue domains. A silicone EP of known material properties was inserted 7mm into the ear canal.

2.2 Material properties

Implementation of a coupled linear elasto-acoustic model of the external ear requires Young's moduli, Poisson's ratios, densities and loss factors of all tissue domains to be known. To examine the sensitivity of the model's response variable (OE) to varying material property configurations, representative high and low levels must be set for each experimental factor (material properties). This task is very challenging as the material properties of the external ear have, to date, only been sparsely determined. Central values (0) of experimental factors were drawn from the literature. Material properties that could not be found in the literature were approximated using identical tissues in close anatomical proximity which were assumed to be exposed to similar mechanical loading. Afterwards, high (+1) and low levels (-1) were defined by varying each factor ± 20% about its central value.

2.3 Boundary conditions

Outer circumferential (parallel to center axis of ear canal) boundaries of the skin and cartilage domains are fixed. Medial boundaries of the bone (which corresponds to the petrous part) and the skin tissue that surrounds the tympanic membrane are fixed. The lateral surface of the skin tissue that surrounds the ear canal entrance is modeled as free. For open ears, the ear canal entrance is modeled as acoustic radiation impedance of a baffled flat disc. For occluded ears, the lateral earplug boundary is left free. Finally, the eardrum impedance is modeled using a two piston-model (Shaw & Stinson, 1981; Shaw, 1977). A structure borne excitation (1N total force) is applied normally to the circumferential boundary of the bony tissue in both open and occluded models.

2.4 Fractional factorial design (2^k-p)

A 2^12-7 folded saturated factorial design with one replicate and one single block was implemented. The chosen design is of resolution IV. The generators were chosen so that the design is of highest possible resolution and of minimum aberration (Montgomery, 1997). All aforementioned steps were carried out using STATGRAPHICS (STATPOINT Technologies, Inc., USA)

2.5 Simulation of OEs and statistical analysis

The obtained sheet was imported into COMSOL 4.2 (COMSOL, Inc., Sweden). Parametric sweeps were carried out for all property configurations. Open and occluded models were meshed according to the four elements per wavelength criterion using quadratic tetrahedral elements. Obtained transfer function levels between excitation and acoustic pressure at the center of the tympanic membrane were calculated and 1/3 octave band filtered in MATLAB (MathWorks®, USA). Filtered results were imported into STATGRAPHICS TM for post-processing.

3. PRELIMINARY RESULTS

Analysis of variance indicates significant single factor effects for skin (Poisson's ratio, Young's modulus, loss factor), soft (Poisson's ratio, Young's modulus) and bony (Poisson's ratio) tissues. Estimated effect magnitudes vary as a function of frequency. Single factors of skin and soft tissues tend to have positive effects on simulated OEs. Overall, Poisson's ratio and Young's modulus of the skin tissue were estimated to have the biggest effects on the OE data. Statistically significant two factor interactions occurred throughout the entire frequency range. Again, estimated effect magnitudes vary over frequency. Overall, the interaction between Young's modulus and Poisson's ratio of the skin tissue tends to have the greatest estimated effect.

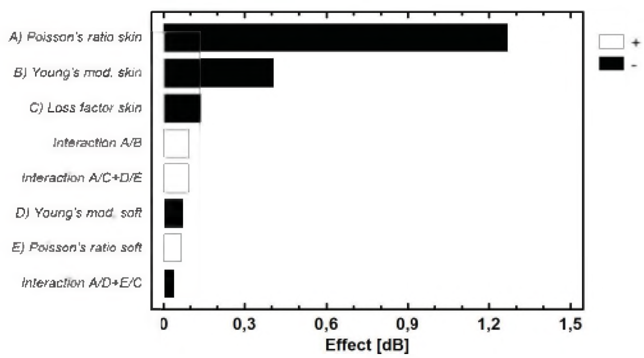


Figure. 1. Pareto chart for OE at 100Hz. All effects were significant (p-Values < 0.05). Filled bars denote positive effects. Empty bars denote negative effects.

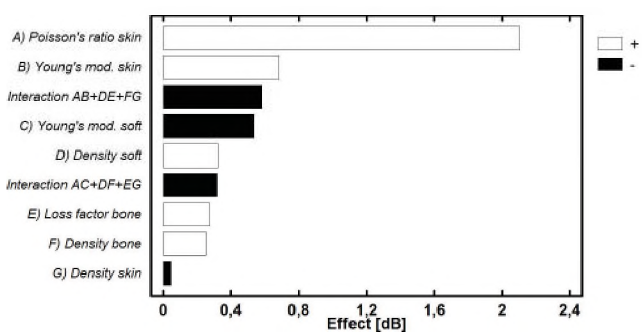


Figure. 2. Pareto chart for OE at 1000Hz. Apart from the skin density, all effects were significant (p-Values < 0.05). Filled bars denote positive effects. Empty bars denote negative effects.

In Figures 1 and 2 Pareto charts are provided. Estimated main effects correspond to the difference between the OE at the high and low levels of the factor when all other factors are kept at their central values. Apart from the density of the skin tissue at 1 kHz, all factors and two factor interactions are statistically significant ($\alpha = 0.05$).

4. DISCUSSION AND CONCLUSION

Skin, soft and bony material properties were found to contribute significantly to simulated OEs. Yet, mostly Poisson's ratio and Young's modulus of the skin tissue tend to exhibit effect estimates which are large enough to cause relevant variations in simulated OE data. These preliminary results suggest an important role of the external ear's skin tissue. This trend is in agreement with experimental findings (Stenfelt et al., 2003) which showed that the non-osseous external ear tissues represent the greatest source of ear canal sound pressure in open ears for frequencies below 2 kHz. Future studies should examine the contribution of the ear canal skin to ear canal sound pressure levels. It should be noted that the authors are anticipating a study to determine material properties of the external ear tissues experimentally. In this context, OE measurements with and without ear canal skin could be carried out. Furthermore, it would be desirable to examine whether relevant higher order factor interactions are existent. This can be achieved via factorial designs of higher resolution or through full factorial designs that consider only the most influential material properties.

REFERENCES

Berger, E. H. (2000). Hearing Protection Device Utilization Around the World. *Spectrum*, *17*(Suppl. 1), 18.

Brummund, M., Sgard, F., Petit, Yvan, & Laville, F. (2011). A simplified axi-symmetric finite element model of the human outer ear to determine the earplug induced auditory occlusion effect. *The Journal of the Acoustical Society of America* (Vol. 130, No. 4, p. 2469). Presented at the 162nd Meeting Acoustical Society of America, San Diego, CA, USA.

Montgomery, D. C. (1997). *Design and analysis of experiments*. New York: John Wiley and Sons.

Organisation mondiale de la santé (OMS). (2001). Le bruit au travail et le bruit ambiant.

Shaw, E. A. ., & Stinson, M. R. (1981). Network concepts and energy flow in the human middle ear. *J. Acoust. Soc. Am.*, *69*(1), 43.

Shaw, E. A. G. (1977). Eardrum representation in middle-ear acoustical networks. *J. Acoust. Soc. Am.*, *62*(1), S 12.

Stenfelt, S., & Reinfeldt, S. (2007). A model of the occlusion effect with bone-conducted stimulation. *International Journal of Audiology*, *46*(10), 595 – 608.

Stenfelt, S., Wild, T., Hato, N., & Goode, R. L. (2003). Factors contributing to bone conduction: the outer ear. *The Journal of the Acoustical Society of America*, *113*(2), 902–913.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support received from the IRSST.

EVOLUTION OF AUDIOMETRY: CLINICAL TESTING OF A NEW TABLET AUDIOMETER

Nicolas N Ellaham[1], Guy-Vincent Jourdan[1], Sandra Champagne[2], Jeff Yeung[2], and Matthew Bromwich[2]
[1]School of Electrical Engineering and Computer Science, University of Ottawa, 800 King Edward Av, ON, Canada, K1N6N5
[2] Div. of Pediatric Otol. Head & Neck Surgery, Children's Hospital of Eastern Ontario, 401 Smyth rd, ON, Canada, K1H8L1
nellaham@uottawa.ca, gvj@site.uottawa.ca, schampagne@cheo.on.ca, jeffyeung@gmail.com, mbromwich@cheo.on.ca

1. INTRODUCTION

Early detection of hearing loss in young children is critically important as hearing plays an essential role in the development of speech and communication skills [1]. Typical hearing tests, whereby sound is administered and the patient presses a button upon hearing it, pose a particular challenge for the paediatric population: getting children to cooperate and retaining their attention for the duration of the test is difficult. Conditioned play audiometry (or CPA) is a standard testing technique designed to address this problem. In a CPA test a child is conditioned to respond to hearing sounds at different frequencies with specific play behaviour: for example, when the sound is heard, the child is allowed to drive a car down a race track. However, CPA is resource intensive and only available in select locations. It is more time consuming than conventional audiometry, and requires special training and the availability of two audiologists to administer the listening test.

The development of tablet technology and the widespread popularity of the touch interface, especially among children, have made possible a new iOS-based application designed to measure hearing thresholds in this population [2]. The iPad® (Apple® Inc., Cupertino, CA) application helps overcome some challenges associated with CPA by introducing a novel paradigm to automate certain tasks. The benefits of automating audiometric tests using computer-based systems have been documented in recent studies [3]–[8]. Moreover, owing to the portability of the tablet environment, the application opens the door to low-cost, automated hearing diagnosis for children in new locations, such as developing countries where hearing loss is prevalent [9].

In this work, we present a clinical study to determine the accuracy of audiometric thresholds measured using tablet audiometry. The results are compared to those obtained with accepted conventional CPA.

2. METHODS

2.1. Tablet audiometry: application overview

Under the new paradigm, control over the presentation and pace of the sound stimuli is driven by the patient rather than the audiologist. The 'game' consists of presenting the subject with a series of objects (*e.g.* eggs) to be sorted in two containers: 'sound-producing' objects in one container (*e.g.* a chicken coop), and 'silent' objects in another (*e.g.* an egg carton). Thus, using a simple dragging motion, the child is able to navigate his/her own way through the test by responding in a yes/no fashion to each stimulus. The sound intensity decreases with each presentation until the child is not able to sort the object reliably, at which point the sound intensity is increased. This process allows the hearing thresholds to be determined using an up/down bracketing procedure [10]. Several "silent objects" are also presented randomly to provide a measure of internal consistency (reliability). The application interface provides different test stimuli (pure-tones, warble tones, and narrowband noise) for testing at the standard audiometric frequencies from 125 to 8000 Hz. Finally, once all selected frequencies are tested, a standard audiogram is obtained.

2.2. Test subjects

The present study was conducted on a population of 85 patients with normal or abnormal hearing, aged 3–16, at the Audiology Clinic at the Children's Hospital of Eastern Ontario. Patients were identified by a staff otolaryngologist (MB) after a review of their patient record. Informed consent was obtained from the parent/guardian at the time of enrollment. Of these patients, 15 were discarded after the assessment due to technical/game-play issues, behavioural issues or questionable reliability defined as incorrectly assigning silent objects more than 50% of the time. The demographics for the 70 subjects retained for this study are summarized in Table 1.

2.3 Study design

All participants completed two audiometric evaluations, one with the tablet audiometer and one with conventional CPA. The order of the test was determined at random. In both cases, measurements of warble-tone thresholds were performed inside a double-walled sound booth. For simplicity and speed, unmasked air-conduction thresholds at 4 test frequencies (500, 1000, 2000, and 4000 Hz) were measured for this study. Sound stimuli were presented using TDH-39 headphones for both assessments, testing the left and right ears separately. Children who were not amenable to wearing headphones (14 subjects) were tested in a sound-field using the sound booth speaker system or the tablet speaker. All equipments (tablets, headphones and speakers)

Table 1. Population demographics and reliability statistics.

	Abnormal hearing	Normal hearing
Number of subjects	15	55
Mean age	5.81 (range 3–13)	5.06 (range 3–9)
Reliability (%)	92.0 (SD: ±10.8)	90.4 (SD: ±22.4)
SD denotes standard deviation.		

Table 2. Comparison of tablet audiometry and conventional CPA

Tablet audiometry \ Play audiometry	**Abnormal hearing (15)**	**Normal hearing (55)**
Abnormal hearing (17)	14	3
Normal hearing (53)	1	52
Sensitivity: 93.3% (95%CI = 71.7–99.6%) Specificity: 94.5% (95%CI = 88.6–96.3%) Positive Predictive Value: 82.3% (95%CI = 63.3–87.9%) Negative Predictive Value: 98.1% (95%CI = 92.0–99.9%) Positive Likelihood Ratio: 17.1 (95%CI = 6.31–26.7)		

used in this study were calibrated according to ANSI S3.6-1996 (R2010) [11]. A trained audiologist accompanied the participants in both assessments to provide motivation.

3. RESULTS

Table 2 summarizes the results of the analysis based on 70 patients. In this study, a patient with abnormal hearing is defined as one who scored a threshold of 30 dB or more for at least one of the audiometric frequencies tested. Overall, 55 subjects were identified by the conventional CPA test to have normal hearing. Of these, 52 were correctly identified with tablet audiometry, the remaining 3 children scoring slightly outside the parameters defined for normal hearing. This appeared to be the results of the child moving too quickly through the game, or the presentation timing out before the child could make a decision. Moreover, a total of 53 patients were identified by tablet audiometry to have normal hearing, of which one child was found to have a true mild hearing loss. This child appeared to understand the game, but scored a low reliability of 75%, which, however, is still above the criteria for exclusion (<50%).

Preliminary statistical measures to evaluate the performance of tablet audiometry are also shown in Table 2. The data, especially the narrow confidence intervals which suggest sufficient statistical power, reveal that the tablet audiometer produces warble-tone thresholds that are in agreement with conventional CPA. Moreover, for patients who had separate left-right ear assessments (54 patients), a repeated measures model for the threshold in each ear at each frequency was fitted using linear mixed effects modelling. The model showed no significant effect of the assessment modality (tablet versus conventional CPA) for all patients, as well as for each group (normal versus abnormal hearing).

4. DISCUSSION AND CONCLUSION

This work summarizes the results of the first trial study of a new tablet audiometer. It is, to our knowledge, the first tablet-based semi-automated play audiometer to be used in a paediatric setting. The study is aimed at testing the novel interactive play algorithm used in the tablet audiometer: the user-directed paradigm requires more action (and decision) from the user than in standard audiometry, where no action is required when the user does not hear a sound. Despite this challenge, this work shows the tablet audiometer to be a child-friendly application, as the majority (82%) of children aged 3 and up were able to understand the concept of the game and complete the hearing assessment. Moreover, the 82% represents a conservative estimate since 4 out of the 15 patients excluded from the study failed to complete the assessment due to technical issues related to the audiologist. It is also worth noting that, of the remaining 11 patients excluded, 10 had abnormal hearing.

The data in this study demonstrate that air-conduction thresholds measured using the tablet audiometer were not significantly different from those obtained by standard CPA. With a strong predictive value for normal hearing, and high sensitivity for hearing loss, the portable tablet audiometer is shown to be an efficient and clinically accurate instrument for hearing assessment in children.

REFERENCES

[1] Canadian Association of Speech-Language Pathologists and Audiologists (2000), "Hearing health for children: Fact sheet", available from: http://www.speechandhearing.ca/files/children_hearing_fact_sheet.pdf. Accessed: July, 30, 2012.

[2] Ellaham NN, Yilma Y, Jourdan GV, and Bromwich M (2011): "A new iPad application for hearing screening in children", *Canadian Acoustics*, 39(3): 118–119.

[3] Margolis RH, Morgan DE (2008) "Automated pure-tone audiometry: an analysis of capacity, need and benefit", *Am. J. of Aud.*, 17(2), 109–113.

[4] Margolis RH, Glasberg BT, Creeke S, Moore BCJ. AMTAS® (2010): "Automated method for testing auditory sensitivity: validation studies", *Int. J. Aud.*, 49(3), 185–194.

[5] Margolis RH, Frisina R, Walton JP. AMTAS® (2011): "Automated method for testing auditory sensitivity: II. Air conduction audiograms in children and adults", *Int. J. Aud.*, 50(7), 434–439.

[6] Margolis RH, Moore BCJ. AMTAS® (2011): "Automated method for testing auditory sensitivity: III. Sensorineural hearing loss and air-bone gaps", *Int. J. Aud.*, 50(7), 440–447.

[7] McPherson B, Law MMS, Wong MSM (2010): "Hearing screening for school children: comparison of low-cost, computer-based and conventional audiometry", *Child: Care and Health Development*, 36(3) 323–331.

[8] Liao WH, Young ST, Lien CF, Wang SJ (2011): "An audio-meter to monitor progressive hearing change in school-aged children", *J. Med. Screening*, 18(1), 8–11.

[9] Tucci D, Merson MH, Wilson BS (2010): "A summary of the literature on global hearing impairment: current status and priorities for action", *Otology and Neurotology*, 31(1):31–41.

[10] Westlake W, HD H (1944): "Manual for program outline for rehabilitation of aural casualties both military and civilian", *Trans. Am. Academy Opht. and Otol.*, 48(supp.), 1–15.

[11] American National Standards Institute (1996), "Specification for audiometers" (ANSI S3.6-1996 (R2010)), New York: Acoustical Society of America.

Human Cochlear Maps

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

1. INTRODUCTION

In the present context, the term "map" designates a function $x(f)$, where f is the frequency of a sinusoidal tone and x is a related distance from the cochlear base, measured along the cochlear basilar membrane. Presumably, the term "cochlear map" is commonly used because "cochlear function" might be misunderstood as "the way in which the cochlea works". The purpose of cochlear maps is to contribute to a better understanding of cochlear function.

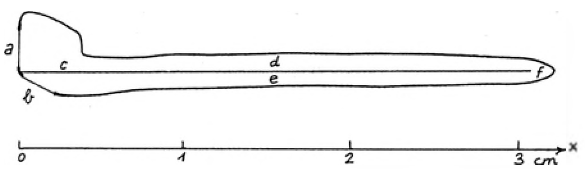


Figure 1. Profile of the unrolled human cochlear channel; a) oval window, b) round window, c) basilar membrane, d) scala vestibuli and scala media (Reissner's membrane is not shown), e) scala tympani, f) helicotrema.

The human cochlea is a liquid-filled helical channel in the temporal bone. An unrolled profile of that channel is shown in Fig. 1. For an adequate description of cochlear physiology, four cochlear-map categories, as specified in Subsections 1.1–1.4, are required. Fig. 41.1 of Frosch (2010a), reproduced in Fig. 2, contains preliminary human cochlear maps. Modified human cochlear maps, first presented in Fig. 4 of Frosch (2010b), are reproduced in Fig. 3. The aim of the present note is to justify that modification.

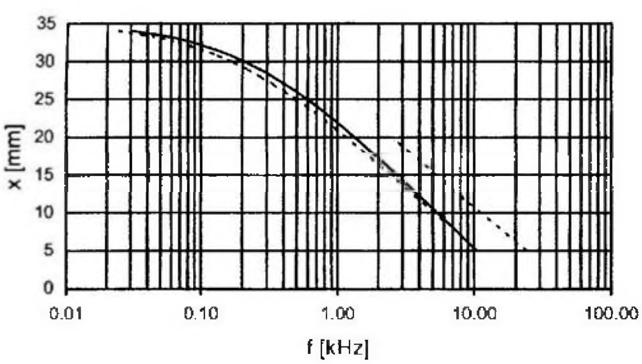


Figure 2. Preliminary human cochlear maps, reproduced from Fig. 41.1 of Frosch (2010a); see text.

1.1 The passive-peak (PP) map

Passive human cochleae (i.e., cochleae in which the mechanical energy generated by active outer hair cells is absent) were studied, e.g., in the post-mortem experiments of von Békésy (1960), who found that during a stationary sinusoidal tone there is a travelling wave in the cochlear channel. In most of those experiments, wave energy was fed into the cochlea through the oval window. That wave energy (kinetic and potential energy of the liquid particles and of the cochlear-partition elements) travels in the $+x$-direction (see Fig. 1). In spite of the fact that this travelling-wave energy starts to be transformed into frictional heat at the cochlear base already, the velocity amplitude of the basilar-membrane (BM) oscillation increases with x. At the passive-peak place $x_{PP}(f)$, that increase stops, and at greater x the amplitude quickly drops to small values. At high [low] f, $x_{PP}(f)$ is small [large], as shown in Fig. 2, where $x_{PP}(f)$ is given by the lower dashed curve, ranging from 0.025 to 6 kHz.

1.2 The low-level active-peak (AP) map

At high f and low sound-pressure level (SPL), the active BM oscillation velocity peak can be much higher and sharper than the just described passive peak. The active peak arises because in a healthy cochlea the outer hair cells (OHCs) near $x_{PP}(f)$ feed mechanical energy into the cochlear travelling wave. The AP-map place $x_{AP}(f)$, shown by the solid curve in Fig. 2, is defined to be the active-peak place in a healthy cochlea at SPL < 20dB. The maximal power that healthy OHCs can feed into cochlear travelling waves is so small that its influence on the place of maximal BM oscillation velocity is negligible at SPL > 100dB; i.e., above 100 dB the BM oscillation velocity peak is close to the passive peak even in a healthy cochlea.

1.3 The basilar-membrane resonator (BMR) map

That map, i.e., the function $x_{BMR}(f)$, shown in Fig. 2 by the upper dashed curve, ranging from 2.5 to 24 kHz, gives the place of that organ-of-Corti slice (BM element and attached cells) which in a cochlea without liquid above and below the partition would have a free-oscillation frequency of f. During the oscillations of the BM resonator, the center-of-mass of the considered organ-of-Corti slice oscillates in a direction perpendicular to the BM. A detailed discussion of the BMR map is given in Chapter 34 of Frosch (2010a).

1.4 The internal organ-of-Corti resonator (IOCR) map

The IOCR-map $x_{IOCR}(f)$, similarly to the BMR map, defines the place of that element of the internal organ-of-Corti (IOC) resonator which in a cochlea without liquid above and below the partition would have a free-oscillation frequency of f. During the oscillations of the IOC resonator, the *shape* of the considered organ-of-Corti slice varies periodically. As discussed in Chapter 41 of Frosch (2010a), the human IOCR map is conjectured to coincide, for f > 1 kHz, with the PP map (lower dashed curve). The IOC resonator enables the outer hair cells to feed energy into cochlear travelling waves and thus to give rise to the active BM oscillation velocity peak. The main source of that energy is conjectured to be the electric current which flows through the outer hair cells and is modulated by the stereocilia of those cells.

