ISSN 0711-6659

canadian acoustics
acoustique canadienne

SEPTEMBER 1998 | SEPTEMBRE 1998
Volume 26 -- Number 3 | Volume 26 -- Numéro 3

Jérémie Voix

Blachford

"The ABC's of noise control"

H.L. Blachford's Comprehensive Material Choices

Noise treatments can be categorized into three basic elements: Vibration Damping, Sound Absorption and Sound Barriers.

Vibration Damping

It is well known that noise is emitted from vibrating structures or substrates. The amount of noise can be drastically reduced by the application of a layer of a vibration damping compound to the surface. The damping compound causes the vibrational energy to be converted into heat energy. Blachford's superior damping material is called ANTIVIBE and is available either in a liquid or a sheet form.

ANTIVIBE DL is a liquid damping material that can be applied with conventional spray equipment or troweled for smaller/thicker application.

It is water-based, non-toxic and provides economical and highly effective noise reduction from vibration.

ANTIVIBE DS is an effective form of damping material provided in sheet form for direct application to your product.

Sound Barriers

Sound Barriers are uniquely designed for insulating and blocking airborne noise. The reduction in the transmission of sound (transmission loss or "TL") is accomplished by the use of a material possessing such characteristics as high mass, limpness, and impermeability to air flow. Sound barriers can be a very effective and economical method of noise reduction.

Blachford Sound Barrier materials:

BARYMAT

Limp, high specific gravity, plastic sheets or die cut parts. Can be layered with other materials such as acoustical foam, protective and decorative facings to achieve the desired TL for individual applications.

Sound Absorption

Blachford's CONASORB materials provide a maximum reduction of airborne noise through absorption in the frequency ranges associated with most products that produce objectionable noise. Examples: Engine compartments, computer and printer casings, construction equipment, cabs,...etc.

Available with a wide variety of surface treatments for protection or esthetics. Material is available in sheets, rolls and die-cut parts – designed to meet your specific application.

Suggest Specific Material or Design

Working with data supplied by you, H.L. Blachford will make recommendations or treatment methods which may include specific material proposals, design ideas, or modifications to components.

A Quality Supplier

The complete integration of:

- Experience
- Quality-oriented manufacturing technology
- Research and development
- Problem solving approach to noise control

Our Mississauga Plant is ISO-9001 CERTIFIED

Result in:

Comprehensive Noise Control Solutions

MISSISSAUGA (905) 823-3200

MONTREAL (514) 938-9775

VANCOUVER (604) 263-1561

APPLICATION OF MULTICHANNEL ACTIVE NOISE CONTROL IN A LARGE CIRCULAR INDUSTRIAL STACK

André L'Espérance (1), Martin Bouchard (2), Bruno Paillard (1), Alex Boudreau (1)

(1) Faculty of Applied Sciences, Université de Sherbrooke, Sherbrooke (Que.) Canada, J1K 2R1
andre.lesperance@gme.usherb.ca

(2) School of Information, Technology and Engineering, University of Ottawa, Ottawa (Ont.) Canada, K1N 6N5

1. INTRODUCTION

This paper exposes the solutions found to realize one of the first multichannel of active noise control (ANC) system on a large industrial chimney stack. This application has been realized at the Lauralco Aluminum plant (Deschambault, Québec) on a 40 m high and 1.8 m wide chimney stack (fig.1) that generated a 320 Hz pure tone.

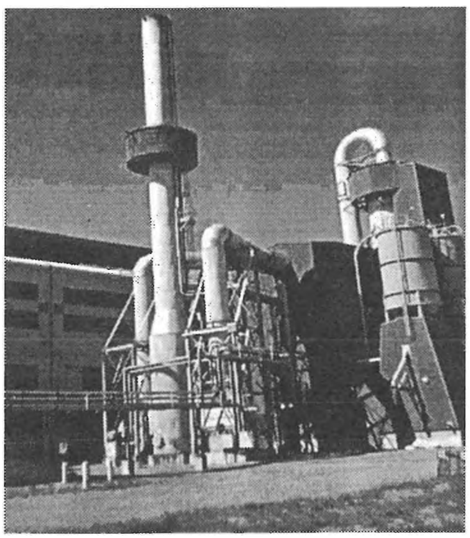


Figure 1 Chimney stack of the aluminum plant.

The first part of this paper reviews the error sensors plane concept recently proposed[1] to obtain an efficient control in the case of a complex sound propagation in a circular duct. The second part describes the software and hardware used for the industrial installation, and the special features used to ensure the effectiveness, robustness and easy maintenance of the system.

2. ANC of high mode using the error sensor plane concept

According to common knowledge in ANC, *M* error sensors have to be used to control *M* modes, and those sensors should not be located at the nodal lines[2]. However, in circular ducts, the location of the nodal lines changes along the duct axis[3], which may explain why ANC of high order modes in circular or irregular ducts appears to be difficult[4].

To overcome this problem, an alternative strategy of locations of the error microphones has recently been proposed, the *error sensor plane* concept[1]. The objective of this error sensor plane is to create a *quiet cross-section* in the duct so that, based on the Huygen's principle, the noise from the primary source cannot propagate over this cross-section.

According to this method, 10 error sensors were necessary to reduce the 320 Hz pure, one error sensor being located at the center of the duct axis, and 9 at 0.21 meters from the duct wall, at 40° intervals.

3. Operation of the system

The controller developed is based on a VME-RACK mount with a PC and a C-40 DSP board. The algorithm used is a multi-channel filtered-X LMS. The hardware of the control system (controller, amplifier, error sensors and control sources) have been installed in an existing platform located at a height of 20 m on the stack (figure 2). When the power of the system is turned on, the controller system automatically proceeds the following steps :

1. Load the control code
2. Make a calibration/compensation of the acoustic sensors
3. Identify the 100 transfer function (10 x10 systems)
4. Start the active noise control
5. Continuously verify the efficiency of the system

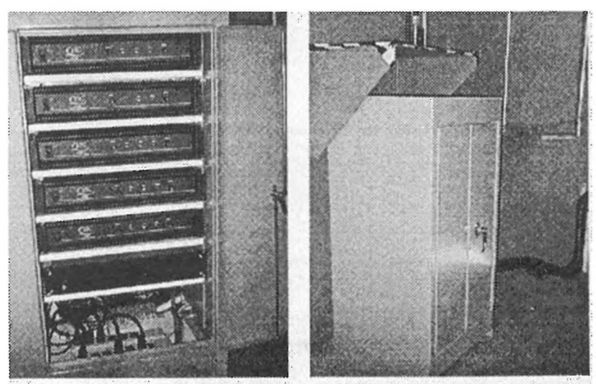


Figure 2 Active noise control system at the sampling platform of the chimney.

3.1 In situ calibration/compensation process

Because it was unthinkable to request an operator to do the calibration of the 10 error sensors regularly to ensure the efficiency of the system, an automatic calibration/ compensation process was developed to calibrate the probes at their operating location.

This procedure consists in generating a number of pure tones with the control speaker and in measuring the value of those pure tones at each error microphone. This process is repeated with each speaker, and the mean value of some of the pure tones at each error sensor are then compared to compensate the response of the error sensors.

3.2 Continuous verification of the system's efficiency

To verify the performance of the system, the mean sound pressure level of the 320 Hz at all error sensors with control is continuously compared to the level without control on the front panel of the controller (fig.3). The reference level with control is evaluated from measurement made at different occasions. For example, the control may be stopped for 30 seconds every day and the reference level estimated from the mean value of the last 10 measurements.

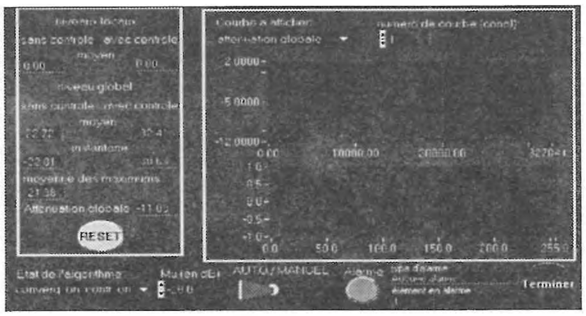


Figure 3 Main front panel of the controller

3.3 Automatic failure detection system

An automatic *failure detection system* has been included in the control process to ensure its effectiveness and trustworthiness. For example, if the noise at the error sensor plane increases compared to the reference level, then the control is immediately stopped. The adaptation filters are then re-initialized and the control process is restarted. If successive divergences are detected, the control process is restarted at step 2, i.e. a new calibration/compensation of the error sensors is done, as a new identification of the transfer function.

The divergences may be due, for example, to a malfunction of an error sensor or speaker. Then, during the calibration/compensation process, different tests are done to verify the functionality of each of those elements. If there is a malfunction of one of them, then an error sight is shown on the front panel of the controller display and the element in fault is identified.

Note finally that the daily performance of the system as well as the status of the controller (calibration, control on/off, divergence, misfunction an specific element) are recorded as function of the time. Using this information, one can verify the performance and trustworthiness of the active noise control system. In the actual application, this information is automatically transferred to the plant exploitation system, so that the personnel can easily follow-up the operation of the active noise control.

3.4 Noise reduction

Figure 4 gives the mean spectra measured at 200 m of the stack, with and without the active noise control system on. The contribution of the 320 Hz pure is reduced to the background noise level, which corresponds to a 10 dB noise reduction. The noise reduction objective has thus been achieved.

The actual active noise control system is in continuous operation since April 1997, and no failure has occurred until now.

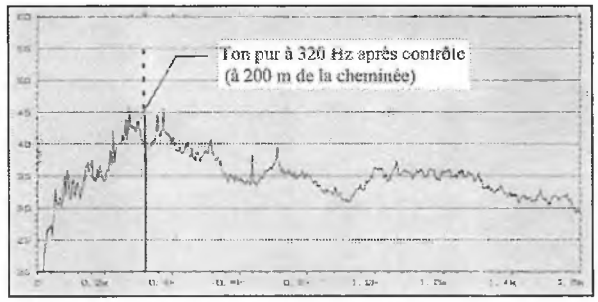


Figure 4 Typical frequency response of the speakers with the teflon membrane

4. Conclusion

An active noise control system has been developed in order to reduce the propagation of the 320 Hz emitted by an industrial stack. A 10 MIMI control algorithm using the error sensor plane concept was developed and installed in the stack. The system is fully automatic so that, when the power is turned on, the control algorithm does all the steps to achieve a control, and a verification process is continuously in operation to ensure the efficiency of the system. This system has been installed on the industrial stack, and the objective of a 10 dB noise reduction of the 320 Hz pure tone in the neighborhood of the industrial plant has been reached. The actual active noise control system is in continuous operation since April 1997 and no failure has occurred so far.

ACKNOWLEDGMENTS

The authors are grateful to Aluminerie Lauralco Inc. for its financial support and to its personnel, namely Jean-Claude Dubé and Lise Sylvain, for their assistance and enthusiasm during this project.

REFERENCES

[1] A. L'Espérance et al. "*Method and apparitus for active noise control of high order modes in ducts*". United States Patent no 5,748,750, May 1998.

[2] Baumann D. C. et al., "*Number of error microphones for multi-modal cancellation*", Proc. Inter-noise92, Toronto, Canada p. 345-351(1992).

[3] C.L. Morfey, "*Rotating pressure patterns in ducts: their generation and transmission*", J. Sound & Vib.***1***, p.60-87 (1964).

[4] S. Laugesen, "*Active control of multi-modal propagation of tonal noise in ducts*", J. Sound & Vib, **195**(1), 33-56 (1996).

Active Noise Control on a High Pressure Fan

Alex Boudreau, André L'Espérance, Bruno Paillard and Stephane Boucher
Faculty of Applied Sciences, Université de Sherbrooke, Sherbrooke (Que.) Canada, J1K 2R1
alex.boudreau@gme.usherb.ca

1. INTRODUCTION

High pressure fans offer very interesting aerodynamic specifications for several domestic applications (vacuum cleaners and fans). However, those fans generate a great pure tone caused by the rotation of the blades. The adding of passive silencers in the ducts to reduce the noise of the blade rotation unfortunately brings non negligible pressure drops, which affects the performance of the systems.

Active noise control is an interesting noise reduction alternative when the pressure drops are an important factor in the conception (see ref. 1 and 2). However, active noise control, which is still a rather recent technology, may seem too expensive. The objective of our work was to develop a compact and less expensive active noise control system for high pressure fans for domestic use.

2. DESCRIPTION OF THE FAN

The fan used in this study has 11 blades, and its rotation frequency is about 56 Hz. It generates a pure tone at the following frequency :

$$f = nbr.pale \bullet f_{rot}$$

where f_{rot} is the rotation frequency of the wheel

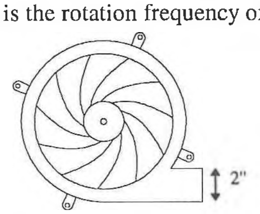


Figure 1 Scheme of the high pressure fan

Since the rotation speed varies little in function of the load, a 615 Hz pure tone comes out from 20 to 30 dB of the noise spectrum (see Figure 4). And, this type of pure tone is very disagreeable in terms of acoustical discomfort. With a 2 inches diameter duct, the cut off frequency of the first mode is about 4000 Hz. Thus, in this case, a single channel control system is sufficient.

3. THE DEVELOPED SYSTEM

Firstly, several control systems have been evaluated (sensor, actuator and algorithm), disregarding the cost constraint in order to define the problematic of the studied case. The control algorithms which were used have been implemented on a DSP C31 put on a PC. For the feed-back or the feed-forward type control algorithms, there must be a measurement of the acoustical pressure in the duct. And, since the flow speeds are high in the duct (100 km/h), it is difficult to take a noise measurement directly in the flow. Expensive and cumbersome protective barriers have to be used. Other geometrical constraints also have to be considered. The control system must be compact : it must be applied the closest possible to the noise source, thus near the fan outlet.

From those preliminary tests, and considering several geometrical and cost constraints, the following system has been developed (see Figure 2):

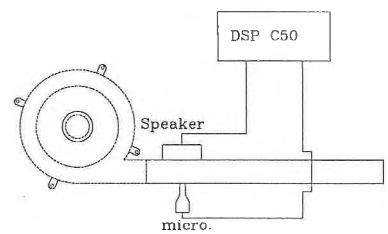


Figure 2 Scheme of the developed control system

The algorithm :

The type of algorithm was chosen following the geometrical constraints and the type of noise to be controlled. The feed-back type is more compact than the feed-forward type, since it only uses one microphone. The feed-forward type uses a second reference microphone upstream. The causality problem between the error microphone and the noise to be controlled leads to some geometrical constraints. Moreover, the feed-back type gives a great performance for the harmonic noise control. Thus, a feed-back type algorithm with perturbation reconstruction has been implemented on an independent DSP C50. The algorithm has been modified in order to meet the needs of the application (see Figure 3).

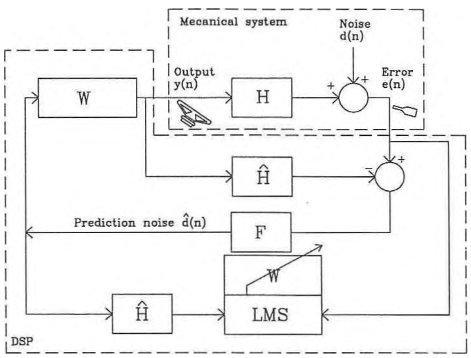


Figure 3 Scheme of the control algorithm

A pass-band filter F around the tone has been implemented in order to minimize the number of coefficients for the control filter. In this way, less computation power is needed for an equivalent performance. Moreover, the F filter ensures a great controller performance on the tone, even if the spectral content around the tone changes. Indeed, the resonance on either side of the tone (see Figure 4) due to the longitudinal interference patterns inside the duct may vary depending on the geometrical conditions. Furthermore, the system had to be completely independent; thus, an automatic procedure has been developed to detect the divergences of the system as well as possible mechanical component failures.

Actuator :

The actuator used in this study is a low cost typical full range speaker, just like the ones that come with portable radios. Several tests showed that the linearity of that type of speaker was sufficient to ensure a good control in the ranges of frequency and sound pressure levels needed for this application. Indeed, the required sound power to generate the opposite wave is little, since the interior of the duct is very enclosed, and it is easy to generate high levels of noise. For example, with 0.2 watt, the control is appropriate on the tone at 615 Hz. The speaker has been installed very close to the duct, brushing against it, so as not to disrupt the flow.

Error sensor :

The error measurement is done with a low cost condenser microphone also fixed brushing against the wall but protected with a thin film. This technique allows to avoid static pressure problems related to high flow speeds. It also showed great performances compared to the expensive and cumbersome protective barriers.

4. RESULTS

The typical performance of the developed system is showed on Figure 4. The performance is of about 25 dB to 35 dB on the tone in function of the emergence of the tone. Note that the shape of the spectrum to be attenuated may change depending on the configuration and on the length of the duct. For all the configurations evaluated in laboratory, the system gives the same very good performance : a complete elimination of the tone.

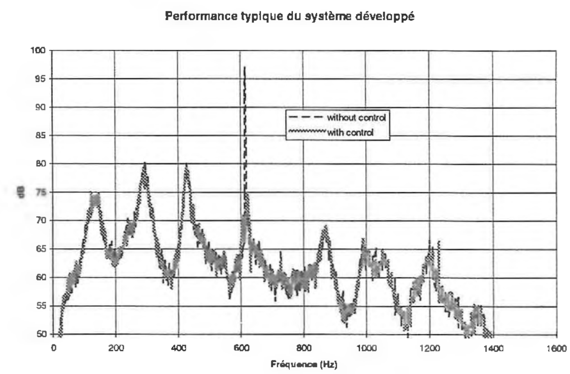


Figure 4 Typical performance of the control system

The computation power required to carry out a control is about 1/10 of the one taken by the DSP C50. This allows to consider the use of a lower cost DSP (a C20 for example). A preliminary cost analysis shows that a card containing a preamplifier for the error microphone, a 2 watt amplifier for the speaker, and the DSP would allow to have a complete system for less than 75 dollars.

5. CONCLUSION

The active noise control system developed for this type of application shows good performances. The noise component that causes most nuisance is eliminated without perturbing or restraining the flow. Moreover, the cost restraint has been considered : the complete system costs more or less 75 dollars.

6.0 REFERENCES

[1]BESOMBES M., MICHEAU P. (1995) *Active control device integrated in a centrifugal turbo-machine : a compact solution for the periodic noise control,* ACTIVE 95, Newport Beach, p. 263-274.

[2]CHRISTENSON T.N. (1996) *Active noise cancellation in a duct with highly turbulent airflow,* Noise-con 96, Seattle Washington, p. 369-374.

New concept of active control of transformer noise, part 1 : The active envelope

André L'Espérance, Bruno Paillard, Stéphane Boucher, Philippe Michaud
Faculty of Applied Sciences, Université de Sherbrooke, Sherbrooke (Que.) Canada, J1K 2R1 andre.lesperance@gme.usherb.ca

1. INTRODUCTION

In the past few years, active control systems for transformer noise have been developed, and some systems are currently available. However, noise specialists have to be involved in each application to evaluate the radiation pattern of the transformer. Moreover, the error microphones are located in the field far from the transformer, which requires cable installation. Those systems are thus complex and expensive to use. To overcome those limitations, a new concept called the *active envelope* is investigated.

The first part of the paper presents the principle of the active envelope. The second part presents a theoretical model and an experimental setup realized to evaluate the potential and efficiency of this principle.

2. General concept of the acoustic envelope

The classical objective of an acoustical ANC system is to drive the output of an actuator (ex. a speaker) so that the sound pressure at the error sensor (ex. a microphone) is minimized. Minimizing the pressure at the error microphone creates a '*zone of quiet*' around the error microphone[1,2]. The objective of the acoustic envelope is to create a quiet *skin* over the transformer using an appropriate arrangement of control units.

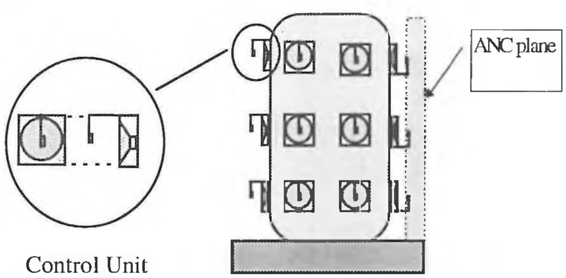


Figure 1 Schema of the active envelope over a transformer.

Also, for its simplicity and cost-effectiveness, it was decided to develop independent control units instead of a real multi-channel (MIMO) controller. However, this may cause instability of the overall system. A theoretical model was developed and an experimental setup was realized to evaluate not only the efficiency of the principle of the active envelope, but also its stability.

3. Theoretical model

A theoretical model has been developed in order to analyze the effect of the distance between the control units (L) and the effect of the distance between the actuator and the sensors (d) on the overall noise reduction.

In this model, the transformer is considered as a vibrating structure of N subsurfaces S which generate a sound field in a semi-infinite space according to the Raleigh integral :

$$P(M) = -\rho_0 w^2 \int_S \frac{W(P)\exp(jkr)}{2\pi r} dP$$

W(P): amplitude of the vibration at point P on the vibration structure
P(M):Sound pressure at a distance r of the plate

This formulation is used to calculate the primary sound field generated in front of the transformer and, among others, at the error sensor of each control unit. The sound field generated by the actuator of each control unit being:

$$P(M) = -y_j \frac{\exp(jkr)}{2\pi r}$$

where y_j is the optimal amplitude and phase needed at the actuator to minimize the sound field due to the transformer. The optimal command for the actuator is calculated with :

$$y_{opt}(\omega_n) = -C^{-1}(\omega_n)\, d(\omega_n)$$

Figure 2 gives a typical example of the radiated sound field in front of the transformer with ANC and noise reduction obtained for various configurations of the active noise control units.

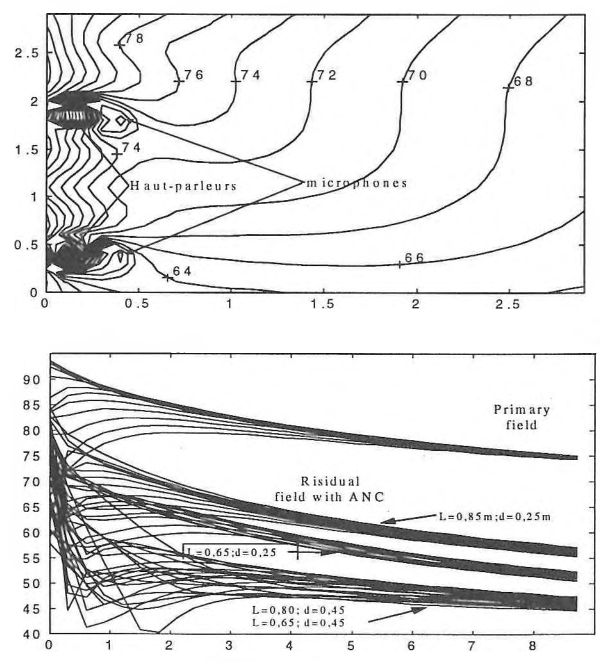


Figure 2 A) Sound field in front of the transformer with ANC; B) Noise reduction with ANC.

From those results, it was found that distances L= 0,70m between each control unit and d=0,30 m between the actuator and the sensors can provide 15 to 20 dB of global attenuation at 120 Hz.

In an application, the secondary sound field generated by a given control unit generates a residual sound field at the error sensor of the adjacent units. Those adjacent units will then adapt their response to minimize the noise at their error sensors and obviously generate a residual sound field to the error microphone of adjacent control units, and so on. This phenomenon can thus generate instability, since each control unit is independent in the proposed system. To analyze these effects, an experimental setup was necessary.

4. Experimental setup

A system of 24 control units was developed to verify the efficiency and stability of the system. Each control unit is composed of an error microphone, a DSP (TMS320C50) an amplifier and a speaker. A feedback control algorithm (see 2nd paper on this subject) is loaded on each DSP using a serial link with a PC[3]. Once the control is turned on, each control unit works to reduce the noise at is own error microphone.

After a set of tests in laboratory, it was found that the stability of the system is dependent, among others, on the distance between the error sensors and the speakers. For a distance of L=0,70 m between the control units, a distance of d=0,20 m allows an efficient noise reduction and a suitable stability.

Using those results, an experiment was then conducted on a real transformer (figure 3) on the first harmonic. A noise reduction of 10 dB was obtained (figure 4). In terms of perception, this reduction appears to be very significant for a person who is on the site.

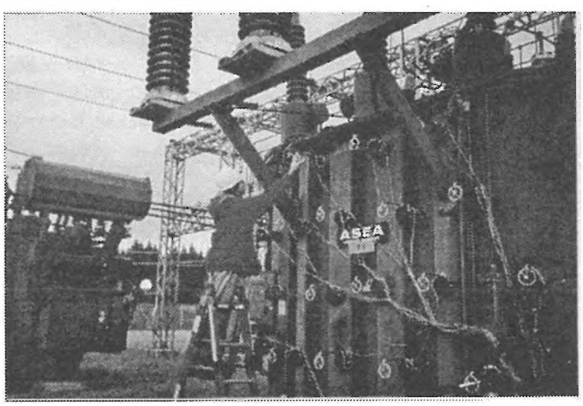


Figure 3 Experimental setup on a real transformer. Distance between the control units is 0,70 m.

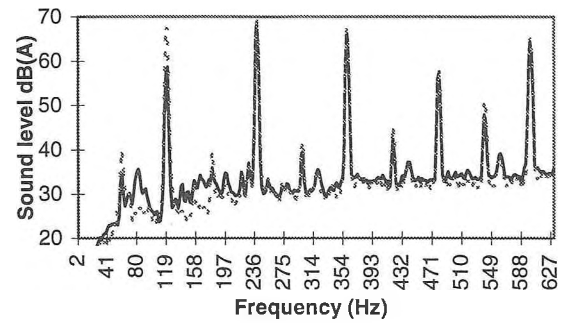


Figure 4 Sound spectra measured with and without control.

5. Conclusion

A new concept of active noise control for transformers has been proposed. This concept uses a set of independent control units distributed on the transformer surface.

A theoretical model has been developed to determine the optimum distribution of the control unit on a transformer surface. An experimental setup composed of 24 control units has also been realized. This system allowed to determine that a distance of about 0,20 m between the error sensor and the speaker of each control unit ensures a good stability of the system.

An experiment realized on a real transformer gave a noise reduction of 10 dB of the first harmonics.

The system is actually under development to obtain a noise reduction of the first 2 harmonics (120 Hz and 240 Hz).

ACKNOWLEDGMENTS

The authors are grateful to Hydro-Quebec and its personnel for their support in this project, namely Blaise Gosselin (Vice-president Environment) and Marc Foata (Institut de Recherche en Électricité du Québec).

REFERENCES

[1] A. David, S.J. Elliott, "*Numerical studies of Actively Generated Quiet Zones*", Applied Acoustics, **41**, 63-79, (1994).

[2] M. Miyoshi et al., "*On arrangement of noise-controlled points for producing larger quiet zones with multi-point active noise control*", Proceedings of Inter-noise 94, 1299-1304, (1994).

[3] Stéphane Boucher, Bruno Paillard, André L'Espérance, Alex Boudreau. "*A new concept for active control of transformer noise, part II : control algorithm*", Accepted for Canadian Acoustic Association meeting, London, Ontario, October 1998.

A new concept of active control of transformer noise, Part 2 : Control Algorithm

Stéphane Boucher, André L'Espérance, Bruno Paillard, Alex Boudreau
Faculty of Applied Sciences, Université de Sherbrooke, Sherbrooke (Que.) Canada, J1K 2R1

The new concept of active envelope for the control of transformer noise[1] needed the implementation of identical independent control units. The controller used for those units is briefly described here along with the problems and solutions associated with the placement of independent units side by side.

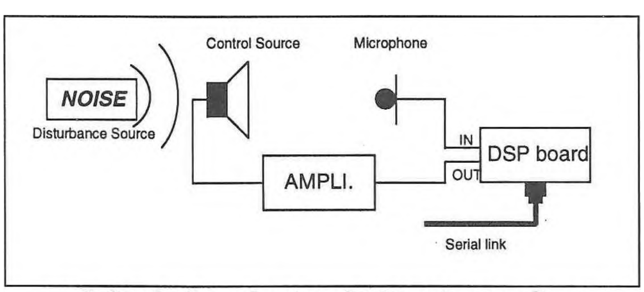
The controller

Figure 1 illustrates an independent control unit. It consists of a mono-channel feedback active adaptive noise control system and it is physically similar to the one presented by Olson and May in 1953 [2]. The choice of a feedback configuration leads to a very simple system (Fig. 1) composed of a few low cost components. The problem to be solved is the generation of the acoustic interference to obtain a reduction of the noise at the microphone and thus, to create a « zone of quiet » around the microphone. In the feedback ANC system shown in Fig. 1, the disturbance is not available because it is intended to be canceled by the microphone signal.

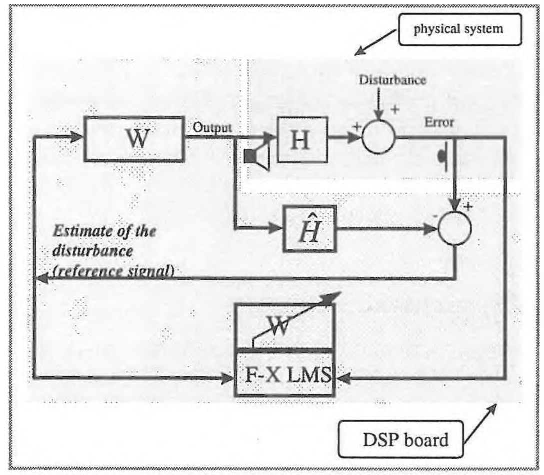
The feedback algorithm that we use (Fig. 2) is said to « use a feedforward approach » [3] because it estimates the disturbance and use it as a reference signal for the control filter. On this figure, « H » represents the transfer function of the plant, which is the relation between the signal sent to the loudspeaker (the output signal y(n)), and the signal measured at the microphone (the error signal e(n)). A model of this transfer function (denoted $\hat{H}$ in figure 2) allow to approximate the contribution of the control signal at the microphone. This approximation is digitally subtracted from the microphone signal, yielding the signal $\hat{d}$ (n), which represents an estimate of the unwanted signal alone (the contribution from the control speaker has been subtracted). This estimate can then be used as the input of the control filter « W ». The output of this filter, the control signal, is sent to the loudspeaker in order to produce the destructive interference.

The adaptive filter « W » is adapted with the filtered-X LMS algorithm [4]. This filter has two purposes :

- It predicts the value of the perturbation signal, some time in the future, from the present and past sample values. The prediction time corresponds roughly to the propagation delay between the loudspeaker and the microphone.



1. ***Single channel active adaptive noise canceller***



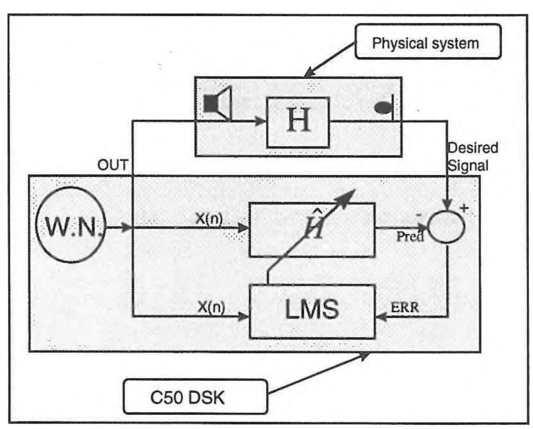
2. ***Feedback controller using a feedforward approach***

- From this predicted value, it generates the control sample to be sent to the control speaker. This generation takes into account the transfer function between the speaker and the microphone.

Actually, both filters are combined into a single control filter W.

Identification of H

The filtered-X LMS and the calculation of the estimate of the disturbance requires both a model of the secondary transfer function « H ». This model can be obtained prior to the control itself with an identification process. This identification process is illustrated in Fig. 3. A finite impulse response filter (FIR) is used to model « H ». This filter is optimized by an LMS algorithm in order to make the signal predicted by the model as equal as possible to the desired signal returned by the microphone. A white noise signal is used as input to the system because it allows an equally fast modeling for the frequency band of interest. This white noise is generated by the DSP, and driven to the speaker for a few seconds. After this phase, the control algorithm of figure 2 is engaged, and the filtered-X LMS algorithm adapts W so as to minimize the residual energy at the error microphone.



3. ***Identification process of the plant***

Problems

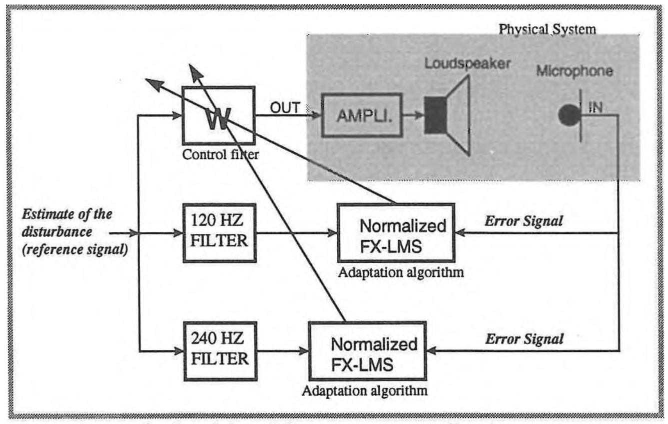
A 40 dB reduction at 120 and 240 Hz can be observed at the microphone when a unit is working on its own (no other units around trying to control). However, the concept of active envelope involves the placement of independent units side by side. The consequence of this tight network of control units is that the control signal of one unit is picked up by the microphone of neighbor units. The error signal of each unit is then influenced by the control signal of neighbors units. Therefore, the estimate of the disturbance is not only function of the disturbance but is also function of the control signal of neighbor units, resulting in a potential instability of the system and in difficulties to control more than one frequency component at a time. Changes have been made to the classical active adaptive control algorithm we used, in order to solve those problems. Spatial normalization as well as frequency normalization of the control algorithm have resulted in a drastic improvement of the behavior of the system.

Spatial normalization

For a non-normalized version of the filtered-X LMS algorithm, the convergence speed of a control unit depends on the energy of the reference signal. A unit converges very fast if the energy of the estimate of the disturbance is high and vice-versa. In our case, many units are placed side by side. Each unit has it's own estimate of the disturbance which is dependent of the noise at it's microphone and will thus converge at it's own speed. The consequence of this lack of uniformity in the convergence speed is a system that can be unstable. We solved this problem with a normalization of the adaptation algorithm (normalized FXLMS)[4]. This « spatial » normalization force the units to converge at the same speed, no matter the level of noise present at each unit.

Frequency normalization

While the spatial normalization allowed to standardize the convergence speed between the control units, it didn't allow the convergence speed between the frequency components to be uniform. Actually, the convergence speed for one frequency component depends on the energy of the reference at this frequency. The control filter is adapted faster for components with more energy in the reference and vice-versa. For example, if the energy of the frequency component at 120 Hz is 10 dB (10 times) superior to the energy of the frequency component at 240 Hz, the control filter will be adapted 10 times more rapidly for the 120 Hz component than for the 240 Hz component. This would not be a big problem if the ratio between the components could stay between 0 and 20 dB. It is usually the case when a unit works on its own since the reference signal represents well the disturbance. However, the placement of many units side by side distorts the estimate of the disturbance in such a way that the ratio between frequency components is not anymore the same as the disturbance itself. Moreover, this distorted ratio can be several dB over/under the real ratio, resulting in a convergence speed that can be very different from one frequency component to the other. Actually, we could observe for some control units a distort ratio of approximately 40 dB which resulted in a convergence speed 100 times faster for the most energetic component. In that case, the less energetic component could take several hours to converge. Sometimes, it will not converge at all, du to the limited calculation precision of the DSP.



4. Spatial and frequency normalization

The solution to this problem is to « frequency normalize » the control algorithm. This strategy, illustrated in Fig. 4, consist in making the convergence speed independent of the frequency. In our case, since we want to control two different frequencies (120 and 240 Hz), we use two filters. Each one filters the reference signal in order to separate the two frequency components. The two output signals then serve as inputs for two normalized FXLMS, running in parallel. Since the FXLMS algorithm is normalized, the control filter will be adapted in order to make the two frequency components converge at the same speed.

Conclusion

Spatial and frequency normalization has been applied to the independent control units in order to improve the stability of the overall system and to facilitate the control of more then one frequency component at a time. As anticipated, a drastic improvement of the behavior of the system resulted.

ACKNOWLEDGMENT

The authors are grateful to Hydro-Quebec for their support in this project, namely Blaise Gosselin (Vice-presidence Env.) and Marc Foata from (Institut de Recherche en Electricité du Québec).

REFERENCES

[1] André L'Espérance, Bruno Paillard, Stephane Boucher, Philippe Michaud, A NEW CONCEPT FOR ACTIVE CONTROL OF TRANSFORMER NOISE PART I: THE ACTIVE ENVELOPE, Proceedings of the Acoustics Week in Canada 1998

[2] H.F. Olson and E. G. May, « Electronic sound absorber, » Journal of the Acoustical Society of America 25, pp. 1130-1136, 1953.

[3] S.J. Elliott, T.J. Sutton, B. Rafaely and M. Johnson, « Design of feedback controllers using a feedforward approach, » Proc. of Active 95, pp.863-874, 1995

[4] B. Widrow and S.D. Stearns, Adaptive Signal Processing, Prentice-Hall, 1985.

Active adaptive noise cancellation with a feedforward open headset

Stéphane Boucher, Bruno Paillard, André L'Espérance
Faculty of Applied Sciences, Université de Sherbrooke, Sherbrooke (Que.) Canada, J1K 2R1

Active noise-reduction headsets currently on the market use feedback-type controllers. This choice of configuration results in the active control of only low-frequency harmonics (usually under 1000 Hz). Moreover, no discrimination is made by feedback algorithms and any predictable acoustic signal, within the control band, is attenuated. Consequently, in addition to cancel the noise, the feedback-type headset will affect speech, alarms and other predictable useful signals, resulting in a discomfort sensation for the user. Also, in order to allow those headsets to attenuate higher frequencies, they are often combined with a passive attenuation shell, with the consequence that the user does not hear any useful signal and has almost the feeling of being « deaf ». A solution to this problem is to use a feedforward controller instead of a feedback controller, thus allowing the headset to :

- Actively cancel high frequency components of a harmonic disturbance,
- Cancel only the disturbance without modifying useful signals,
- Be light and comfortable (the headset can be of « open » type, since no passive attenuation is required).

Description of the controller

The figure 1 illustrates the structure of the controller. Because of the feedforward configuration, the input of the control filter « C » is taken directly at the source of noise. The control filter is adapted in order to attenuate, at the microphone, any signal that is correlated with the disturbance. This way, since useful signals are not correlated with the disturbance, they are not attenuated. The input signal (the reference signal) is filtered by « C » in order to produce an output which will, after having been modified by the transfer function between the speaker and the microphone (secondary-path transfer function « S »), reduce the contribution of the disturbance at the microphone. In order for the system to follow the changes in the transfer function between the primary source and the error microphone (the principal transfer function), « C » has to be continuously adapted. Those changes in the principal transfer function occur when the user moves. The filtered-X LMS [1] is responsible of the adaptation of « C ». This algorithm is frequently used in ANC because it compensates for the secondary-path transfer function effect in the adaptation of the control filter.

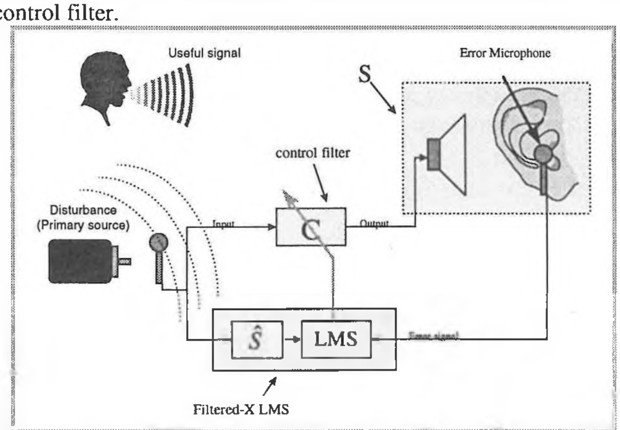


Figure 1 One channel adaptive feedforward controller

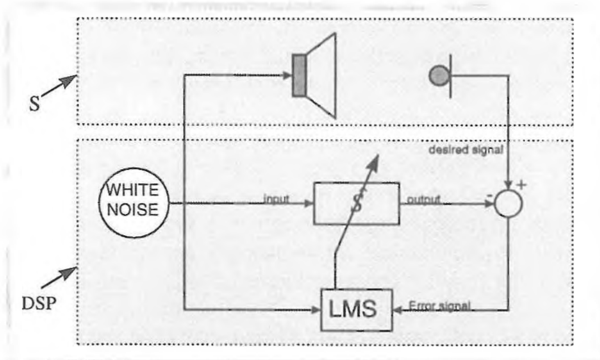


Figure 2 Off-line modeling of secondary-path

Identification of S

The FXLMS algorithm requires knowledge of the secondary-path transfer function « S ». Since this transfer function is different from one user to the other and is dependent of the position of the headset on the head, the modeling of « S » has to take place every time the headset is used and before the control starts. Since « S » is time invariant as long as the user doesn't move the headset from it's initial position on the head, off-line modeling can be used to estimate « S » prior to the control itself (no on-line modeling is necessary).

Figure 2 illustrates the off-line identification process. A finite impulse response filter (FIR) is used to model « S ». This filter is optimized by an LMS algorithm in order to make the output signal of the model as close as possible to the desired signal returned by the microphone. A white noise signal is used as input to the system because it allows an equally fast modeling for the frequency band of interest. In order to obtain a precise enough model of « S » despite the presence of an ambient noise, the duration of the identification has to be automatically adjusted, in regards to the energy of the ambient noise. The louder the ambient noise, the longer the identification phase has to be. In fact, a doubling in the intensity of the ambient noise requires a doubling in the duration of the identification. If the ambient noise is beyond a certain level, it becomes difficult to obtain a good model because of limitations in the computational precision. The range of duration to obtain a good model of « S » goes from 0,1 (low ambient noise) to 1,5 seconds (loud ambient noise).

Implementation

We chose to use two fixed-point TMS320C50 DSPs (one per channel) for the implementation in order to have enough calculation power for the algorithms to run in real time, even at relatively high sampling frequencies. This way, we could expect fast adaptation for the identification and the control process. The choice of fixed-point DSPs was encouraged by the low-cost manufacturing possibility of an eventual commercial product and the low power consumption of this kind of DSP (possibility to power the system from batteries). The down side is that fixed-point DSPs have a more limited computational dynamic and are more difficult to program (normalization) than floating-point DSPs [2].

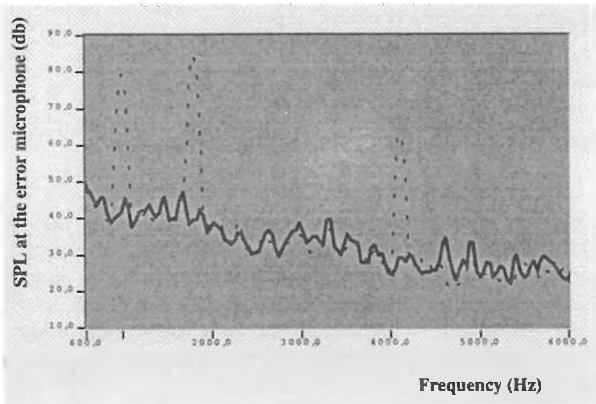


Figure 3 Spectrum of a 3 frequency components signal, before and after control

Performances

The use of two fixed-point DSPs allowed us to reach real time performances for sampling rates up to 20 kHz, thus allowing the control of harmonics as high as 8 kHz. However, even if harmonics up to 8 kHz can be attenuated at the error microphones, the attenuation is only subjectively felt for harmonics under 4 kHz. The reason is that the microphone is too far from the eardrum, compared to the wavelength of the frequency components higher then 4 kHz. To improve this point, we are developing a headset with a microphone that extends to the entrance of the ear canal.

Figure 3 shows the spectrum of a 3 frequency components signal at the error microphone without and with control. We observe an attenuation of 30 to 40 dB per component. The speed at which this attenuation is reached for the 1800Hz component is illustrated in figure 4. It shows a 100 ms segment of the signal at one error microphone when the control is started. We notice that it takes approximately 50 ms for the controller to converge. Perceptually, the transition seems to be instantaneous. This convergence speed corresponds to an adaptation step size small enough to insure almost no sensitivity to useful signals. Actually, even if the controller has a feedforward structure, useful signals can have a negative impact on the adaptation of the coefficients of « C ». A step size too high results in a faster convergence speed but allows useful signals to misadjust the control filter, thus increasing the noise. A compromise was made between the capacity of the controller to follow user's displacements and the misadjustment due to useful signals.

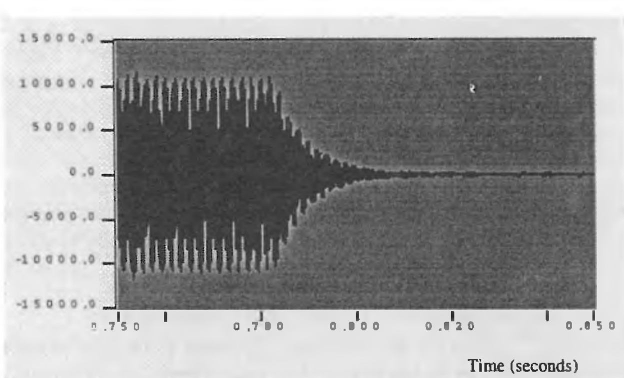


Figure 4 Convergence speed for the 1800Hz component

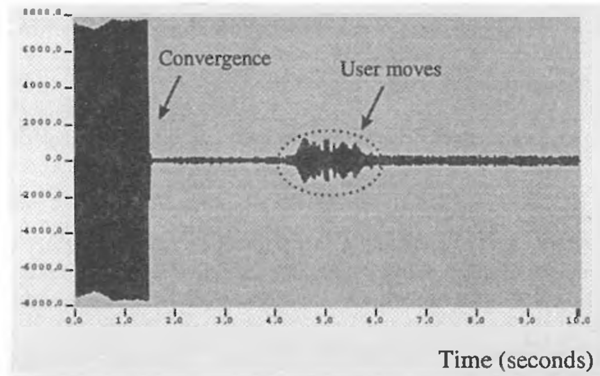


Figure 5 Misadjustment of the controller when the user moves

When the user moves, another misadjustment can sometimes be felt due to the limited tracking speed of the adaptation algorithm. The importance of this misadjustment depends on the convergence speed of the control algorithm, the rapidity of the user's movement, the highest frequency component in the disturbance and the number of frequency components in the disturbance. The highest frequency component present in the disturbance will influence the spatial complexity of the acoustic field. The more the acoustic field is complex, the more the controller has to « follow » rapid changes. Also, the number of frequency components in the disturbance will influence the adaptation speed of the controller. The control filter will adapt more slowly if there are many frequency components in the disturbance. As we can see, the capacity of the controller to follow user's displacement varies with the disturbance characteristics. Figure 5 illustrates the misadjustment of the controller when the user moves at about 2m/s. We notice that despite the misadjustment, the noise level at the microphone while the user moves stays under the initial noise level (before convergence).

Conclusion

The performances achieved in laboratory with the active adaptive feedforward open headset prototype are very encouraging for the future. This kind of headset seems to be a promising solution for many applications, where source noise reduction is not possible.

ACKNOWLEDGEMENT

This project was made possible because of the financial contribution of the NSERC and the IRSST (Institut de Recherche en Santé et Sécurité au Travail).

REFERENCES

[1] B. Widrow and S.D. Stearns, Adaptive Signal Processing, Prentice-Hall, 1985.