2. THE MAPS OF FROSCH (2010a)

These preliminary human cochlear maps are shown in Fig. 2 above. The lower dotted curve (0.025–6 kHz) is taken from Fig. 1 of Greenwood (1990); that Greenwood map, predominantly based on post-mortem measurements, is conjectured to be close to the passive-peak (PP) map.

I have not found published direct experimental determinations of the human active-peak (AP) map. The solid curve in Fig. 2 (0.03–10 kHz) is the result of an attempt to derive the AP map indirectly from published results of psychoacoustic experiments; see Chapter 41 of Frosch (2010a). Briefly, Eq. (10.29) and Fig. 10.10 of Hartmann (1998) yield the "Cambridge" critical-band number $z(f)$; for instance, $z(0) = 0$, and $z(20\text{kHz}) = 41.5$. The solid curve in Fig. 2 is based on the assumption that the function $x(z)$ is exactly linear [Eq. (41.3) of Frosch (2010a)]; e.g., $x(z=0) =$ 35mm, and $x(z=41.5) = 0$. As stated in Frosch (2010a), the just mentioned assumption, and therefore the AP-map curve in Fig. 2 above, cannot be claimed to be accurate.

3. THE MAPS OF FROSCH (2010b)

Thanks to in-vivo experiments, the cochlear maps of several mammals are known better than those of humans; see Chapters 38 (guinea-pig), 39 (chinchilla), and 40 (Mongolian gerbil) in Frosch (2010a). These mammalian maps are similar to those in Fig. 2; the mammalian active-peak (AP) map curves at small x, however, are about halfway between the PP and the BMR map curves, so that at given (small) x the AP-map frequency $f_{\mathrm{AP}}(x)$ is higher, by half an octave, than $f_{\mathrm{PP}}(x)$. The corresponding half-octave difference between $f_{\mathrm{IOCR}}(x)$ [which is postulated to be close to $f_{\mathrm{PP}}(x)$] and $f_{\mathrm{AP}}(x)$ is plausible, since the (–10dB) width of the IOCR resonance peak amounts to about one octave; the mentioned half-octave frequency difference implies that the low-level active BM-oscillation velocity peak is located at the large-x limit of the zone of the strongly active outer hair cells.

The modified human cochlear maps, represented in Fig.3, are based on the hypothesis that the human maps resemble those of the mentioned mammals; in particular, the AP-map frequency $f_{\mathrm{AP}}(x)$ is assumed to be higher by half an octave than $f_{\mathrm{IOCR}}(x)$. The PP-map in Fig. 3 is intended to range from 0.025 to 6 kHz; the IOCR-map is postulated to coincide with the PP-map, but to range from 1 to 6 kHz only. At frequencies to the left and to the right of the curves in Fig. 3 the maps are concluded to be not yet known.

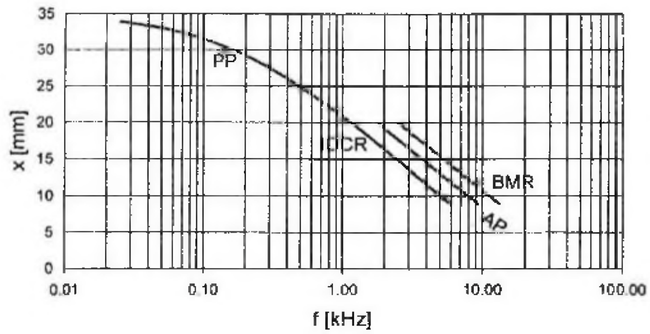


Figure 3. Modified human cochlear maps, reproduced from Fig. 4 of Frosch (2010b); see text.

4. CONCLUSIONS

In the modified human-map diagram (Fig. 3) the active-peak (AP) map, giving the location of the basilar-membrane oscillation velocity peak generated in a healthy cochlea by a sinusoidal tone of low sound-pressure level, differs from that in the previously published diagram (Fig. 2). That change is based on the hypothesis that the cochlear physiology of humans closely resembles that of guinea-pigs, chinchillas, and Mongolian gerbils.

REFERENCES

Von Békésy, G. (1960). Experiments in Hearing. McGraw-Hill, New York, pp. 403-484.

Frosch, R. (2010a). Introduction to Cochlear Waves. vdf, Zurich, pp. 331-386.

Frosch, R. (2010b). Analysis of Human Oto-Acoustic Emissions. Canadian Acoustics, Vol.38, No.3, 88-89.

Greenwood, D.D. (1990). A cochlear frequency-position function for several species – 29 years later. J. Acoust. Soc. Am. 87, 2592-2605.

Hartmann, W.M. (1998). Signals, Sound, and Sensation. Springer, New York, pp. 251, 253.

ANALYSIS OF HUMAN TONE-BURST-EVOKED OTOACOUSTIC EMISSIONS

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

1. INTRODUCTION

In one of my contributions to AWC-2010 in Victoria BC [Frosch (2010a)] it was shown that human click-evoked otoacoustic emissions (OAEs) documented in the literature agree with predictions based on cochlear maps. In a paper presented at the Forum Acusticum 2011 in Aalborg [Verhulst (2011)], human tone-burst-evoked OAEs appreciably different from those in Frosch (2010a) were shown. The purpose of the present study has been to find out if the OAEs in Verhulst (2011), too, agree with cochlear-map-based predictions.

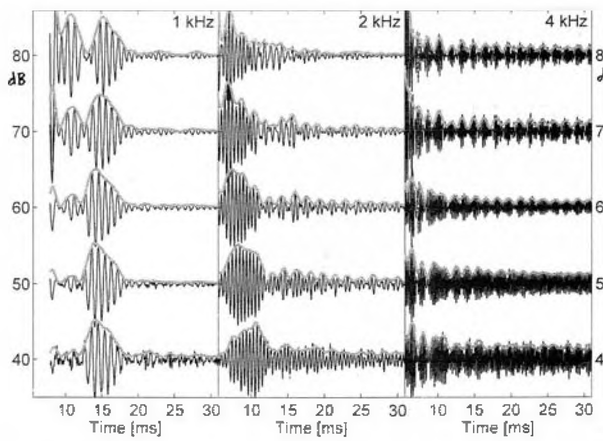


Figure 1. Waveforms of human tone-burst-evoked OAEs, reproduced with permission from Verhulst (2011).

The waveforms shown in Fig. 1 were generated by a tone burst emitted by an earphone inserted in an ear canal, and were recorded by a microphone; see Verhulst (2011) for details.

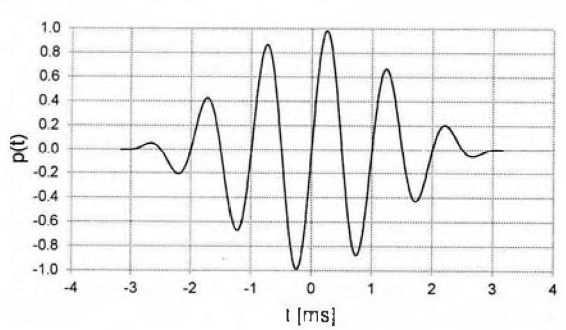


Figure 2. Sound-pressure function of the Hann-windowed tone burst used for the 1-kHz part of Fig. 1

The envelope of the tone burst in Fig. 2 is proportional to the squared cosine of ($\pi t/T$), where T = 6.3 ms.

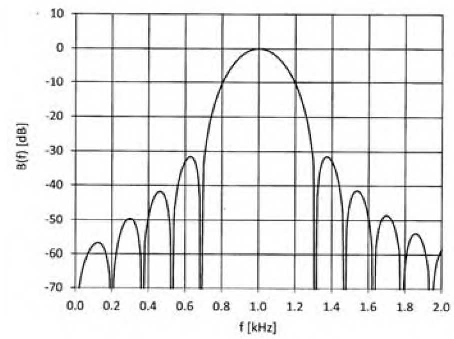


Figure 3. Spectral function of the tone-burst in Fig.2.

The corresponding spectral function, shown in Fig. 3, contains satellite peaks at –31dB, –41dB, etc.

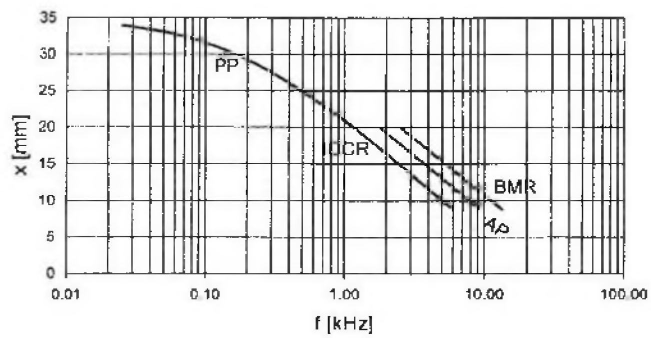


Figure 4. Human cochlear maps, reproduced from Fig. 3 of Frosch (2012).

In Fig. 4, four human cochlear maps are shown, namely the passive-peak (PP) map, the (low-level) active-peak (AP) map, the basilar-membrane resonator (BMR) map, and the internal organ-of-Corti resonator (IOCR) map. The IOCR map is hypothesized to coincide with the PP-map between 1 and 6 kHz, and is concluded to be not yet known at frequencies above and below that range. Cochlear maps are discussed in Chapters 33–41 of Frosch (2010b); a brief introduction is given in Frosch (2012).

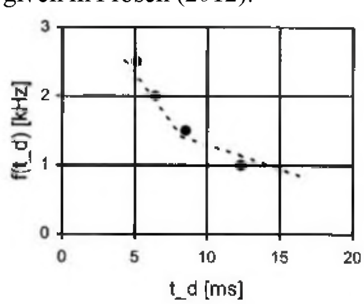


Fig. 5. Instantaneous click-evoked OAE frequency f versus OAE delay t_d; reproduced from Fig. 3 of Frosch (2010a).

The dashed line in Fig. 5 represents the instantaneous frequency of the human click-evoked OAEs according to Fig. 16.20 of Fastl and Zwicker (2007) [reproduced in Fig. 1 of Frosch (2010a)]. The filled circles in Fig. 5 represent the theoretical OAE delay $t_d(f) = 2\tau_{SW} + \tau_{rise}$; here, τ_{SW} is the surface-wave group travel time, from $x = 0$ to $x_{IOCR}(f)$ according to a box-model short-wave formula [Eq. (2) of Frosch (2010a)], and $\tau_{rise} = Q/(\pi \cdot f) \approx 1.3/f$ is the rise time of the forced oscillations of the internal organ-of-Corti resonators (IOCRs), which have a quality factor (resonance frequency divided by –3 dB resonance width) of $Q \approx 4$.

At $f = 1$ kHz, e.g., the IOCR-map place is $x_{IOCR} = 20.9$ mm (Fig.4), the just defined time durations are $\tau_{SW} = 5.5$ ms and $\tau_{rise} = 1.3$ ms, and the resulting OAE delay is $t_d = 12.3$ ms, as indicated in Fig. 5. A more detailed explanation of Fig. 5 is given in Chapter 44 of Frosch (2010b).

2. INTERPRETATION OF FIG. 1

1-kHz part, Time = 12-18 ms: OAEs transported by reverse travelling waves (TWs), emitted by OHCs in the 1-kHz IOCR resonance region (20 mm < x < 22 mm); center of primary tone burst is at Time = 3.15 ms; center of OAE burst is predicted to occur at Time = 3.15 ms + 12.3 ms = 15.45 ms, in agreement with Fig. 1.

1-kHz part, 80dB, Time = 7-14 ms: OAEs transported by reverse TWs, emitted by OHCs in the IOCR resonance regions of the high-frequency satellite peaks (1.37 kHz, 1.54 kHz, etc.; Fig. 3), superimposed on a decaying 1-kHz oscillation generated by the primary tone burst, but not due to an OAE. The satellite peaks of an 80-dB tone burst have sound-pressure levels (SPLs) of 49dB, 39dB, etc., and are therefore strongly amplified by the OHCs in their IOCR resonance regions. A part of the mechanical energy generated by these OHCs is transported back to the cochlear base by a reverse TW and thus causes OAEs.

1-kHz part, 80 dB, Time > 18 ms: Here, the OAEs due to the low-frequency satellite peaks (0.63 kHz, 0.46 kHz, etc.) are expected; they are negligibly weak, however, because the OHCs feed little mechanical energy into low-frequency TWs; see Chapter 35 of Frosch (2010b).

2-kHz part, 80 dB, Time = 11-15 ms: Now, low-frequency satellite peaks are at 1.40 kHz, 1.13 kHz, etc.; at these frequencies, the OHCs do feed significant mechanical energy into TWs. The corresponding low-frequency OAE at Time = 11-15 ms is clearly visible in Fig. 1.

1-, 2-, and 4 kHz parts, 60-80 dB: At large delays, these OAEs form stationary beats, different from the mentioned click-evoked OAEs in Fig. 1 of Frosch (2010a), where a stationary 3-kHz oscillation, attributed to spontaneous OAEs (SOAEs) triggered by the click, was observed. The large-delay beats in the 1-kHz part of Fig. 1 are consistent with being due to the superposition of two SOAEs of 1.4 and 1.6 kHz. As an alternative to Shera (2003), these SOAEs are conjectured to be due to OHCs feeding energy into localized basilar-membrane (BM) oscillations in the 1-kHz IOCR resonance region (20mm < x < 22mm). In that region, the without-liquid BMR resonance frequencies range from 1.7 to 2.4 kHz (Fig. 4). In a real (liquid-filled) cochlea, the just mentioned BM oscillations are thought to involve standing evanescent liquid-pressure waves. As discussed in Frosch (2011), the frequency of such a with-liquid BM oscillation is lower than that of the local without-liquid oscillation by typically 0.24 octave, i.e., by a factor of $2^{-0.24} = 0.85$. Thus in the present case the predicted SOAE frequencies range from 1.4 to 2.0 kHz, and so are consistent with the mentioned experimental SOAE frequencies of 1.4 and 1.6 kHz.

f_0 [kHz]	f_1 [kHz]	f_2 [kHz]
1.0	1.4	1.6
2.0	1.5	2.1
4.0	3.4	4.1

Table 1. SOAE frequencies f_1, f_2 consistent with the stationary beats at large delays in Fig. 1; f_0 is the central frequency of the primary tone burst.

The large-delay beats in the 2- and 4-kHz parts of Fig. 1 are consistent with being due to the superposition of two SOAEs as listed in Table 1. These frequencies f_1, f_2 are below the range expected for SOAEs caused by OHCs feeding energy into BM oscillations in the IOCR resonance regions of f_0; they agree, however, with predictions based on the IOCR resonance regions of the highest satellite-peak frequencies below f_0, which for f_0 = 2 or 4 kHz are so high that the generated TWs are strongly amplified by the involved OHCs.

3. CONCLUSION

The tone-burst-evoked OAEs in Fig. 1, in spite of differing appreciably from the click-evoked OAEs in Fig. 1 of Frosch (2010a), are found to be compatible with predictions based on the human cochlear maps represented in Fig. 4.

REFERENCES

Fastl, H., Zwicker, E. (2007). Psychoacoustics, Facts and Models. Springer, Berlin, p. 335.

Frosch, R. (2010a). Analysis of Human Oto-Acoustic Emissions. Canadian Acoustics, Vol. 38, No. 3, 88-89.

Frosch, R. (2010b). Introduction to Cochlear Waves. vdf, Zurich, pp. 331-386.

Frosch, R. (2011). Cochlear Evanescent Liquid Sound-Pressure Waves During Spontaneous Oto-Acoustic Emissions. Canadian Acoustics, Vol.39, No.3, 122-123.

Frosch, R. (2012). Human Cochlear Maps. Contribution to this conference.

Shera, C.A. (2003). Mammalian spontaneous otoacoustic emissions are amplitude-stabilized cochlear standing waves. J. Acoust. Soc. Am. 114, 244-262.

Verhulst, S. (2011). Investigating the Periodicity of Transient-Evoked Otoacoustic Emission Envelopes. Proceedings of Forum Acusticum 2011, Aalborg, pp. 1235-1240.

AUDITORY SPATIAL ATTENTION IN A COMPLEX ACOUSTIC ENVIRONMENT WHILE WALKING: INVESTIGATION OF DUAL-TASK PERFORMANCE

Sin Tung Lau[1], Jacob Maracle[1], Dario Coletta[1], Gurjit Singh[1,2], Jennifer Campos[2], and M. Kathleen Pichora-Fuller[1,2]
[1]Dept. of Psychology, University of Toronto Mississauga, 3359 Mississauga Road N., Mississauga, ON, Canada, L5L 1C6
[2]Toronto Rehabilitation Institute, 550 University Ave., Toronto, ON, Canada, M5G 2A2

1. INTRODUCTION

In many everyday situations, listeners must attend to one sound while dividing attention to also listen to other sound sources or while trying to ignore irrelevant competing sound sources. Furthermore, they may need to listen while dividing their attention between the task of listening and non-auditory tasks (e.g., driving a car while talking on a cell phone). In an earlier study, we measured the effect of spatial attention on word recognition in younger and older adults using Coordinate Response Measure (CRM; Bolia, Nelson, Ericson, & Simpson, 2000) stimuli presented in a display of three spatially separated "talkers" (Singh, Pichora-Fuller & Schneider, 2008). One important finding was that all participants performed better when they had prior knowledge of callsign identity (Singh et al., 2008).

1.1 Psychoacoustic studies

Psychoacoustic and speech experiments, including the study by Singh et al., are usually conducted in soundbooths. In daily life, however, we encounter more complex, less controlled listening situations. In a pilot study, the testing paradigm employed by Singh et al. was adapted and examined in a new, state-of-the-art Virtual Reality (VR) laboratory facility, called StreetLab. In StreetLab it was possible to incorporate visual and 3D auditory cues in the display (see Maracle, Lau, Coletta, Singh, Pichora-Fuller, & Campos, 2012). Maracle et al. (2012) found that three healthy, young adults had similar performance when tested in a sound-attenuating booth as when tested in StreetLab. The replication in StreetLab of the results obtained earlier in typical soundbooth conditions established a baseline against which we could compare the effects of increasing task demands during listening in even more realistic scenarios.

1.2. Current experiments

Another level of complexity that exists in the real world involves executing and coordinating multiple tasks and competing priorities at the same time. Therefore, the current experiment expands upon the findings of Maracle et al. (2012) by introducing a concurrent walking task. The purpose of this study was to compare word identification and the effects of spatial attention when listening was either the sole task or when listening and walking were concurrent dual tasks. These two tasks are arguably the most ubiquitous of dual tasks performed in everyday life. In particular, the scenario selected for the study was intended to represent the real life situation of waiting to cross or walking across a major city road intersection, while listening to one talker with two competing talkers nearby.

2. METHOD

2.1 Participants

Three undergraduate students, ages 19 to 26 years, with normal pure-tone air-conducted hearing thresholds (≤ 25 dB HL) for frequencies from 0.25 to 8 kHz, performed a word identification task in two experimental conditions. All participants were fluent speakers of English who began learning English prior to the age of 5 (participants A and B were bilingual, while participant C was monolingual).

2.2 Apparatus

Testing was conducted in StreetLab in the Challenging Environment Assessment Laboratory (CEAL) at the Toronto Rehabilitation Institute. A 240^0 horizontal curved projection screen surrounded participants. A high-resolution virtual representation of a major city road intersection was displayed on the screen. No vehicle traffic or pedestrians were simulated. Sound was presented from seven loudspeakers embedded behind the projection screen in the horizontal plane at head height. The distance from the loudspeakers to the participant was 2.14 m. As seen in Figure 1, a treadmill is located in the floor of StreetLab. The display is stationary when the participant is standing still and the treadmill is not activated. When the treadmill is activated, the visual and auditory displays are updated so as to simulate changes in the scene as would occur during walking.

Figure 1: Interior of StreetLab

2.3 Stimuli

The stimuli for the listening task were all sentences recorded by four male talkers for the Coordinated Response Measure (CRM; Bolia et al., 2000). The sentences took the format "Ready [callsign], go to [colour][number] now" (Bolia et al., 2000). There were eight different callsigns and numbers, and four different colours. Each sentence was presented at 60 dB A. Three sentences were presented simultaneously in each trial; one sentence was the target and the other two were competing sentences. Each sentence was displayed at a different simulated location, with the target sentence at 0° azimuth (in front) and the two competing sentences at ±90° (right/left).

2.4 Design and Procedures

All participants completed 8 sessions in each of two conditions; standing and walking. The standing and walking conditions differed in terms of whether or not there was a secondary task (walking) during listening. Testing in the standing condition was completed before testing in the walking condition (two participants had also been tested in the study of Maracle et al., 2012).

Standing condition. In the standing condition, the treadmill was not activated and the visual and auditory displays were spatially static. Participants remained positioned at the southwest corner of the simulated intersection for all trials. On the first test day, participants completed a set of 30 practice trials in the 1.0 location certainty condition, followed by 30 practice trials in the 0.6 condition. On subsequent test days, participants completed 15 practice trials in the 0.6 location certainty condition.

Walking condition. At the start of each testing block in the walking (dual task) condition, participants were positioned at the southwest corner of the simulated intersection and used a joystick to activate the treadmill to begin the simulation of walking across the road. Once they reached the southeast corner of the intersection, the display was repositioned back to the starting point. The treadmill was set at a maximum speed of 1 metre per second (m/s). Prior to testing in the walking condition, participants practiced walking on the treadmill and maneuvering the joystick for 5 minutes. Following this, participants completed 15 practice trials in the 0.6 location certainty condition in the walking condition.

In each session there were four blocks of 30 test trials. On each trial, one of eight possible callsigns was displayed visually on the center of the screen in StreetLab for 3 sec followed by a 2 sec pause, followed by the presentation of a target and two competing sentences. Listeners were told to listen for the sentence containing the callsign which had been visually displayed and to report the colour and number from that sentence. Once they verbally responded, their response was typed into a tablet by the experimenter and the next trial was initiated.