[2] DSP Processor Fundamentals, architectures and features, Phil Lapsley, Jeff Bier, Amit Shoham, Edward A. Lee, IEEE Press, 1997

ACOUSTICAL BEHAVIOR OF LIGHTWEIGHT WALLS

Weixiong Wu
Centre de recherche en aménagement et développement (CRAD)
Pavillon Félix-Antoine-Savard, 1636, Université Laval, Québec, Canada G1K 7P4

INTRODUCTION

An experimental investigation of parameters that may affect the transmission loss of lightweight walls is briefly presented. According to the criteria for the sound transmission insulation STC-50 National Building Code in Canada, and the recommended criteria STC-55 by Canada Mortgage and Housing Corporation (CMHC) in condominium buildings, two groups of gypsum board partition walls (coupled walls and uncoupled walls), with or without sound absorbing material in the gap, were examined at the acoustics laboratory of Laval University. The measurements had been carried out using the standard pressure technique with the verification of intensity and vibration techniques in some partitions. The results from measurements and calculations show very strong agreement of the resonance and critical frequencies. The parameters considered in this study include the influence of panel mass, distance of steel studs, absorbing material in the cavity, thickness, multi-leaf panel, symmetric and asymmetric positions, etc. The main goal of this study is to present an acoustical behaviour of lightweight partitions, especially those using traditional materials and an experimental application to the kind of partition made of gypsum board, which could achieve the acceptable criteria STC-50 in Canadian homes and residential buildings, and the recommended criteria STC-55 by CMHC in condominium buildings.

OBJECT OF STUDY

Reviewing the conditions of the last National Building Code STC-45 for party walls and floors separating dwellings established in 1941, one could learn that this criteria was not appropriate to the sound insulation requirement in Canada. A paper, presented to the Canadian Acoustic Association in October 1986 [1], consisted of a collection of complaints investigated by the National Research Council. The compilation contains a relatively small number of complaints about situations in which separation was better than STC-50 and almost none above STC-55. However, the most frequent complaints were in the category from STC-45 to STC-50. The particular aspect of airborne sound insulation was also reported by Bradly [2]. As part of a large survey to relate subjective and objective measures of party walls between adjacent homes, values of several acoustical quantities were obtained in a large number of Canadian homes. Measured STC values ranged from a low of 30 dB to a high of 60 dB. Thus, in spite of efforts to obtain a broad range of STC values, most walls had STC values between 45 and 55 dB. Another research report had been directed by Migneron [3], concerning the acoustic quality of condominiums in the Quebec area. Measured FSTC values show about 38% of the wall partitions achieved STC-55, and 30% of the wall partitions did not achieve STC-50. Unfortunately, 9% of the wall partitions were below STC-45.

Noise reduction of construction elements is a minimum requirement specified in most building codes and, correspondingly, a tool of legal control. Therefore, the prediction of sound transmission properties of building partitions by laboratory testing is of great importance to architects, builders and, consequently, the occupants of the dwelling. In order to achieve the sound insulation criteria STC-50 and the recommended criteria of STC-55 for lightweight walls, academic research concerning the sound reduction index of certain lightweight walls was carried out at the acoustics laboratory. The main objectives were to assess the acoustic parameters of lightweight walls, their sound insulation values, and finally, to investigate the optimum samples in regard to economy, thinness, light weight, and the accepted sound insulation requirements.

METHOD

All sound transmission loss tests were conducted with standard pressure technique in accordance with ASTM 90-70 [4] and ASTM E-413-73 [5]. The sound insulation measurements were performed in the laboratory with two reverberation rooms of 60 m^3 and 200 m^3. Two types of walls, coupled double wall and uncoupled double wall, were chosen for the research. Thirteen test specimens, constructed with gypsum boards and steel studs, were evaluated by various parameters on mass, distance of cavity space, absorption material, stiffness, material orientation, etc. According to the basic transmission loss (TL) theory, there are three important regions for the TL of the wall. When weight restrictions are critical and substantial TL is required, the single wall is generally not adequate. The double wall has greater benefits in middle and high frequencies because it has an airspace. However, for panels and related thin materials (such as plywood, gypsum board etc.) the double wall resonance and coincidence effect are noticeable and they occur in an important part of the frequency range, typically around 1k-4k Hz. (see Figure 1).

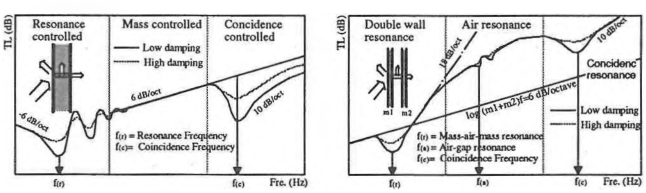


Figure 1. Theory of sound transmission loss

ANALYSIS AND DEVELOPMENT

As shown in Figure 2, the TL increases at low and critical frequencies as the mass of the panel is increased. At middle frequencies (400 Hz to 1600 Hz), TL values experience a small increase, affected by a series of sound coincidence performances which are transmitted by sound bridges in the cavity. There is a significant reduction in TL at the critical frequency. Because the gypsum panels are attached by self-drilled screws, not adhered, the same thickness of gypsum panel has the same coincidence dip at 2562 Hz. Obviously, increasing the mass of the partition can improve STC values. However, with a symmetric orientation such as the N°3 specimen, with two layers of 16 mm gypsum boards on each side, the STC value is less than specimen N°2, since the double-leaf panel resonance is magnified at 82 Hz frequency.

The two TL curves are very different between symmetric partition N°13 and asymmetric partition N°12 (see Figure 3). The results show that the TL can be significantly influenced by the material orientation. For example, the STC value of N°13 is 3 dB lower than N°12, although its internal panel mass per unit area is increased. It can be seen that its mechanical coupling resonance is amplified at 274 Hz. Clearly, using symmetric orientation seems to improve the TL at certain high frequencies, especially at critical frequency (2500 Hz) in this case. However, the vibration and harmonic resonance occur at middle frequencies so that the TL values are significantly reduced in this region.

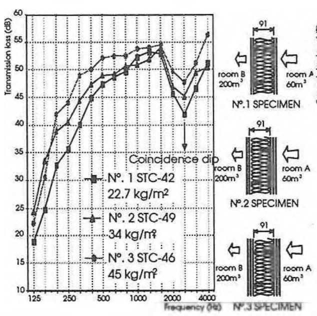


Figure 2. Influence of mass

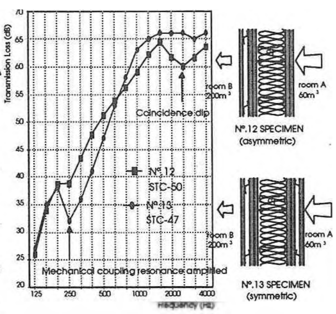


Figure 3. Material orientations

TL of the uncoupled double wall has many benefits at middle and high frequencies (see Figure 4). However, TL was not greatly increased at low frequencies since it is influenced by the resonance of double panel at 105Hz. After the resonance, TL increases about 28 dB per octave until 250 Hz, and 12dB per octave until the air-gap at 798 Hz where the resonance is encountered. It is very important to learn that STC still depends on the TL values at low frequencies. Even though the uncoupled wall has much higher TL values at middle and high frequencies than the coupled wall has, its STC value is improved by only 3 dB.

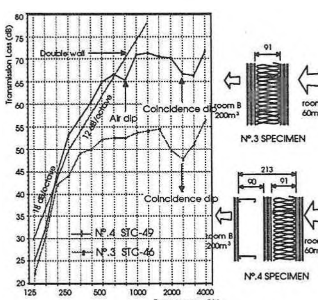


Figure 4. Influence of cavity

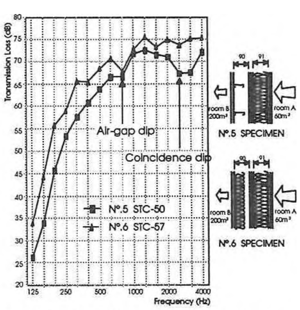


Figure 5. Absorption materials

Figure 5 demonstrates the variation results with and without absorption material in a cavity. The top curve signifies a large TL increase from 5 to 10 dB for all the frequencies except the air-gap frequency. It is interesting to note that the TL value cannot be affected by adding absorption material in the cavity at this frequency. At higher frequencies, the phase of the sound pressure varies with the thickness of the cavity, and the acoustic coupling is weaker. The TL value also significantly increases at the critical frequency 2500Hz. Generally, perimeter absorption has a great effect on TL at most frequencies.

The results are very similar using different cavity space (118 mm, 141 mm, and 166 mm). It seems that to change the space distance in the cavity can vary their STC values a small amount. In addition, the effect of partition stiffness was also measured by reducing the distance of steel studs from 600 mm to 400 mm, and STC improved 1 dB only. On the other hand, the intensity method was used to verify the reliability of STC values in which the TL demonstrated a small difference at certain high and low frequencies. Furthermore, the vibration method was used to identify the frequency of mass-air-mass resonance, air-gap resonance, coincidence effect, etc. Table 1 outlines the information and measured results of thirteen samples [6].

Table 1. Measured results

Sample N°	1	2	3	4	5	6	7	8	9	10	11	12	13
Thick mm	125	140	158	260	260	260	292	317	342	358	184	180	202
Mass kg/m²	22.7	34.1	45.5	53.9	53.9	53.9	56.9	56.9	56.9	68.2	56.9	53.9	62.3
STC dB	42	49	46	49	50	57	62	62	63	66	50	50	47
$/m²*	25.6	29.1	32.6	41.9	42.3	53.9	54.9	54.9	54.9	58.4	37.6	36.9	41.1

* the price valued in 1992

CONCLUSIONS

For coupled walls: They have a high level of agreement between measurements and theoretical calculations in three controlled regions. Adding layers of gypsum board increases STC, but not for the symmetric orientation samples. The TL increases 16 to 18 dB/octave after resonance frequency, 6 to 10 dB/octave at the mass controlled region. The TL increases 2 to 5 dB/octave at middle and high frequencies (400 to 2k Hz) by adding one gypsum panel. The coincidence resonance can be reduced by increasing the mass. The TL increases 15 dB/octave after coincidence frequency. The mechanical coupling resonance is significantly amplified with symmetric position. The coincidence dip can be minimised by using different thickness of gypsum boards and asymmetric position.

For uncoupled walls: They have great benefits at the middle and high frequencies. The TL increases 28 dB/octave after resonance frequency, 12 dB/octave until the first air-gap resonance, 5 dB/octave in the general air-gap resonance region. Adding absorption material to the cavity space increases STC to 7 dB, and considerable improvement occurs at the low and critical frequency bands. The STC values do not show a large difference when placing steel studs 600 mm or 400 mm away from each other, or changing the air-gap space distance from 118 mm to 166 mm.

Finally, specimen N°2 (STC-49) nearly achieved STC-50. It should reach more than STC-50 with the following developed construction model. Specimen N°6 (STC-57) is optimum an ideal STC-55 wall: thin, lightweight, and economical.(see Figure 6).

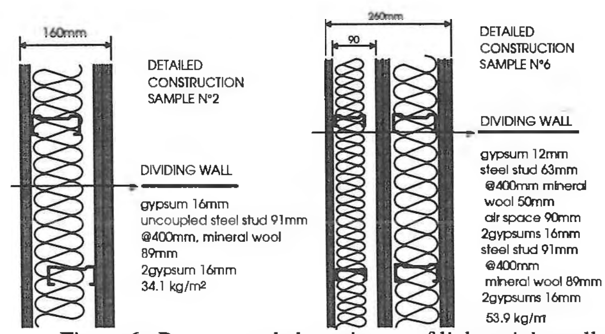


Figure 6. Recommended specimens of lightweight walls

REFERENCES

[1] NORTHWOOD, T.D.: *Criteria for airborne sound insulation between dwelling.* Paper presented at meeting of Canadian Acoustical Association. 1986

[2] BRADLEY, J.S.: *Acoustical Measurements in Some Canadian Homes.* Canadian Acoustics V. 14 N°4 October 1986

[3] MIGNERON, J.-G., et MARTEL, C.: *Qualité acoustique, insolation et intimité des immeubles d'habitation en copropriétés.* 325p. Rapport CRAD/SCHL, contrat 6585/M12-4, Université Laval, 1989

[4] American Society for Testing and Materials ASTM Standard E-90-70 (Laboratory Measurement of Airborne Sound Transmission Loss of Building Partitions)

[5] American Society for Testing and Materials ASTM Standard E-413-73 (Determination of Sound Transmission Class)

[6] WU, Weixiong: *Airborne Sound Transmission Loss of Lightweight Walls.* Master's thesis, Université Laval, 1994

Computer Studies of Optimum Classroom Acoustics

J.S. Bradley and R. Reich, Institute for Research in Construction, National Research Council, Montreal Rd., Ottawa, K1A 0R6.

Introduction

The intelligibility of speech in a classroom depends on both room acoustics effects and on the speech-to-noise ratio. Very high quality speech communication requires both optimum room acoustics conditions as well as a low ambient noise level to provide adequate speech-to-noise ratios.

Several studies have shown that a speech–to-noise ratio of 15 dB or more will provide 100% speech intelligibility (where the speech and noise levels are A-weighted levels). Room acoustics has traditionally been described in terms of reverberation time (*RT*) and various optimum reverberation times have been proposed to maximize intelligibility. (Typical recommended *RT* values are from 0.4 to 0.7 s). More modern work suggests that the effects of room acoustics on speech intelligibility are better assessed in terms of early-to-late sound ratios (C_{50}) or Speech transmission Index (*STI*) values. (*RASTI* is a simplification of the *STI* measure).

Adding sound absorbing material to optimise *RT* or to maximize C_{50}, will also affect speech and noise levels in the classroom. Thus to determine optimum conditions for speech, one must consider both speech-to-noise ratios and room acoustics effects. This can best be done in terms of newer measures that combine both effects into a single measure such as useful-to-detrimental sound ratios (U_{50}) or the speech transmission index (*STI*). When these measures are maximized, the particular combination of room acoustics and speech and ambient noise levels provides the best possible speech intelligibility.

Experiments

A typical classroom was simulated using the ODEON room acoustics modeling program. U_{50} values were calculated from the ODEON output combined with speech and noise levels. Speech intelligibility scores were estimated from the calculated U_{50} values.

In the first series of tests the ceiling absorption was varied and the *RT* corresponding to the maximum intelligibility was determined. In the second series of experiments, the location of the sound absorbing material was varied to determine the location that maximized intelligibility.

As seen in Figure 1, the optimum *RT* corresponds to about 0.5 s but a range of *RT* values from about 0.3 to 0.6 s lead to intelligibility scores within 0.5 % of the maximum. Thus, it is not necessary to achieve exactly the optimum RT.

The choice of absorption configuration could increase C_{50} values by as much as 4 dB. Average speech levels could vary by as much as 3 dB with absorption configuration. Only small overall improvements were possible, because the absorption configurations that maximized C_{50} tended not to maximize speech levels. Figure 2 shows the average U50 values for 9 different absorption configurations. These results suggest that previous recommendations for 'optimum' locations of absorbing material may not be of practical importance.

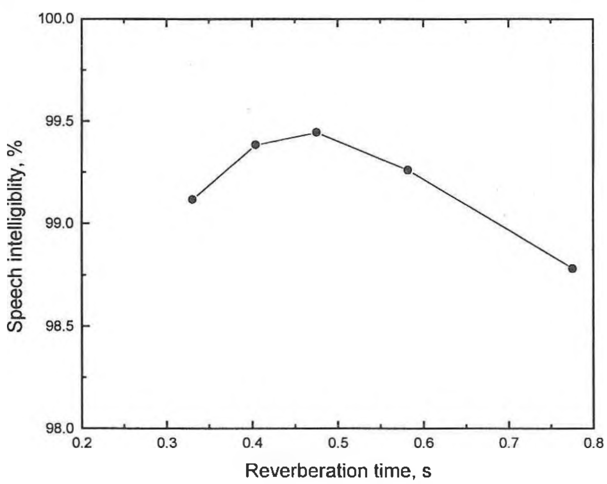


Figure 1. Calculated speech intelligibility versus reverberation time.

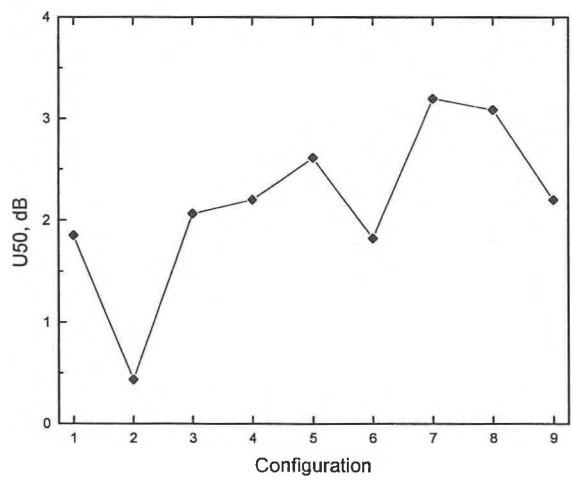


Figure 2. Calculated U_{50} for 9 absorption configurations and omni-directional source.

Sound Transmission Through Floor/Ceiling Assemblies

A.C.C. Warnock
Institute for Research in Construction, National Research Council Canada.

Introduction.

The Institute for Research in Construction Acoustics Laboratory has completed the measurement phase of a study of airborne and impact sound transmission through typical floor constructions used in Canadian housing. As well as IRC/NRCC, the project was supported by a consortium including 18 companies or associations from Canadian and US industry. This paper discusses briefly the results found in terms of sound transmission class (STC) and impact insulation class (IIC) ratings made in accordance with the relevant ASTM standards[1,2,3,4]. A more detailed report is available[5].

Types of floors tested.

About 200 different floor variations were included in the study. Joist types included solid wood, steel, wood I-joists and wood trusses. A few joist floors with concrete toppings and three concrete slabs were also tested.

Repeatability

Most important for comparing test results within a series of measurements in a single laboratory is the concept of *repeatability*. This may be defined as the closeness of agreement between independent results obtained with the identical test specimen in the same laboratory with the same equipment and test method by the same operator within a short time period.

Rebuild repeatability may be defined as the closeness of agreement between results obtained on nominally identical test specimens constructed with nominally identical materials with the same test method in the same laboratory. This repeatability is of most relevance where comparisons are being made among floors that were completely rebuilt and represents the highest uncertainty associated with this project.

To estimate *rebuild repeatability*, nominally identical floors were constructed and tested eight times in the laboratory over a period of about 1 year using new materials each time. Four of the STC ratings obtained were 51 and four were 52. Four of the IIC ratings were 45 and four were 46. It was concluded that a change of more than 1 point in the STC or IIC rating could be taken as significant and attributed to a change in the specimen. A change of only 1 is regarded as not significant unless an examination of the 1/3 octave band data shows significant changes.

Major findings.

In the space available it is only possible to give some of the highlights of the report in point form.

- The major factor controlling the sound insulation of a given type of cavity floor is the sum of the masses per unit area of the floor and ceiling layers.
- Of lesser importance, but still significant, are the thickness and density of the sound absorbing material, the depth and spacing of the joists and the spacing of resilient metal channels. Increasing any of these variables increases sound insulation.
- Joist floors without resilient metal channels do not achieve STC 50 in any practical configuration, with or without sound absorbing material in the cavity.
- Using 22 mm deep U-channels to support the gypsum board gave about the same results as using 19 x 64 mm wood furring. Both are markedly inferior to resilient metal channels.
- Changing the joist length had no effect on the sound transmission.
- The tightness of the screws attaching the subfloor to the joists had no effect on sound transmission.
- Increasing the number of screws attaching the subfloor to the joists by a factor of four greater than normal had no effect on sound transmission.
- Attaching the subfloor to the joists using both construction adhesive and nails gave the same results as attaching it using only screws.
- There were no significant differences in STC or IIC between pairs of floors where a 35 mm thick concrete topping was poured on top and allowed to set or where an existing slab was lifted into place on the floor.
- There was no significant difference between a floor constructed using cross-bracing and one using wood strapping. Floors gave the same sound insulation with or without cross-bracing.
- Putting sound absorbing material in the cavity of a joist floor with a ceiling that is not resiliently suspended provides no significant increase in sound insulation.
- Floors with concrete toppings and no additional resilient surface or support for the ceiling, typically get IIC ratings less than 30.
- Differences among types of sound absorbing material are significant but small.

Regression Analyses

Regression analyses of the data collected permit interpolation and extrapolation of the results to cases that were not actually measured. Developing an analytical model would be more satisfactory but requires much more work. This section presents some of the more useful results of the regression analyses.

A regression analysis of all the measured results as one collection of data would not be fruitful. The many variations in construction that are possible have too great an influence on sound insulation and are not easily dealt with using simple linear regression models. Such models would not easily deal with a collection of data including floors having resilient metal channels separating two layers of gypsum board, or floors with and without resilient metal channels. The data were separated into major categories as follows:

- Solid wood joist floors with resilient metal channels directly attached to the joists and with sound absorbing material in the cavity (70 floors),
- Wood I-joist floors with resilient metal channels directly attached to the joists and with sound absorbing material in the cavity (23 floors),
- All cavity floors with resilient metal channels directly attached to the joists and with sound absorbing material in the cavity (110 floors), and

- All cavity floors with resilient metal channels directly attached to the joists and with *no* sound absorbing material in the cavity (11 floors).

Other categories did not contain enough data to allow meaningful analysis.

For analyses with IIC as the dependent variable, floors with concrete toppings or resilient toppings were excluded from the regression analysis. The resilience of the floor layer struck by the tapping machine strongly influences the level of impact sound generated by the ISO tapping machine. This important variable needs specific measurements for its characterization but the project deliberately did not focus on this aspect of sound insulation; it is a problem sufficiently complex that it needs a separate study.

For all the analyses, the physical variables found to be significant were the mass per unit area of the sub-floor and the ceiling, joist depth and spacing, resilient metal channel spacing, and the thickness and density of the sound absorbing material. Other parameters did not correlate with STC or IIC. In particular, adding the mass of the floor framing as an independent variable or in combination with other variables decreased the correlation. In many cases not all of these variables were significant especially when the number of cases was low.

Some of these groups provided fairly satisfactory regression equations that could be used for prediction within the limits of the variables used to generate the equations. However, regression equations developed from solid wood joist data predicted sound insulation for other types of beams poorly; the I-joists and trusses had depths much greater than the deepest solid joist tested (286 mm). For simplicity, all joist types were assumed to be similar and data for all types (solid wood, wood I-joists, wood trusses and steel joists) were analyzed together. The regression equations found were

$$STC = 7.1 + 23.9 * \log_{10}(Layers) + 0.0086 * JstDepth + 0.0066 * JstSpace + 0.017 * InsThick + 0.0085 * RCSpace + 0.030 * InsDensity,\ r^2 = 0.92,\ 110 \text{ cases}$$

$$IIC = 10.6 + 22.2 * \log_{10}(Layers) - 0.010 * JstSpace + 0.016 * InsThick + 0.012 * RCSpace,\ r^2 = 0.92,\ 102 \text{ cases}$$

where *Layers* is the total mass per unit area of the floor and ceiling layers in kg/m^2, *InsDensity* is in kg/m^3, and all other dimensions are in mm.

It is surprising that the IIC rating shows a negative dependence on joist spacing. There is no obvious explanation to be found in this analysis. More detailed study using one-third octave band data may provide insight.

For STC, 90% of all the predictions fell within ± 1 dB of the measured values, 96% within ± 2 dB, and 94% of the predictions were no more than 1 dB below the measured values. For IIC the corresponding values are 75%, 92% and 89%.

Accuracy of prediction for different types of sound absorbing material

Representative regression equations that are generally applicable are only obtained when there is a reasonably uniform distribution of the values of each predictor variable. Since the majority of the measurements were made using glass fiber batts as the sound absorbing material, the regression equations predict the results for this material well. Not enough data were collected for the other types of sound absorbing material to get the same accuracy of prediction for them. The coefficient for the variable insulation density in the STC regression equation is 0.03. Thus increasing density from 10 to 30 kg/m^3 should increase the STC by 0.6 dB (This corresponds approximately to a change from glass fiber batts to rock fiber batts). There is no dependence on the density of the sound absorbing material in the regression equation developed for IIC. The data in Fig. 1 come from three sets of floors where more direct comparisons were made. This figure suggests that changing from glass fiber batts to rock fiber batts should increase the STC and IIC by about 1 point, if not more. The corresponding plot for IIC also suggests about a 1 point advantage from using rock or cellulose fiber. More measurements are needed to clarify this issue.

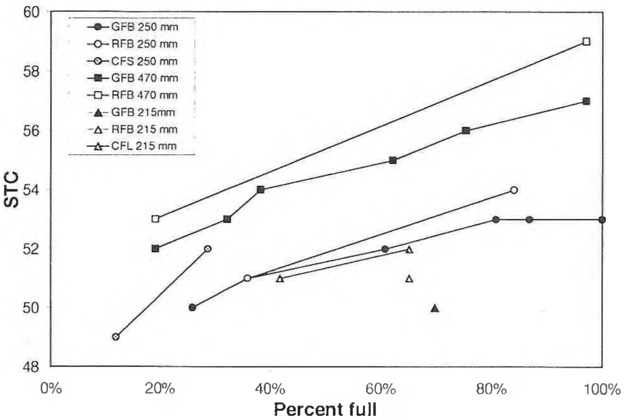


Figure 1: STC for three different floor systems with different joist types and depths with varying amounts of sound absorbing material in the cavity. The floors used solid wood joists (250 mm), steel joists (215 mm) and wood I-joists (470 mm).

References.

[1] ASTM E90 Standard Test Method for Laboratory Measurement of Airborne Sound Transmission Loss of Building Partitions.

[2] ASTM E413 Classification for Rating Sound Insulation.

[3] ASTM E492 Standard Test Method for Laboratory Measurement of Impact Sound Transmission through Floor-ceiling Assemblies using the Tapping Machine.

[4] ASTM E989 Standard Classification for Determination of Impact Insulation Class.

[5] "Summary Report For Consortium On Fire Resistance And Sound Insulation Of Floors: Sound Transmission Class And Impact Insulation Class Results", A.C.C. Warnock and J.A. Birta, IRC-IR-766, 1998.

Sound Transmission Through Ceilings from Air Terminal Devices in the Plenum

A.C.C. Warnock
Institute for Research in Construction, National Research Council Canada.

Introduction.

HVAC sources in ceiling plenums are often major contributors to the noise level in occupied spaces below. This paper presents results from ASHRAE RP-755 [1, 2, 3, 4] and discusses primarily the sound attenuation through ceiling systems and the sound pressure levels in the room below. The focus of RP-755 was on the interactions between terminal units (positioned above and close to a lay-in ceiling), the ceiling panels, the plenum and the room below. The intent was to evaluate the calculation methods used in ARI 885-90 [5] and suggest improvements where necessary.

At the beginning of the project it was known that transmission loss results obtained in reverberation room tests did not apply to this situation because of the close coupling between the source and the ceiling panels and the absence of a diffuse sound field in typical plenums. As well, the information used to prepare ARI 885 was provided by a few manufacturers, but it did not form a consistent set based on a standard test procedure or accepted method of measurement.

Summary of the investigation.

Test Room. The room acoustics test (RAT) room, where the measurements were made, is a rectangular parallelepiped 4.71 m wide and 3.6 m high. One end wall can be moved but for most of the experiments, the length was set at 9.2 m giving a room volume of 156 m^3. The T-bar system for supporting tiles was installed so the distance from the face of the supporting surface of the T-bar to the true ceiling of the room was 740 mm. To provide some scattering, 8 sheets of 16 mm gypsum board measuring 1.22 x 1.22 m were hung on the walls or placed on the floor and inclined against the walls.

Sound Sources. The four terminal types used in the project were an air-to-air ceiling induction unit, a VAV shutoff unit, a series flow fan-powered VAV unit, and a parallel flow fan-powered VAV unit. To provide more convenient reference sources with good repeatability, two metal boxes each containing two loudspeakers radiating random noise were also used as sources above the ceiling. One was positioned near the middle of the room, close to the devices being tested. The second was placed in one corner of the plenum.

Ceiling Types. Six ceiling panel types laid on a standard T-bar grid were tested. No clips or other devices were used to hold the panels down. The types of panel and the coded identifiers used for brevity are given in Table 1. Sound transmission loss (ASTM E90) and sound absorption (ASTM C423) with the specimen mounted on an E400 frame (ASTM E795) were measured for each ceiling type in NRC's reverberation rooms.

Table 1: Ceiling types and codes used for identification.

Code	Ceiling panel type
A895	16 mm thick mineral fiber tiles
A755B	16 mm thick lightweight mineral fiber tiles
G13	13 mm vinyl-faced gypsum board
FGvin	50 mm thick glass fiber tile with perforated vinyl face
FGTL	50 mm thick glass fiber tile with perforated vinyl face and metal foil backing
A2910	16 mm thick glass fiber tiles with vinyl face randomly perforated with fissures

Measurements. Sound pressure levels were measured in the RAT room for each source in combination with each ceiling type. As well, for each ceiling the reverberation times and sound pressure level as a function of distance from a nominally omni-directional source were measured in the room.

Major Results.

The difference between the sound power, L_w, of a given device placed in the plenum and the average sound pressure level, $\langle L_p \rangle$, in the room below measures the combination of the "plenum/ceiling effect" and the average "space effect". These terms are defined in ARI 885 as

Plenum/Ceiling Effect: The difference between the octave band sound power level from the source located in the plenum/ceiling cavity and the sound power level transmitted to the occupied space.

Space Effect: The difference between the octave band sound power level entering the occupied space and the resulting octave band sound pressure level at a specific point in the space.

Ceiling attenuations. The attenuations for all the sources used were averaged to get the average $L_w - \langle L_p \rangle$ for each type of ceiling tile (See Fig. 1). The interesting feature of this graph is the small differences among tile types with the exception of the G13 and A2910 tiles.

One might have expected that the heavier G13 tiles would have given much lower levels in the room than the lighter tiles. The conclusion drawn from this result was that for most of the tiles used, the dominant path through the ceiling is the leakage between the edges of the tiles and the T-bars. For the mineral fiber and glass fiber tiles there will be different relative amounts of sound power transmitted due to leakage, absorption on the rear face and at the edges, and transmission through the body of the tile.

The light A2910 tiles provide little attenuation through the body of the tile at high frequencies and for the gypsum board tiles, there is no sound absorbing material to offset the effects of the leaks around the edges of the tiles. These two types of tiles are quite different from the others but are perhaps not typical of products used below air terminal units.

One conclusion that can be drawn from this figure is that since most normal tiles give about the same result, there is little point in creating a test procedure to rate the effectiveness of ceiling tiles as attenuators of sound from air terminal units. The sound powers of the devices tested all decreased fairly rapidly as frequency increased. So, even the poor attenuation of the G13 and A2910 tiles at high frequencies is not likely to be important. On the other hand, mounting systems for the tiles other than standard T-bars, give more attenuation [6] so it may still be deemed advisable to create a test procedure.

The attenuations measured for the different ceiling types, did not agree well with those predicted from values given in ARI 885.

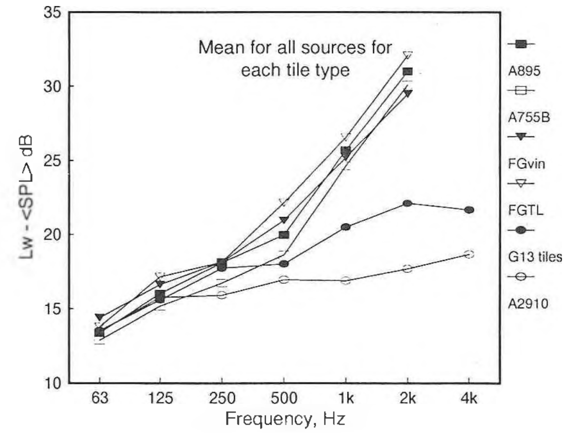


Figure 1: Average of L_w - $\langle L_p \rangle$ for all sources for each type of ceiling tile.

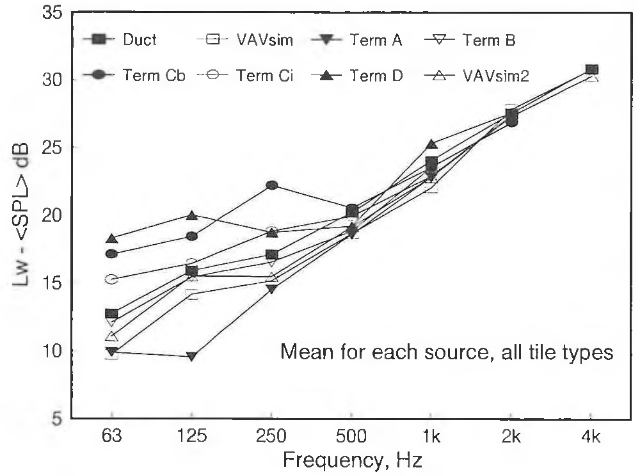


Figure 2: Average of L_w - $\langle L_p \rangle$ for all ceilings for each type of source.

Differences among sources. Figure 2 shows for each source the average attenuation for all the ceiling types used. If the sound power emitted by the device were not altered by the presence of the ceiling, if the attenuation provided by the ceiling were constant, and if there were no interaction between the device and the ceiling, then all of these curves would be approximately the same. There are, however, quite significant differences at and below 250 Hz among the devices. The conclusion drawn from this graph is that the coupling between the ceiling tiles and the source influences the sound power radiated into the room below the ceiling. This makes it difficult to accurately predict the sound pressure level in the room below using only sound power levels for the device measured according to standards and some fixed insertion loss values for the ceiling tiles.

Dependence of ceiling attenuation on source area. Examination of the data revealed a fairly strong correlation between the effective ceiling attenuation and the area of the surface of the source closest to the ceiling. As the area increased, so did the ceiling attenuation. This correlation explains much of the scatter at low frequencies in Fig. 2. No physical model or analytical expression has been found to explain this dependence. The empirical model developed is:

$$SPL(f) = P(f) - A(f) + m(f) \times (S - 0.83).$$

where

- f is the mid-band frequency of the octave band, Hz,
- $SPL(f)$ is the average sound pressure level in the room, dB,
- $P(f)$ is the power emitted by the terminal unit when tested according to standards [7], dB,
- $A(f)$ is the nominal attenuation of the ceiling tiles, dB,
- $m(f)$ is the slope of the regression of attenuation on area, dB/m^2,
- S is the area of the lower face of the terminal unit, m^2, and
- 0.83 is an empirical constant determined from the measured data.

The values of m found from experiment are given in Table 2.

Table 2: Values of coefficient m.

f, Hz	63	125	250	500	1k	2k	4k
m, dB/m^2	4.4	4.1	2.5	0	0	0	0

Spatial attenuation. For a source placed in the room *below* each ceiling the Schultz [8] formula predicts an attenuation of 3 dB/distance doubling independent of room absorption. In this work, however, the attenuation depended on the reciprocal of the room reverberation time according to

$$\text{Attenuation} \approx (0.9/RT + 0.5) \text{ dB/dd}.$$

The dependence on RT, while quite clear, is not very important in practice in typical rooms.

When the source was *above* the ceiling in the plenum, the sound field in the room below varied very little with distance from the source. Except at 2000 and 4000 Hz, the attenuation is less than 1 dB/dd. ARI 885 specifies the use of the Schultz formula and so was inaccurate in this respect.

This work showed that insertion losses for ceiling systems cannot readily be obtained from standard measurements in reverberation rooms. Based on the project new procedures for calculating the sound pressure level in a room below a terminal unit were recommended to ARI and ASHRAE.

References.

[1] Warnock, A.C.C., "Sound Transmission Through Ceilings from Air Terminal Devices in the Plenum", Final report ASHRAE RP755. January 1997.

[2] Warnock, A.C.C., "Sound pressure level versus distance from sources in rooms," ASHRAE Transactions 1998, V.104, Pt.1.

[3] Warnock, A.C.C., "Environmental correction factors for predicting room sound pressure levels", ASHRAE Transactions 1998, V.104, Pt.1.

[4] Warnock, A.C.C., "Transmission of sound from air terminal devices through ceiling systems", ASHRAE Transactions 1998, V.104, Pt.1.

[5] ARI Standard 885-90, Air-Conditioning & Refrigeration Institute, Arlington, Virginia, 1990.

[6] F.P. Mechel, "Theory of Suspended Ceilings", Acustica, **81**, p491, 1995

[7] ARI Standard 880, "Air terminals", Air-Conditioning & Refrigeration Institute, Arlington, Virginia, 1990.

[8] Schultz, T., 'Relationship between sound power level and sound pressure level in dwellings and offices', RP-339, ASHRAE Transactions 1985, V. 91, Pt. 1.

Comparisons of Computer Simulations of Acoustical Conditions in Classrooms

Sylvio R. Bistafa* & John S. Bradley
Institute for Research in Construction, National Research Council, Montreal Road, Ottawa, K1A 0R6

Introduction

Acoustical specifications for classrooms are usually established solely on the basis of reverberation time. Although in general the conditions for speech communication improve as an optimum reverberation time is approached, reverberation time by itself does not provide the best direct measurement of speech intelligibility. Many metrics have been developed and used to measure speech intelligibility such as the Speech Transmission Index (STI), Definition (D), Early/Late Ratios (C_{50}), and Useful/Detrimental Ratios (U_{50}). All of them can be measured in real rooms, but predictions of these parameters can best be obtained by computer simulations. This work consists of comparisons of computer simulations of acoustical conditions in classrooms for a number of newer parameters and using two different computer models.

Computer Programs and Simulated Classroom

Many room acoustical computer simulation programs are available today that can be used for this purpose. To study the relation among the newer speech intelligibility metrics, as well as to obtain some indications about the reliability of this type of room acoustic simulation tool, we use two computer programs available to us namely: Odeon 2.6 and Raynoise 2.1A. Both programs use so-called hybrid models in that they combine different procedures for calculating the earlier and later parts of the impulse response. The modeled classroom had dimensions of 7.6m x 10.0m and 3.3m height. Four different sound absorption material configurations, using mineral wool, were simulated to investigate the relation between reverberation and speech intelligibility. The ceiling and the back wall surfaces were respectively covered: 1: 100%, 100%; 2: 35% (outer ring), 100%; 3: 0%, 100%; 4: 0%, 0%, in each configuration.

Speech Intelligibility Metrics

Both Odeon and Raynoise give values for Reverberation Time and Definition, and Raynoise also gives the Sound Transmission Index. With Definition defined as:

$$D = \frac{\int_0^{50ms} p^2(t)dt}{\int_0^{\infty} p^2(t)dt}, \qquad (1)$$

where p(t) is the room impulse response; C_{50} and U_{50} were calculated using the formulas:

$$C_{50} = 10\lg\left[D/(1-D)\right], \qquad (2)$$

and

$$U_{50} = 10\lg\left\{\frac{D}{1-D+\left[10^{(\mathbf{Noise}-\mathbf{SPL})/10}\right]}\right\}, \qquad (3)$$

where **Noise** is the overall background noise level in dBA and **SPL** is the octave band speech sound level .

Results and Discussion

The speech intelligibility metrics and reverberation time were determined at eight different microphone positions uniformly distributed inside the classroom. The classroom was supposed to be empty with no pupils or furniture, and the overall background noise level was 32 dB(A). The speech level and directivity was that of a male talker with a normal vocal effort. The results from the eight microphone positions were averaged in each octave band. The final results are displayed for a specific metric as an octave band spectrum for each sound absorption configuration. Figure 1 shows the results for the Reverberation Time. As expected, RT decreases with the increase of sound absorption. The agreement between the results furnished by both programs is good with the exception of configuration 4. This corresponded to a classroom with no sound absorbing material. For this very reverberant room, Figure 1 also shows the RT as given by the Sabine equation. It is seen that neither Odeon nor Raynoise estimated the expected RT, as given by Sabine equation very acurately. For Odeon, changing the coefficient of diffusion of the room surfaces did not result in significant differences of RT. For Raynoise, diffusion is not taken into account for calculations at specific microphone positions (IMAGE Option Calculation). This situation is said to have been changed in Raynoise Rev3.0. It was found during the simulations that, for both programs, diffusion does not seem to have a significant effect on the speech intelligibility metrics. Figure 2 shows comparative results for D, as given by both computer programs. Figure 3 shows calculated values of C_{50}, using Equation (2). Figure 4 shows calculated values for U_{50}, using Equation (3), with the speech SPL values as given by the programs. The agreement is quite good and shows that speech intelligibility improves as the room sound absorption increases, but an upper limit might exist as shown by the U_{50} values on Figure 4. As can be seen, U_{50} begins to decrease at high frequencies due to the reduction in the speech levels with added absorption inside the classroom. The same fact can be seen on Figure 5, which shows values of STI, as given by Raynoise, for the four different absorption configurations. An upper limit on STI seems to have been reached a 4000 Hz for the absorption configuration 4. Figure 6 shows values for the overall U_{50} in dBA calculated using the frequency band levels given in Figure 4. Figure 7 shows microphone position-averaged overall STI as given by Raynoise. The same trend is observed on these overall results; that is, there is an increase on the values of U_{50} and STI with sound absorption. However this might vary if the ambient noise level was increased.

Conclusions

The simulations showed that the speech intelligibility metrics have been consistently estimated using two different room acoustic computer programs. The prediction of reverberation time seems to deviate from expected values at high frequencies, with both programs, for a very reverberant room. The position-averaged results were in quite good agreement, and the general trend seems to have been correctly predicted; that is, an increase of speech intelligibility occurs with added sound absorption, with an upper limit at high frequencies for the most absorbing sound configuration.

* Visiting Scientist, University of Sao Paulo, Brazil.

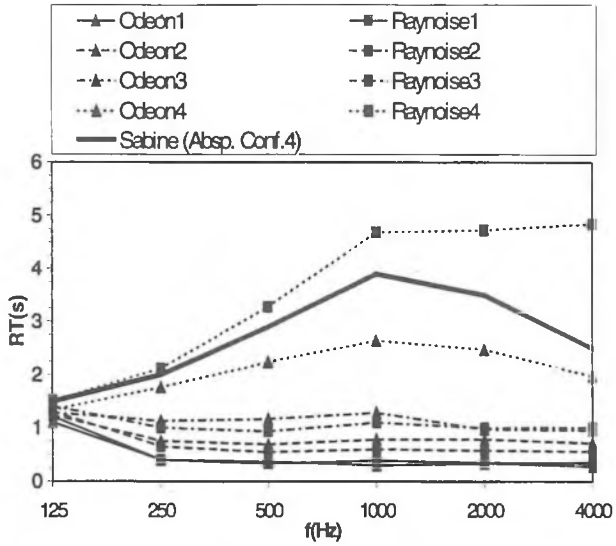


Figure 1: Reverberation Time in Octave Frequency Bands for Four Sound Absorption Configurations as Given by Odeon and Raynoise.

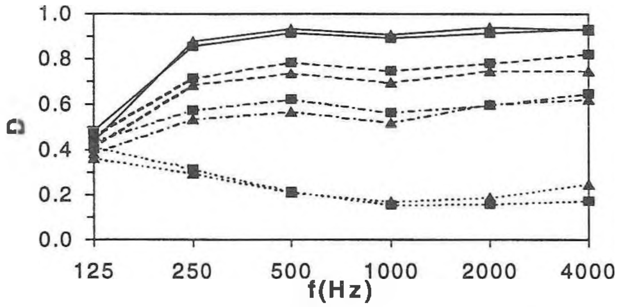


Figure 2: Definition in Octave Frequency Bands for Four Sound Absorption Configurations as Given by Odeon and Raynoise.

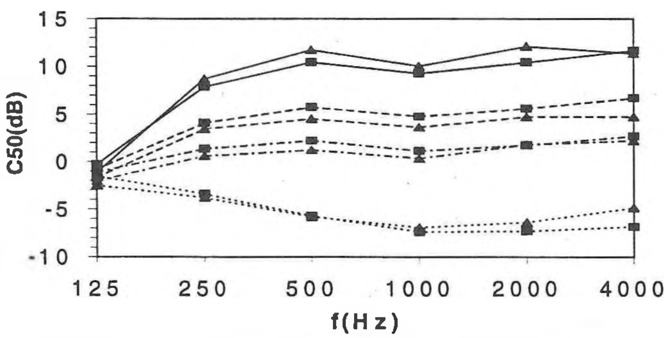


Figure 3: C_{50} in Octave Frequency Bands Calculated Using Equation (2) for Four Sound Absorption Configurations.

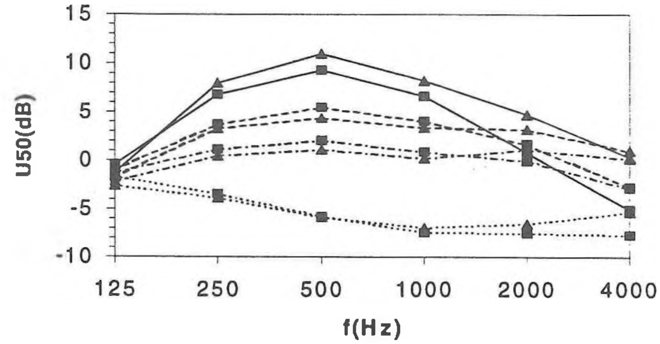


Figure 4: U_{50} in Octave Frequency Bands Calculated Using Equation (3) for Four Sound Absorption Configurations.

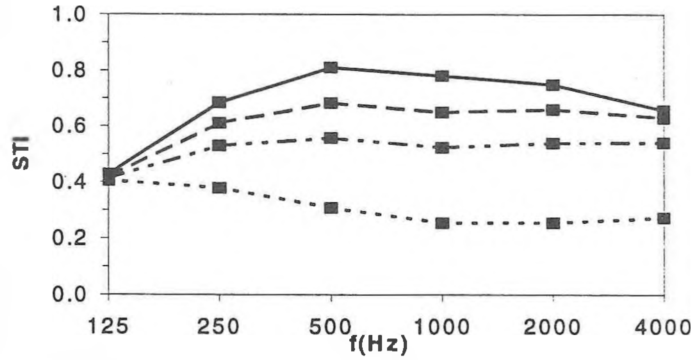


Figure 5: STI in Octave Frequency Bands as Given by Raynoise for Four Sound Absorption Configurations.

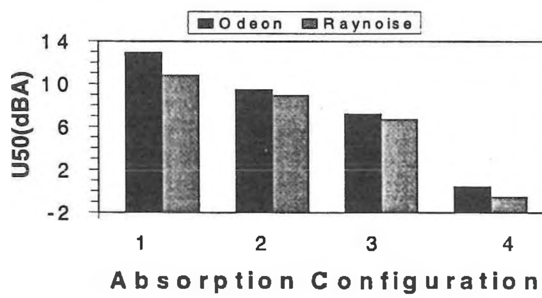


Figure 6: Overall U_{50} Calculated Using the Frequency Band Levels of Figure (4) for Four Sound Absorption Configurations.

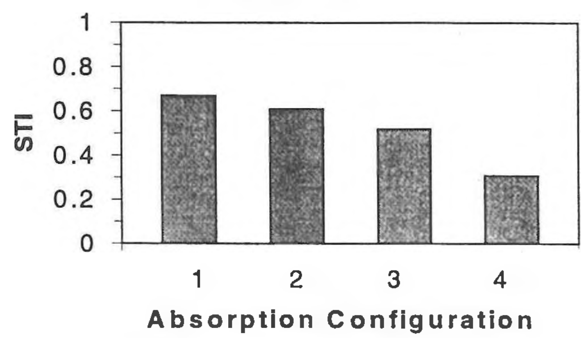


Figure 7: Raynoise Overall STI for Four Sound Absorption Configurations.

NOISE REDUCTION FOR A SHEET METAL SHEAR – A PROGRESSIVE APPROACH

H. M. Williamson, Hugh Williamson Associates, Ottawa, Ontario, Canada
J. C. S. Lai, M. A. Burgess and A. Bahrami, Acoustics and Vibration Unit, University College, University of New South Wales, Australian Defence Force Academy, Canberra, ACT, Australia
C. Speakman, Sound Control Pty. Ltd., Brisbane, Queensland, Australia

Introduction

The control of industrial noise, and the reduction of noise induced hearing loss among workers, remains a serious challenge to industry and to the engineering and acoustics communities. This is particularly the case in smaller manufacturing facilities where practical and effective noise control measures need to be retrofitted to existing machinery. This paper reviews a case study[1-4] where noise control measures were progressively developed for a manufacturing facility which produced roll-formed sheet metal for the building industry. The study demonstrates how noise control goals can be achieved through a step by step development process with the effectiveness of each measure being assessed at each stage of development.