Four location certainties were tested (1.0, 0.8, 0.6, 0.33) corresponding to the likelihood that the target stimulus would originate from each of the three simulated spatial locations. In the 1.0 location certainty condition, the target was presented from the "front" location 100% of the time; whereas in the 0.8 location certainty condition, there was an 80%, 10%, and 10% probability that the target would originate from the front, left, and right locations, respectively. At the start of each block of trials, participants were informed verbally of the certainty condition to be tested. Within a session, each of the four location certainty conditions was tested in a random order, with one block used to test each certainty condition.

The tester entered participant's responses on a tablet computer and the participant was informed of the accuracy of their response by a "correct" or "incorrect" that was displayed on the screen after each trial.

3. RESULTS

In all conditions, word identification accuracy decreased as certainty about the location of the target decreased. The pattern of results was virtually identical in the standing and walking conditions. Furthermore, the results of the participants tested in the current study all fall within one standard deviation of the group means found previously (Singh et al., 2008) for 24 young adults tested in a typical soundbooth environment (see Figure 1).

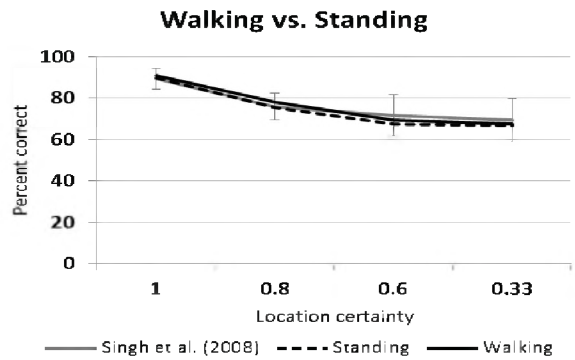


Figure 2. Percent correct as a function of location certainty.

4. DISCUSSION

The high degree of similarity in listening task performance between the standing and walking conditions suggest that the listening abilities in multi-talker environments of these participants were not affected by the inclusion of a walking component. The dependent measure in this study was target identification accuracy in the listening task. It would be of interest to include measures of gait in future studies to investigate whether maintaining a similar level of listening performance in the walking condition was associated with deviations in gait. Parameters of interest related to gait include measures of walking stability such as cadence, speed, step length, step width and head sway. In addition, it would be interesting to examine the performance of healthy older adults and younger and older adults with hearing on this dual task.

REFERENCES

Bolia, R., Nelson, W., Ericson, M., & Simpson, B. (2000). A speech corpus for multitalker communications research. *Journal of the Acoustical Society of America*, *107*, 1065-1066.

Maracle, J., Lau, S., Coletta, D., Singh, G., Pichora-Fuller, M., & Campos, J. (2012). A comparison of spatial listening in a soundbooth versus an immersive virtual environment. *Canadian Acoustics*. This issue.

Singh, G., Pichora-Fuller, M.K., & Schneider, B. (2008). The effect of age on auditory spatial attention in conditions of real and simulated spatial separation. *Journal of the Acoustical Society of America*, *124*, 1294-1305.

ACKNOWLEDGEMENTS

The study was funded by an NSERC grant to author MKPF. The Canadian Foundation for Innovation partly funded the StreetLab in the iDAPT (Intelligent Design for Adaptation, Participation and Technology) Centre for Rehabilitation Research; www.torontorehab.com/Research/Facilities/Labs/StreetLab.aspx). Thank you to Bruce Haycock for programming assistance.

A COMPARISON OF SPATIAL LISTENING IN A SOUNDBOOTH VERSUS AN IMMERSIVE VIRTUAL ENVIRONMENT

Jacob Maracle[1], Sin Tung Lau[1], Dario Coletta[1], Gurjit Singh[1,2], M. Kathleen Pichora-Fuller[1,2], and Jennifer Campos[2]
[1]Dept. of Psychology, University of Toronto, 3359 Mississauga Rd N, Mississauga, Ontario, Canada, L5L 1C6
[2]Toronto Rehabilitation Institute, 550 University Ave., Toronto, Ontario, Canada, M5G 2A2

1. INTRODUCTION

The soundbooth is usually considered to be the ideal location for psychoacoustic studies. Listening studies that have involved complex or multi-talker listening tasks have also typically been conducted in the controlled acoustical space a soundbooth provides. Transposing a psychoacoustic study from its original controlled soundbooth environment to a more complex one has always been challenging. As the environment becomes more complex, testers need to be concerned about maintaining a level of control over extraneous variables. There is, however, great advantage in conducting listening studies under conditions that more closely approximate real world interactions. Such conditions would allow investigators to better understand whether the results of soundbooth studies can be generalized to real life.

1.1. Earlier Research

In earlier research, a multi-talker listening task was completed in a soundbooth. Participants were asked to identify words in a target sentence when it was presented simultaneously with two other sentences, each sentence originating from a separate spatial location (Singh, Pichora-Fuller & Schneider, 2008; see also Kidd, Arbogast, Mason & Gallun, 2005). Participants' performance was affected by the certainty with which the target was presented from a known location. When location certainty was high (i.e., 100% chance of the target being presented from a specified location), identification performance was higher than when the location was uncertain (i.e., 33% chance of it being presented from each of three possible spatial locations).

1.2. Current Experiments

The objective of our study was to see if the results from Singh et al. (2008) could be replicated in an immersive, virtual reality (VR) environment. The first part of this study was completed in the controlled environment of a soundbooth at the University of Toronto at Mississauga (UTM) campus. The second part of this study was identical to the first part, but took place in the Toronto Rehabilitation Institute's new state-of-the-art VR lab – "StreetLab".

2. METHODS

2.1. Participants

The participants were three English speaking, healthy young adults with normal hearing (pure-tone air-conduction thresholds < 25 dB HL at 0.25, 0.5, 1, 2, 3, 4, and 8 kHz).

2.2. Stimuli

A target sentence and two competing sentences drawn from the sentences recorded by four males for the Coordinate Response Measure (Bolia et al., 1999) were presented concurrently to the listeners from three different locations (either to the left, right, or in front of the listener). The sentences had the format "Ready [callsign], go to [colour] [number] now", where there were eight different callsigns and numbers, and four different colours.

2.3. Procedures

All participants completed 8 sessions in each of two conditions. The two conditions differed in terms of the test environment (soundbooth or VR). Each session consisted of 4 blocks of 30 trials. There were four location certainty conditions (1.0, 0.8, 0.6, 0.33). Within a session, each of the four location certainty conditions was tested in random order, with one block used to test each certainty condition. At the start of each trial, two pieces of information were presented visually to the participant. The first was the callsign, which the listener used to identify the target sentence. The second was one of the four possible probability specifications, which indicated the likelihood of the target being presented from the left, centre, or right locations, respectively (0-100-0; 10-80-10; 20-60-20; 33-33-33). For example, "10-80-10" indicated that there was an 80% chance of the target being presented from in front of the listener and a 10% chance of the target being presented from either the left location or the right location. The listener's task was to report the colour and number from the sentence that contained the target callsign. In both the soundbooth and VR conditions, feedback regarding the percentage of correct trials was provided at the end of each block. Each sentence was presented at 60 dB A. Participants completed two practice blocks at the 1.0 and 0.6 location certainty conditions at the beginning of each testing day.

<u>Condition 1: Soundbooth</u>

Condition 1 was a replication of the real spatial separation condition in the study of Singh et al. (2008). Participants were seated in a soundbooth and told to face forward for the duration of testing. The three loudspeakers in the soundbooth were positioned at ±54° and 0° azimuth at a distance of 1.83m from the participant.

Condition 2: Virtual Reality Environment (StreetLab)

Condition 2 was conducted in the Toronto Rehabilitation Institute's StreetLab (See Fig. 1). A 240^0 horizontal curved projection screen surrounded participants. A high-resolution virtual representation of a major city road intersection was displayed on the screen. No traffic or pedestrians were included in the simulation. Sound was presented from seven loudspeakers embedded behind the projection screen in the horizontal plane at head height. Auditory stimuli were presented at ±90° (to the right or left) or 0° azimuth (in front). The distance from the loud-speakers to the participant was 2.14 m. Participants remained standing in one position for the duration of the experiment and the visual scene was spatially static.

Figure 1: Interior of StreetLab

3. RESULTS

3.1. Soundbooth vs. StreetLab

Overall, similar results were observed in both the soundbooth and StreetLab conditions (see Fig. 2). In both test environments, performance was best when listeners were certain about the location of the target and performance declined as location certainty decreased. Furthermore, the largest difference in performance between the two test environment conditions was only 4.6 percentage points (for the 0.8 location certainty condition). Finally, the mean results from the current study are within 1 SD of those found in the previous experiment conducted in a soundbooth (Singh et al., 2008).

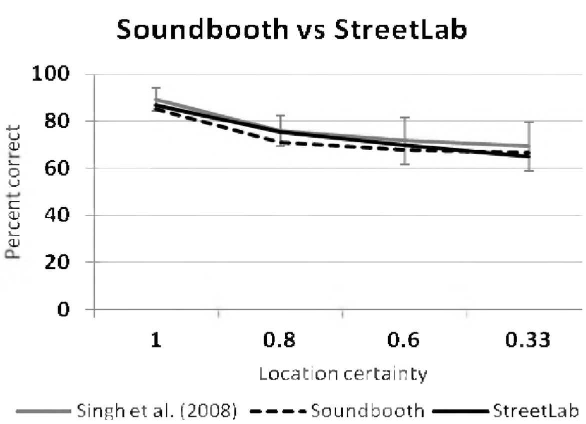


Figure 2. Percent Correct as a function of Location Certainty.

4. DISCUSSION AND CONCLUSIONS

The goal of this study was to determine whether the findings of the study of Singh et. al. (2008) could be replicated in a realistic, yet acoustically less controlled, VR laboratory environment (StreetLab). Investigators are now beginning to understand the importance of taking the knowledge that has been accumulated through traditional, highly controlled laboratory studies and moving towards evaluating the implications of these findings for real world interactions. Without understanding whether baseline performance measures in well-validated tasks under highly controlled conditions can be reproduced within highly realistic, ecologically valid conditions in simulated real world scenarios, differences in findings between the two in future experimental paradigms will be difficult to explain. In the current study, the results obtained in the soundbooth were replicated in an acoustically less-controlled space where additional non-auditory information (i.e., a complex visual scene) was presented. We were able to demonstrate that the findings are robust and likely translate to more complex situations. These results also suggest that future studies within this type of VR simulation may provide a valid testing environment for other psychoacoustic studies. Finally, this study provides a baseline for future studies in which we plan to test increasingly complex, real world challenges, such as multi-tasking activities requiring, for instance, listening and/or talking while walking with varying concurrent demands in terms of negotiating the sights and sounds of pedestrian and/or vehicle traffic.

REFERENCES

Bolia, R.S., Nelson, W.T., Ericson, M.A., Simpson, B.D. (1999). A speech corpus for multitalker communications research. *Journal of the Acoustical Society of America, 107,* 1065-1066.

Kidd, G., Arbogast, T.L., Mason, C.R., and Gallun, F.J. (2005). The advantage of knowing where to listen. *Journal of the Acoustical Society of America, 118,* 3804-3815.

Pichora-Fuller, M.K. (2003). Processing speed and timing in aging adults: psychoacoustics, speech perception, and comprehension. *International Journal of Audiology, 42,* 59-67.

Singh, G., Pichora-Fuller, M.K., and Schneider, B.A. (2008). The effect of age on auditory spatial attention in conditions of real and simulated spatial separation. *Journal of the Acoustical Society of America, 124,* 1294-1305.

ACKNOWLEDGEMENTS

This research was conducted by the first two authors for a 4th year undergraduate independent research course. The study was funded by an NSERC grant to author MKPF. StreetLab was partially funded by the Canadian Foundation for Innovation as part of the iDAPT (Intelligent Design for Adaptation, Participation and Technology) Centre for Rehabilitation Research; www.torontorehab.com/Research/Facilities/Labs/StreetLab.aspx). Thank you to Bruce Haycock for programming assistance.

ESTIMATION OF NOISE LEVELS AND HPD ATTENUATION IN THE WORKPLACE USING MICROPHONES LOCATED IN THE VICINITY OF THE EAR

Hugues Nélisse[1], Cécile Le Cocq[2], Jérôme Boutin[1], Frédéric Laville[2] and Jérémie Voix[2]
[1]Institut de Recherche Robert Sauvé en Santé et Sécurité du Travail (IRSST), Montréal Québec, Canada
[2]Dépt. de Génie mécanique, École de technologie supérieure (ÉTS), Montréal, Québec, Canada

1. INTRODUCTION

With the increase popularity of individual fit testing and miniaturization of electronic components and microphones, the Field-microphone-in-real-ear approach (F-MIRE) is becoming more appealing and well suited for estimating hearing protection devices (HPD) attenuation both in laboratory (Berger et al., 2008; Voix and Laville, 2009) or in "real" occupational conditions (Nélisse et al., 2012). The F-MIRE approach utilizes two miniature microphones to simultaneously measure the sound pressure levels in the ear canal under the hearing protector, as well as outside of the device. The location of this outside microphone is of primary importance for accurate and precise measurement of the HPD attenuation. A previous study by the present authors (Le Cocq et al., 2011) focusing on earmuffs in laboratory settings was performed in semi-anechoic and reverberant conditions where various earmuffs, microphone positions, source locations and subjects were tested. The study allowed making recommendations regarding an "optimal" position of the external microphone on the earmuffs for attenuation measurement purposes. Additionally, the results also suggested that recordings from this microphone could also be used to assess the sound pressure levels that would exist in the absence of the subject. Such data could be of interest if one is interested in performing noise survey while measuring, simultaneously, HPD attenuations in the workplace.

The present study focuses primarily on examining how the data obtained in the previous study can be used to propose a simple approximation for estimating the noise levels that would exist without the subject's presence. The first part presents the methodology used in a laboratory to establish this approximation and the second part presents the method proposed to validate it, i.e a simulation of "controlled" workplace conditions. Examples of results illustrating the findings are finally shown and discussed.

2. METHODS

In the first part, explained in more detailed in the previous study (Le Cocq et al., 2011), tests were conducted in a diffuse field in a reverberant room and in a free field in a semi-anechoic room on 4 subjects with 5 different earmuffs and without earmuffs. In free field conditions, twelve sound directions from -150° to 180° with 30° steps were considered. Six miniature microphones were fixed on each earmuff. The sound pressure level (SPL) was measured in third octave bands. To take into account the individual frequency responses of the 12 microphones, the free-field SPL measurements were normalized by the microphone responses obtained in a diffuse field where all microphones were located at the same location in the reverberant room. The tests in both rooms were realized with a pink noise of about 85 dB overall SPL. The results were then used to compute the difference between the SPL at the microphone location and the SPL at the center of the head without the subject. This difference, noted Δ in the rest of the paper, is expressed in dB. Values of Δ were then used to establish an approximation which can be used to estimate the SPL in the absence of the subject from microphones located on the earmuff.

In the second part, a "controlled" simulation of a workplace environment was performed to validate the approximation,. As illustrated in figure 1, a background sound field was created using four speakers placed in a mechanical shop where two machines were running at the same time. The objective was to create a sound field which could be more directional at certain location. A subject was asked to stand still at 9 different locations in the room with 4 different head orientations. Binaural recordings were performed by using two microphones (one per ear) located on the upper part of an earmuff. This choice of microphone position was made after analysis of the results of the first part. Microphone recordings were also made at the nine subject locations without the subject presence, at the head center positions.

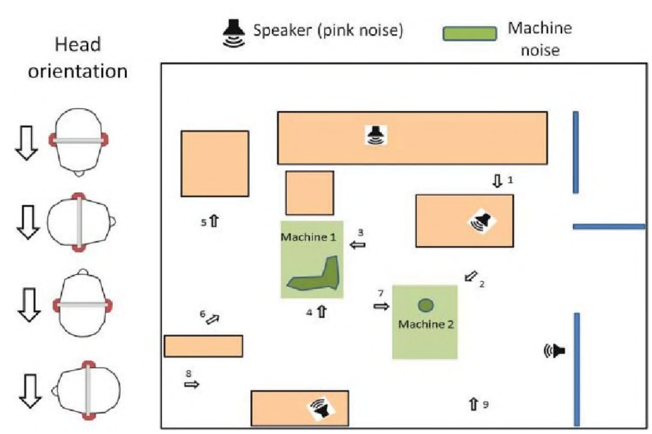


Figure 1: Overview of the room and speaker setup used for the workplace simulation

One objective was to apply the proposed correction factor to the microphone measurements made on the earmuffs and then compare these estimates to the measurements done without subjects.

3. RESULTS

3.1 Approximation for Δ

To establish a simple and practical approximation, the values of Δ were computed for four different zones of noise source location: Zone 1- Source in the front/back of the head (-30°<θ<30° or -150°<θ<150°); Zone 2 – Ear/mic facing the source; Zone 3 – Ear/mic in the shadow zone (source facing the opposite ear); Zone 4- diffuse field. Figure 2 shows the mean and standard deviation for the difference Δ computed, in each zone, for all subjects and earmuffs tested. Simple regression lines which can be used as approximation curves are also plotted. Results for zone 4 (diffuse field) are not shown here because they were found to be very similar to zone 1 results. Consequently, the three regression lines can be used to correct the earmuff microphone readings as long the source location can be determined.

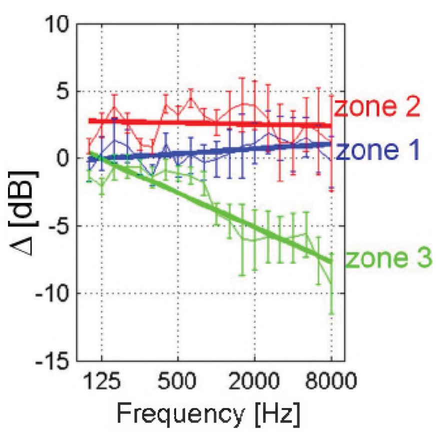


Figure 2: Mean SPL differences Δ for the three source location zones

3.2 Validation with workplace simulation

To estimate the source location when a subject was placed in the workplace simulated environment, binaural recordings were used as follow: 1) compute the overall SPL at each ear; 2) compare the overall levels between left and right ears; 3) if the difference is below a certain threshold, consider the source to be either in front of the head, in the back of the head or to be a diffuse field. If the left/right difference is positive (negative) and above the threshold level, consider the source to be facing the left (right) ear; 4) apply the appropriate correction to the earmuffs microphone SPL readings using the regression lines proposed above. This procedure was applied for all subjects, earmuffs, locations in the room and head orientations. Figure 3 presents the results for all instances where the earmuff microphone was considered in zone 2 (facing the source) using a threshold of 1.5 dB to select source zone. Values of Δ are plotted as a function of frequency. The upper line represents the bare “uncorrected” data while the line with the circle symbol represents the corrected estimate. While a difference of 0 dB represents the target to achieve, it is found that applying the proposed correction Δ allows to obtain a reasonable estimate within a ±2 dB range.

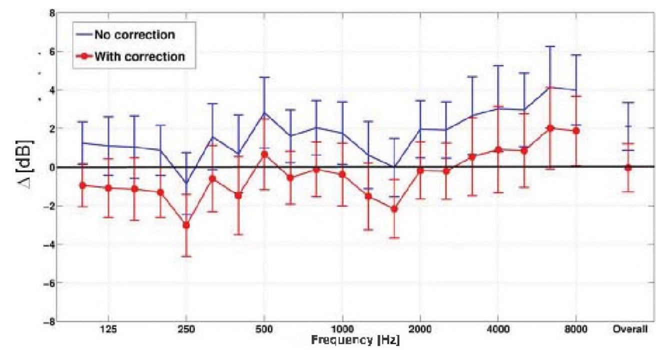


Figure 3: Comparison of the SPL difference Δ with and without correction (microphone determined to be in zone 2)

4. DISCUSSION AND CONCLUSIONS

Results of the study suggested that values of Δ on average did not exceed 2.5 dB when the microphone was well “seen” by the noise source. When the microphone was in a “shadow” zone, significant differences were observed. A simple and practical preliminary approximate correction could then be derived. It allows obtaining, using simple binaural recordings, fairly good estimates of the SPL values that could be measured without the presence of the subject. In the context of assessing HPD attenuation at the same time as performing noise survey, work is currently under way to refine the approximation and to repeat the experiment with earplugs. In this case, larger differences are expected as different elements of the external ear (eg. pinna) will be affecting more the sound field.

REFERENCES

Berger EH, Voix J., Hager LD (2008) Methods of fit testing hearing protectors, with representative field test data. Hearing loss: 9th International Congress On Noise as A Public Health Problem (ICBEN). Foxwoods, CT, p. 8

Le Cocq C, Nélisse H, Boutin Jérôme, et al. (2011) Influence of source location, subjects and HPD size on the sound field around earmuffs. Semaine canadienne d’acoustique. Québec, pp. 98–99

Nélisse H, Gaudreau M-A, Boutin J., et al. (2012) Measurement of Hearing Protection Devices Performance in the Workplace during Full-Shift Working Operations. Annals of Occupational Hygiene 56(2)

Voix J., Laville F. (2009) The objective measurement of individual earplug field performance. The Journal of the Acoustical Society of America 125(6): 3722

The Development of VOT Perception in School-Aged Children

Nicole Netelenbos[1], Fangfang Li[2]
[1]Dept. of Psychology, University of Lethbridge, AB, Canada, nicole.netelenbos@uleth.ca
[2]Dept. of Psychology, University of Lethbridge, AB, Canada, fangfang.li@uleth.ca

1. INTRODUCTION

Voice onset time (VOT) refers to the time elapse between the release of the oral constriction and the onset of vocal cord vibration during articulation of pre-vocalic stops. Since the seminal study by Lisker & Abramson (1964), VOT has frequently been demonstrated to be the primary acoustic correlate for voicing contrast in stop production and perception in English (Docherty, 1992; Klatt, 1975). VOT has also been employed in studying infant speech perception. By varying VOT in equal steps to form a continuum between /ba/ and /pa/, Eimas, Siqueland, Jusczyk, & Vigorito (1971) discovered that 1- and 4-month old infants are able to discriminate the voicing contrast in a way similar to adults. However, such seemingly innate discriminatory abilities in infants are only suggestive of auditory perceptual acuity but not category recognition; the latter requires years of linguistic experience to form. Subsequent studies on children indicate that speech perception in identifying stop categories based on VOT continues to progress beyond infancy (Zlatin & Koenigsknecht, 1975; Flege & Eefting, 1986). In particular, Zlatin & Koenigsknecht (1975) compared two-year olds and six-year olds in their identification performance of three stop cognates: /b/-/p/, /t/-/d/, and /k/-/g/. They found the greatest age-related differences in children's perception of the velar pair (/k/-/g/), such that two–year olds exhibited longer phoneme boundary VOT values than six-year olds and adults. Further, a more recent study by Hazan & Barrett (2000) suggests that twelve-year olds have yet to reach adult-like phonemic contrast proficiency, as their identification curve of the /k-/g/ pair is not as consistent as that of adults.