BHP Building Products obtains coils of pre-coated sheet steel from a central production facility, however roll-forming operations are decentralised at smaller plants at many locations throughout Australia. The roll forming production line discussed in this case study was one of approximately 100 similar roll-forming lines operated by the company. The sheets, metal thickness approximately 0.45 mm, were roll-formed to a sinusoidal or corrugated profile typically used as a roofing material. Sheets were approximately 900 mm wide and cut to various lengths to suit customer requirements using a flying shear, Figure 1. The high level of impact noise produced by the shear was the most significant source of noise on the production line.

Noise Source Assessment and Measurement

Prior to the development of control measures, it was necessary to identify which components of the operation were the major radiators of noise. This was done by recording the sound pressure signal using a PC based data acquisition system so that sound emissions could be correlated with mechanical operations during the shearing process, Figure 2. Additionally, an accelerometer was used to measure the vibration of various machine components, simultaneously with noise measurements, to further assist in identifying the noise sources. These preliminary measurements

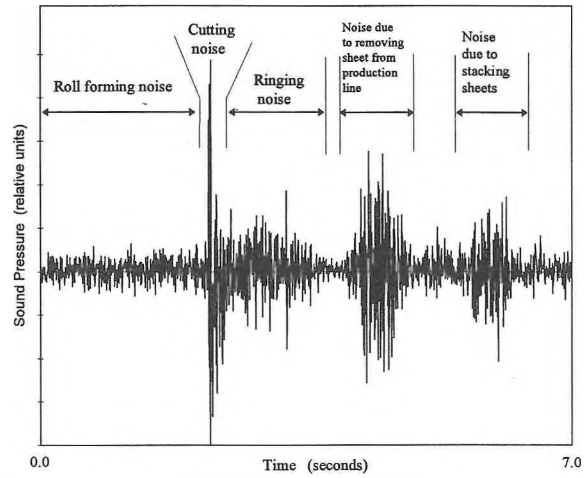


Figure 2 Typical sound pressure during roll-forming

showed that the shearing process resulted in high levels of transient vibration in the sheet product which were the major source of impact noise for the process. It was clear that the control of these transient vibrations travelling upstream and downstream from the shear would be the key to effective noise control.

The noise measurement procedure allowed the impact noise due to shearing to be separated from noise due from other sources such as material handling. The analysis was primarily based on the A-weighted L_{max} and L_{eq} for just the shearing part of the operation. It was found that the noise produced during a shear cut varied by several dB depending on the length of sheet being cut. Hence a standard product length of 5 m was used so that the effectiveness of noise control measures could be evaluated.

Noise Control Measures

Having established a suitable measurement procedure, a series of noise control measures were applied to the shear. The results in terms of A-weighted L_{max} and L_{eq} are shown in Figure 3.

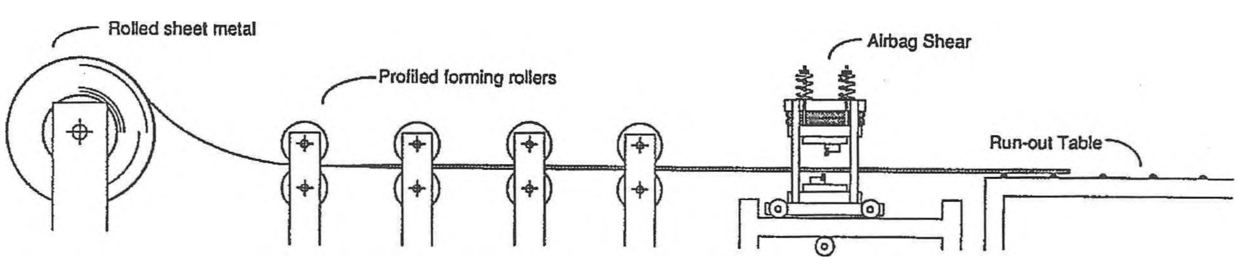


Figure 1 Roll-forming production line

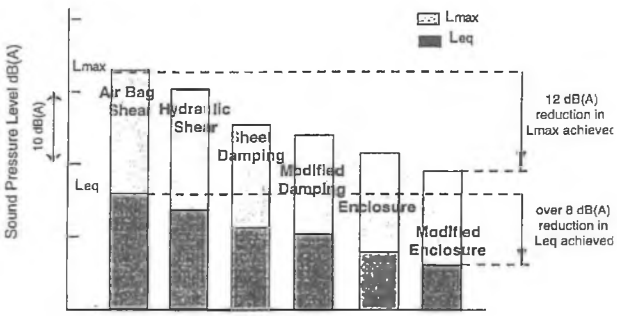


Figure 3 Results of noise control measures

Hydraulic Shear At other production facilities the pneumatic (air bag) shear operating mechanism had been replaced by hydraulic operation resulting in an improved control of the shear stroke. When this modification was applied to the shear under study approximately 2 to 3 dBA reductions in noise levels were achieved. These reductions were attributed to a better controlled and slower shear blade movement.

Sheet Damping To reduce the sheet vibrations induced by shearing, sheet dampers were developed. It was found that these dampers needed to be located as close as possible to the shear blades, say within 10 to 15 mm, in order to be effective. Polyurethane dampers were developed which were profiled to the shape of the sheets and operated in conjunction with the motion of the shear blades. As can be seen in Figure 3, these dampers, particularly after some modifications, resulted in reductions of 5 to 6 dBA.

Enclosure Following the above measures, an enclosure was built over the shear with entry and exit apertures for the product. Double glazed windows and a door gave the operator good visual and physical access to the shear. Heavy vinyl sheeting was used to block air-borne sound transmission at the entry and exit apertures. The further sound reduction achieved, 2 to 3 dBA, was beneficial but was far less than the 10 to 15 dBA, say, which might be expected from a full enclosure. This lack of effectiveness is due to two factors. Firstly it is not practical to completely seal the entry and exit apertures to air-borne noise. Secondly the transient vibrations readily travel along the sheet and out of the enclosure resulting in structure-borne impact noise being radiated by the sheet.

Modified Enclosure Powered rollers were installed at the exit aperture of the enclosure. As well as assisting in transporting the cut sheet out of the enclosure, the rollers reduced vibrations in the downstream portion of the sheet. This modification led to a further 2 to 3 dBA reduction in noise. In essence, structure-borne noise, due to vibrations travelling downstream along the sheet, was reduced by the rollers. (Similar measures were not applied at the entry side of the enclosure because the sheet was already well constrained there by the roll-former. Accelerometer measurements had shown that the vibrations were considerably larger in the sheet downstream of the shear. Once the cut has occurred, the cut sheet was virtually unconstrained.)

Following all these modifications, measurements of the noise exposures of plant operators were carried out under a variety of production circumstances. The exposures included noise from the shear plus other sources such as material handling noise. These measurements showed that the factory could be operated ear-muff-free under most production conditions.

Further research and development work[5] has studied improved blade profiles as a means to reduce shear noise. New blade designs have been developed which give a much smoother shearing action and have the potential for further noise reductions of the order of 6 dBA.

Conclusions

This case study shows how relatively low-cost retrofit noise reduction measures can be successfully applied to existing machinery through a progressive or evolutionary development process. In many instances this will be more practical and cost effective than replacing existing machinery with less noisy new equipment. In many specialised processes, there may be no obvious suppliers of less noisy equipment and a developmental approach may be preferred in order to control costs and avoid risks.

Essential steps in this approach are as follows.

- Establish the noise source and noise generating mechanisms
- Establish noise measurement techniques which can evaluate the effectiveness of retrofit measures
- Develop retrofit measures in collaboration with plant management, operators and maintenance personnel in order to ensure that measures are practical and cost-effective. At each stage of development, retrofit measures are evaluated for effectiveness.

Acknowledgements

This study was jointly carried out by the Acoustics and Vibration Unit at the Australian Defence Force Academy and by engineering staff at BHP Building Products. The authors express their appreciation for the support and encouragement of Worksafe Australia, BHP Building Products and the University of New South Wales.

References

1. Boyley, M., Burgess, M. A. and Williamson, H. M., Experiences in Implementing a Noise management Program in the Sheet Metal Industry, Occupational Injury Symposium, Sydney, Australia, February 1996.
2. Bahrami, A. and Williamson, H. M., *Sources of Noise in Sheet Metal Shears*, Australian Acoustical Society Annual Conference, Brisbane, December 1996.
3. Lai, J. C. S., Speakman, C. and Williamson, H. M., *Effectiveness of Retrofit Noise reduction Techniques Applied to a Roll-former Shear*, Inter-Noise 98, Christchurch, New Zealand, November 1998.
4. Burgess, M. A., Williamson, H. M. and Speakman, C., *Some Experiences with Noise Control for a Roll-Former Shear*, Inter-Noise 98, Christchurch, New Zealand, November 1998.
5. Bahrami, A., Williamson, H. M. and Lai, J. C. S., *Control of Shear Cutting Noise: Effect of Blade Profile*, Applied Acoustics, Vol. 54, No. 1. Pp. 45 – 58, 1998.

The Influence of Simple Room Geometries on Acoustical Parameters

John O'Keefe
Aercoustics Engineering Limited
Toronto, Ontario

Introduction

In most rooms the Early Decay Time (EDT) is shorter than the Reverberation Time (RT). This often means that the room does not sound reverberant enough. For acousticians, RT has been the predominant quantifier of sound since it was first developed by Sabine, at the beginning of this century. One reason for this is that it can be easily calculated. If you know the enclosed volume a room and the amount of acoustical absorption, you know the RT. Nothing could be easier!

Unfortunately, things aren't quite that simple. It turns out that the subjective significance of RT is not as important as was originally thought. The EDT is a measure of the early part of the decay, those few fractions of a second that separate musical notes. In the 1960s it was found to correlate much better with the subjective assessment of reverberation.[1] Unfortunately, the only way to predict EDT is with a computer or scale model.

Hypothesis

One possible explanation for the difference between EDT and RT has been suggested by Hodgson[2]. It came about during scale model experiments on the Queen Elizabeth Theatre (QET) in Vancouver. The EDT/RT ratio is very low in the QET. The QET is a bit different from other theatres in that it has a relatively low ceiling. Traditionally, opera houses or concert halls are high and narrow. The QET is flat and wide. When this was pointed out, Hodgson suggested that it might be the reason for its low EDT/RT ratio.

The hypothesis can be explained as follows:

1. A theatre or concert hall, in its simplest form, can be thought of as six sided box with acoustical absorption on only one of the six sides, i.e. the floor.
2. One might expect the early reflected sound (and hence the EDT) to be influenced by the sides of the box that are closest to each other. In a narrow shoe box shaped room, this would be the two non-absorbent side walls.
3. In a flat and wide room like the QET, the closest pair of sides is the ceiling and the floor. The latter, of course, is the only acoustically absorbent surface in the "box".

Experimental Procedure

To test this hypothesis, a number of experiments have been performed using computer models of six sided shoe box and fan shaped rooms. In all cases the floors and ceilings were flat, the acoustical absorption was limited to the floor and the rooms were 40 m in length. The height of the rooms was varied from 1/8th of the width to twice the width. Two room widths were tested, 20 m and 40 m. The angles of the fan shaped rooms were 8.5° and 16.7° for the 20 m and 40 m rooms respectively. Calculations were performed at five receiver locations in each of the four computer models. A single source location was used, situated at the front of the room, stage left of the centre line. The computer program employed for the experiments was CATT Acoustic Version 7. The method of images algorithm was set to 5th order with a truncation time of 300 ms and diffuse reflections commencing after the 1st order. The ray tracing algorithm was set to 12,000 rays and a truncation time of 6000 ms.

Early Decay Times

The computer model studies confirm the hypothesis. Please see Figure 1. For both the shoe box shaped and fan shaped rooms, the height to width ratio has a strong influence on the EDT/RT ratio. For height to width ratios greater than 1.0, the EDT/RT ratio is perfectly efficient, i.e. there is no compromise in Early Decay Time for a given Reverberation Time. If the height to width ratio is less than 1.0 there is a degradation of the Early Decay Time and hence the perceived reverberance in the room.

An interesting aspect of Figure 1 is that there appears to be no difference between shoebox and fan shaped rooms. This is intriguing because in most other acoustical aspects, the shoebox shaped format is superior to a fan shape. Indeed, measurements in existing auditoria have demonstrated low EDT/RT ratios in fan shaped rooms. The explanation for this apparent discrepancy is found in the geometry of the fan shaped format. Unlike our computer model, actual fan shaped auditoria have sloped floors and ceilings. The reality of the format is that the majority of seats are at the back of the room and, hence, closer to the ceiling. The effective height to width ratio for a fan shaped room is therefore quite low. For example, the Hummingbird Centre in Toronto has a height to width ratio of 0.24 (when measured in individual seats). Compare this to the 0.38 on the orchestra level of the QET (measured in individual seats) and 0.88 in the tall and narrow Musikvereinssaal in Vienna (gross average).

The Effect Of Balconies and Side Wall Boxes

These findings prompted three further computer model experiments. In the experiments, two levels of balconies and side wall boxes were introduced into the standard six sided shoebox room, i.e. 40 x 20 x 20 m (l-w-h). The balconies were 3 m deep and were wrapped around the two side walls and the wall opposite the stage.

In the first experiment, the vertical distances between the two balconies was varied from 3 to 7 m. Merely introducing these shallow balconies into the shoebox shaped model reduced the EDT/RT ratio by almost 30%. Contrary to received wisdom, the height between balconies had little influence on the measured acoustics. Parameters that were investigated included RT, EDT, Strength, 80ms Clarity, Early Lateral Fraction and the EDT/RT ratio.

In the second part of the balcony experiments, the importance of the facia height was examined. In the experiment the first balcony was 5 m above the orchestra level and second was another 5 m above that. The height of facia was varied from 0 m (i.e. no facia) to 4.5 m. It turns out that facia height may have a marginal effect on the EDT/RT ratio. The EDT/RT ratio is in the range of 65% for facia heights less than 1.0 metre. A larger facia, for example 2 metres or higher, results in a ratio of 70% to 74%, an improvement of almost ten percent. Difference limen for Reverberance are thought to be in the range of 0.1 seconds.[3]

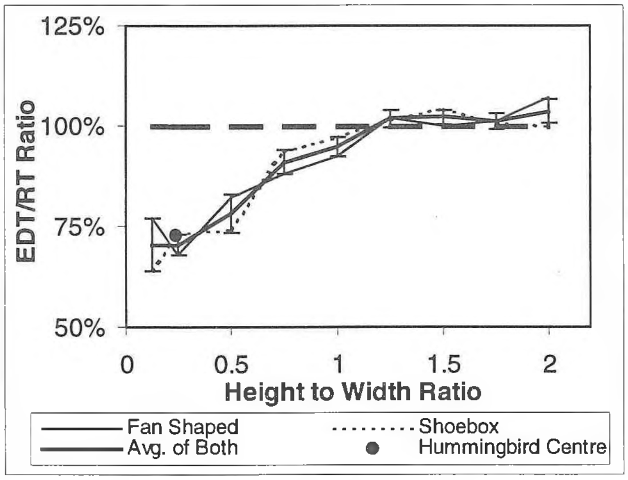


Figure 1 EDT/RT versus Height/Width Ratio.

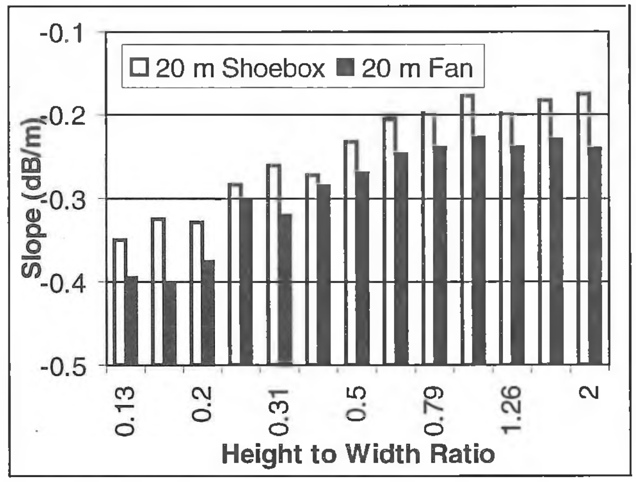


Figure 2 Slope of Strength versus Height/Width Ratio.

In the last of the three experiments in this series, the depth of the balcony overhang investigated. As before experiments were performed in both the 20 m and 40 m wide shoebox shaped rooms. In both cases the rooms were 20 m high and 40 m long. The depth of the overhang ranged from 1 to 8 m in the 20 m wide room and 2 to 16 m in the 40 m wide room. As expected, the EDT/RT ratio is reduced significantly as the overhang is increased.

Clarity

The balconies and boxes increase Clarity quite a bit. In the 20 m wide shoebox shaped room, Clarity is about 0 dB without the boxes. Introducing boxes on the side walls increases the Clarity by approximately 3 dB. The difference limen for Clarity is 0.67 dB[4]. A change in Clarity of 3 dB - more than four times the minimum noticeable difference - would surely be heard by audience members.

The explanation for the increased clarity proves interesting. Acoustical Clarity is a simple ratio of early to late reflected sound. One might expect that the reason for increased Clarity is because the side wall boxes provide stronger early sound to the listeners. The computer model study suggests otherwise. When balconies are introduced into the 40 m wide shoebox the strength of the early reflected sound remains essentially the same, regardless of the height of the balconies. In a 20 m wide room, the early energy goes up slightly, about 1.0 dB. However, in both rooms, the late reflected energy is reduced by approximately 3 dB when the balconies are added. Once again, the vertical distance between balconies does not appear to influence late reflected sound. In other words, contrary to expectations, Clarity is increased not by stronger early reflected sound but by weaker late reflected sound.

Strength

Measurements in the 1980s established that acoustic Strength decreases towards the back of a hall and that the rate of decrease is in the range of 0.1 to 0.2 dB/m[5,6]. Some room shapes, for example, a fan shape, tended to have higher rates of decrease.

These computer based experiments confirm that finding. Figure 2 shows the slope of Strength versus the Height to Width ratio predicted in the six sided boxes without balconies. The solid bars represent the fan shaped room and they can be seen to be consistently lower than the shoe-box shaped room. Note however that the Height to Width ratio of the room has a greater effect on Strength than its shape in Plan.

Wall and Floor Angles

Starting with a 40x20x20 m box (without balconies), the angle of the side walls and floor were varied systematically; in Plan and Section respectively. Increasing the side wall angle from a shoebox shape to a very broad 36° fan had the anticipated effect on parameters like Strength and Early Decay Time, both of which were lower than statistical calculations. The Early Lateral Fraction, of course, was found to decrease as the angle of the side walls increased. Interestingly, the EDT/RT ratio is effected more by the angle of the floor than by the angle of the side walls. This suggests again that the shape of a room in Section is as important as its shape in Plan.

Acknowledgements

Part of this work was supported by the Industrial Research Assistance Programme, National Research Council of Canada

References

[1] Atal, B.S., Schroeder, M..R. & Sessler, G.M., "Subjective Reverberation Time and its Relation to Sound Decay", Proc. 5th International Congress on Acoustics, Liege, Paper G32

[2] Hodgson, Murray, Private Communication

[3] Cremer, L., Muller, H.A. & Schultz, T.J. *Principles and Applications of Room Acoustics*, Applied Science, 1978 p. 505

[4] Cox, T.J., Davies, W.J. and Lam, Y.W. "The Sensitivity of Listeners to Early Sound Field Changes in Auditoria" Acustica 79, pp 27-41 (1993)

[5] Barron, M., Lee, L.J. 'Energy relations in concert auditoriums. Part I', J. Acoust. Soc. Am. 84, pp 618-628 (1988)

[6] Bradley, J.S. 'Hall Average Conditions of 10 Halls', Proc. 13th Int. Congress on Acoustics 2 pp 199-02, Belgrade (1989)

Standardized prediction methods for estimating sound insulation in buildings

Trevor Nightingale, Institute for Research in Construction, National Research Council Canada, K1A 0R6
Thomas Alber, Fachhochschule Stuttgart-Hochschule für Technik, Department of Building Physics, 70174 Stuttgart Germany

Introduction

In Europe, "objective" or "performance" based building codes have spurred the development of standardized methods to predict the sound insulation of buildings. These models are reviewed and their suitability for North American constructions discussed. In this summary we retain the terminology used in the standards and note that "sound reduction" is synonymous with "transmission loss".

Using either measured sound reduction data or information on the material properties, a model should be able to predict "apparent sound reduction" which determines the subjective perception of sound privacy. The apparent sound reduction, R', is the sum of all the transmission paths and includes the direct transmission through the nominally separating wall or floor as well as indirect transmission paths or flanking paths. Figure 1 shows some of the common flanking paths that can exist. In the following sections the ability of three European models to predict the sound reduction of these paths is discussed.

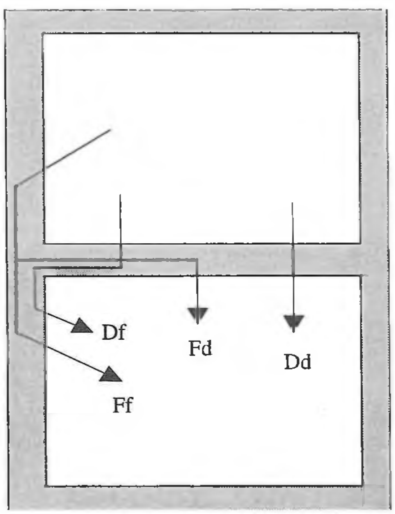


Figure 1: Direct (Dd) and possible flanking transmission paths for a joint.

CEN method

The CEN model (prEN 12354-1) was published as a draft standard in April 1996. It is based on the principle of energy balance[1] and enables calculation of sound insulation of all flanking paths that involve only a single joint. This means that there can be 12 flanking paths and one direct path between rooms separated by a floor or a wall. Figure 1 shows the flanking paths for a single junction. The sound reduction of the individual flanking paths can be calculated using,

$$R_{ij} = \frac{R_i + R_j}{2} + K_{ij} + 10\lg\frac{S_{Partition}}{l_0 \cdot l_{ij}} \qquad (1)$$

where the subscript i indicates the source surface and j the receive surface, R is the one-third octave band sound reduction, K_{ij} is the vibration reduction index describing the joint attenuation between plates i and j, S is the area of the partition and l is the joint length.

Values of the joint vibration reduction index, K, can be obtained from Annex E of the standard which lists seven joint details and provides an empirical value based on measured data. Alternately, the K_{ij} will have to be computed from measured quantities using,

$$K_{ij} = \frac{D_{v,ij} + D_{v,ji}}{2} + 10\log\frac{l_{ij}}{\sqrt{a_i a_j}} \qquad (2)$$

where D_v is the velocity level difference and $a_{i,j}$ is the equivalent absorption length of element. The apparent sound reduction index R' for the assembly is determined by the sum of the direct path plus all flanking paths which is given by:

$$R' = -10\log\left[10^{-R_{Dd,w}/10} + \sum_{F=f=1}^{n} 10^{-R_{Ff,w}/10} + \sum_{f=1}^{n} 10^{-R_{Df,w}/10} + \sum_{F=1}^{n} 10^{R_{Fd,w}/10}\right] \qquad (3)$$

Thus, in theory, the CEN model provides a reasonably general description of flanking transmission between two rooms. However it is restricted to paths involving only a single joint.

Normally, equations 1 and 2 are evaluated for all possible paths and at all one-third octave band frequencies of interest and inserted in equation 3 to obtain the apparent one-third octave sound reduction. The standard refers to this as the "detailed model." A "simplified model" is also given that uses only single number ratings for the sound reduction. The simplified model has been widely used in heavy, monolithic constructions and typically exhibits good agreement with measured results. Unfortunately, the simplified method is only valid for heavy monolithic constructions where the single number ratings are determined by resonant transmission through the assembly (i.e., the critical frequency is at or below the lowest frequency of interest). In North American buildings employing double-leaf construction, the single number ratings are often determined by transmission in the 125 and 160 Hz one-third octave bands, namely non-resonant transmission, since the critical frequency for lightweight building materials (OSB, plywood and gypsum board) is usually in the range 1600 and 3150 Hz. Consequently, the CEN "simplified model" is not appropriate for wood frame constructions and will not be considered further.

DIN method

The German standard DIN 4109 includes both a prediction model for heavy, monolithic constructions in "Beiblatt 1" as well as a calculation procedure for wood-frame linings that may be used as partition walls in heavy monolithic buildings.

Unlike the CEN model, the DIN method considers only the flanking path Ff (Figure 1) and assumes that paths Df and Fd that involve the partition wall are insignificant. This assumption may be a reasonable approximation for junctions between lightweight interior partition walls and heavy monolithic concrete floors and exterior walls. In this case, joint attenuation between a lightweight wall and a monolithic wall (Df and Fd) will be very large with respect to the almost negligible attenuation for the monolithic element (Ff). This means that the velocity level difference between plates F and f will be near zero and leads to a very simplified model given by,

$$R'_{L,w,R,i} = R_{L,w,R,i} + 10\lg\frac{S_T}{S_0} - 10\lg\frac{l_i}{l_0}\,dB \qquad (4)$$

where $R'_{L,w,R,i}$ is the sound reduction for the flanking path Ff expressed as a single number rating and $R_{L,w,R,i}$ is the sound reduction for the flanking path obtained from a listing in the standard (again a single number rating), S_T is the area of the partition element, S_0 is the reference area (for walls $S_0 = 10m^2$),

l_l is the common length of the partition and flanking element and l_0 is the reference length (for floors and ceilings $l_0 = 4{,}5$m; for walls $l_0 = 2{,}8$m). This is in essence the same result that one would get if the CEN K_{ij} were set to zero (which would happen if walls F and f were identical and very massive with respect to wall D).

In lightweight wood frame constructions, it is not likely that paths Fd and Df will be insignificant with respect to Ff. This may be viewed as significant impediment if the DIN standard where applied to wood frame constructions. The second impediment is that it assumes all flanking paths past a floor/ceiling assembly have a sound reduction of 65 dB.

ÖNorm method

The ÖNORM 8115 is the national standard of Austria for the prediction of the sound insulation offered by heavy masonry or cast-in-place constructions. This model will be of limited use for North American wood frame constructions due to the very different behaviour of the constructions.

Application of the CEN Model to a Wood Frame Construction

Two of the more serious difficulties of the CEN model are discussed by examining the flanking sound reduction for the paths from the floor and party wall in Room B to the party wall of room D of the assembly shown in Figure 2. Figure 3 shows the measured and predicted K_{ij}'s for these paths. The measured K_{ij}'s were calculated from measured velocity level differences and structural decay times while the predictions were obtained from the listed assemblies in Annex E of the standard.

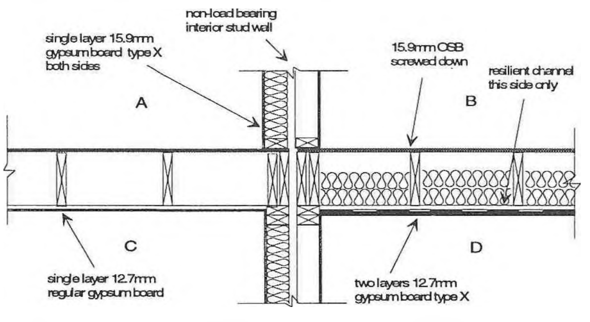


Figure 2: Section through a potentially common wood frame construction.

Figure 3 indicates that Annex using either equation E7 or E8 would grossly underestimate the vibration reduction index for these paths. The error could be in excess of 35 dB. Equation 1 shows that the underestimation in K will cause a corresponding underestimation in the sound reduction for the flanking path. Thus, it is possible to underestimate the flanking sound reduction by as much as 35 dB. Fortunately model codes in Canada have standardized building material dimensions and practices making it possible to create a catalogue of common joint details and Kij's for wood frame constructions.

Another difficulty occurs in determining the correct value of the sound reduction for the building elements involved in the flanking paths. It has been shown[2] that only the resonant component of the sound reduction should be used in equation 1. An estimate of the error that can occur if the sound reduction contains non-resonant transmission can be seen in Figure 4 by comparing the predicted resonant and non-resonant transmission for the OSB subfloor. It can also be seen that ASTM E90 or E336 data can not be used as transmission is dominated by non-resonant transmission below the critical frequency (about 2000 Hz) and would also lead to a significant underestimation of the flanking sound reduction. Computing the resonant components of the OSB or gypsum board surfaces involved in the flanking paths between Rooms B and D is quite simple. However, computing the resonant transmission through the party wall is very difficult and would be required in determining the sound reduction of the flanking paths between Rooms A and B. The standard does not give a method for predicting resonant transmission through a wall.

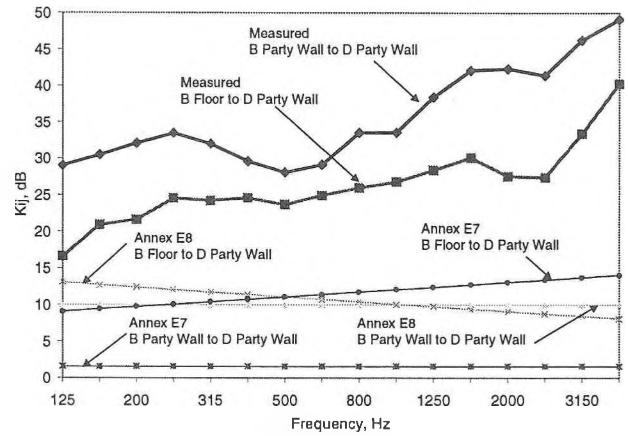


Figure 3: K_{ij}'s obtained from measured results and from Annex E of the CEN standard.

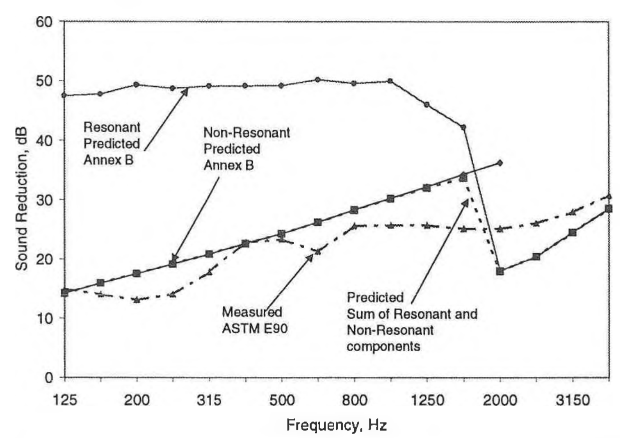


Figure 4: Sound reduction data obtained by measurement and calculation according to Annex B for the OSB floor decking.

Conclusions

Wood frame constructions are considerably more difficult to model than the heavy monolithic assemblies commonly found in Europe. Consequently, the three prediction models examined would be of limited use in their present form. The CEN model is perhaps best equipped to be adapted for wood frame construction but this will require considerable effort.

[1] Gerretsen, E., "Calculation of Airborne and Impact Sound Insulation between dwellings," Applied Acoustics **19** (1986) 245-264.

[2] Nightingale, T.R.T., "Application of the CEN draft building acoustics model to a lightweight double leaf construction," Applied Acoustics **46** (1995) 265-284.

Telephony acoustics at Mitel

Philippe Moquin
Mitel Corp., Kanata, ON

1. Introduction

This is an overview of Mitel's acoustical research and development in the desktop terminals group. Telephones and operator consoles are the primary products designed by the group.

2. Handset performance

Basic performance of the handset is strongly standardised however there are a few areas that still affect the sound quality: wind noise and receiver real ear performance.

The microphone wind noise during explosive consonants proved to be a significant problem on a new handset and due to time constraints a co-operative research project with Dr. Stinson of NRC was set up. Results from the research yield an interim design that significantly improved the performance. Further research provided us an opportunity to reduce the wind noise to a level that is among the lowest in the industry. An accurate analytical model was also developed. Having a good physical understanding of the noise generation mechanism proved to be very valuable in the redesign.

In standardised tests the artificial ears used accurately replicate a handset perfectly sealed to the handset. Unfortunately few people hold a handset that tightly to their ear. The air leak between the ear and the handset creates a significant loss of low frequency sensitivity. The majority of existing receivers are designed for the sealed ear condition. However, we anticipate that, with the advent of high quality personal music systems, clients will soon demand better performing receivers.

3. Speakerphones

All of Mitel's telephones are made in Kanata and feature high quality speakerphones. Three years ago Mitel had no acoustical modelling capability. It was decided that a university research project could provide that capability. The project has started and we are sponsoring a doctoral candidate under Dr. Laville's direction at l'école de Technologie Supérieure.

We are looking to model the transducer frequency response in the plastic housings and microphone to speaker acoustic separation. The approach taken is rather novel as it integrates analytical, empirical and numerical methods to provide a computationally efficient and reasonably accurate model.

High quality speakerphones today necessarily imply full duplex operation. This unfortunately brings up the "barrel effect" where the received signal is retransmitted to the far end talker providing an annoying echo.

The present preferred solution is to use an adaptive FIR filter that mimics the room impulse response. This has some significant draw backs.

Firstly, there is the computational requirements to model a reasonably live room. There is much work being done to develop more efficient algorithms.

Secondly, there is the finite performance possible with large FIR filters. New approaches to solve the problem are needed that do not have the performance limitations inherent in conventional NLMS FIR algorithms. This requires a good understanding of both telephony and acoustics of rooms.

Thirdly, this in no way addresses the microphone to talker distance. Methods for reducing the effective talker to microphone have to be developed. We feel that microphone arrays such as those described by Drs. Ryan and Stinson will be useful.

Fourthly, these algorithms assume absolute linearity. Non-linearity due to button rattle and other distortions can be controlled by physical design but some non-linearity's would be most useful.

4. Speech recognition

Mitel is not proposing to do significant research in speech recognition since we believe that we will be able to licence this type of technology. However, the acoustical design of telephones will have to accommodate speech recognition. Obviously the challenge is greater in the speakerphone mode. The primary problem is that of providing the best signal to noise for the talker signal. Again the belief is that directional microphones or arrays will provide significant performance enhancement.

5. Voice over IP

The problem with transmitting voice or audio visual stream over computer network is that you don't have a dedicated channel as one has in a telephone network. Packets are sent out serially as in the telephone network but the time to get to the other end is not known. The packet may never get there, it may arrive late and even if they do get the in order the timing between one packet and the next may vary significantly.

Presently, there are only theories as to the best way to deal with these problems. Large buffers can be used but this means long time delays which can inhibit communication. Lost or late packets can be replaced by silence or noise but this may not be acceptable to users of wired telephones while it works fine in the wireless world. Alternative coding schemes with built in redundancy could help but then one is locked into one vendor's solution.

We believe that using conventional technology with a good understanding of speech perception creative solutions providing toll quality voice over IP networks is achievable. This area obviously requires significant research effort.

On the use of the FNTF algorithm in subband acoustic echo cancellation

Benoît Champagne and François Caron
INRS-Télécommunications, Université du Québec, 16 place du Commerce
Verdun, Québec, Canada, H3E 1H6; champagne@inrs-telecom.uquebec.ca

1. Introduction

Acoustic echo cancellation (AEC) is an effective approach for the control of acoustic echoes generated by hands-free terminals [1]. Even though RLS adaptive filtering achieves faster convergence, the NLMS algorithm is used in most AEC applications because of its low complexity of $2N$ multiplies per iteration (mpi), where N is the number of filter taps. The identification of the long echo paths found in teleconference applications (e.g. $N = 1000$ at 8kHz sampling rate) further requires the use of subband processing to bring down the computational complexity within acceptable limits [2].

Recently, the fast Newton transversal filter (FNTF) algorithms have been proposed in an attempt to bridge the performance gap between the NLMS and the fast RLS (FRLS) algorithms [3]. FNTF models the excitation signal as an AR(M) process, where $0 \leq M \leq N$, and achieves a complexity of $2N + 12M$. In [4], FNTF is shown to be an attractive candidate for AEC applications in the mobile context (short filters, $N = 250$ at 8kHz sampling) since significant improvements in convergence speed over NLMS can be obtained with small values of M. This conclusion does not hold for teleconference applications, where FNTF may lead to a loss of performance.

The investigation in [4] is limited to the use of a single (i.e. full-band) transversal adaptive filter. In the case of long impulse responses, FNTF might benefit from a subband implementation because of the reduced adaptive filter length in subbands, as a result of downsampling. In this work, we investigate the performance of FNTF in a subband AEC structure, with emphasis on the identification of long echo paths typical of teleconference applications.

2. The FNTF algorithm

Application of adaptive system identification to AEC is illustrated in Fig. 1. The unknown system $\mathcal{H}$ consists of the loudspeaker, acoustic medium and microphone. The system input is the far-end signal $u(n)$, where $n \in \{1, 2, ...\}$ is the discrete-time, and the system output is the microphone signal $d(n)$, which contains additive noise and possibly local speech. The unknown system is modeled by an adaptive transversal FIR filter operating on $u(n)$. The time-varying coefficients of the FIR filter are denoted by $h_k(n)$, $k = 0, 1, ..., N-1$, and the filter output is computed as $\hat{d}(n) = \mathbf{h}(n)^T\mathbf{u}(n)$, where $\mathbf{h}(n) = [h_0(n), ..., h_{N-1}(n)]^T$ and $\mathbf{u}(n) = [u(n), u(n-1), ..., u(n-N+1)]^T$. The filter weight vector is recursively adjusted in real-time so as to minimize the power of the error signal, defined as $e(n) = d(n) - \hat{d}(n)$. Practical operation of the adaptive filter requires the use of a double-talk detector (not considered in this study).

The FNTF algorithms belong to a modified class of stochastic Newton (SN) adaptive algorithms:

$$\mathbf{c}_N(n) = -\lambda^{-1}R_N^{-1}(n-1)\mathbf{u}(n), \quad \gamma(n) = 1 - \mathbf{c}_N^T(n)\mathbf{u}(n) \quad (1)$$

$$e(n) = d(n) - \mathbf{h}^T(n)\mathbf{u}(n), \quad \epsilon(n) = e(n)/\gamma(n) \quad (2)$$

$$\mathbf{h}(n+1) = \mathbf{h}(n) - \epsilon(n)\mathbf{c}_N(n) \quad (3)$$

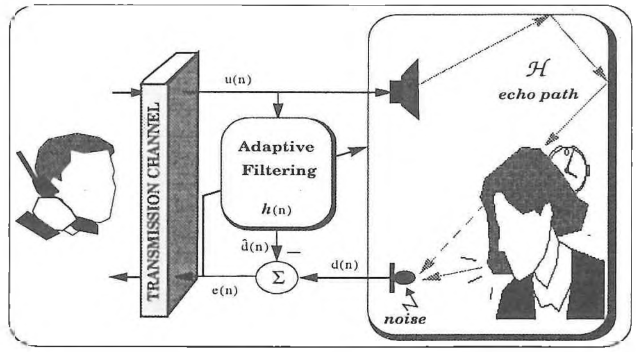


Figure 1: Adaptive identification applied to AEC

where $R_N(n)$ is an estimate of the data covariance matrix, $0 < \lambda \leq 1$ is a forgetting factor, $\mathbf{c}_N(n)$ is a generalized dual Kalman gain, $e(n)$ is the *a priori* estimation error, $\epsilon(n)$ is the *a posteriori* error, and $\gamma(n)$ is a conversion factor. Both NLMS and FRLS can be obtained from (1)-(3) with a proper choice of $R_N(n)$ in (1). In [3], an additional AR(M) assumption on $u(n)$, where $0 \leq M \leq N$, is exploited to derive time-order recursions for the extension of an $(M+1)$th order covariance matrix into the desired N-order one, i.e. $R_N(n)$. Upon substitution of these extension formulae in (1), three distinct FNTF versions are obtained with complexity $2N + O(M)$.

In the context of AEC, practical considerations point to the use of FNTF Version 1 [4]. The latter, used in our work, is summarized below: (a) Using a FRLS forward predictor of order M applied to $u(n)$, update the forward predictor weight vector $\mathbf{a}_M(n-1)$, the residual error $e_M^f(n)$ and the error power $\alpha_M^f(n-1)$ and compute

$$\mathbf{s}_{M+1}(n) = \frac{e_M^f(n)}{\lambda\alpha_M^f(n-1)}\begin{bmatrix} 1 \\ -\mathbf{a}_M(n-1) \end{bmatrix} \quad (4)$$

(b) Using a FRLS backward predictor of order M applied to $u(n_d)$, where $n_d = n - N + M$, update the backward weight vector $\mathbf{b}_M(n_d - 1)$, the residual error $e_M^b(n_d)$ and the error power $\alpha_M^b(n_d - 1)$, and compute

$$\mathbf{t}_{M+1}(n_d) = \frac{e_M^b(n_d)}{\lambda\alpha_M^b(n_d-1)}\begin{bmatrix} -\mathbf{b}_M(n_d-1) \\ 1 \end{bmatrix} \quad (5)$$

(c) Update the dual Kalman gain $\mathbf{c}_N(n)$ and $\gamma(n)$:

$$\begin{bmatrix} \mathbf{c}_N(n) \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{c}_N(n-1) \end{bmatrix} - \begin{bmatrix} \mathbf{s}_{M+1}(n) \\ \mathbf{0}_{N-M} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{N-M} \\ \mathbf{t}_{M+1}(n_d) \end{bmatrix} \quad (6)$$

$$\gamma(n) = \gamma(n-1) + s_{M+1}^1(n)e_M^f(n) - t_{M+1}^{M+1}(n_d)e_M^b(n_d) \quad (7)$$

(c) Filtering part: Same as (2)-(3) above.

To define initial conditions for this algorithm, a soft constraint approach is described in [3]. Assuming that two distinct FAEST algorithms [5] are used in steps (a) and (b), the total complexity of FNTF is $2N + 12M$.

3. Weaver SSB subband structure

Subband adaptive filtering offers several advantages over a conventional full-band approach, including reduction of computational complexity and improved signal conditioning [2].

---

[1] Support for this work was provided by FCAR.

Design of analysis/synthesis (A/S) filter banks for a subband AEC system is a complex problem involving several trade-offs/requirements: near-perfect reconstruction, low processing delays, oversampling in the subbands, low complexity. For ease of implementation, it is also desirable that the subband signals be real. Based on these considerations, we have found it convenient to use Weaver SSB A/S banks in our subband filtering structure; the latter is illustrated in Fig. 2.

The loudspeaker signal $u(n)$ and the microphone signal $d(n)$, with sampling rate F_s, are each decomposed into B real subband signals by analysis filter banks based on Weaver SSB modulators. Each bank consists of B band-pass filters followed by decimators by $K \leq B$. The corresponding subband signals are denoted by $u_b(m)$ and $d_b(m)$, where $b = 0, ..., B-1$ and m denotes sampling-time at the lower rate $F_s' = F_s/K$. In each subband, an FNTF adaptive algorithm operating at the reduced rate F_s' is used to identify the corresponding subband component of the echo path. The subband error signals $e_b(m)$ are finally recombined by a dual synthesis bank to produce a full-band error signal $e(n)$, at the original rate F_s.

In our implementation, the digital spectrum $[-\pi, \pi]$ is divided into B real subbands, with bandwidth $\omega_\delta = \pi/B$. The center frequency of the bth subband (positive sideband) is given by $\omega_b = (b + \frac{1}{2})\omega_\delta$, $b = 0, 1, ..., B-1$. Each narrow-band filter in the analysis bank is a Weaver modulator with center frequency ω_b; a corresponding Weaver demodulator is used in the synthesis bank (see [6] for details). In theory, subband aliasing may be avoided with critical downsampling, i.e. $K = B$, provided an ideal low-pass filter $h(n)$ with cut-off $\omega_\delta/2$ is used. In practice, because of non-ideal filters, oversampling is necessary, i.e. $K < B$; we chose $K = B/2$. A window technique is used to design $h(n)$. For $B = 16$, the resulting A/S bank has the following properties: processing delay of 16ms, amplitude distorsion within $\pm 0.05dB$, linear phase.

4. Results and discussions

The sampling rate is set to $F_s = 8$kHz. Different loudspeaker signals $u(n)$ are used in the experiments. Here, we show results for the composite source signal (CSS), a speech-like signal [7]. The echo path $\mathcal{H}$ in Fig. 1 is simulated with a synthetic room impulse response of duration 2048 samples. To produce the microphone signal, this response is convolved with $u(n)$ and independent white noise is added. The convergence performance is evaluated in terms of the short term power of $e(n)$ (32ms window, dB relative to the echo level). The parameter values (N, M, λ) are the same for all the subband FNTFs. Following [4]: λ is set to a large value ($\lambda = 1 - 1/2N$), an acceleration mechanism is used in (3) and the filtering part of FNTF is initially frozen (for 0.5s). Note that for $M = 0$, FNTF corresponds to a modified form of NLMS, while in the case $M = N$, FNTF corresponds to FAEST (FRLS family).

Fig. 3 shows convergence curves of the full-band FNTF for $N = 2048$ (256ms), SNR= $30dB$ and different M. The top and bottom curve represent the microphone signal $d(n)$ and the additive noise. It is seen that a large value of M (≥ 256), is needed to obtain a performance comparable to RLS. Fig. 4 shows convergence curves of the subband FNTF for $N = 256$ (256ms), SNR= $30dB$ and different M. Again, a relatively large value of M (about 128), is required to achieve RLS-like performance, despite the fact that in subbands, the effective prediction order of $u(n)$ is smaller due to decimation.

Our study points to the following general conclusions: (1) the selection of the parameter M in FNTF can not be based strictly on the necessary AR modeling order of the source signal $u(n)$; (2) the use of subband processing is not effective in reducing the ratio M/N necessary for efficient operation of FNTF with long filters; (3) subband FNTF appears to be of limited practical value for AEC in teleconference applications.

References

[1] E. Hansler, *Signal Processing*, vol. 27, pp. 259–271, 1992.
[2] A. Gilloire, *Proc. IEEE Int. Conf. ASSP*, Dallas, Texas, 1987, pp. 2141–2144.
[3] G. V. Moustakides *et al.*, *IEEE Trans. on SP*, vol. 39, pp. 2184–2193, Oct. 1991.
[4] T. Petillon *et al.*, *IEEE Trans. on SP*, vol. 42, pp. 509–518, March 1994.
[5] G. Carayannis *et al.*, *IEEE Trans. ASSP*, vol. 31, pp. 1394–1402, Dec. 1983.
[6] R. E. Crochiere *et al.*, *Multirate Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1983.
[7] H. Gierlich, *Signal Processing*, vol. 27, pp. 281–300, 1992

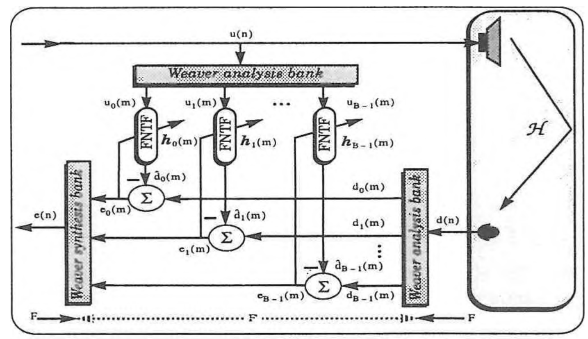


Figure 2: Weaver SSB subband structure for AEC.

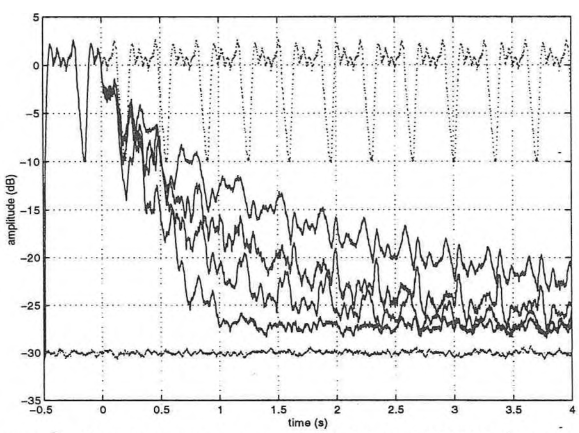


Figure 3: Convergence curves for fullband FNTF (top to bottom: echo; FNTF with M=0,16,256,2048; noise).