To our knowledge, no previous study has examined stop phonemic perception in all three places of articulation (labial, alveolar, and velar) in school-aged children. The present study presents preliminary results of an on-going research project that aims to investigate the development of pre-vocalic stop categorization in children aged 4 to 9. The purpose of our study is twofold: 1) to examine whether there exists age-related perceptual differences in school-aged children 2) to investigate whether developmental differences in phonemic categorization vary as a function of stop place of articulation. Based on previous research, we predict that children continue to refine their identification functions throughout childhood. Further, we expect to find an interaction between stop consonantal place of articulation and developmental curves. Specifically, the approximation towards an adult-like pattern should occur first in labial and alveolar stops and last in velar stops.

2. METHODS

Natural stimuli were recorded and edited to create VOT continua for three minimal pairs: *bear/pear*, *deer/tear*, and *goat/coat*. The VOT values ranged from -70 ms to 70 ms, separated by 10 ms steps, yielding a total of 45 tokens (15 tokens * 3 pairs). In addition, approximately 10% of the stimuli were included as repetitions to gauge inra-subject reliability. All stimuli were randomized over trials.

A total of 48 children aged 4 to 9 participated in the study (5 four year olds, 8 five year olds, 8 six year olds, 5 seven year olds, 17 eight year olds, and 4 nine year olds). All participants self-reported (questionnaire addressed to parents) as normal hearing individuals with no known language, speech, or learning delays. An identification task was employed to examine participants' speech perception. Children sat in front of a computer monitor while speech stimuli were played over computer speakers. The child was asked to click on the image (one of six) matching the token that they heard.

3. RESULTS

Findings revealed that the boundary location and slope of perceptual identification function varied with age; older children demonstrated a sharper transition between phonemic categories as well as a more precise phonemic boundary location. In addition, this developmental trend observed in relation to the slope of identification and boundary location, differentiated according to the three places of articulation of the perception stimuli (labial, alveolar, and velar). In terms of the labial (*bear/pear*) stop discrimination, the VOT boundary identification across the age groups was fairly consistent, falling around the 20-30ms range at the 50% crossover, while, the slope of identification demonstrated a clear developmental trend becoming steeper with the accession of age, reaching its steepest point around 9 years of age (see figure 1). The alveolar (*tear/deer*) discriminatory pair revealed less consistency in boundary identification across the age groups, fluctuating between the 40-60ms range at the 50% crossover, with the identification function reaching its steepest point at age 9 (see figure 2). Lastly, the velar pair (*goat/coat*) revealed far less consistency in location and precision of boundaries. Boundary location identification was located between 30-70ms at the 50% crossover. The identification curve becomes progressively steeper as children's age increases (see figure 3).

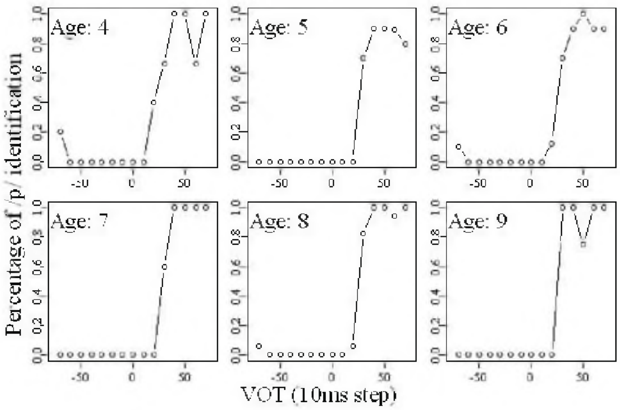


Figure 1. Mean identification scores for the labial pair for children aged 4 to 9

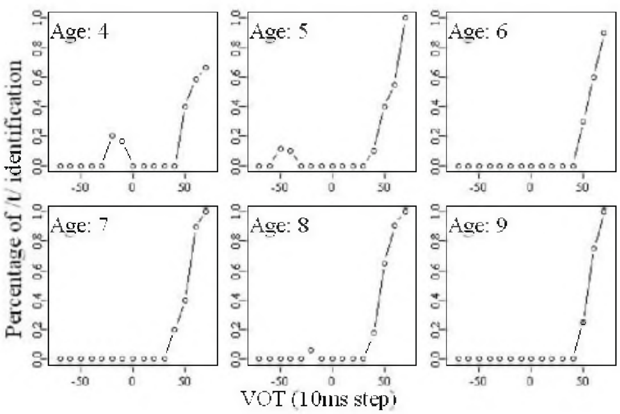


Figure 2. Mean identification scores for the alveolar pair for children aged 4 to 9

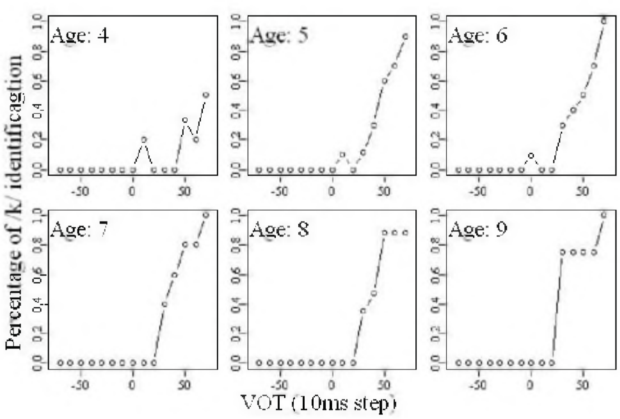


Figure 3. Mean identification scores for the velar pair for children aged 4 to 9

4. DISCUSSION

The results of the present study suggest that the perception identification of certain phonemic contrasts is still attuning in school-aged children. Even in the older children, the VOT parameter of minimal pair discrimination does not appear to have been fully established for all phonemic contrasts, specifically for those of the velar tokens. Consistent with previous research (e.g. Zlatin & Koenigsknecht, 1975), more developmental difference was found in the velar pair than in the labial and alveolar pair, suggesting later mastery of perceptual norms for the velar stops. As a whole, the slope of identification became sharper with age, revealing more uniformity among the older children. One explanation for this developmental perceptual progress is the role of language experience where language abilities, such as identifying speech token contrasts, generally improve with longer exposure. Moreover, linguistic boundary solidification in VOT perception development observed in the present study parallels the reported production development sequence, with the phonemic categorization pattern being consistent with the suggested sequence of emergence of these articulators, where children have been reported to produce labial stops first, followed by alveolar and velar stops (e.g. Smit, Hand, Freilinger, Bernthal, & Bird, 1990). This sequence of phoneme development may be interpreted by anatomical ease of articulatory gestures, with the stop discriminatory pairs (i.e. /p/-/b/, /t/-/d/, /k/-/g/) generally being acquired at different stages of phonological development, respectively, possibly owing to differences in articulatory complexity. These findings have theoretical implications in broadening our understanding of the development of VOT perception of phonemic contrasts in mid to late childhood.

REFERENCES

Docherty, G. (1992). *The timing of voicing in British English obstruents*. Berlin, New York: Foris.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., Vigorito, J. (1971). Speech perception in infants. *Science, 171*, 303-306.

Flege, J. E., & Eefting, W. (1986). Linguistic and developmental effects on the production and perception of stop consonants. *Phonetica, 43*, 155-171.

Hazan, V., & Barrett, S. (2000). The development of phonemic categorization in children aged 6-12. *Journal of Phonetics, 28*, 377-396.

Klatt, D. H. (1975). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of speech and Hearing Research, 18*, 686-706.

Lisker, L. & Abramson, A. (1964). Cross-language study of voicing in initial stops: acoustical measurements. *Word*, 20, 384-422.

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., Bird, H. (1990). The Iowa articulation norms project and its Nebraska Replication. *Journal of Speech and Hearing Disorders, 55,* 779-798.

Zlatin, M. A., & Koenigsknecht, R. A. (1975). Development of the voicing contrast: perception of stop consonants. *Journal of Speech and Hearing Research, 18,* 541- 553.

DO URBAN SOUNDSCAPES INFLUENCE VISUAL ATTENTION?

Tristan D. Loria and Frank A. Russo
Dept. of Psychology, Ryerson University, Toronto, Ontario, Canada.
tristan.defrancescol@ryerson.ca / russo@ryerson.ca

1. INTRODUCTION

Attention is vital to the survival of all organisms. Regardless of size or strength, every organism must adaptively shift its attention in order to survive. Research by Easterbrook (1959) highlights the importance of arousal as a determinant for alternating between broad based, and locally based visual attention. His hypothesis states that increases in arousal lead to attentional selectivity, which in turn supports appropriate responding. However, the modern city often provides a cacophonous auditory environment saturated with opportunities for attentional selectivity, and hence distraction. In the current investigation we use an empirically validated behavioural measure to assess whether urban soundscapes have an influence on visual attention.

1.1. Soundscapes effects: Short-term and Long-term

The urban soundscape has been identified as a source of deleterious effects on human health and performance. In particular, aversive soundscape exposure over the short term has been linked to increases in heart rate (Raggam et al. 2006) and delayed stress recovery following the offset of an accute stressor (Ulrich et al. 1991).

The effect of long-term exposure to aversive soundscapes has also been assessed empirically. Shield and Dockrell (2008) demonstrated that exposure to aversive soundscapes in the classroom can be detrimental to school performance. In addition, long-term exposure to aversive soundscapes has been linked to a plethora of health problems including, but not limited to: cardiovascular health, sleep disturbances, hypertension, and hearing loss (Aydin & Kaltenbach 2007; Basrur, 2001). Taken together, these results suggest that soundscapes can influence auditory and non-auditory processes.

1.2. Purpose and Hypothesis

The purpose of this study was to determine whether soundscapes influence visual attention. Motivated by previous research on the subject of attentional fluctuations as mediated by affect and arousal levels (Gable and Harmon-Jones, 2010), we hypothesized that soundscapes would influence the focus of attention, and that the effect would vary with type of exposure.

2. METHOD

2.1 Participants

68 undergraduate students at Ryerson University (53 females) were recruited for participation. The average age of participants was 21 years (SD=6.10). Participants were randomly assigned to one of three exposure conditions: concurrent, prolonged, and brief. Participants in the *concurrent* condition (n=14) heard the soundscapes continuously, making a visual judgment while the soundscape played. Participants in the *prolonged* condition (n=19) heard the soundscapes for forty-five seconds, followed by a visual judgment. Participants in the *brief* condition (n=35) heard the soundscapes for fifteen seconds, followed by a visual judgment.

2.2 Procedure

Thirty two-minute soundscapes were recorded binaurally to a TAS-DR-1 Tascam Portable Solid State Digital Audio Recorder using SP-BMC-20 Audio Technica subminiature omnidirectional microphones (with windscreens) clipped above each ear. Sound pressure levels were also recorded simultaneously using an Extech 407768 sound level meter with dB(A) weighing. All soundscapes used in this study as well as data about accompanying sound levels are available as part of the torontosoundmap.com.

The main study was conducted in a double-walled IAC sound attenuated chamber. A Mac MINI computer presented auditory stimuli over SONY MDRXD200 circumaural headphones. Visual stimuli were presented using an ACER X243w computer screen situated 35 cm away from the participant. After hearing one of thirty soundscapes, the participant was asked to make a visual discrimination on global or local features of a Navon stimulus (1977). A Navon stimulus is a large letter made up of smaller letters. For example, a large "H" composed of small "T"s. In a global trial involving this example, participants might be asked to respond "H" or P" (in this case H is correct). In contrast, for a local trial involving this example, participants might be asked to respond "O" or "T" (in this case T is correct). While accuracy rates tend to be near ceiling, the reaction times are indicative of the participant's focus of visual attention.

2.3. Subjective and Acoustic Analysis

Using likert-scales, participants provided subjective appraisals of each soundscape. These appraisals included valence (pleasant/unpleasant), arousal (calm/excited), stress (not at all / extremely). In addition, participants indicated the maximum exposure (in minutes) they could tolerate each soundscape. Objective analyses of soundscapes were conducted using MIRtoolbox, a freely available MATLAB toolbox, which supports extraction of low and mid-level acoustic features from audio recordings (Lartillot and Toivianinen, 2007). The most predictive feature derived from this analysis was spectral irregularity (1), defined as amplitude variation throughout the successive peaks of the spectrum (Jensen, 1999):

$$\left(\sum_{k=1}^{N}(a_{k-}\,{}_{a_{k+1}})^2\right)/\sum_{k=1}^{N}a_k^2$$

3. RESULTS AND DISCUSSION

For each participant, a local bias estimate was determined for each soundscape by subtracting global from local reaction time (RT). Local bias estimates were subjected to a 3-way analysis of variance (acute, prolonged, or concurrent). There was a main effect of soundscape, F (29, 1885) = 2.1, $p < .001$, which suggests that the influence of the soundscapes assessed was variable. Mean RTs for local and global targets (collapsed across exposure types) are plotted in Figure 1.

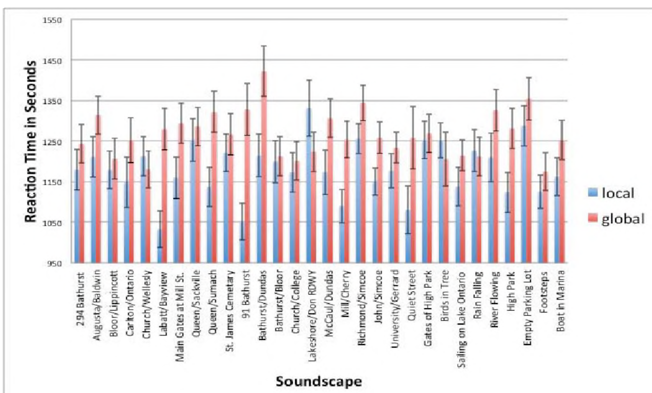


Figure 1: Mean RT for Local and Global Targets.

The main effect of exposure type and its interaction with soundscape were not significant, which suggests that soundscapes tend to have an immediate influence over visual attention.

To better understand the main effect of soundscape we correlated average local bias estimates (drawn from the prolonged exposure group) with subjective appraisals and acoustic features. Although none of the correlations involving acoustic features was significant, local bias estimates were significantly correlated with subjective variables. Increases in spectral irregularity were correlated with increases in stress, decreases in valence, and decreases in maximum exposure judged to be tolerable. Future work will incorporate physiological measures to better understand the subjective findings reported here. Our motivation for this ongoing program of research is to inform an emerging dialogue between urban planners, architects, and acousticians, concerning the sound quality of our cities.

REFERENCES

Aydin, Y., & Kaltenbach, M. (2007). Noise perception, heart rate and blood pressure in relation to aircraft noise in the vicinity of the Frankfurt airport. *Clinical research in cardiology: German Cardiac Society*, *96*(6), 347-58. doi:10.1007/s00392-007-0507-y

Basrur, V. S. (2000) Health Effects of Noise. City of Toronto, Community and Neighbourhood Services, Toronto Public Health, Health Promotion and Environment Protection Office

Easterbrook, J.A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66,* 183–201.

Gable, P., & Harmon-Jones, E. (2010). The Blues Broaden, but the Nasty Narrows: Attentional consequences of Negative Affects Low and high in Motivational Intensity. Psychological Science, 21(2), 1-5. doi: 10.1177/0956797609359622

Jensen, K.K. (1999). Timbre models of musical sounds. Report 99/7, University of Copenhagen. ISSN: 0107-8283.

Lartillot, O., & Toiviainen, P. (2007). A Matlab Toolbox for Musical Feature Extraction From Audio, *International Conference on Digital Audio Effects*, Bordeaux.

Navon, D. (1977). Forest Before Trees: The Precedence of Global Features in Visual Perception. Cognitive Psychology, 9, 353-383. doi:10.1016/0010-0285(77)90012-3

Raggam, B. R., Cik, M., Holdrich, R. R., Fallast, K., Gallasch, E., Fend, M., Lackner, A., & Marth, E. (2007). Personal noise rankings of road traffic: Subjective estimations versus physiological parameters under laboratory conditions. *International Journal of hygiene and Environmental Health*, 210, 97-105. doi: 10.1016/j.ijheh.2006.08.007

Shield, B. M., & Dockrell, J. E. (2008). The effects of environmental and classroom noise on the academic attainments of primary school children. The Journal of the Acoustical Society of America, 123(1), 133-44. doi:10.1121/1.2812596

Ulrich, S. R., Simons, F. R., Losito, D. B., Fiorito, E., Miles, A. M., & Zelson, M. (1991) Stress recovery during exposure to natural urban environments. Journal of Environmental Psych., 11, 201-230. doi:10.1016/S0272-4944(05)80184-7

THE EFFECT OF AGING ON COCHLEAR AMPLIFIER: A SIMULATION APPROACH USING A PHYSIOLOGICALLY-BASED ELECTRO-MECHANICAL MODEL OF THE COCHLEA

Amin Saremi[1], Stefan Stenfelt[2]

[1,2] Department of Clinical and Experimental Medicine, Division of Technical Audiology, Linnaeus Centre HEAD, Linköping University, SE 581 85, Sweden, email : amin.saremi@Liu.se

1. INTRODUCTION

Age-related hearing loss is one of the top three most common chronic health conditions affecting individuals aged 65 years and older (Frisina, 2010). Models have been exploited as powerful analytical tools to sharpen our understanding of various mechanisms and systems. Many of the existing models for human peripheral auditory system use computational elements such as filters, etc. to represent each stage of the system regardless of tiny biological details (Irino and Patterson, 2001); therefore, they cannot be used for differentiating the underlying physiological roots of various cochlear hearing impairments.

Here the electrical, acoustical, and mechanical elements of the cochlea are explicitly integrated into a transmission-line model. The aim is to choose these elements to have a physiological interpretation of the human cochlea insofar as is known[a]. As a result, the model enables fundamental simulation of specific cochlear lesions such as metabolic presbyacusis.

The prevailing hypothesis that high-frequency hearing loss in older adults may be due to metabolic presbyacusis whereby there is a reduction in the endocochlear potential (EP) as a result of age-related cellular degeneration in the lateral wall of the cochlea (Schmiedt, 1996; Dubno *et al.*, 2008; Lang *et al.*, 2010). Besides of aging, similar degenerations are observed in patients with diabetes type 2 millitus (Fukushima *et al.*, 2006). This physiologically-based model is used to quantitatively validate this hypothesis.

2. MODELING METHOD

A sound pressure field in the air is transmitted via the outer and middle ear to the inner ear. It causes the stapes to vibrate resulting in a travelling wave along the organ of Corti propagating from base towards apex (Bekesy, 1960). The organ of Cori rests on the basilar membrane (BM); the displacements on BM trigger an active electro-mechanical element - outer hair cell (OHC) – which boosts the passive vibrations. The amplified vibration is converted to neural signals by the inner hair cells (IHCs).

[a] In spite of the new measurement techniques, this is still a challenge remembering that human cochlea is covered by bones and is rather inaccessible for real-time measurements; therefore, animal data plays an important role for creating reasonable assumptions of physiological details.

In order to study this complex mechano-electro-acoustical system the cochlear duct is assumed to be 'N' discrete partitions extending from the stapes on the left to the apex on the right. These partitions are structurally coupled to each other via longitudinal components; forming a transmission line as shown in fig 1. Each cochlear partition consists of passive loads (BM, RL and TM) together with an active component (OHC). Each passive load (BM, RL and TM) is modeled by a classical second-order 'mass-spring-damper'.

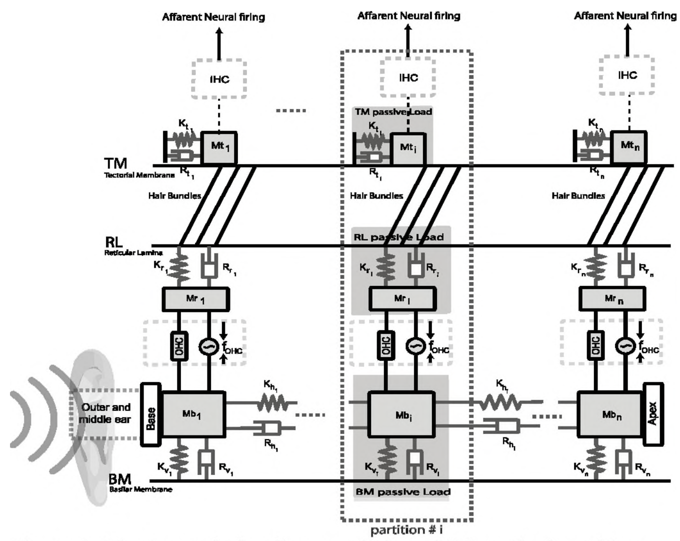


Figure 1: The transmission line consisting of 'N' vertical partitions.

2.1 Active processes: Cochlear amplifier

The OHC stretches against RL on one side and BM on the other side (Kim and Mountain, 2011) pulling them together. This molecular force generator -acting between BM and RL- is known as somatic motility. Somatic motor is triggered by the receptor current produced by the MET mechanism. Stria vascularis, in turn, supplies the necessary charge for MET by secreting K+ into the cochlear endolymph of scala tympani. This 3-stage system is depicted in fig 2; it is then solved for the parameters.

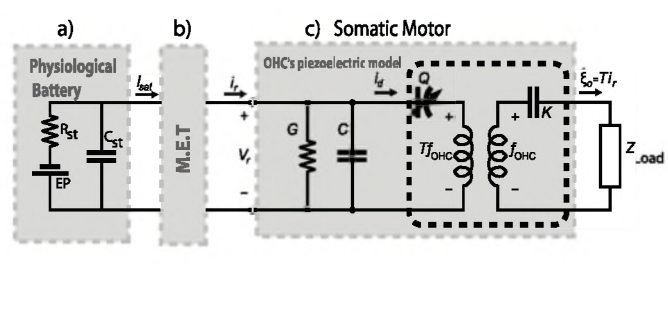


Figure 2: *a)* stria vascularis is modeled by its body resistance and capacitnce (Rst and Cst respectively) together with an ideal voltage source (EP). Stria vascularis functions as a 'battery' providing the whole cochlear amplifier with the necessary electrical charge *b)* The MET mechanism which plays the role of a motor driver for the somatic motor. *c)* Circuit model of the OHC somatic motor according to piezoelectric model of Liu and Neely (2009).

2.2 Metabolic presbyacusis

As a result of age-related cellular degeneration in the lateral wall of the cochlea, stria vascularis is not able to maintain the maximum EP (89mV) anymore. As the EP decreases, the MET produces less receptor current which, according to the 3-stage circuit of fig 2, eventually leads to a decline in the force/displacement generated by the somatic motor.

3. RESULTS

Figure 3 illustrates how the cochlear amplification is affected by a decrease in EP. The results are illustrated at 30% and 70% of cochlear length from the base to compare how the amplification declines while the EP decreases from 89mV to its half value (44.5mv).

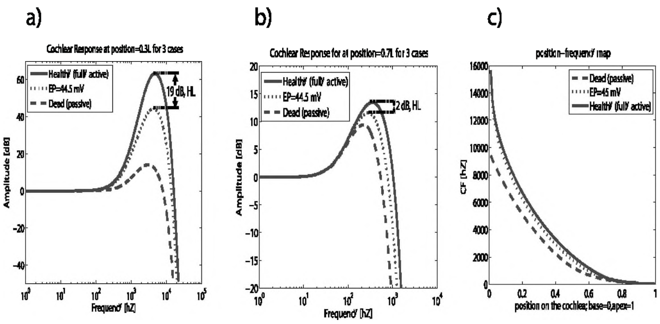


Figure 3: The effect of EP degeneration on magnitude and sensitivity of cochlear amplification (a, b) and position-frequency (c). Note that these are simulated at a low intensity corresponding to a 120 nm displacement of RL, according to small signal analysis. *a)* magnitudes at location 30% of total cochlear length from the base. *b)* at 70% of total length from the base *c)* illustrates how the CFs are shifted backwards as a result of EP degeneration.

When the EP is decreased to its half value, there is a 19-dB loss at the peak; CF is also shifted downwards from 4700 Hz to 4212 Hz. Once the active processes are totally dead (passive cochlea) the curve is flattened, the amplification is significantly decreased to only 14.12 dB and the CF is further shifted backwards to 3026 Hz. A similar decline is seen in fig 3.b however there is only a 2-dB loss at the peak of the curve and the CF is shifted backwards from 340 Hz to 271 Hz as a result of a 50% decrease in EP.

4. DISCUSSION

Comparing fig 3.a with fig 3.b indicates that the decrease in EP affects cochlear amplifications in higher frequencies more than it does in lower frequencies. This is consistent with Dubno *et al.* (2008) where they argue that the reason why presbyacusis is associated with high frequencies is that cochlear amplifications is greater in higher frequencies than in lower frequencies therefore the lack of amplification is significantly felt in higher frequencies. Figure 3 also reveals another aspect of cochlear amplification: The cochlear position-frequency map changes when the active processes are deteriorated. The CFs of the curves tend to move backwards in a presbyacusis cochlea. This is consitent with Robles and Ruggero (2001) where they indicate that in a passive cochlea (EP=0) the CFs are shifted backwards.

Clinical tests can be performed to further improve the parameters and predictions of the model. Moreover, this biologically-based model is developed with the aim of being connected to models of nervous pathways and more central parts of the auditory system in a signal-cognitive manner as suggested by Stenfelt and Ronnberg (2009).

REFERENCES

Bekesy, G. V. (**1960**). *Experiments in Hearing* (McGraw-Hill, New York, USA).

Dubno, J. R., Lee, F. S., Matthews, L. J., Ahlstrom, J. B., Horwitz, A. R., and Mills, J. H. (**2008**). "Longitudinal changes in speech recognition in older persons," The Journal of the Acoustical Society of America **123**, 462-475.

Frisina, S. G.-S. R. D. (**2010**). *The Aging Auditory System* (Springer, USA).

Fukushima, H., Cureoglu, S., Schachern, P. A., Paparella, M. M., Harada, T., and Oktay, M. F. (**2006**). "Effects of type 2 diabetes mellitus on cochlear structure in humans," Archives of otolaryngology--head & neck surgery **132**, 934-938.

Irino, T., and Patterson, R. D. (**2001**). "A compressive gammachirp auditory filter for both physiological and psychophysical data," The Journal of the Acoustical Society of America **109**, 2008-2022.

Kim, D. J., and Mountain, D. C. H., A. E. (**2011**). "How Much Do Somatic and Hair Bundle Motility Contribute to Cochlear Amplification?," American Institute of Physics:WHAT FIRE IS IN MINE EARS: Proceedings of the 11th International Mechanics of Hearing Workshop **1403**, 632-637.

Lang, H., Jyothi, V., Smythe, N. M., Dubno, J. R., Schulte, B. A., and Schmiedt, R. A. (**2010**). "Chronic reduction of endocochlear potential reduces auditory nerve activity: further confirmation of an animal model of metabolic presbyacusis," Journal of the Association for Research in Otolaryngology : JARO **11**, 419-434.

Liu, Y. W., and Neely, S. T. (**2009**). "Outer hair cell electromechanical properties in a nonlinear piezoelectric model," The Journal of the Acoustical Society of America **126**, 751-761.

Robles, L., and Ruggero, M. A. (**2001**). "Mechanics of the mammalian cochlea," Physiological reviews **81**, 1305-1352.

Schmiedt, R. A. (**1996**). "Effects of aging on potassium homeostasis and the endocochlear potential in the gerbil cochlea," Hearing research **102**, 125-132.

Stenfelt, S., and Ronnberg, J. (**2009**). "The Signal-Cognition interface: Interactions between degraded auditory signals and cognitive processes," Scandinavian journal of psychology **50**, 385-393.

CONSTRUCTION VIBRATIONS IN THE CITY OF EDMONTON

Clarence Stuart, QEP
Environmental Scientist, City of Edmonton, 11404 – 60 Avenue, Edmonton, AB, Canada, T6H 1J5
clarence.stuart@edmonton.ca

1. Introduction

The City of Edmonton conducts vibration monitoring in response to vibration complaints or damage claims from the public due to vehicular traffic on uneven road surfaces. Over time, vibration monitoring services have expanded to include construction vibrations during neighborhood renewal projects (residential), streetscape reconstruction in commercial areas, building demolition and pile driving. The following paper presents some of the data collected from traffic vibration monitoring and over 500 construction-related vibration tests which includes a variety of construction equipment collected from demolition, neighborhood reconstruction, and commercial high rise buildings.

2. Method

The City's Roadway Maintenance department has adopted an "annoyance" value of 2.5 mm/s Peak Vector Sum (PVS) to evaluate whether further assessment or action is prudent. The City of Edmonton has adopted the frequency-based US Bureau of Mines USBM RI 8507 criteria to evaluate the potential risk to residential buildings due to vibration (Dowding 1996). For commercial buildings, both USBM-R18507 and the German DIN4150 guidelines are referred to for context.

An Instantel Mini Mate Pro 6 triaxial vibration sensor with dual geophones and datalogger was used to collect data. Vibration monitoring requires measurement of particle velocities and frequencies in 3 directions: vertical, longitudinal and transverse. The vibration source, transmission path and the potential receiver are considered in sensor placement and test duration.

Monitoring equipment is set in histogram combo mode which allows peak values to be displayed throughout the assessment at set intervals. This mode also generates a waveform during histogram recording if the signal exceeds the trigger level.

Vibration test results are presented as Peak Vector Sum (PVS) since this method reflects the effect of all three directional components. Individual waveforms provide the Maximum Peak Particle Velocity (PPV) and frequency level for each axis. The most dominant frequencies for each axis are also reported. The PVS typically occurs concurrently with the PPV of one of the components, but the addition of the two components slightly increases its magnitude.

The USBM-RI8507 guideline is represented graphically, and considers both the PPV and frequencies (Figure 1 and Figure 2). The USBM-RI8507 "simplified approach" provides criteria at two specific frequencies, < 40Hz and > 40 Hz. The guideline levels for modern drywall homes and older plaster homes are most commonly referred to (19.1 mm/s and 12.7 mm/s at <40Hz, respectively). Guideline levels increase to 50.8 mm/s at > 40 Hz (Siskand 1980, Kalinski 2007).

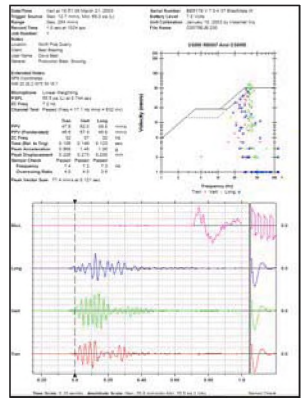


Figure 1

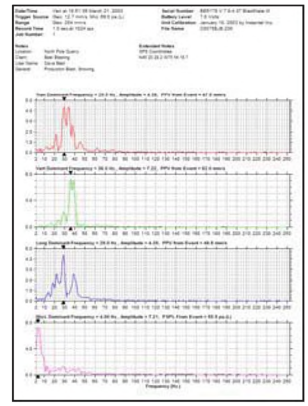


Figure 2

The equipment set-up protocol and sensor coupling method depends on the nature of the assessment. Sensor coupling methods include:

1. 12-inch steel probe with a horizontal sensor-mounting plate, pounded into the ground; or
2. Sand bag placed on top of sensor; or
3. Wall bracket, where the sensor is mounted horizontally.

3. Roadway Maintenance Complaints

Numerous residential vibration complaints come to the city help line each year, usually due to vehicular traffic (primarily large vehicles such as buses). Large cracks, potholes and infrastructure repair are the most common root-cause. Only some of the complaints progress to vibration testing; many of the immediate concerns are corrected by road resurfacing where the problems are obvious. The City conducts approximately twenty vibration tests a year.

Table 1: Roadway maintenance-related vibration assessments on residential properties (Sample of 26 homes tested)

	Maximum Peak Vector Sum (PVS) – mm/s Frequency level at most prominent axis							
	> 10		5 to 10		2.5 to 5		< 2.5	
Frequency	<40Hz	>40Hz	<40Hz	>40Hz	<40Hz	>40Hz	<40Hz	>40Hz
Sensor at property line	-	-	2	3	5	1	12	3
Sensor at foundation	-	-	-	-	5	-	20	1

Typical Distance: Curb to Property Line ~2.0 meters
Curb to Foundation ~10 meters

Note: the simplified USBM RI8507 suggests that a PPV of 12.7 mm/s could result in minor damage to plaster-on-lath interiors of older structures for vibration frequencies of <40Hz

Historically, one sensor was placed at the property line. With the purchase of the Mini Mate Pro 6 Vibration Unit an additional vibration sensor is now placed at the foundation. All triggered events are examined over a 24 hour collection period.

4. Neighborhood Renewal

The City's Neighborhood Renewal Program focuses on upgrade and replacement roads, sidewalks, streetlights and gutters. Reconstruction processes include sidewalk removal and demolition, utility work, street light installation, sidewalk and subgrade reconstruction, pouring of concrete for sidewalks, asphalt pulverization and re-compaction, curb construction, and paving. Numerous construction equipment and different construction processes are monitored for vibration impacts.

Vibration data was collected at selected test houses of typical construction and age for the neighborhood. Baseline vibration levels due to normal traffic were collected prior to construction. The equipment trigger level was set to 1mm/s. The highest vibration events were tabulated for each test site during each major construction process or equipment usage. Using this selection method, all of the events evaluated exceeded the perception threshold of 0.5 – 1 mm/s, meaning that a resident standing near the sensor would notice the vibrations.

Of the 233 peak events selected for evaluation, 8% were above 10 mm/s, which are considered to be "disturbing" and approaches the USBM guideline value of 12.7 mm/s for potential risk of cosmetic damage in older and plaster-and-lath interiors. The data demonstrates that the backhoe with breaker generated the greatest magnitude and greatest percentage of peak vibration events, followed by the vibrating steel roller, and the vibrating padfoot roller (Table 2).

Table 2: Construction equipment vibrations during neighborhood renewal.

Type of equipment	# of times used	Typical Frequency Hz	Maximum Peak Vector Sum (PVS) (mm/s)			
			> 10	5 to 10	2.5 to 5	< 2.5
Backhoe with breaker or bucket	34	<40	6	12	8	8
Vibrating steel roller	53	<40	9	24	16	4
Vibrating pad foot roller	18	<40	2	11	5	0
Hand tamper (1000 lb steel plate)	10	<40	1	3	3	3

Both sensors at foundation: foundation ~10 meters to curb

The removal of sidewalks, catch basins and curb-and-gutter generate significant peak vibration events (Table 3). Most of the concrete removal work was conducted using the backhoe with breaker, identified as one of the three pieces of equipment that generated the highest vibrations during this study. Contact the author for details about the numerous types of equipment used.

Table 3: Vibrations during neighborhood renewal processes.

Type of process	Maximum Peak Vector Sum (PVS)mm/s			
	> 10	5 to 10	2.5 to 5	< 2.5
Asphalt removal	0	0	0	4
Road compaction	1	2	3	4
Sidewalks, catch basins, curb & gutter, demo & reconstruct	7	20	12	13
Pulverization	2	13	7	10
Foaming	3	11	2	12
Paving	4	10	10	32

5. Down Town (Commercial)

Streetscaping activities in the downtown core were monitored, where the construction activities (vibration sources) are typically in close proximity to large commercial properties including high rises and the underground Light Rail Transit system.

Table 4: Construction vibrations during commercial streetscaping

	Maximum Peak Vector Sum (PVS) – mm/s Frequency level at most prominent PPV / Axis									
	>50		10 to 50		5 to 10		2.5 to 5		< 2.5	
Frequency	<40 Hz	>40 Hz	<40 Hz	>40 Hz	<40 Hz	>40 Hz	<40 Hz	>40 Hz	<40 Hz	>40 Hz
Drainage Pipe Replacement	-	2	4	6	-	-	2	-	1	
Road Construction	-	-	2	2	-	1	7	7	10	3

Both sensors at building foundation, 2 to 3 meters from curb

A large-track backhoe with a 4" pin hammer produced the highest vibrations and frequencies. Equipment size and the short distance to the building had a dramatic impact on magnitude.

6. Other Vibration Project

Vibrations were monitored during a major downtown demolition project. The proximity of the neighboring high-rise created space constraints, and particular business concerns necessitated the development of a custom sensor coupling method. A steel wall-mount bracket was fabricated and installed into an interior concrete wall (~ground level) of the adjacent high rise, with the geophones mounted horizontally onto the bracket. Vibration data was collected during demolition which entailed a top-down dismantling of the building, where the debris was dropped down the elevator shaft. Large equipment movements and pile removal also took place.

A large backhoe with a pin-hammer generated the highest PPV measurements: 13.9 mm/s at 3.3 Hz, 69.2 mm/s at >100Hz and 38.2 mm/s at >100 Hz in the transverse, vertical and longitudinal direction respectively. The PVS was 78.9 mm/s. In general the vertical sensor axis was most impacted during this assessment.

7. Conclusions

Equipment selection and distance from source to receiver are major factors influencing the magnitude of measured vibrations. The backhoe with breaker during concrete demolition generated the highest measured vibrations. Historic vibration data combined with real-time data on various equipment and processes provides valuable context and a better understanding of potential risks to buildings, which is invaluable where pre-construction inspections of impacted property were not possible or practical.

REFERENCES

Dowding, Charles H. 1996. Construction Vibrations. ISBN 0-9644313-1-9

DIN4150-3 (1999). Structural vibration Part 3: Effects of vibration on structures. Deutch Norm (German Standard) DIN4150-3, February 1999. ICS 91.120.25

Kalinski, 2007 USBM RI 8507 Vibration Criteria.

Siskind, D.E., et al. 1980. Structure response and damage produced by Ground vibration from surface mine blasting. United States Bureau of Mines Report of Investigations 8507 (USBM RI 8507).

CSA S304 TECHNICAL COMMITTEE ON OCCUPATIONAL HEARING CONSERVATION

Tim Kelsall

Director of Noise and Vibration, Hatch, 2800 Speakman Drive, Mississauga, ON L5K 2R7 - tkelsall@hatch.ca

The Canadian Standards Association S304 Technical Committee on Occupational Hearing Conservation consists of one subcommittee writing a standard on Hearing Conservation Management, five subcommittees covering such aspects as Hearing Protection, Noise Exposure Assessment and Control, Hearing Surveillance, and Vibration Exposure. This paper will give the history of its formation, its current activities and discuss how people can get involved in our work.

1. INTRODUCTION

For over 30 years, the Canadian Standards Association (CSA) Technical Committee Z107 was the main body involved with acoustical standards in Canada, (with TC Z94.2 handling a single standard on hearing protection). In 2009, CSA decided to focus their standards efforts on occupational noise and handed over other aspects of acoustics to the Canadian Acoustical Association.

A new CSA Technical Committee, TC S304 on occupational noise, was formed which combined the former Z107 standards related to occupational noise with the Z 94.2 standard on hearing protectors.

The work is done by a series of subcommittees discussed below. Those interested in working on these issues are invited to contact the author (who chairs S304) or the subcommittee chairs.

2. SC1 - HEARING PROTECTION DEVICES – ALBERTO BEHAR

S304 is still responsible for the Z94.2 standard on Hearing Protection. For many years this standard has advocated the use of type A, B, and C hearing protectors. These ratings are not widely used, because most protectors are also sold in the US, where by law they must be labelled with the protector's NRR rating, which is thus much more widely known.

The NRR system, put in place in 1974, meanwhile, is now recognised by most experts as quite flawed because the number shown grossly overstates the actual protection received. The new draft Z94.2 is advocating the NIOSH approach to derating the NRR ratings which derates the NRR by 25-70% depending on the type of protector (minus 7 dB to account for the difference between dBA and dBC).

At this point a new draft of the standard has been written, but there is a concern that if CSA comes out with it before the US decides how they are going to address these ratings, we could again be out of step with that large market for hearing protection.

3. SC2 NOISE EXPOSURE ASSESSMENT AND CONTROL – STEPHEN BLY

This subcommittee is responsible for several standards inherited from the former TC Z107:

3.1 CSA Z107.56

The most widely used of the Z107 series, Z107.56 covers the measurement of occupational noise exposure. A new version is now being proposed which will extend its scope to cover noise exposure under headsets. The new approach encompasses measurements with probe microphones in real ears, measurements using mannequins and artificial ears and a new calculation method using the NR of the headset and the measured sound level outside the headset.

The use of probe microphones and mannequins is covered by Australian and international standards, to which the new version refers. However the calculation method is new. It is intended to be a low cost initial assessment compared to the other systems. It is based on research indicating that most people adjust the volume on a headset to be about 15 dB above the existing background sound inside the protector.

3.2 CSA Z107.58

This standard describes in one location all that Canadians need to know to navigate the variety of standards, codes and regulations which make up the system whereby the sound produced by machinery is documented and available to prospective buyers and users. Health Canada has recently recommended its use by Canadian industry and a new version is expected which will update the constantly changing standards on which the system is based. This system helps industry buy quiet equipment and helps manufacturers provide purchasers with information about sound levels produced by their equipment

4. SC 3 HEARING SURVEILLANCE – CHRISTIAN GIGUERE

SC3 looks after the former Z107.6 standard on Pure Tone Air Conduction Threshold Audiometry. Theyu also will be providing guidance on hearing testing for the Hearing Conservation Management standard.

5. SC 4 VIBRATION EXPOSURE ASSESSMENT & CONTROL – TONY BRAMMER

This subcommittee provides a liaison with the committee responsible for ISO 2631, which Tony Brammer also chairs and will be responsible for writing the vibration section of the Hearing Conservation Management standard.

6. SC5 HEARING CONSERVATION MANAGEMENT – JEFF GOLDBERG

CAALL-OSH, the Occupational Safety and Health Committee of the Canadian Association of Administrators of Labour Legislation, agreed to fund the development of a new Canadian standard on Hearing Conservation Management, which gave a large impetus to the new Technical Committee. This new standard would be part of CSA´s OHS Management systems standards series. It would encompass prevention of occupational hearing loss, control of noise in the working environment and be applicable to all occupational sectors and to all workers and occupations. This work was undertaken by SC1 chaired by Jeff Goldberg, who have just completed the first draft of the standard.

COMMUNICATION HEADSET USE AND NOISE MEASUREMENT IN THE WORKPLACE

Flora Nassrallah[1], Christian Giguère[2,3], and Hilmi R. Dajani[3]
[1]Institute of Population Health, U. of Ottawa, 1 Stewart St., Ontario, Canada, K1N 6N5
[2]School of Rehabilitation Sciences, U. of Ottawa, 451 Smyth Rd., Ontario, Canada, K1H 8M5
[3]School of Electrical Engineering and Computer Science, U. of Ottawa, 800 King Edward Ave., Ontario, Canada, K1N 6N5

1. INTRODUCTION

Noise in the workplace is a significant risk factor causing on average 16% of adult-onset hearing loss worldwide [1]. According to the U.S. National Health and Nutrition Examination Survey, 22 million American workers are exposed to hazardous noise on the job and approximately 10 million Americans suffer from hearing loss as a result [2].