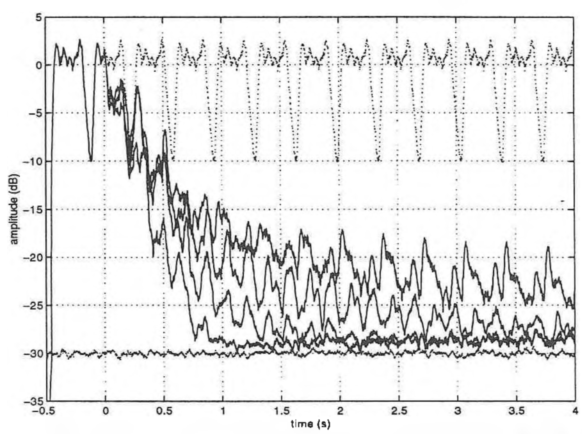


Figure 4: Convergence curves for subband FNTF (top to bottom: echo, FNTF with M = 0,8,128,256; noise).

OPPORTUNITIES FOR ACTIVE NOISE CONTROL IN COMMUNICATION HEADSETS

A. J. Brammer and G. J. Pan

Institute for Microstructural Sciences, National Research Council, Ottawa, Ontario K1A 0R6

1. Introduction

Almost all the current generation of commercial circumaural headsets incorporating active noise reduction (ANR) employ feedback control, with fixed filters and analogue signal processing. This combination provides, at best, a conditionally stable device. Good performance can be obtained at the lower speech frequencies, with some headsets reducing environmental noise at the ear at frequencies of up to 1 kHz. Inconsistent control of environmental noise results, however, when air leaks occur in the seal between the earmuff and the head.[1] These can occur, for example, with poor fitting of the device, or head rotation when wearing the headset.

The variation in coupling between the ear and the secondary source located within the ear cup when the device is displaced relative to the ear suggests an adaptive control system is required that is capable of optimizing its performance while the headset is used.

The signal processing algorithm and device that best meet the requirements of a communication headset remain to be established. The purpose of this paper is to describe the potential for improved speech communication and reduced noise-induced hearing loss in an ANR headset by employing adaptive feedforward control. Details of the ANR produced by the device have been described elsewhere [2,3].

2. Adaptive Feedforward ANR Headset

Earmuff. The earmuff consists of a rigid ear cup lined with sound absorbing material, together with a circumaural cushion that provides an acoustic seal to the head. A microphone is attached to the outside of the muff, senses the sound field surrounding the ANR headset and provides a reference input signal X to the digital controller (see Fig. 1). A miniature loudspeaker is used to generate the secondary sound field in the volume enclosed by the earmuff, and is driven by signal U, which is derived from signal X. A communication signal may also be reproduced by the loudspeaker. A second microphone, located close to the ear canal entrance, provides the error signal, E, used by the algorithm to optimize the control filter.

Control System. A block diagram of the control system is shown in Fig. 1. Environmental noise is transmitted through the acoustic plant, consisting of the earmuff with air leaks, to become the primary noise in the volume enclosed by the muff. Adaptation of the control filter is performed using the filtered-X LMS algorithm.

A dual-rate sampling structure is used to provide short control system latency.[3] This involves over-sampling the acoustic signals at the reference and error microphones, and updating the control signal to the secondary source at a decimated rate of the sampling frequency. The technique effectively reduces the total signal delay in the control path and, at the same time, permits the low-frequency performance of the digital control filter to be improved.

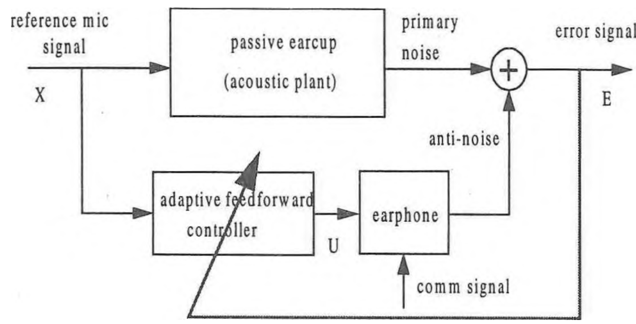


Fig. 1: Control structure.

A separate control system has been constructed for each ear using TMS320C32 floating-point digital signal processors and 16-bit A/D and D/A converters.

3. Speech Communication

In contrast to feedback control, a feedforward approach does not derive the control signal directly from the sound in the volume enclosed by the earmuff (i.e., from the error microphone), which contains both the undesired environmental noise and speech reproduced by the secondary source. In the former approach, the speech component sensed by the microphone close to the ear must be removed from the control signal, which can be an imprecise process with a tendency to introduce distortion. It is evident from the feedforward control structure (Fig. 1) that only the component of the sound under the earmuff correlated with the reference signal can be reduced. The extent to which reproduction of a speech signal by the secondary source at the same time as it is controlling environmental noise degrades ANR performance has been demonstrated in the noise environment of a Leopard tank. For this experiment, the subject was immersed in a simulated diffuse sound field. The ANR systems were configured to control environmental noise at frequencies below 400 Hz, with input/output signals digitized at a frequency of 33 kHz, a control frequency of 3 kHz, and an adaptive filter containing 400 coefficients.

While the control systems were operating, a pre-recorded speech signal was fed into the ANR system for the right ear, and replayed by the miniature loudspeaker at a level that could be clearly understood by the subject. The speech signal was not introduced to the ANR system for the left ear, which consequently continued to function as an unperturbed active noise control system. The speech consisted of a male voice repeating sentences with a brief pause between each sentence. The pause was not of sufficient duration for the active control system to re-adapt to the sound under the earmuff without speech, but was long enough for the ANR to be measured.

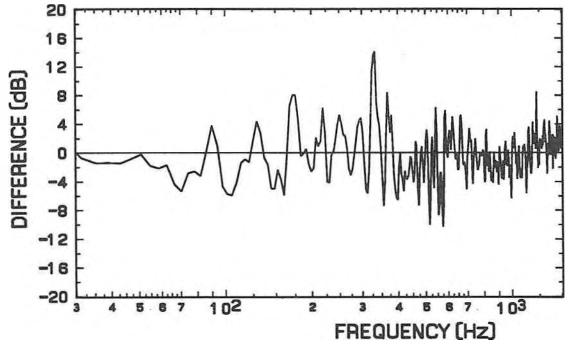


Fig. 2: Difference in ANR.

The difference between the ANR recorded at the error microphones for the left and right ears under these conditions is shown in Fig. 2. It is evident from the difference spectrum recorded with the two ANR systems operating [(right ANR, with speech) - (left ANR, no speech)] that the noise reduction displays little systematic degradation with the presence of the communication signal (i.e., the difference is not generally negative). There is an indication of a minor reduction in ANR at frequencies below 80 Hz, and a cyclical variation in ANR with frequency which may reflect the tonal nature of the noise source. Thus, the expected lack of dependence of the ANR performance on the presence, or absence, of a communication signal is confirmed for this adaptive controller. In addition, the feedforward control structure does not perturb the communication signal, and so offers the potential for higher fidelity reproduction.

4. Intelligibility and Hearing Loss

It is well known that the risk of noise-induced hearing loss is related to the sound level at the ear, when the latter is expressed in terms of the A-weighted sound pressure level. It is also possible to relate speech intelligibility to the speech signal to noise ratio, when both are expressed as A-weighted sound pressure levels.[4] In view of these considerations, the possibility of operating the adaptive feedforward ANR headset with a frequency-dependent target convergence function has been explored.

A critical feature of this application of feedforward control is the short distance between the reference and error microphones, and the correspondingly short time delay in which the controller must function to maintain causality. Once this condition has been achieved, as in the present feedforward system,[2] introducing additional phase shift with a frequency-dependent target convergence function is possible. In contrast, feedback control systems do not lend themselves to the introduction of additional phase shifts.

A demonstration of the potential for controlling band-limited white noise (150-700 Hz) so as to produce an A-weighted noise spectrum at the ear has been conducted by introducing an analogue filter in the error signal path with approximately this frequency response. The experiment employed one earmuff mounted on a flat-plate coupler in a small enclosure.[5] The control system employed an input/output frequency of 40 kHz, a control frequency of 10 kHz, and an adaptive FIR filter containing 200 coefficients.

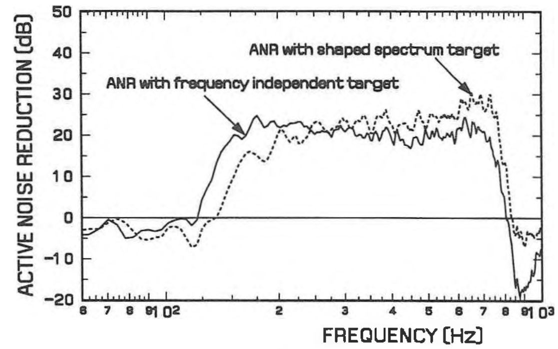


Fig. 3: ANR with different target functions.

The difference between the ANR recorded at the error microphone when the system was operating with, and without, a frequency-selective target convergence function can be seen from Fig. 3. It is evident from these results that the ANR is not compromised by selecting the A-weighted frequency network as the target convergence function. Thus it is believed that the adaptive feedforward ANR headset may be operated at close to optimum for maintaining speech intelligibility and the preservation of hearing, at least for environmental noise at frequencies above 150 Hz.

5. Acknowledgment

This work was done in collaboration with the Defence and Civil Institute of Environmental Medicine, Toronto.

6. References

[1] R.B. Crabtree, "Constraints in the application of personal active noise reduction systems", in *Audio Effectiveness in Aviation*, AGARD-CP-596, 15.1-15.6 (1997).

[2] G.J. Pan, A.J. Brammer, J. Zera, and R. Goubran, "Application of adaptive feedforward active noise control to a circumaural hearing protector," *Proc. Active 95, Newport Beach*, 1319-1326 (1995).

[3] A.J. Brammer, G.J. Pan, and R.B. Crabtree, "Adaptive feedforward active noise reduction headset for low-frequency noise," *Proc. Active 97, Budapest*, 365-372 (1997).

[4] International Organization for Standardization, "Acoustics - The construction and calibration of speech intelligibility tests," ISO/TR 4870-1991(E), 1-21 (1991).

[5] J.G. Ryan, E.A.G. Shaw, A.J. Brammer, and G. Zhang, "Enclosure for low-frequency attenuation assessment of active noise reducing circumaural headsets and hearing protectors," *Canadian Acoustics*, **21**, No. 4 19-20 (1993).

SOUND PICKUP IN THE PRESENCE OF DIFFRACTION EFFECTS

Michael R. Stinson, James G. Ryan and Gilles Daigle

Institute for Microstructural Sciences
National Research Council
Ottawa, Ontario K1A 0R6, Canada

INTRODUCTION

Microphones are an essential part of telephony. With the growing popularity of teleconferencing, videoconferencing, and handsfree telephony, the appropriate placement of microphones in a product can be important: the effects of scattering and diffraction of sound can lead to significant spatial variations of the sound pressure level (SPL). Furthermore, for many applications, a directional microphone response is required to reduce the effects of room reverberation and background noise. Whether directional microphones with two ports on the surface of the device or an array of several microphones are used to achieve this directionality, the placement of the microphones and allowance for diffraction effects is even more important.

An example of a situation where spatial effects can be important is shown in Fig. 1. We consider an array of microphones built into the top frame of a computer monitor. The use of multiple microphones permits the preferential pickup of speech signals from a zone located at the position of the user's head, significantly reducing the auditory effects of ambient noise and room reverberation.

Consider first a single microphone. Because of the effects of scattering of incident sound and diffraction there will be spatial variations of the sound pressure.[1] Different locations of a microphone will lead to different signals. These physical effects depend strongly on frequency. Hence, coloration of the frequency response curves will occur. Some locations may be more sensitive than others. For the example shown, the distance between the top edge of the monitor and the microphone position should not be selected too casually.

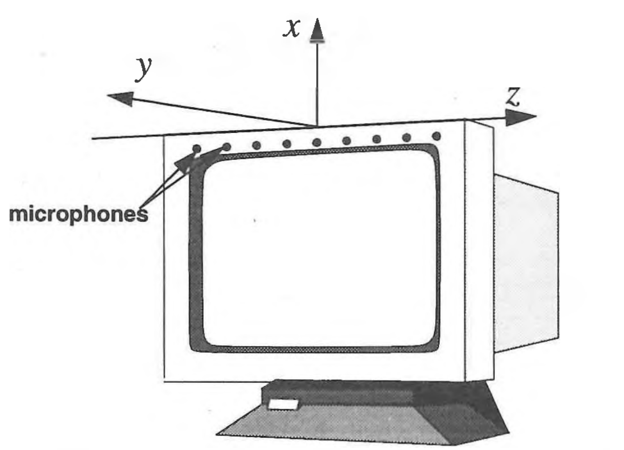


Figure 1. Sketch showing a microphone array intended for computer telephony. The sound field about the monitor can have large spatial variations which need to be accounted for in the beamforming procedure

When beamforming using multiple microphones, it cannot be assumed that all microphones receive the acoustic signal with the same magnitude. With a source in the near field[2,3], there will be different magnitude responses because of different propagation distances and diffraction effects will not necessarily affect all microphones the same. For example, the microphones closest to the sides of the monitor will experience a different sound field because of effects due to the side edges of the monitor.

SOUND FIELD CALCULATION

The variability of the SPL due to diffraction will be illustrated here. We consider microphone positions near the vertical center plane of the monitor, i.e., near $z = 0$ on Fig. 1. Above about 300 Hz, these positions will be essentially independent of the side edge conditions. We can then apply the exact theoretical formulation of Hadden and Pierce[4] for a rigid wedge to calculate the SPL. Choosing a wedge angle of 90°, four integrals, corresponding to direct and diffracted paths from actual and image sources, need to be evaluated. The quadrature is straightforward although some care needs to be taken to ensure convergence.

The results of a calculation are shown in Fig. 2. A point source radiating sound of frequency 1000 Hz is assumed to be located 50 cm in front of the "monitor" and 10 cm below the top edge (a typical location for a user's mouth). The sound pressure level, relative to the free field sound pressure, is calculated at a number of points distributed around the monitor and presented graphically in the 2D grayscale plot of Fig. 2

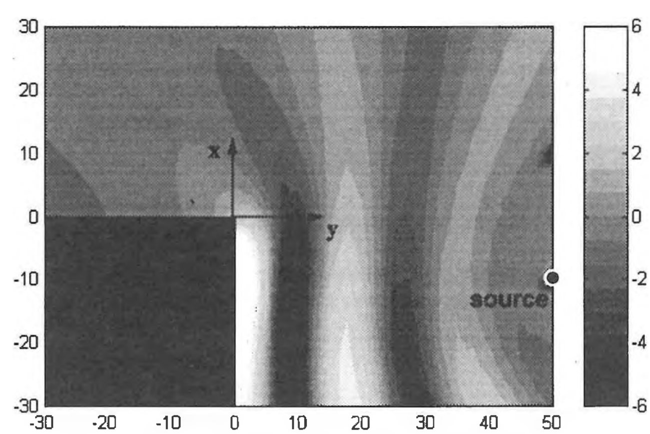


Figure 2. Grayscale plot showing the variation in sound pressuure level around a right-angled wedge, representing the $z = 0$ slice through the monitor of Fig. 1. The panel on the right shows the SPL (dB) corresponding to the different gray levels. The source is at (-10 cm, 50 cm).

Along the front face of the wedge ($-x$ axis), except near the vertex positon, the SPL is 6 dB, as indicated by the white area on the front face in Fig. 2. This is pressure doubling due to reflection of the incident acoustic signal. Out from this face a distance corresponding to a quarter wavelength is seen the dark interference null between incident and reflected waves. The positions along the top of the wedge ($-y$ axis) are in the acoustic shadow of the source and the sound pressure level drops smoothly. Near the vertex of the wedge, the SPL changes quite rapidly.

The pressure variation along the front face, particularly near the vertex, is examined more closely in Fig. 3. The relative sound pressure level is plotted as a function of the position x, for $y = 0$. For $x \gg 0$, far from the monitor, the diffraction effects are minimal and the relative SPL is nearly 0 dB. For $x \ll 0$, pressure doubling gives a level 6 dB above free field. The rapid transition between the two limiting regimes, through the vertex at $x = 0$, is evident. Right at the vertex, the SPL is 2.5 dB. Oscillations in the curve, due to interference between the incident and the edge-diffracted waves, are also noted.

The effects of diffraction are known to be dependent on the sound frequency so it is not surprising if anomalous frequency variations are introduced. For Fig. 4, we consider a small number of positions on the surface of the wedge (all for $z = 0$) and compute the SPL as a function of frequency. The source location is the same as that used in the previous figure. The labels on the various curves give their (x, y) coordinates. The frequency response for the vertex position $(0, 0)$ is found to be absolutely flat. For positions on the top surface of the wedge, in the diffractive shadow, a high frequency rolloff is found that increases with distance from the vertex. If the source was lower, these positions would be deeper in the shadow and levels would be lower still. On the front face, the transition from 0 to 6 dB, observed in Fig. 3, is found to occur at different frequencies for different positions and oscillations in the frequency response are noted.

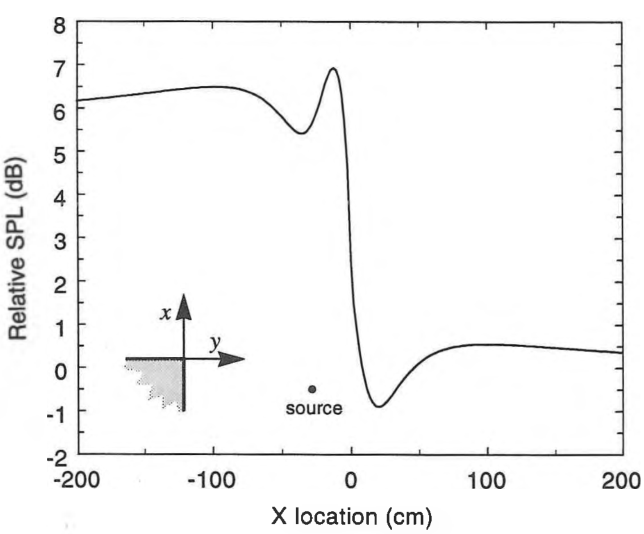


Figure 3. Relative sound pressure level along the front face of a right-angled wedge, with the same source location and frequency (1000 Hz) as in Fig. 2.

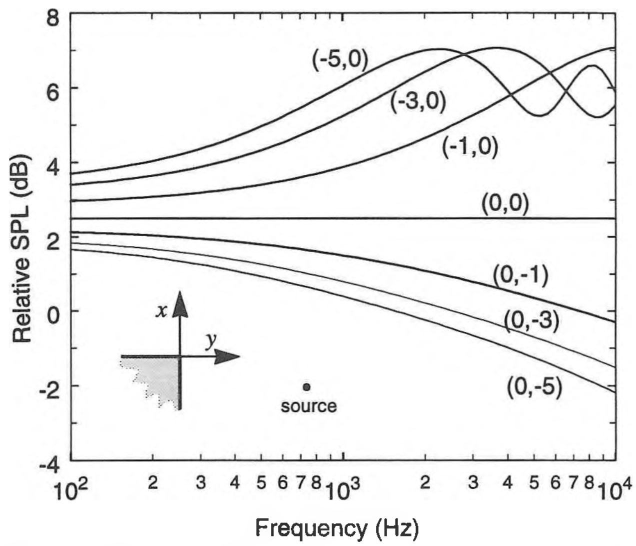


Figure 4. Frequency response functions for various microphone positions on a right-angled wedge, with same source location as in Fig. 2. Coordinate pairs next to the curves indicate the microphone positions.

DISCUSSION

The calculations here demonstrate the considerable variation in SPL near the vertex of a right-angled wedge. For a device such as a computer monitor, different microphone locations can lead to differences of several dB, particularly true at higher frequencies. If a directional receiver with two ports is used, account must be made of possible differences in received signal strength or the assumed directionality could be lost. Similarly, for an array of receivers, the effects of diffraction at each receiver location need to be determined to ensure effective beamforming and desired performance.

The use of the right-angled wedge was convenient here for the computer monitor application. For other applications, analytical formulations for other simple shapes, including wedges of different angles, spheres, and cylinders, can be applied. More complicated shapes or environments can also be considered but will require numerical techniques such as the boundary element approach.

REFERENCES

1. J.J. Bowman, T.B.A. Senior and P.L.E. Uslenghi (1969). *Electromagnetic and Acoustic Scattering by Simple Shapes*, North-Holland, Amsterdam.

2. J.G. Ryan and R.A. Goubran (1997). "Near-field beamforming for microphone arrays", *Proceedings of IEEE ICASSP*, May 1997.

3. J.G. Ryan (1998). "Criterion for minimum source distance at which plane-wave beamforming can be applied", J. Acoust. Soc. Am. **104**, 595-598.

4. W.J. Hadden, Jr., and A.D. Pierce (1981). "Sound diffraction around screens and wedges for arbitrary point source locations", J. Acoust. Soc. Am. **69**, 1266-1276.

Least-Squares Inverse Filtering of Speech

Bradford Gover†‡and James G. Ryan‡
†Department of Physics, University of Waterloo, Waterloo, ON N2L 3G1, Canada
‡Institute for Microstructural Sciences, National Research Council, Ottawa, ON K1A 0R6, Canada.

1 Introduction

When speech propagates through a room and is received by a microphone, the microphone output signal sounds unnatural and suffers reduced intelligibility (the so-called "barrel effect"). It is necessary to employ some sort of enhancement technique to process the microphone output so that the subjective impression of the subsequently reproduced speech is improved.

If the speech originates from a loudspeaker, then there are means to approximately invert the room's effect (see, for example, [1] and references therein). Most of these depend on the knowledge of the room impulse response as measured from the loudspeaker to the microphone, and the subsequent design of an appropriate inverse filter.

If, however, the source of the speech is a human talker, then it is not possible to determine this impulse response exactly. One way to attempt to enhance the live speech is to replace the talker with a loudspeaker, measure the impulse response, design an inverse filter, and then use it for the live speech. This introduces the question of the suitability of inverse filters for sources other than those from which they are designed.

This paper will illustrate some of the problems with such an approach. First, the technique of least-squares inversion will be summarized. Next, some actual room measurements using a loudspeaker and a mannequin will be presented. The mannequin is used to represent a real talker, and the results of its attempted dereverberation with the filter designed from the loudspeaker measurement will be presented and discussed.

2 Least-Squares Inverse Filtering

A least-squares inverse filter, when applied to a system impulse response, yields a result which is the best least-squares approximation to an ideal impulse. To invert a known impulse response $g(n)$, a filter $h(n)$ is sought such that $g(n) * h(n) = \delta(n - n_0)$, where n_0 is some delay. This can be written

$$\mathbf{Gh} = \mathbf{d}, \tag{1}$$

where $\mathbf{G}$ is the convolution matrix of $g(n)$, and the elements of the vectors $\mathbf{h}$ and $\mathbf{d}$ are the values of the time signals $h(n)$ and $\delta(n - n_0)$. The least-squares solution is obtained by minimizing the L_2-norm of the error vector $\mathbf{e} = \mathbf{Gh} - \mathbf{d}$. This occurs when it is orthogonal to the subspace spanned by the columns of $\mathbf{G}$ [2], so that

$$\mathbf{G}^T\mathbf{e} = \mathbf{G}^T\mathbf{Gh} - \mathbf{G}^T\mathbf{d} = 0, \tag{2}$$

or

$$\mathbf{Rh} = \mathbf{z}, \tag{3}$$

where $\mathbf{R} = \mathbf{G}^T\mathbf{G}$ is the correlation matrix of $g(n)$ and $\mathbf{z} = \mathbf{G}^T\mathbf{d}$ is the cross-correlation vector of $g(n)$ and the desired signal $\delta(n - n_0)$. The solution of Eq. (3) yields the desired filter.