Amongst various factors, noise exposure in the workplace can result from the use of communication headsets or other wearable listening devices. With the increasing use of wired and wireless communication headsets, they are now commonly found in many occupational settings, such as call centers, fast food outlets, airport ground and control tower operations, industrial and construction sites, military sites, etc. Some workers wear noise-excluding headphones to attenuate the very noisy background and enhance the received communication signal [3]. Other workers use hands free communication headsets or low attenuation devices in an environment where background noise is not as significant [3]. In all cases, workers are exposed to the surrounding workplace noise and to the internal acoustic signals from the communication headset. Methods for noise exposure assessment under headsets vary widely in complexity, and relatively little is known about measurement reliability and the degree of agreement between methods.

2. EXISTING METHODS

Several challenges present themselves when conducting noise measurements under communication headsets. Firstly, measurements under occluded ears are dependent on the acousto-mechanical properties of the outer ear [4]. Secondly, in-ear measurements must be converted to free field measurements in order to enable a comparison with occupational standards [4]. Finally, measurement of noise under communication headsets involves three interdependent factors: the ambient noise, the internal signal from the headset which varies in duration, and the headset attenuation. Consequently, because of the complexity of conducting noise measurements under occluded ears, specialized equipment and techniques are required for direct and indirect sound measurements under communication headsets. The equipment must 1) capture the acoustic signal path from the headset and the external background noise path passing through the device, 2) consider the headset attenuation if applicable, and 3) allow the worker to operate normally while the measurements are being taken.

2.1 Standardized Methods

The ISO 11904 describes two procedures for the direct measurement of sound under communication headsets.

For the Microphone in a Real Ear (MIRE) technique (ISO 11904-1) [5], a specially trained individual must insert a miniature or probe microphone in the worker's ear canal. While the worker is exposed to close and far sound sources, sound pressure measurements in the ear canal are obtained in 1/3 octave bands. Measurements are then converted back to the free-field levels. Although this technique provides the most direct estimate of sound exposure to the worker, its invasive nature could restrict his/her movement, and the positioning of the microphone or probe can be problematic.

The ISO 11904-2 [6] specifies sound measurements on an acoustic manikin. The standardized manikin simulates the changes that happen to sound waves as they pass a human head and torso. Sound pressure levels are measured by the artificial ear embedded in the manikin in 1/3 octave frequency bands. For direct measurements in the field, one headset is worn by the worker and a second one is placed on the manikin positioned near the worker. An alternative is to record the electrical signal from the worker's headset and to either playback the signal on a manikin in the laboratory or filter the signal according to the electrical-to-acoustic transfer function of the headset [7]. In both cases, the worker is free of any measuring device or microphone and carries out tasks in the most natural way while the noise is being measured. Unfortunately, the required equipment and instrumentation is not widely available, and the differences in shape and mechanical properties between the worker and manikin's ear can be a limiting factor.

Since the manikin method can be cumbersome to use in the workplace, the Australian and New Zealand Standards AS/NZ 1269.1 proposed a general-purpose artificial ear procedure [8]. Although this method is less expensive, more practical, and more accessible, it was not initially designed to conduct sound measurements under communication headsets and must therefore be used with care.

2.2 Calculation Method

An indirect calculation method has also been proposed that can serve as a basic tool to predict noise exposure in workplace settings where communication headsets are worn [9]. It is based on the assumption that to hear and understand the headset signal, workers must adjust the volume of the signal depending on the level of the surrounding ambient noise (L_N) and the headset attenuation (ATT). The equation below comes from the formula used to sum sound levels from two independent sources and considers the signal from the surrounding background noise (term 1) and the communication signal from the headset (term 2) corrected for the listening duration (t) over the noise assessment period (T), assuming a listening signal-to-noise ratio (SNR) of 13.5 dB [10]:

$$L_{ex,T} = 10\log\left(10^{(L_N - ATT)/10} + \frac{t}{T}10^{(L_N - ATT + SNR)/10}\right)$$

This prediction tool offers a simple approach that industrial hygienists or safety officers can use in the early stages of a noise assessment program, or in environments where there is no access to more complex acoustical measurement equipment.

3. A FIELD MEASUREMENT GUIDE

Despite the increase of communication headsets in the workplace, guidelines are lacking regarding their safe use. Moreover, health and safety officers often have limited access to resources to conduct noise measurements under these devices. The overall goal of this project is therefore to examine, compare, and develop knowledge on the best measurement tools to evaluate the noise exposure of individuals who wear communication headsets in the workplace. Another purpose is to provide valid and practical measurement tools for health and safety officers. A four step study is suggested in order to attain the proposed goals.

3.1 Questionnaire

Due to the lack of information on communication headset use in the workplace, a questionnaire will be administered to document the extent of the problem by determining the distribution of the type of devices found in the workplace, the type of working environments and occupations where these devices are used, and the actual practices in terms of selection and usage. Health and safety officers and other stakeholders in hearing loss prevention will be surveyed on their knowledge of noise exposure from these devices, their awareness of the problem, their ability to measure sound levels, and their access to the proper measuring equipment.

3.2 Comparison of the direct measurement methods

Although both the acoustic manikin and artificial ears are suggested to take noise measurements under communication devices, the artificial ears are more compact and accessible. However, there is no data comparing the acoustic manikin and artificial ears for the measurement of noise under headsets. In a laboratory setting, the test-retest measurement variability and the degree of agreement for each of these methods will be determined.

3.3 Refinement of the indirect calculation method

Due to the complexity of conducting measurements with artificial ears and manikins, reliance on these methods may limit the workplaces where noise exposure under communication headsets is evaluated. The calculation method is an easy alternative to indirect measurement of noise under communication headsets. However, it requires further work to determine if the average SNR of 13.5 dB observed in other studies varies in different conditions. In a laboratory setting simulating the workplace environment, different conditions will be tested such as: background noise types, background noise levels, headset types, and headset configurations.

3.4 Field measurements

Field measurements will be carried out to compare the various measurement methods as they would be used in real working conditions. While conducting these measurements, it will be possible to evaluate the constraints and logistics associated with each method. A significant challenge associated with taking noise measurements under communication headsets is the necessity for the workers to continue their work duties as naturally as possible while the measurements are being conducted.

4. FINAL REMARKS

This study will help establish detailed guidelines to assist health and safety officers in selecting the best measurement method to assess noise exposure under headsets according to the available resources and best practices, and in promoting safe use of communication headsets in the workplace.

REFERENCES

[1] Nelson DI, Nelson RY, Concha-Barrientos M, Fingerhut M. The global burden of occupational noise-induced hearing loss. Am J Ind Med. 2005; 48(6): 446–458.
[2] Centers for Disease Control and Prevention. Noise and hearing loss prevention [Internet]. 2012 [updated 2011 July 28; cited 2012 July 25]. Available from: http://www.cdc.gov/niosh/topics/noise/.
[3] Williams W, Presbury J. Observations of noise exposure through the use of headphones by radio announcers, Noise Health. 2003; 5(19): 69-73.
[4] Dajani H, Kunov H, Seshagiri B. Real-time method for the measurement of noise exposure from communication headsets. Appl. Acoust. 1996; 49(3): 209-224.
[5] ISO 11904-1. Acoustics – Determination of sound immission from sound sources placed close to the ear: Technique using a microphone in a real ear (MIRE technique). Geneva, Switzerland: ISO; 2002.
[6]. ISO 11904-2. Acoustics – Determination of sound immission from sound sources placed close to the ear: Technique using a manikin (Manikin technique). Geneva, Switzerland: ISO; 2004.
[7] OHSA. Evaluating noise exposure of employees wearing sound-generating headsets. Washington, DC: Occupational Safety and Health Administration; 1999. Report No.: TED 01-00-015.
[8] AS/NZS 1269.1. Occupational Noise Management Part 1: Measurement and Assessment of Noise Immission and Exposure. Wellington, New Zealand: AS/NZ; 2005.
[9] Behar A, Giguère C, Kelsall T. Measurement of noise exposure from headsets. Proceedings of Noise-Con; 2008 July 28-31; Dearborn, MI.
[10] Giguère C, Dajani HR. Noise exposure from communication headsets: the effects of external noise, device attenuation and effective listening signal-to-noise ratio. Proceedings of Inter-Noise; 2009 August 23-26; Ottawa, Canada.

Abstracts for Presentations without Proceedings Paper
Résumés des communications sans article

Speech Communications / Sciences de la parole

Vowel-inherent spectral change in the spontaneous speech of Nova Scotians
Michael Kiefte & Terrance Nearey

The purpose of this research was to examine patterns of vowel-inherent spectral change [Andruski and Nearey, J. Acoust. Soc. Am. 91, 390 (1991)] in both spontaneous and read speech. Recordings were made from 197 long-time residents of Nova Scotia. Participants were asked to read a series of sentences in the context "Please say the word /hVd/ again." with each of the 14 vowels found in standard English. In addition, participants were also provided a monologue on a topic of their choosing. Formant frequencies were measured for stressed vowels and were normalised for speaker vocal-tract size [Adank et al., J. Acoust Soc. Am. 116, 1729 (2004)] and corrected for place-of-articulation effects from neighbouring consonants. Similar to other dialects of English [Hillenbrand et al., J. Acoust. Soc. Am. 97, 3099 (1995)], large formant transitions were observed for all vowels with the exceptions of /i/ and /É'/. In contrast to data reported by Labov et al. [The Atlas of North American English (2006)], the merger of /É'/ and /É"/ was not observed in conversational speech. In addition, /Ã¦/ was raised and fronted relative to that reported for other parts of Canada indicating that speakers in this region are not participating in the Canadian Vowel Shift [Clarke et al. Lang. Variation, Change 7, 209 (1995)]. In spontaneous speech, there is a general trend toward vowel centralisation for all vowels as well as a large degree of fronting for /u/ and /ÊŠ/.

Compensatory response to second formant shifted feedback in French and English speakers
Takashi Mitsuya, Fabienne Samson, Lucie Ménard & Kevin Munhall

Past studies have shown that when speakers receive auditory feedback with altered formant frequencies, they change their formant production, trying to offset the difference between the feedback and the intended sound. This behavior was thought to reduce an overall acoustic error, and thus not be influenced by the language that is being spoken. However, a recent study by Mitsuya et al. (2011) demonstrated that the magnitudes of compensation for altered F1 varied for two language groups depending upon the direction of perturbation (increased/decreased F1 in Hz). These results are inconsistent with the hypothesis that formant compensatory behaviour is a simple acoustic adjustment. Rather, the data suggest that the target vowel is represented in relation to the vowels around it in the F1/F2 acoustic space. The current study attempted to generalize Mitsuya et al.'s hypothesis by examining 1) different language groups (English speakers [ENG] and French speakers [FRN]) and 2) a different acoustic parameter (F2). In the experiment, participants produced the vowel /ε/. The acoustic feedback was gradually modified by decreasing the frequency of F2. This perturbation made the vowel /ε/ sound more like /œ/ (phonemic in French, but not phonemic in English). The results showed that both groups increased their F2, but FRN 1) started compensating to smaller perturbations and 2) showed greater maximum compensations than ENG. In addition, the threshold of compensation of the groups seems to mirror their perceptive /ε-œ/ boundary. Finally, FRN also increased their F3 (associated with lip rounding) while ENG did not.

The synergy between laryngeal constriction and open vowels: An examination of acoustic observations from natural speech and computational modeling of the relationship
Scott Moisik & John Esling

Several researchers have observed that the epilaryngeal tube has consequences for vocal fold function on both mechanical (Laver 1980) and aero-acoustic levels (e.g. Titze 2008). Laryngoscopy data also supports the view that the epilaryngeal tube plays a key phonetic role in the production of phonemes traditionally identified as pharyngeal consonants as well as in glottal stop and creaky voice production (Edmondson & Esling 2006). We believe that it also plays a more general role in vowel production that has not yet been fully explicated.
Our contention is that the consequences of epilaryngeal stricture are encountered in everyday speech with specific distributional regularities that can be attributed directly to physiological aspects of the relationship between the epilarynegal tube, the tongue, and the vocal folds. The key argument is that the articulatory synergy of the lingual and laryngeal structures required to produce epilaryngeal stricture is such that there is a greater tendency for laryngeal constriction to occur in the context of relatively more open vowels with relatively higher levels of F0, meaning that the tongue and laryngeal configurations will favour the engagement of epilaryngeal stricture.

In support of our hypothesis, we first present several case studies of English speech data obtained from various sources, including the popular media, online media resources (such as lectures and reviews), and recordings of people playing board games (which frequently feature lively discussion). Our purpose with these data is to illustrate the acoustic and auditory dimensions of epilaryngeal constriction as it occurs during natural and performed speech and to draw comparisons to similar data obtained in the laboratory setting where laryngeal constriction can be visually confirmed. Laryngeal constriction in these examples overwhelmingly occurs with open vowels such as [ɛ æ ɑ ɔ] and is most likely to be found on emphatically stressed tonic syllables.

The second component of our study features two computational models used to illustrate the biomechanical consequences of epilaryngeal stricture. First, we demonstrate the action of epilaryngeal stricture on the lingual-laryngeal system as a whole using a model developed in ArtiSynth (Stavness, Lloyd, Payan & Fels 2011). This model provides insight into how the progressive narrowing and compaction of the epilaryngeal tube ultimately leads to vocal-ventricular contact. We then present a lumped-element model that seeks to simulate the effects of this vocal-ventricular contact in the form of weak mechanical coupling. Although further work is required in both of the domains examined here, the results suggest that a move away from the linguo-centric view of vowels is warranted. Our work shows that the epilaryngeal tube is a physiological system that interconnects the tongue and larynx in a way that makes their mutual, synergistic engagement inevitable even in the most everyday contexts.

References:
Edmondson, J.A., Esling, J.H. 2006. The valves of the throat and their functioning in tone, vocal register, and stress: laryngoscopic case studies. Phonology 23. 157-191.
Laver, J. 1980. The phonetic description of voice quality. Cambridge: CUP.
Stavness, I., Lloyd, J.E., Payan, Y., Fels, S. 2011. Coupled hard-soft tissue simulation with contact and constraints applied to jaw-tongue-hyoid dynamics. International Journal for Numerical Methods in Biomedical Engineering 27. 367-390.
Titze, I. R. 2008. Nonlinear source-filter coupling in phonation: Theory. Journal of the Acoustical Society of America 125 (5). 2733-3749.

Modeling the effects of vowel context in assessing vowel-inherent spectral change

Terrance Nearey & Michael Kiefte

Vowel production research has often focused on the formant frequencies of a single time frame taken near the vowel centre. However, vowel-formant trajectories across time are affected by at least two important factors: First, in some dialects of North American English, there is strong evidence of vowel-inherent spectral change (VISC) reflected in formant variation in a wide range of vowel categories, including vowels usually regarded as monophthongs [Andruski and Nearey J. Acoust. Soc. Am. 91, 390 (1991)]. Second, there are strong influences from adjacent consonants on vowel formant frequencies [Hillenbrand et al., J. Acoust. Soc. Am. 109, 748 (2001)] that may obscure VISC patterns. This paper presents two models of /CVC/ coarticulation that can be used to better extract underlying vowel-related formant change by factoring out the effects of adjacent consonant contexts. The first model is an elaboration of one proposed by Broad and Clermont [J. Acoust. Soc. Am. 81, 155 (1987)], that incorporates VISC by modeling formant trajectories throughout the duration of the vowel. The second model specifically exploits the near-linear relation between vowel and onset/offset formant frequencies using a method related to locus equations [e.g., Sussman, et al., J. Acoust. Soc. Am. 90, 1309 (1991)]). This simpler method can be used to obtain vowel formant-frequency information at given time frames corrected for consonantal context. Results will be presented 1) comparing alternative models on naturally produced recordings and 2) testing the efficacy of the simpler model on synthetic stimuli generated by more elaborate trajectory models.

Apport des méthodes multifactorielles pour l'analyse de la variabilité spectrale des données phonatoires

Abderrazak Rougab, Mhania Guerti & Mohamed Seddik Rougab

Les méthodes multifactorielles ont fait la sujette de plusieurs recherches notamment, la méthode d'analyse en composantes principales (ACP). Dans ce travail, nous utilisons l'ACP d'une part afin de localiser la corrélation entre les spectres des productions vocaux des nos locuteurs. Et d'autre part, pour réduire les dimensions du corpus de données dans le but d'aboutir à une représentation parcimonieuse. En outre une classification par la méthode de Clustring a été aussi impliquée dans le but de grouper par catégories les locuteurs possédant des caractéristiques spectrales phonatoires similaires.
Les résultats obtenus ont confirmé la possibilité de réduire les dimensions des données de (locuteurs*1025) à (locuteurs*2) et cela grâces au choix précieux des composantes principales les plus dominantes avec un taux d'informativité de l'ordre de 80%. En plus la représentation par l'ACP et par Clustring des spectres montre une bonne détection des similarités phonatoires chez les locuteurs.

Mapping perceptual distances between sibilant and palatal fricatives
Corey Telfer

Sibilant fricatives, such as s and ʃ, are generated by directing an oral airstream at an obstacle in the vocal tract in order to generate high intensity, high frequency noise (Shadle 1985, 1991). These speech sounds are relatively complex, and have proven difficult to characterize in both acoustic and articulatory terms (Boersma & Hamann 2008, Toda, Maeda & Honda 2010). Sibilants are typically quantified using global acoustic measures such as centre of gravity, peak frequency and spectral tilt, but these values are not precise enough to give a detailed description of these sounds, nor are they adequate for accurate speech synthesis. The perception of sibilants is likewise poorly understood, and understanding how listeners differentiate these sounds will help to reveal their internal structure.
This study was designed to investigate the relative perceptual similarity of the sibilant fricatives s, ʃ, ɕ, and the palatal fricative ç. Phonetically representative versions of these four speech sounds were synthesized by filtering white noise. The filter shapes were then interpolated linearly between pairs of the prototype stimuli to create a thirteen-step continuum. Listeners were given an AX discrimination task where they were presented with pairs of interpolated stimuli and asked to determine if each pair was the same or different. The participants' reaction times were analyzed with multidimensional scaling and hierarchical clustering. Results of these analyses suggest that stimuli near the prototype s stimulus were easiest to discriminate. Possible implications of these results for theories of phonetic and phonological distinctiveness will be discussed.

Identifying reduced speech in prepositional phrases
Benjamin Tucker & Dallin Mendenhall

Researchers have found that factors such as phonetic context (e.g., Podlubny et al., 2011, Warner et al., 2010) and semantic/ syntactic context (e.g., Ernestus et al., 2004, Ven et al., 2011) affect the processing of spontaneous (reduced) speech. The present study focuses on the contribution of predictability in prepositional phrases in reduced speech, e.g. on the table (high frequency phrase). We created a list of 246 prepositional sequences varying in frequency from high to low. These phrases were recorded so that the degree of reduction varied (e.g. items with shorter duration contained reductions and items with longer duration contained fewer or no reductions). We predict that the reduced items would be more difficult to process, due to the decreased acoustic information and that Frequency would have a large impact on the unreduced items and a smaller influence on the reduced items. In a lexical decision task with 58 native English speaking listeners we find that response times to real words are indeed influenced by Frequency and Reduction. However, preliminary analysis shows that reduced forms are responded to more quickly (when stimulus duration is controlled) than unreduced items. We do find, as predicted, that frequency of the whole phrase has a stronger effect on the unreduced items (higher frequency items are responded to more quickly) than the reduced items. These results are discussed in light of previous results and how contextual predictability affects word identification in reduced speech.

Bioacoustics / Bioacoustique

The North East Pacific Time-series Undersea Networked Experiments (NEPTUNE Canada): A comparative study of sound production in two marine environments
Francis Juanes, Rodney Rountree & Corinne Pomerleau

The North East Pacific Time-Series Undersea Networked Experiments (NEPTUNE) is part of the Ocean Networks Canada (ONC) Observatory. NEPTUNE Canada is the world's first regional-scale cabled observatory network and is located in the Pacific Ocean, off the west coast of Vancouver Island, British Columbia. The network extends from the rocky coast to the deep abyss across the Juan de Fuca plate and it gathers live data from a rich constellation of instruments deployed in a broad spectrum of undersea environments. Data are transmitted via high-speed fibre optic communications from the seafloor to an innovative data archival system at the University of Victoria. As part of NEPTUNE Canada, we developed a project related to sound production in the marine environment. Hydrophones were deployed at two NEPTUNE Canada node sites; Barkley Canyon located on the continental slope (1000 m) and Folger Passage on the continental shelf (100 m) and ambient sounds were recorded simultaneously at those two sites over the course of one year (2010-2011). Our main objectives were to develop a library of sounds collected at NEPTUNE sites including anthropogenic and biotic sounds (eg. whales and fishes) and to determine spatial (across locations and depths) and temporal (daily, seasonal) patterns in sound production, and correlate sound production patterns to environmental sensor data including anthropogenic noise levels. The ultimate goal of this ongoing research is to advance the study of deep-sea ecosystems by providing the first directed study of deep-sea fish sound production and to monitor their evolution related to climate change.

Proximal mechanisms for sound production in male Pacific walruses
Ole Larsen & Colleen Reichmuth

The songs of male walruses during the breeding season have been noted to have some of the most unusual characteristics that have been observed among mammalian sounds. In contrast to the more guttural vocalizations of most other carnivores, their acoustic displays have impulsive and metallic features more similar to those found in industrial work places than in nature. The patterned knocks and bells that comprise male songs are not thought to be true vocalizations, but rather, sounds produced with structures other than the vocal tract and larynx. To determine how male walruses produce and emit impulsive and metallic signals, we conducted a series of in situ studies with two captive adult male Pacific walruses that were trained for voluntary participation in bioacoustic research. Through a combination of observational, acoustic, endoscopic, and ultrasonic methods, we confirmed the probable anatomical origins of knocking and bell sounds and gained a mechanistic understanding of how these sounds are generated within the body and transmitted to the environment. These pathways are illustrated with acoustic and video data and considered with respect to the unique biology of this species.

Tracking multiple marine mammals using widely-spaced hydrophones
Eva-Marie Nosal

Recent interest in methods to track marine mammals using passive acoustics stems from concern over anthropogenic impacts on the marine environment, and considerable progress has been made in the field over the past several years as a consequence. Nevertheless, significant challenges persist. One of the more tricky situations occurs when widely spaced hydrophones record multiple animals that make similar calls with short inter-call-intervals. In this case it can be very difficult to establish how many animals are calling, let alone to track the individuals. Conventional tracking methods require that the calls are pre-associated for an individual so that times of arrival (TOAs) or time-differences of arrival (TDOAs) can be used to locate the animal. This talk will discuss a different approach, in which association and localization/tracking are performed simultaneously. A tracking algorithm that relies on TOAs or TDOAs and that uses joint association/tracking will be described. Results from simulations and application to real data will be presented. Work supported by ONR and NSF.

Multivariate classification of echolocation clicks of Commerson's dolphins
María Vanesa Reyes Reyes, Miguel Iñíguez, John Hildebrand & Mariana Melcón

Commerson's dolphin's (Cephalorhynchus commersonii) inhabits exclusively the Southern hemisphere and its distribution is restricted to coastal waters of Southern South America and Kerguelen Islands. Little information exists on the acoustic signals emitted by this species in the wild. We recorded 478 min containing dolphins' vocalizations, using a single hydrophone in Bahía San Julián, Patagonia Argentina. Signal parameters were calculated and a cluster analysis was made on a random subsample of 850 echolocation clicks, using two parameters as variables. We obtained nine clusters where the different clusters differed in peak frequency and/or 10dB minimum frequency (ranging from 121 to 141 kHz peak frequency, and 115-125 kHz 10dB minimum frequency). Buzz clicks were analyzed separately. They had mostly two or three peak frequencies between 129 and 140 kHz, with minimum differences in sound level. This study provides the first evidence that Commerson's dolphins may produce a variety of echolocation clicks, instead of only one type of stereotyped narrow-band high frequency clicks. This information could be useful while doing passive acoustic monitoring.
We also detected other signals in a lower frequency range of clicks which may serve for communication. These novel results will be discussed.

Time versus frequency controversy in the bottlenose dolphin auditory analysis
Gennadi Zaslavski

Studies on a dolphin's sonar target recognition are greatly focused on acoustic cues associated with individual echoes returned by ensonified targets. Theoretical time resolution of a bottlenose dolphin echolocation click is as high as 15 -20 microseconds. Given a very broad frequency range of a bottlenose dolphin hearing extending from 5-10 kHz to as far as 130-140 kHz, the dolphin sonar time resolution could be as high as the theoretical time resolution of the echolocation clicks. In response to a brief echolocation click a target usually returns a multiple highlights echo having rippled energy spectrum. Theoretically, the time domain as well as the frequency domain cues both can be used to discriminate targets. In the psychophysical experiment discussed in this paper a bottlenose dolphin discriminated two pairs of clicks having rippled energy spectra. The pair spectra differed in location of the ripples within the dolphin hearing range. The locations of the energy peaks in one spectrum coincided with locations of the energy troughs in the other. The dolphin was able to discriminate the pairs having the ripple modulation smaller than 1 dB and interclick intervals as small as 10 microseconds. Surprisingly, the dolphin continued to discriminate the signals even when a third click was symmetrically added to each pair transforming them into the time reversed triplets with identical energy spectra. We do not have clear cut explanation to this phenomenon yet.

Noise and Noise Control / Bruit et contrôle du bruit

Upgrade of a multi-channel active noise control system for an industrial chimney
Louis-Alexis Boudreault, André L'Espérance & Alex Boudreau

Active Noise Control has been studied in the 90s as an innovative to reduce the noise in some specific situations. Some applications are well known today and some of them found commercial success like the noise-cancelling headphones. The use of active noise control in industrial applications is however more complex thus being an uncommon solution in this environment. Active noise control for an industrial chimney stack is one of these applications. One of the first large-scale implementation has been done at the end of the 90s. In this case, a 10-Channel active noise control system has been installed in a 1.8 m wide chimney to attenuate a 320 Hz pure tone. At that time an 8 dB noise reduction has been achieved.
15 years later, it has been decided to update the system to take advantage of signal processing improvements during this period using the latest generation of DSP (Digital Signal Processor). It allows obtaining a real-time optimization and a better tracking capability for the controller. This paper presents the multi-channel active noise controller developed for this application and the noise reduction achieved. In addition, laboratory experiments were performed to determine the importance of several aspects on the system performance like the speed and precision of the controller and the geometrical location of the error sensors.

Compliance and vegetated-barrier acoustical testing in a purpose-built sound-transmission suite
Murray Hodgson, Shira Daltrop, Rick Peterson & Paul Benedict

The transmission loss and absorption coefficient of a vegetated noise barrier of Criblock™ construction were measured in a sound-transmission suite built specifically for the purpose. The suite was tested for compliance with ASTM E90-09, and found to be substantially, but not completely, in compliance. It was found that the transmission loss of the vegetated barrier ranged from 42 dB at low frequencies to 66 dB at 1000 Hz; above 1000 Hz, only a lower limit of the TL could be determined ? values of 57-62 dB were found. These values are at least 25 dB higher than recommended. The absorption coefficients of the unplanted and planted barriers were measured; the plants decreased the absorption slightly, from NRC 0.42 to 0.37.

On the make-up and applications of acoustic arrays
Martha Moore & Richard Peppin

An acoustic array is used to locate sources in a far field. Often, due to the complexity or locations of the sources, one can't find source contributions or even source locations with sound level meter or acoustic intensity. The array makes this simple. This paper explains the basics of arrays, gives some examples of use. While focusing on one brand, we want to emphasize that most brands are similar, all with their own advantages and limitations.

Acoustical characterisation of NRC Construction's 150 mm and 200 mm concrete slabs
Norbert van Lier, Ivan Sabourin, Berndt Zeitler & Stefan Schoenwald

NRC Construction has two concrete reference slabs (150 mm and 200 mm thick) available for research and fee for service purposes. The most common standard tests conducted on these slabs with and without floor treatments (covering and/or toppings) include the airborne sound transmission loss tests ASTM E90/ ISO 10140-2 and impact sound transmission tests ASTM E492/ ISO 10140-3. In this paper the bare concrete slabs are characterised through comparison of the airborne transmission loss and the impact sound transmission with each other as well as with the ISO 15712 prediction model. Input data for the ISO prediction model is gained through other measurement methods, such as structural reverberation according to ISO 10848.

Underwater Acoustics / Acoustique sous-marine

Trans-dimensional dispersion inference in porous sediments with strong transition layers
Jan Dettmer, Stan Dosso & Charles Holland

This paper applies a trans-dimensional (trans-D) hierarchical model to Bayesian inversion of seabed reflection-coefficient data to infer sound-wave dispersion in a layered seabed which exhibits strong velocity and density gradients. The porous seabed is described by a Buckingham's viscous grain-shearing model which obeys causality and accounts for attenuation-frequency dependence and sound-velocity dispersion. The sediment dispersion relationships are inferred over large bandwidth (500--6300 Hz) and angular range (10--80 degrees) resulting in high-resolution, low-uncertainty parameter estimates. The trans-D approach treats number and positions of layers as unknown and implicitly accounts for uncertainty due to the model parametrisation. In addition, both abrupt discontinuities and transition layers (arbitrary gradients) are intrinsically addressed by the trans-D model without requiring prior specification. To ensure sufficient convergence of the inversion algorithm, interacting Markov chains are applied to sample the non-linear trans-D posterior distributions. The data were collected in the Straits of Sicily, Mediterranean Sea, using a seismo-acoustic boomer sound source and a single hydrophone mounted above to the sea-floor. The approach estimates sound-velocity dispersion and attenuation-frequency dependence for the complex sediment at this site and quantifies the uncertainties. The results indicate a highly layered sediment, including several strong transition layers, which appear to be well addressed by the trans-D hierarchical model. Such sediments would be difficult to address by prior specification of a fixed dimensional model. Sound-velocity and density profile marginal distributions are compared to core estimates taken at the experiment site. [Work supported by the Office of Naval Research]

Low frequency shallow water fine-grained sediment attenuation measurements
Charles Holland & Stan Dosso

Attenuation is perhaps the most difficult sediment acoustic property to measure, but arguably one of the most important for predicting passive and active sonar performance. Measurement techniques can be separated into 'direct' measurements (e.g., via sediment probes, sediment cores and laboratory studies on 'ideal' sediments) which are typically at high frequencies, O(10^4 -10^5) Hz and 'indirect' measurements where attenuation is inferred from long-range propagation or reflection data, generally O(10^2-10^3) Hz. A frequency gap in measurements exists in the 500 - 5000 Hz band and also a general acknowledgement that much of the historical measurements on fine-grained sediments have been biased due to a non-negligible silt and sand component. A shallow water measurement technique using long range reverberation is critically explored. An approximate solution derived using energy flux theory shows that the reverberation is very sensitive to depth-integrated attenuation in a fine-grained sediment layer and separable from most other unknown geoacoustic parameters. Simulation using Bayesian methods confirms the theory. Reverberation measurements across a 10 m fine-grained sediment layer yield an attenuation of 0.009 dB/m/kHz with 95% confidence bounds of 0.006-0.013 dB/m/kHz. This is among the lowest values for sediment attenuation reported in shallow water.

Range-dependent reverberation and target echo calculations using a Gaussian beam model
Sean Pecknold, Dale Ellis & Diana McCammon

The Bellhop Gaussian-beam underwater acoustic propagation model, developed by Porter [Porter MB, Bucker HP. Gaussian beam tracing for computing ocean acoustic fields. J. Acoust. Soc. Am. 82(4):1349–1359 (1987)] has been extended to include reverberation and signal excess calculations for active sonar problems. Here, we use this model to predict reverberation and target echo for several benchmark sonar test problems. These problems, defined for the 2006 and 2008 US Reverberation Modeling Workshop and the 2010 UK Symposium on Sonar Performance Assessment Tools, include both flat-bottom and range-dependent cases. We then compare the predictions from Bellhop to those made using other models, including an adiabatic normal mode model, in an attempt to gain insight into their limits of validity.

Extending the capability of the complex effective depth approximation
David Thomson & Gary Brooke
Zhang and Tindle's effective depth approxim
ation [J. Acoust. Soc. Am. 93, 205–213 (1993)] for treating shallow water propagation over a lossy, low shear speed, solid half-space has been generalized to accommodate a layered sediment structure and a rough sea-surface. As a result of this new capability, the complex effective depth paradigm provides an efficient iterative algorithm for determining the exact complex eigenvalues for the normal modes in an isovelocity sea over a layered sea-bottom that bypasses the need for numerical searches in the complex plane. The method is straightforward to code and results in fast and accurate modal solutions of shallow water propagation that also includes the effects of leaky modes. Coherent forward scattering from a rough sea-surface is easily handled in the context of the Kirchhoff approximation. Several numerical examples are provided to illustrate the enhanced capability of the method.

Architectural and Building Acoustics / Acoustique architecturale

Noise control in multi-family buildings: updating the building code?
Bradford Gover, David Quirt, Stefan Schoenwald & Berndt Zeitler

The 2010 National Building Code of Canada (NBCC) currently has requirements intended to limit the transmission of airborne sound into a dwelling unit from spaces elsewhere in the building. A Task Group addressing these requirements for airborne sound insulation in Part 5 and Part 9 of the NBCC has developed a proposed revision for the 2015 NBCC. The fundamental technical change would focus the requirements on acoustical performance of the complete building system (ASTC between adjoining spaces), in place of the simplistic focus on just that part of the sound transmitted through the separating wall or floor (STC of that assembly). With this change, the NBCC would address the actual sound insulation perceived by occupants. This paper explains the implications and significance of the proposed change for various construction types, and defines some potential approaches for demonstrating compliance.

Prediction of apparent sound transmission class (ASTC) in typical Canadian buildings
Bradford Gover, Berndt Zeitler, Stefan Schoenwald, Stefan & David Quirt

ISO 15712-1 provides a set of procedures for calculating sound transmitted via direct and flanking paths, using data from standardized laboratory measurements to characterize the separating assemblies and their junctions. These measurements for wall and floor assemblies and their junctions are defined in ISO 15712-1 by reference to ISO measurement standards. To facilitate use in Canada, where industry relies primarily on acoustic test data conforming to ASTM standards, essentially equivalent ASTM standards may be used. The calculation method is ideal for buildings whose structural wall and floor assemblies are of homogeneous concrete and/or masonry, but for different types of construction including lightweight elements framed with wood or steel, or mixed types of construction, specific constraints and variants on the method are recommended. In these latter cases, direct and flanking sound transmission should be measured according to ISO 10848. Measurements and calculations conducted at the National Research Council Canada have been successfully used to develop a powerful software calculation tool, which can be used by designers to estimate acoustical performance for many designs. This paper discusses the approach, and highlights some limitations in the methods and gaps in the available data.

Prediction of the effect of a sound-masking system in an open-plan office, including the Lombard effect
Murray Hodgson & Yizhong Lei

Sound masking has been developed as an effective method for improving indoor speech privacy, by increasing background noise levels in the space that masking distracting speech sounds. However, the Lombard effect indicates that an increase of background noise level would also result in an increase of people's voice levels, and thus, makes the speech privacy worse, reducing or cancelling the benefit of the sound masking system. A model of an existing open-plan office was created using CATT-Acoustic. Ambient background-noise levels of 30, 40 and 45 dBA were considered; the room contained 1, 4, 7 or 10 pairs of talkers and listeners. Prediction results of reverberation time and speech intelligibility index were compared to measurements for model verification. The model was used to compute the speech intelligibility index at 1 and 4 m from a speech sourceand to predict the effect of a sound masking system, with and without considering the Lombard effect. With 30 dBA ambient noise, SII values at 1/4 m were ??/?? without the sound masking system, ??/?? with it but ignoring the Lombard effect, and ??/?? considering the Lombard effect; taking the Lombard effect into consideration reduced the benefit of the SMS by 10/20%. With higher ambient noise, the SMS was less effective; taking the Lombard effect into consideration reduced the benefit of the SMS by 20/40% at 40 dBA and by 50/70% at 45 dBA,

The Multifunctionality Degree as a parameter to evaluate the sound quality of multi-purpose halls
Lineu Passeri

Starting with an overview of the acoustic demands of theaters and auditoriums dedicated to specific types of shows, the author makes an analysis of the needs of multi-purpose halls, with emphasis on variable acoustic features, so that they meet the most diverse uses. Then, he addresses the quality aspect of room acoustics through a conceptual review of their subjective attributes and objective parameters for analyzing their sound quality, and how it is related to the question of multifunctionality. At last, the author proposes subsidies for designing multi-purpose halls, including the Multifunctionality Degree of rooms and other mathematical formulae relating the sound source, the observer, the direct sound and the reflected sound.

Analysis and diagnosis of the acoustics of three multipurpose halls to support modifications to meet the demands of the predominant programming
Lineu Passeri & Sylvio Reynaldo Bistafa

Starting with an overview of the acoustic characteristics of auditoriums for different types of programs, an analysis is made of the needs of multipurpose rooms, so that they can meet diverse uses. The sound quality aspects are addressed through a conceptual review of elicited subjective attributes and related objective parameters, and how they fit into the multipurpose room. These issues are confronted against measurement and simulated results of eight objective parameters (RT60, EDT, D50, C80, LEF, BR, TR, and ST1) in three multipurpose auditoriums case studies in São Paulo, Brazil. By means of calibrated computer models of these rooms, architectural adjustments were simulated in the virtual rooms, as a basis for proposing modifications in the real rooms, in order to optimize the sound quality for the predominant programming in each auditorium.

The effect of heavy toppings and ceiling designs on wood framed construction
Ivan Sabourin, Berndt Zeitler, Stefan Schoenwald

During a parametric study at NRC Construction to reduce low frequency impact sound, heavy toppings and different ceiling designs were tested on wood joist floors to determine their effects on direct transmission using airborne and light impact excitation sources. The second part of the study, the flanking study, focuses on the significance of flanking on the overall impact sound transmission between two rooms, one above the other, and on how a topping affects direct and flanking impact sound transmission differently.

Physio- and Psycho-Acoustics / Physio et Psychoacoustique

Effect of head movements on the spatial localization of reverse alarms with and without hearing protection
Véronique Vaillancourt, Chantal Laroche, Christian Giguère, Anne Gravel, Jérémie Chiasson

Despite the widespread use of reversing alarms on heavy vehicles, accidents still occur. Among others, poor spatial localization is a major concern with conventional tonal alarms consisting of a single pure tone. The broadband alarm technology was introduced into the marketplace to overcome such shortcomings of tonal alarms. Indeed, it is generally accepted that signals with a broader frequency spectrum are easier to localize, as more cues are available. A previous study compared the performance of tonal and broadband alarms on measures of sound detection, equal loudness, urgency ratings and sound localization, with and without the use of passive hearing protection (earplugs and earmuffs), in individuals with normal hearing. Despite similar results in spatial localization with and without earplugs, earmuffs where however shown to significantly disrupt localization cues, particularly in the front/back dimension. The present paper extends on the previous study by investigating the effects of head movements on localization performance and the inclusion of a communication headset. Overall, better localization results are obtained with the broadband technology over conventional tonal alarms, with and without hearing protection, and head movements significantly contribute to improving performance, particularly by reducing front/back confusions. In addition to superior spatial localization, other advantages of the broadband technology over tonal reversing alarms include a more uniform sound field behind heavy vehicles, lower sound pressure levels to achieve equal loudness and reduced environment nuisance.

Vibration, Engineering and Physical Acoustics / Vibrations, Génie et Physique acoustique

Diffraction correction for anechoic wind tunnel
Werner Richarz, Rafik Chekiri & Rob Jozwiak

The Anechoic Wind Tunnel (AWT) at the University of Toronto has been refurbished by Bombardier Aerospace, Aercoustics Engineering and the University of Toronto Institute for Aerospace Studies. The facility is now used to support various aero-acoustic projects. The AWT is of the open jet type, which has a mean flow speed of up to 60m/s. Sound generated by test objects immersed in the open jet is usually measured in the quiescent region outside the flow. As the shear layer refracts the sound waves, it is necessary to apply appropriate corrections so that better estimates of source directivity and power can be deduced from the measurements. A set of controlled tests were conducted to establish the corrections for a number of source-observer positions, frequencies and flow speeds. To this end a streamlined 'point source' was placed in the jet and the radiated sound measured with and without air flow. Emphasis was placed on cross-spectral densities of the input to the sound source and the sound measured at a field point. Changes in the magnitudes and phases of the cross-spectral densities permit one to establish key features of a propagation model. Cross-correlations between selected field positions were used to determine the virtual position of the source. This then permits one to match the sound field in the quiescent region with the (inferred) sound field in the moving stream.

Thermo-acoustic investigations of non-linear parameters of inorganic salts with fructose
Arun Upmanyu, D. P. Singh & Arvind Sharma

Thermo-acoustic investigations of the non linear properties of fructose in aqueous solution of NaCl and MgCl2 have been done in the temperature range, 298.15K - 308.15K at various concentrations using volume expressivity data. Several non-linear parameters such as Moelwyn-Hughes parameter (C1), Bayer's non-linear parameter (B/A), reduced volume( ), reduced isothermal bulk modulus (â),Sharma's constants (S0, S*, S*0), Huggins parameter (F), isobaric acoustical parameter (K), isochoric acoustical parameter (K"), isothermal acoustical parameter(K'), fractional free volume(f), repulsive exponent (ç), reduced volume expansivity (å*), internal pressure (Pi), cohesive energy density (åi), critical temperature (T*) and thermo acoustical parameter such as A* and B* have been studied. The non-linearity behavior of these parameters has been explained in terms of the concentration and thermal changes in the systems under investigation. The present analysis provides a deep insight about the nature and types of interactions prevalent in the solutions of electrolytes and fructose. The obtained results provide a reliable basis to understand the action of electrolytes and bio-molecules in the bio-chemical processes undergoing in a bio-physical system.

Acoustic Standards / Normalisation

A new hearing loss prevention program standard
Jeffrey Goldberg

CSA has been asked by the federal, provincial, and territorial OHS regulatory authorities (CAALL-OSH) to develop a standard for the Management of Hearing Conservation Programs in Canada. The development of this new standard, Z1007 Hearing Conservation Program Management, is now well underway. It is to be designed to permit persons not particularly familiar with the technical aspect of hearing conservation to design and manage programs where their workplaces noise levels require them. Further, the standard is being developed as a model that some jurisdictions might use as a basis for legislation or a “Best Practices” document. The regulators present on the CSA technical committee supervising this development have requested that the final document stand alone and not require other standards for its use. In this way the standard will tell those reading what to do with information rather than how to develop it. This session will expose the development done to date, the thinking behind the standard, and some of the issues the committee is facing in writing this document. Input would be valuable from all those with experience in this area.

Call for Papers

A Special Issue of the

International Journal of Industrial Ergonomics

OCCUPATIONAL NOISE EXPOSURE:EXPOSURE ASSESSMENT AND CONTROL

Exposure to occupational noise is related to hearing loss, discomfort, fatigue and several other health and safety risks among the exposed workers. Although the research efforts over the past few decades have evolved into valuable guidelines and standards to protect workers from excessive exposures to noise, the subject of health effects, assessment and control continues to pose an array of multi-disciplinary challenges. The objective of this special issue is to compile recent research and development efforts in the field, including characterization and assessments, industrial noise control, the state of the art in the associated supporting technologies, hearing protection and perspectives on future developments and applications.

The specific topics of interest within the scope of this special issue include (but not being limited) the following:

- Characterization and assessments of workplace noise environment and noise sources;
- Hearing protection;
- Audiological and non-audiological health risks;
- Communication in noisy environments and safety issues;
- Comfort and perception issues related to workplace noise and hearing protection;
- Epidemiology;
- Standards: applications and limitations;
- Ergonomic interventions for risk control;
- Techniques for noise mitigation and industrial noise control, active noise control;
- Effect of noise on human performance;
- Analytical and numerical methods for noise assessment and control.

Prospective authors are invited to submit their original works within the scope of the special issue. The authors should follow the journal guidelines (http://ees.elsevier.com/ergon/) for preparing their manuscripts, and submit electronically to the journal website using the web-based submission tools. Each manuscript will be reviewed in accordance with the journal requirements.

SCHEDULE FOR SUBMISSIONS (tentative)

Manuscript Submission Deadline:	15 March 2012
Reviewers' reports and decision:	30 April 2012
Final Manuscript Due on:	30 June 2012

GUEST EDITORS

R. Ramakrishnan, DSc., P.Eng
Associate Professor, Architectural Science
Ryerson University
350 Victoria Street
Toronto, Ontario, CANADA M5B 2K3
Email: rramakri@ryerson.ca

P. Marcotte, Ph.D.
IRSST, Research Department
505 boul. de Maisonneuve West
Montreal, Quebec H3A 3C2
Canada
Email: marcotte.pierre@irsst.qc.ca

Goodbye / Adieu

The September 2012 issue of the *Canadian Acoustics Journal* will be my last issue as the Editor-in-Chief of the journal. It is time for me to step down and hand over the reins to others who can take the journal into the digital age. I had a good run of 14 years. I would like to thank the Executive, and the Board of the Canadian Acoustical Association for their trust and support over these years. I was always provided with an adequate budget to do my job properly and satisfactorily. The board was behind me every time I wanted to try out something new, such as the June French issue, or the publishing of salient articles (peer reviewed) from Underwater Acoustic conferences. I am grateful for the support. Finally, to all the members of CAA and the readers of *Canadian Acoustics Journal*, my heartfelt thanks for giving me an opportunity to serve you all. Exciting and excited new members are waiting to take over my job, and I am very happy to hand over the responsibilities to them.

Ramani Ramakrishnan
Department of Architectural Science, Ryerson Univeristy, Toronto

Central Energy
Plant Louvers

Cooling Tower
Barrier Wall System

Environmental Noise Control

Community friendly solutions for chillers and cooling towers

- Over 50 Years of Proven Design and Performance
- Independently Tested Products
- On Grade or Roof Top Applications
- Maximum Noise Reduction
- Low System Pressure Loss

KINETICS
Noise Control

The

AUSTRALIAN ACOUSTICAL SOCIETY

invites you to attend

INTERNOISE 2014

in MELBOURNE

16 – 19 November, 2014

www.internoise2014.org

ICSV20
PRESS RELEASE

The 20th International Congress on Sound and Vibration (ICSV20) will be held 7-13 July 2013 in Bangkok, Thailand. ICSV20 is sponsored by the International Institute of Acoustics and Vibration (IIAV) and Faculty of Science; Chulalongkorn University, the Acoustical Society of Thailand and the Science Society of Thailand; the ICSV20 is organized in cooperation with: the International Union of Theoretical and Applied Mechanics; the American Society of Mechanical Engineers International and the Institution of Mechanical Engineers. The ICSV20 Congress will be held at Imperial Queens Hotel, Bangkok, Thailand.

Theoretical and experimental research papers in the fields of acoustics, noise, and vibration are invited for presentation. Participants are welcome to submit abstracts and companies are invited to take part in the ICSV20 exhibition and sponsorship. For more information, please visit http://www.iiav.org.

For immediate release, May 17, 2012 2012. For further details, please contact Margarita Maksotskaya at mam0066@auburn.edu or visit www.icsv20.org.


For Digital Recorders
Introducing
For USB A/D Systems
PHANTOM POWER
7052PH
Measurement Mic System
7052H Type 1.5TM
Titanium Diaphragm
3Hz to >20 kHz
<20 dBA > 140 dBSPL
MK224 (<14 dBA to >134 dBSPL) Optional
4048 Preamp
Superior
IEC 1094 Type 1
Long-term Stability
Temperature and Humidity
Performance
Now in Stock
Phantom
to IEPE/ICP
Adaptor
Supplies 3-4 mA
Power
Accelerometers
Microphones
ICP1248
ACOustics Begins with ACO
TM
MATTtm
Family
MATT TM
Mic
Attenuator
Handle Higher Sound Pressure Levels
ACO Pacific, Inc., 2604 Read Ave., Belmont, CA 94002
Tel:(650) 595-8588 FAX: (650) 591-2891 E-Mail: sales@acopacific.com
Web Site: www.acopacific.com

WHAT'S NEW in Canada ??

Promotions
Deaths
New jobs
Moves
Retirements
Degrees awarded
Distinctions
Other news

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

QUOI DE NEUF en Canada??

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

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

Jeremie Voix - Email: voix@caa-aca.ca


Concerto
4-Channel Multi-Function Acoustic Measuring System
All you need in one system:
- 4-channel SLM Class 1
- RT60, EDT, C80, D50 & Ts
- 4-channel Data Logger
- 4-channel Spectrum Analyzer
- Building and Human Vibration
- Monitoring Station with Remote Access
Custom Modules Available
See demo : www.softdb.com/concerto.php
Remote access via iPhone or PC


I-Track
Automatic Real-Time Sound Mapping
Automotive
Industrial
Laboratory
Architectural
5-Minute Mapping
Freehand Scanning Without Grid
See demo : www.softdb.com/itrack.php

INSTRUCTIONS TO AUTHORS FOR THE PREPARATION OF MANUSCRIPTS

Submissions: The original manuscript and two copies should be sent to the Editor-in-Chief. The manuscript can also be submitted electronically.

General Presentation: Papers should be submitted in camera-ready format. Paper size 8.5" x 11". If you have access to a word processor, copy as closely as possible the format of the articles in Canadian Acoustics 39(1) 2011. All text in Times-Roman 10 pt font, with single (12 pt) spacing. Main body of text in two columns separated by 0.25". One line space between paragraphs.

Margins: Top - 0.75"; bottom - 0.75" minimum; sides - 0.75".

Title: Bold, Times New Roman 14 pt with 14 pt spacing, upper case, centered.

Authors/addresses: Names and full mailing addresses, 10 pt with single (12 pt) spacing, upper and lower case, centered. Names in bold text.

Abstracts: English and French versions. Headings, 12 pt bold, upper case, centered. Indent text 0.5" on both sides.

Headings: Headings to be in 12 pt bold, Times-Roman font. Number at the left margin and indent text 0.5". Main headings, numbered as 1, 2, 3, ... to be in upper case. Sub-headings numbered as 1.1, 1.2, 1.3, ... in upper and lower case. Sub-sub-headings not numbered, in upper and lower case, underlined.

Equations: Minimize. Place in text if short. Numbered.

Figures/Tables: Keep small. Insert in text at top or bottom of page. Name as "Figure 1, 2, ..." Caption in 9 pt with single (12 pt) spacing. Leave 0.5" between text.

Line Widths: Line widths in technical drawings, figures and tables should be a minimum of 0.5 pt.

Photographs: Submit original glossy, black and white photograph.

Scans: Should be between 225 dpi and 300 dpi. Scan: Line art as bitmap tiffs; Black and white as grayscale tiffs and colour as CMYK tiffs;

References: Cite in text and list at end in any consistent format, 9 pt with single (12 pt) spacing.

Page numbers: In light pencil at the bottom of each page. For electronic submissions, do not number pages.

Reprints: Can be ordered at time of acceptance of paper.

DIRECTIVES A L'INTENTION DES AUTEURS PREPARATION DES MANUSCRITS

Soumissions: Le manuscrit original ainsi que deux copies doivent être soumis au rédacteur-en-chef. Le manuscrit peut être aussi acheminé par voie électronique.

Présentation générale: Le manuscrit doit être soumis avec mise en page en format de publication. Dimension des pages, 8.5" x 11". Si vous avez accès à un système de traitement de texte, dans la mesure du possible, suivre le format des articles dans l'Acoustique canadienne 39(1) 2011. Tout le texte doit être en caractères Times-Roman, 10 pt et à simple (12 pt) interligne. Le texte principal doit être en deux colonnes séparées d'un espace de 0.25". Les paragraphes sont séparés d'un espace d'une ligne.

Marges: Haut - 0.75"; bas - minimum 0.75" ; côtés,- 0.75".

Titre du manuscrit: Caractères gras, Times New Roman 14 pt,avec espace interligne de 14 pt, lettres majuscules, texte centré.

Auteurs/adresses: Noms et adresses postales. Lettres majuscules et minuscules, 10 pt à simple (12 pt) interligne, texte centré. Les noms doivent être en caractères gras.

Sommaire: En versions anglaise et française. Titre en 12 pt, lettres majuscules, caractères gras, texte centré. Paragraphe 0.5" en alinéa de la marge, des 2 cotés.

Titres des sections: **T**ous en caractères gras, 12 pt, Times-Roman. Premiers titres: numéroter 1, 2, 3, ..., en lettres majuscules; sous-titres: numéroter 1.1, 1.2, 1.3, ..., en lettres majuscules et minuscules; sous-sous-titres: ne pas numéroter, en lettres majuscules et minuscules et soulignés.

Équations: Minimiser le nombre et les numéroter. Insérer directement dans le texte les équations très courtes.

Figures/Tableaux: De petites tailles. Les insérer dans le texte au haut ou au bas de la page. Les nommer "Figure 1, 2, 3,..." Légende en 9 pt à simple (12 pt) interligne. Laisser un espace de 0.5" entre le texte.

Largeur des traits: La largeur des traits sur les schémas techniques doivent être au minimum de 0.5 pt pour permettre une bonne reproduction.

Photographies: Soumettre la photographie originale sur papier glacé, noir et blanc.

Figures numérisées: Doivent être au minimum de 225 dpi et au maximum de 300 dpi. Les schémas doivent être en format bitmap tif. Les photos noir et blanc doivent en format tif sur une échelle de tons de gris et toutes les photos couleurs doivent être en format CMYK tif.

Références: Les citer dans le texte et en faire la liste à la fin du document, en format uniforme, 9 pt à simple (12 pt) interligne.

Pagination: Au crayon pâle, au bas de chaque page. Ne pas paginer si le manuscrit est envoyé par voie électronique.

Tirés-à-part: Ils peuvent être commandés au moment de l'acceptation du manuscrit.

The Canadian Acoustical Association
l'Association Canadienne d'Acoustique

Application for Membership

CAA membership is open to all individuals who have an interest in acoustics. Annual dues total $80.00 for individual members and $35.00 for Student members. This includes a subscription to ***Canadian Acoustics***, the Association's journal, which is published 4 times/year. New membership applications received before August 31 will be applied to the current year and include that year's back issues of ***Canadian Acoustics***, if available. New membership applications received after August 31 will be applied to the next year.

Subscriptions to *Canadian Acoustics* *or* Sustaining Subscriptions

Subscriptions to ***Canadian Acoustics*** are available to companies and institutions at the institutional subscription price of $80.00. Many companies and institutions prefer to be a Sustaining Subscriber, paying $350.00 per year, in order to assist CAA financially (no voting privileges at AGM). A list of Sustaining Subscribers is published in each issue of ***Canadian Acoustics***. Subscriptions for the current calendar year are due by January 31. New subscriptions received before August 31 will be applied to the current year and include that year's back issues of ***Canadian Acoustics***, if available.

Please note that electronic forms can be downloaded from the CAA Website at **caa-aca.ca**

Address for subscription / membership correspondence:

Name / Organization ______________________________

Address ______________________________

City/Province ____________ Postal Code ________ Country ____________

Phone ____________ Fax ____________ E-mail ____________

Address for mailing Canadian Acoustics, if different from above:

Name / Organization ______________________________

Address ______________________________

City/Province ____________ Postal Code ________ Country ____________

Areas of Interest: (Please mark 3 maximum)

1. Architectural Acoustics
2. Engineering Acoustics / Noise Control
3. Physical Acoustics / Ultrasound
4. Musical Acoustics / Electro-acoustics
5. Psychological / Physiological Acoustic
6. Shock and Vibration
7. Hearing Sciences
8. Speech Sciences
9. Underwater Acoustics
10. Signal Processing / Numerical Methods
11. Other

For student membership, please also provide:

________ ________ ________ ________

(University) *(Faculty Member)* *(Signature of Faculty Member)* *(Date)*

I have enclosed the indicated payment for:
[ ] CAA Membership $ 80.00
[ ] CAA Student Membership $ 35.00

Corporate Subscriptions:
[ ] $80 including mailing in Canada
[ ] $88 including mailing to USA,
[ ] $95 including International mailing

[ ] Sustaining Subscriber $350.00
includes subscription (4 issues /year)
to ***Canadian Acoustics.***

Payment by: [ ] Cheque
[ ] Money Order
[ ] Credit Card *(VISA only)*

Credit card number ______________________________

Name on card ______________________________

Expiry date ______________________________

____

(Signature) (Date)

Mail application and attached payment to:

Executive Secretary, Canadian Acoustical Association, PO Box 74068, Ottawa, Ontario, K1M 2H9, Canada

l'Association Canadienne d'Acoustique
The Canadian Acoustical Association

Formulaire d'adhésion

L'adhésion à l'ACA est ouverte à tous ceux qui s'intéressent à l'acoustique. La cotisation annuelle est de 80.00$ pour les membres individuels, et de 35.00$ pour les étudiants. Tous les membres reçoivent ***l'Acoustique Canadienne***, la revue de l'association. Les nouveaux abonnements reçus avant le 31 août s'appliquent à l'année courante et incluent les anciens numéros (non-épuisés) de ***l'Acoustique Canadienne*** de cette année. Les nouveaux abonnements reçus après le 31 août s'appliquent à l'année suivante.

Abonnement à la revue *Acoustique Canadienne* et abonnement de soutien

Les abonnements pour la revue ***Acoustique Canadienne*** sont disponibles pour les compagnies et autres établissements au coût annuel de 80.00$. Des compagnies et établissements préfèrent souvent la cotisation de membre bienfaiteur, de 350.00$ par année, pour assister financièrement l'ACA (sans droit de vote à l'AGM). La liste des membres bienfaiteurs est publiée dans chaque numéro de la revue ***Acoustique Canadienne***. Les nouveaux abonnements reçus avant le 31 août s'appliquent à l'année courante et incluent les anciens numéros (non-épuisés) de ***l'Acoustique Canadienne*** de cette année. Les nouveaux abonnements reçus après le 31 août s'appliquent à l'année suivante.

Pour obtenir des formulaires électroniques, visitez le site Web: **caa-aca.ca**

Pour correspondance administrative et financière:

Nom / Organisation ____

Adresse ____

Ville/Province ____________________ Code postal __________ Pays ____________________

Téléphone ______________ Téléc. ____________ Courriel ____________________

Adresse postale pour la revue Acoustique Canadienne

Nom / Organisation ____

Adresse ____

Ville/Province ____________________ Code postal __________ Pays ____________________

Cocher vos champs d'intérêt: (maximum 3)

1. Acoustique architecturale
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 80.00 $
[ ] Membre étudiant(e) 35.00 $

Abonnement institutionnel
[ ] 80 $ à l'intérieur du Canada
[ ] 88 $ vers les États-Unis
[ ] 95 $ tout autre envoi international

[ ] Abonnement de soutien 350.00 $
(comprend l'abonnement à *L'acoustique Canadienne*)

Méthode de paiement:
[ ] Chèque au nom de l'Association Canadienne d'Acoustique
[ ] Mandat postal
[ ] VISA seulement

Numéro carte de crédit ______________________

Nom sur la carte ______________________

Date d'expiration ______________________

____
(Signature) (Date)

Prière d'attacher votre paiement au formulaire d'adhésion. Envoyer à :
Secrétaire exécutif, Association Canadienne d'Acoustique, CP 74068, Ottawa, K1M 2H9, Canada

The Canadian Acoustical Association
l'Association Canadienne d'Acoustique

SUSTAINING SUBSCRIBERS / ABONNES DE SOUTIEN

The Canadian Acoustical Association gratefully acknowledges the financial assistance of the Sustaining Subscribers listed below. Their annual donations (of $350.00 or more) enable the journal to be distributed to all at a reasonable cost.

L'Association Canadienne d'Acoustique tient à témoigner sa reconnaissance à l'égard de ses Abonnés de Soutien en publiant ci-dessous leur nom et leur adresse. En amortissant les coûts de publication et de distribution, les dons annuels (de $350.00 et plus) rendent le journal accessible à tous nos membres.

ACI Acoustical Consultants Inc.
Mr. Steven Bilawchuk - (780) 414-6373
stevenb@aciacoustical.com - Edmonton, AB

ACO Pacific Inc.
Mr. Noland Lewis - (650) 595-8588
acopac@acopacific.com - Belmont, CA

Acoustec Inc.
Dr. J.G. Migneron - (418) 834-1414
courrier@acoustec.qc.ca - St-Nicolas, QC

ACOUSTIKALAB Inc.
Jean Laporte - (514) 692-1147
jlaporte@acoustikalab.com - Montréal, QC

AECOM
Frank Babic - (905) 712-7054
frank.babic@aecom.com - Mississauga, ON

Aercoustics Engineering Ltd.
Mr. John O'Keefe - (416) 249-3361
johno@aercoustics.com - Toronto, ON

ARMTEC
Ron Galloway - (905) 521-0999
ron.galloway@armtec.com - Hamilton, ON

Bruel & Kjaer North America Inc.
Mr. Andrew Khoury - (514) 695-8225
andrew.khoury@bksv.com - Pointe-Claire, QC

Conestoga-Rovers & Associates
Tim Wiens - (519) 884-0510 x2352
twiens@craworld.com - Waterloo, ON

Dalimar Instruments Inc.
Mr. Daniel Larose - (514) 424-0033
daniel@dalimar.ca - Vaudreuil-Dorion, QC

Dessau Inc.
Jacques Boilard - (418) 839-6034
jacques.boilard@dessau.com - Québec, QC

DuraSystems Barriers Inc.
Fred Woo - (905) 660-4455
fred.woo@durasystems.com - Vaughan, ON

Eckel Industries of Canada Ltd.
- (613) 543-2967
eckel@eckel.ca - Morrisburg, ON

G.R.A.S. Sound & Vibration
- (330) 425-1201
sales@gras.us - Twinsburg, OH

H.L. Blachford Ltd.
Duncan Spence - (905) 823-3200
amsales@blachford.ca - Mississauga, ON

Hatch Associates Ltd.
Mr. Tim Kelsall - (905) 403-3932
tkelsall@hatch.ca - Mississauga, ON

HGC Engineering Ltd.
Mr. Bill Gastmeier - (905) 826-4044
bgastmeier@hgcengineering.com;janstey@hgcengineering.com - Mississauga, ON

Hydro-Québec
- (514) 879-4100 x5309
gosselin.blaise@hydro.qc.ca - Montréal, QC

Integral DX Engineering Ltd.
Mr. Greg Clunis - (613) 761-1565
greg@integraldxengineering.ca - Ottawa, ON

J.E. Coulter Associates Ltd.
Mr. John Coulter - (416) 502-8598
jcoulter@on.aibn.com - Toronto, ON

Jacobs & Thompson Inc.
Chris Brand - (416) 749-0600
cmaida@jacobs-thompson.com - Toronto, ON

Jade Acoustics Inc.
Ms. Dalila Giusti - (905) 660-2444
dalila@jadeacoustics.com - Concord, ON

JASCO Research Ltd.
Mr. Scott Carr - (902) 405-3336
scott@jasco.com - Dartmouth, NS

Kinetics Noise Control Inc.
Mr Mehrzad Salkhordeh - 905-670-4922
msalkhordeh@kineticsnoise.com - Mississauga, Ontario

Mc SQUARED System Design Group
Mr. Wade McGregor - (604) 986-8181
info@mcsquared.com - North Vancouver, BC

MJM Conseillers en Acoustique Inc.
M. Michel Morin - (514) 737-9811
mmorin@mjm.qc.ca - Montréal, QC

Novel Dynamics Inc.
Stan Thompson - (613) 598-0026
stan@noveldynamics.com - Ottawa, ON

OZA Inspections Ltd.
Mr. David Williams - (800) 664-8263 x25
oza@ozagroup.com - Grimsby, ON

Peutz & Associés
M. Marc Asselineau - +33 1 45230500
m.asselineau@peutz.fr - Paris,

Pliteq Inc.
Wil Byrick - (416) 449-0049
wbyrick@pliteq.com - Toronto, ON

Pyrok Inc.
Mr. Howard Podolsky - (914) 777-7770
info@pyrok.com - Mamaroneck, NY

RWDI AIR Inc.
Peter VanDelden - (519) 823-1311
peter.vandelden@rwdi.com - Guelph, ON

Scantek Inc.
Mr. Richard J. Peppin - (410) 290-7726
peppinr@scantekinc.com - Columbia, MD

SNC-Lavalin inc., division Environnement
M. Jean-Luc Allard - (514) 393-1000
jeanluc.allard@snclavalin.com - Longueuil, QC

Soft dB Inc.
M. André L'Espérance - (418) 686-0993
contact@softdb.com - Sillery, QC

Sound & Vibration Solutions Canada, Inc.
Mr. Andy Metelka - (519) 853-4495
ametelka@cogeco.ca - Acton, ON

Soundtrap Inc.
Roger Foulds - (705) 357-1067
roger@soundtrap.ca - Sunderland, ON

Stantec Consulting Ltd.
Mrs. Zohreh Razavi - (604) 696-8472
- Vancouver, BC

State of the Art Acoustik Inc.
Dr. C. Fortier - (613) 745-2003
cfortier@sota.ca - Ottawa, ON

Tacet Engineering Ltd.
Dr. M.P. Sacks - (416) 782-0298
mal.sacks@tacet.ca - Toronto, ON

True Grit Consulting Ltd.
Ina Chomyshyn - (807) 626-5640
ina@tgcl.ca - Thunder Bay, ON

Valcoustics Canada Ltd.
Dr. Al Lightstone - (905) 764-5223
solutions@valcoustics.com - Richmond Hill, ON

Vendatech Inc.
Behrou Ghazizadeh - (416) 787-8797
behrou@vendatech.com - Toronto, ON

Vibro-Acoustics
Mr. Tim Charlton - (800) 565-8401
tcharlton@vibro-acoustics.com - Scarborough, ON

Wakefield Acoustics Ltd.
Mr. Clair Wakefield - (250) 370-9302
clair@wakefieldacoustics.com - Victoria, BC

West Caldwell Calibration Labs
Mr. Stanley Christopher - (905) 595-1107
info@wccl.com - Brampton, ON

Wilrep Ltd.
Mr. Don Wilkinson - (905) 625-8944
info@wilrep.com - Mississauga, ON

Xscala Sound & Vibration
Jim Ulicki - (403) 274-7577
caa@xscala.com - Calgary, AB