3 Room Impulse Response Inversion

A room impulse response measuring system usually involves a signal chain comprising: pre-amplifiers, a loudspeaker, the room, a microphone, and post-amplifiers. In this case, the measured response, $g_{meas}(n)$, contains the responses of the amplifiers and transducers,

$$g_{meas} = g_{pre} * g_{spkr} * g_{room} * g_{mic} * g_{post}, \tag{4}$$

with the obvious definitions of the impulse responses on the right-hand side. Clearly for an inverse to have any chance of being independent of the system used to measure it, it must be an inverse of the room response, $g_{room}(n)$ only. If an inverse of the room response, say $h_{room}(n)$, is found, then the result of filtering the measured response is

$$g_{meas} * h_{room} = g_{pre} * g_{spkr} * g_{mic} * g_{post}, \tag{5}$$

which is exactly the anechoic response of the measuring system. Therefore, to design an inverse of the room response only, the least-squares criterion Eq. (1) should be changed to

$$\mathbf{Gh} = \mathbf{s}, \tag{6}$$

where s is the vector corresponding to $s(n - n_0)$, the delayed anechoic response of the measuring system. The solution for such a filter is found by solving Eq. (3), with $\mathbf{z} = \mathbf{G}^T\mathbf{s}$.

It is known that the room impulse response, $g_{room}(n)$, depends not only on the positions of the source and receiver, but also on the source directivity [3]. It is further known that due to the nonminimum-phase nature of the room impulse response, an exact inverse does not exist [4]. It is explored herein how detrimental the consequences of these facts are on dereverberation by inverse filtering.

4 Measurements

Room impulse response measurements were made with MLSSA, a maximum-length sequence analyzer. The sampling rate was 22 050 Hz and the MLSSA anti-alias filters (8th-order Chebyshev) had a bandwidth of 7.35 kHz. The MLS was 65 535 points long (16th order) and the measured impulse responses were 32 768 points long.

The room used for measurements was rectangular in shape with dimensions 4.65 × 6.70 × 2.44 m, smooth walls and ceiling, and carpeted floor. The room contained a typical amount of furniture, and had a reverberation time of approximately 350 ms. A Panasonic 6 mm electret microphone was located at 2.71 × 4.33 × 1.14 m. Measurements were taken with two sources: a PSB Alpha SE loudspeaker and a B&K HATS mannequin, which is intended to reflect the directional characteristics of a human talker. The sources were positioned at 2.07 × 2.31 × 0.82 m.

Additionally, the same equipment was set up in the same arrangement in an anechoic chamber and the measurements repeated.

The measured room and anechoic impulse responses and corresponding magnitude spectra are shown in the top two curves of Figures 1 (loudspeaker) and 2 (mannequin).

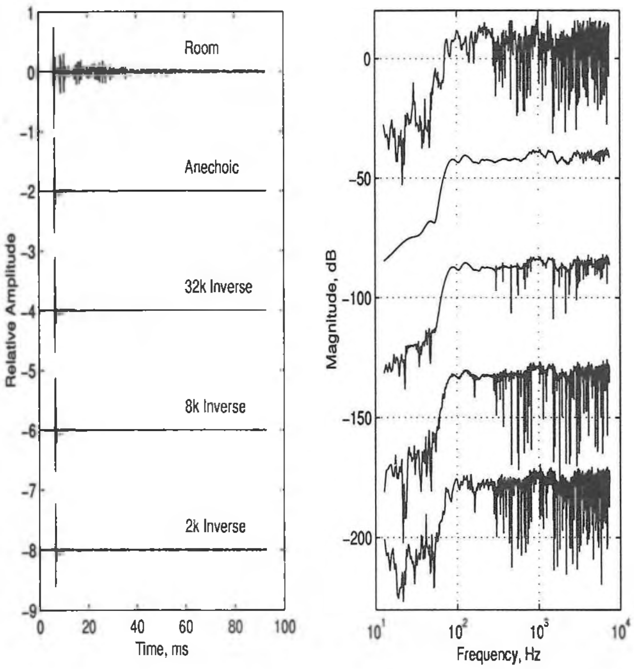


Figure 1: Room and Anechoic Impulse Responses and Magnitude Spectra for Loudspeaker (top two curves) and Equalized Outputs for Different Lengths of Inverse FIR.

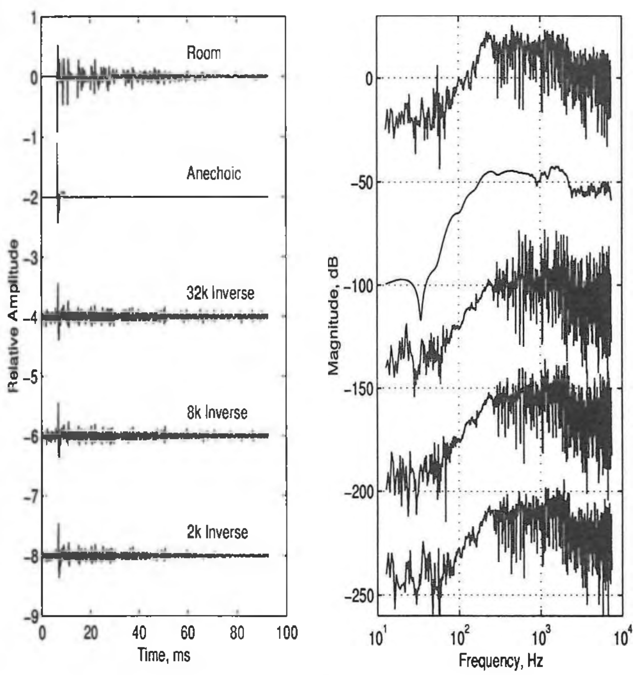


Figure 2: Room and Anechoic Impulse Responses and Magnitude Spectra for Mannequin (top two curves) and Equalized Outputs for Room Inverse Filter Designed with Loudspeaker.

5 Results

The lower three curves in Figures 1 and 2 show the result of post-filtering the impulse response as measured in the room, for the loudspeaker and the mannequin respectively, with inverse filters designed from the loudspeaker measurements.

For the loudspeaker, it is expected that longer inverse filters will yield a lower error [1], and this is supported by Figure 1. Notice, however, that some deep nulls in the room transfer function are not removed; this is not surprising since an all-zero (i.e., FIR) filter is not capable of removing zeros from a system transfer function, in particular nonminimum-phase zeros.

It is clear from Figure 2 that application of the inverse designed from the loudspeaker measurements does not effectively invert the room response for the mannequin measurements in any case. This mismatch is due to variations in the room response $g_{room}(n)$ due to different excitations.

Notice that the longer "inverse" filters are worse for dereverberation of the mannequin signals than the shorter ones. This is opposite from Figure 1, but is understandable since the longer filters are more sensitive to the fine details of the response from which they were designed, as opposed to the short ones which can equalize only gross distortions (such as room resonances), which may be common to both room responses.

6 Summary

The task of dereverberating received speech in a hands-free system is quite daunting. It has been shown that even for the case where the impulse response from source to receiver is known, the dereverberation is difficult and imperfect. Furthermore, an inverse filter designed from measurement of the room impulse response does not appear usable with a different source. In fact, if too long a filter is used, the speech can be degraded far more than if no equalization had been attempted at all. The usefulness of the least-squares inversion technique for dereverberation appears limited. Other techniques, such as beamforming, are known to be much less sensitive to source characteristics, and may prove to be the preferred choice.

Acknowledgements

The authors would like to thank Dr. Gilles Daigle at NRC and Dr. John Vanderkooy at UW for helpful criticisms and suggestions.

References

[1] J.N. Mourjopoulos, "Digital Equalization of Room Acoustics," *J. Aud. Eng. Soc.*, **42**(11), pp.884–900, 1994.

[2] S. Haykin, Adaptive Filter Theory, 3rd ed., Prentice-Hall, Upper Saddle River, NJ, USA, 1996.

[3] J. Mourjopoulos, "On the Variation and Invertibility of Room Impulse Response Functions," *J. Sound Vib.*, **102**(2), pp.217–228, 1985.

[4] S.T. Neely and J.B. Allen, "Invertibility of a Room Impulse Response," *J. Acoust. Soc. Am.*, **66**(1), pp.165-169, 1979.

Fast and accurate 3D acoustic propagation and inversion in layered media environments

James W. Wiskin, D. T. Borup, S. A. Johnson, Dept. Bio-engineering, Univ. of Utah, Salt lake City, Utah, 84112, wiskin@ee.utah.edu

An important problem in marine science is the scattering of sound by objects buried in underwater sediments. A model is desired which accounts for all orders of multiple reverberation between a scattering object and the sediment-sea interface, and is very computationally efficient. This paper will discuss a method for forward scattering from large objects in a 3D layered environment. In particular, we detail the use of a "layered Green's function" for the solution of the forward scattering problem, and can be considered as a generalization of the numerical method in [1]. There exist presently, many methods of forward scattering from an arbitrary object, such as Finite Element Methods, Finite Difference methods, Boundary Element methods, and Volume Integral equation methods. The FEM is very popular presently but has the disadvantage of requiring a large number of spatial samples per wavelength relative to our method, in order to ensure accuracy. A 2D version of this forward problem is implemented in an optimization based inversion scheme [1] developed at the Center for Inverse Problems, Imaging and Tomography (CIPIT) at the University of Utah.

We show the form of the 3-D Layered Green's function, and show how it is utilized efficiently using FFT's. After developing the 3-D Lippmann-Schwinger equation for layered media and listing its benefits and disadvantages compared with other numerical methods, several examples of forward scattering are computed using the integral equation with G_L and an independent method. The geometry relevant to the forward problem is displayed in figure 1. The grid is $nx \cdot ny \cdot nz$ pixels in size, where nx ny, nz are the number of pixels in the x, y and z directions, respectively. An inclusion is buried within the layered medium, which is completely contained within one layer. In the forward problems, the acoustic wavespeed, density, thickness, and attenuation of the layers in which the object is buried are known. We determine the total field that results when a known incident field is projected upon a scatterer with known parameter values (wavespeed etc.).

The careful development of the Green's function reveals that the BiStab-FFT method developed for the free space scattering case is applicable. The development for the Elastic equation is very similar to the acoustic case conceptually, but more difficult by virtue of its vector character, it will be addressed in a future publication. We show the layered media analogue of the Lippmann-Schwinger equation and the layered medium Green's function, we discretize the layered medium Green's function and the integral equation and discuss a fast numerical method of solution. Finally results of simulations and comparison with FDTD codes are shown and discussed.

In the case of an inhomogeneous body residing within a layered matrix, the governing equation for acoustic wave propagation in a 3D medium is:

$$\nabla^2 f(x,y,z) + k^2(x,y,z) f(x,y,z) = S_\omega(x,y,z)$$

where S_ω is the source function. The total field is broken up into scattered and incident field components: $f(x,y,z) = f^{inc}(x,y,z) + f^{sc}(x,y,z)$. The wave equation for the incident field is

$$\nabla^2 f^{inc}(x,y,z) + k_L^2(z) f^{inc}(x,y,z) = S_\omega(x,y,z)$$

Subtracting (3) from (2) the equation for the scattered field is found to be,

$$\nabla^2 f^{sc}(x,y,z) + k_L^2(z) f^{sc}(x,y,z) = -k_s^2(x,y,z) f(x,y,z)$$

where $c_o(z)$ is the speed of sound in the layered background medium, $k_L^2(z) = \omega^2 / c_L^2(z)$ is the wavenumber for the background layered medium, and $k^2(x,y,z) = \omega^2 / c^2(x,y,z)$ is the wavenumber corresponding to the spatially dependent speed of sound, $k^2(x,y,z) \equiv k_s^2(x,y,z) + k_L^2(z)$, where $k_s(x,y,z)$ is the part of the wavenumber that is due to the scattering object alone. See fig. 1. Now since $k_s^2(x,y,z) = \dfrac{\omega^2}{c_{sc}^2}\left(\dfrac{c_{sc}^2}{c^2(x,y,z)} - \dfrac{c_{sc}^2}{c_L^2(z)}\right)$, defining: $\gamma_L(x,y,z) \equiv \left(\dfrac{c_{sc}^2}{c^2(x,y,z)} - 1\right)$ where c_{sc} is the constant speed of sound of the layer which contains the inhomogeneity, gives:

$$\nabla^2 f^{sc}(x,y,z) + k_L^2(z) f^{sc}(x,y,z) = -k_{sc}^2 \gamma_L f(x,y,z)$$

The corresponding Lippmann-Schwinger Integral equation (LSIE):

$$p_{\omega\theta}^{inc}(r) = p_{\omega\theta}(r) - k_0^2 \iiint \gamma(r) p_{\omega\theta}(r) G_\omega(|r - r'|) dx'dy'dz'$$

where $c(x,y,z) = \omega / k(x,y,z)$ is the spatially varying speed of sound, $k_0^2 = \omega^2 / c_{sc}^2$ is the background wavenumber in the layer containing the scattering inhomogeneity, ω is the frequency of the interrogating field, $k(x,y,z)$ is the spatially dependent wavenumber of the body to be imaged, $p(r)$ is the pressure field, with the spatial dependence on the vector $r=(x,y,z)$ shown explicitly, and $\gamma(r)$ is the object function, which will be reconstructed from the scattered data, and we have used the subscripts $\omega\theta$ to indicate the solution dependence upon frequency (ω) and direction of incident field (θ). We also derived the Green's function which incorporates the special structure of the ambient media, i.e. $G_\omega(|r - r'|)$. The discretized version of this Green's function is obtained by convolution with the "sinc" basis functions: $\mathrm{sinc}(x) \equiv \sin(\pi x) / \pi x$ in the horizontal direction and "tent" functions in the vertical direction: $\Lambda(z) \equiv 1/\delta \begin{cases} \delta + z & -\delta \le z \le 0 \\ \delta - z & 0 \le z \le \delta \end{cases}$. denoted $G_\omega(|n|\delta)$, for $n=-nx,..0, .,nx-1$. The Green's function for the layered environment is broken up into a correlational and a convolutional part: $G_\omega(|n|\delta) \equiv G_V(|n|\delta) + G_R(|n|\delta)$ For n=0 $G_V(0)=\int_{u=0}^{u_\delta} J_o(0)\frac{2C(u)f(u)}{i\omega b_{sc}(u)}du$, where u_δ is an upper bound introduced by the band-limiting effect of convolution with the sinc functions. Furthermore $f(u)=\mathbf{R}^{-}\mathbf{R}^{+}\{z_1-1\}+\left\{1-e^{-i\omega b_{sc}\delta}z_1\right\}$, $\mathbf{R}^{+}$, and $\mathbf{R}^{-}$ are generalized reflection coefficients from layers above and below the scattering object, $z_1\equiv\left(e^{i\omega b_{sc}\delta}-1\right)/\,i\omega b_{sc}\delta$, and $C(u)\equiv\left(1-\mathbf{R}^{-}\mathbf{R}^{+}\right)^{-1}$. For n≠0,

$$G_V(|n|\delta)=\int_{u=0}^{u_\delta} J_o(u\omega|n|\delta)2D(u)\frac{g(|n|\delta)}{\omega^2 b_{sc}^2\delta}du$$

where

$$D(u)\equiv\left(1-\mathbf{R}^{-}\mathbf{R}^{+}\right)^{-1}\{1-\cos(\omega b_{sc}\delta)\},$$

and

$$g(|n|\delta)\equiv\left\{e^{-i\omega b_{sc}|n\delta|}+\mathbf{R}^{-}\mathbf{R}^{+}e^{i\omega b_{sc}|n\delta|}\right\}.$$

The correlational part is very similar, however, it is stored in a different order so that the correlation can be applied via a FFT. The discretized integral equations are used in an inversion procedure as follows. It may be the case that we have limited views available to us due to experimental limitations. If so we must rely upon multiple frequencies to increase the well posedness of the problem. To solve the multiple view and multiple frequency problem we minimize the residual between the predicted and observed fields at the receiver positions. This minimization can be carried out with either a nonlinear conujugate gradient procedure, or by an incomplete Newton method. The forward problem for an incident field in the direction θ and frequency ω is denoted by $\phi_{\theta\omega}$. This operator yields the scattered field at all receiver positions, given the incident field at angle θ, frequency ω, and the present estimate for the object function γ. $\phi_{\theta\omega}:\gamma\rightarrow\mathbf{f}_{\theta\omega}$. It turns out to be advantageous to attempt to solve the holomorphic (in γ) system

$$\mathbf{r}_{\theta\omega}^{(n)}\equiv\phi_{\theta\omega}\left(\gamma^{(n)}\right)-\mathbf{d}_{\theta\omega}=0\quad\forall\;\begin{matrix}\theta=1,\ldots,\Theta\\ \omega=1,\ldots,\Omega\end{matrix}$$

via a Newton Raphson procedure, where, as usual, the θ refers to the multiple views and the ω to the multiple frequencies available to us. That is, we assume that the noise level in the system is zero. We apply the Newton-Raphson procedure to this problem with the knowledge that applying a Gauss Newton algorithm to the associated least squares problem would give the same result. This leads to the overdetermined system

$$\left(\frac{\partial\phi_{\theta\omega}}{\partial\gamma}\right)\cdot\delta\gamma^{(n)}\approx-\mathbf{r}_{\theta\omega}^{(n)}\quad\forall\;\begin{matrix}\theta=1,\ldots,\Theta\\ \omega=1,\ldots,\Omega\end{matrix}.$$

which must be solved for the γ-update $\delta\gamma$. Again, we can use the *complex analytic* version of the Hestenes overdetermined conjugate gradient algorithm, adapted for least squares problems to iteratively solve this system. This is equivalent to finding the minimum norm solution of the equations:

$$\left(\frac{\partial\phi_{\theta\omega}}{\partial\gamma}\right)\cdot\delta\gamma^{(n)}\approx-\mathbf{r}_{\theta\omega}^{(n)}\quad\forall\;\begin{matrix}\theta=1,\ldots,\Theta\\ \omega=1,\ldots,\Omega\end{matrix}$$

The formula for the Jacobian in the layered medium situation in the presence of multiple frequencies, is,

$$\left(\frac{\partial\phi_{\theta\omega}}{\partial\gamma}\right)=\left((\mathbf{I}-\mathbf{G}_\omega\cdot[\gamma])^{-1}(\mathbf{G}_\omega\cdot[\mathbf{f}_{\theta\omega}])\right)$$

where G_ω is now the *Layered* Green's function for the frequency ω (for simplicity of notation – which would otherwise get out of hand) I will suppress the L superscript on the Layered Green's function in the above and that which follows. Therefore, in effect, to determine the γ-update, $\delta\gamma$, we merely solve the multiple view problem for each particular frequency, that is, we solve the overdetermined system:

$$\left([\mathbf{I}-\mathbf{G}_k\cdot[\gamma]]^{-1}\mathbf{G}_k\right)\otimes\mathbf{I}_{\Theta\times\Theta}\begin{bmatrix}[\mathbf{f}_{1k}]\\ \vdots\\ [\mathbf{f}_{\Theta k}]\end{bmatrix}\delta\gamma^{(n)}=-\begin{bmatrix}[\mathbf{r}_{1k}^{(n)}]\\ \vdots\\ [\mathbf{r}_{\Theta k}^{(n)}]\end{bmatrix},$$

which in component form is:

$$[\mathbf{I}-\mathbf{G}_\omega\cdot[\gamma]]^{-1}\mathbf{G}_\omega[\mathbf{f}_{\theta\omega}]\delta\gamma^{(n)}=-\left[\mathbf{r}_{\theta\omega}^{(n)}\right]$$

We end the presentation with several reconstructions based on computer generated data. The accuracy of the forward problem is guaranteed by comparison with several Bessel function expansion solutions derived analytically. We show that this method is very effective for propagating fields over large distances in layered environments to a computational grid that contains a complicated scattering object. The reconstructions are carried out using an efficient conjugate gradient and Newton based optimization method. The numerical efficiency of the method is preserved despite the presence of arbitrary layers both above and below the scattering grid by the break-up of the Green's function into correlational and convolutional parts, and the careful use of the Fast Fourier Transform. Also we discuss the various advantages and disadvantes of the method in various geometries. It is compared with the Split step Fourier method, and Finite Difference Time domain methods.

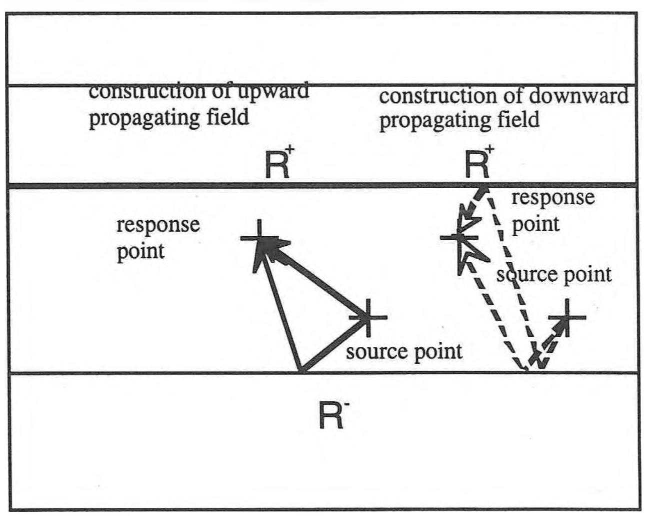


Figure 1

[1] Wiskin, J. W., Borup, D.T., Johnson, S. A., "Inverse scattering from arbitrary two-dimensional abjects in stratified environments via a Green's operator", *JASA 102, No. 2, Pt.1, pp. 853-864*, Aug. 1997

VIBRATION ISOLATION – EXPLORING THE FOUNDATIONS

H. M. Williamson, Hugh Williamson Associates, Ottawa, Ontario, Canada
J. C. S. Lai and J. Dai, Acoustics and Vibration Unit, University College, University of New South Wales, Australian Defence Force Academy, Canberra, ACT, Australia
C. Norwood, Defence Science and Technology Organisation, Melbourne, Victoria, Australia
Y. Li, PBR Automotive Pty. Ltd., Melbourne, Victoria, Australia

Introduction

Effective vibration isolation of machinery is important in many applications. Springs or rubber like materials are often used as mounts for vibrating machinery in order to reduce the transmission of vibrations to the supporting structure. These isolators are normally selected by ensuring that the natural frequency of the mounted system is well below the frequencies of the exciting forces which arise when the machinery is operating. This design method is based on a simple model[1] which approximates the machine as a lumped mass supported on a rigid foundation by an idealised linear spring and damper, Figure 1.

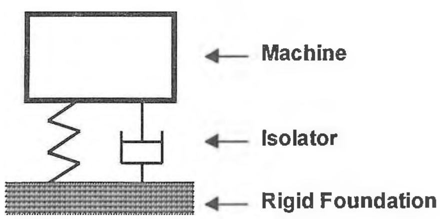


Figure 1 Simple Model of Vibraion Isolation

While this simple approach to the design machinery isolation works well in many instances, there are applications where a more sophisticated approach is needed in order to achieve desired levels of isolation. Some factors which limit the application of the above simple model are given below.

- Foundations are often far from being perfectly rigid. For example a rooftop air conditioning unit is often placed on a relatively light and flexible roof leading to poor isolation.
- The behaviour of the isolator may not be well described by the simple spring and damper. At higher frequencies, the mass of the isolator becomes important and can lead to wave effects and internal resonances within the isolator.
- The machine itself may be far from rigid. The attachment points of the isolators may have significant flexibility, especially at higher frequencies.
- It may not be appropriate to approximate a machine as a point-like mass. Any machine is three-dimensional and may have unevenly distributed mass. It can be vibrating in several directions at once and may have rotational modes of vibration as well. This aspect of isolation design can sometimes be treated by considering each independent mode of linear or rotational vibration separately with repeated applications of the basic theory.

This paper describes research work which explores the first of the above factors, the flexibility of the foundation. In particular, the importance of the size of the contact area between an isolator and its foundation is examined.

The measure of flexibility used here is mobility, M, defined as the ratio of the complex amplitudes of the time varying velocity, V, to the time varying force, F, at the point of application of the force on

$$M = \frac{V}{F}$$

a body. All terms are expressed as functions of frequency so that in general the mobility of a structure or foundation varies in both amplitude and phase with the frequency of the applied excitation. When a foundation is near rigid then the mobility approaches zero, that is, the velocity induced by any force is very small. The relationships between foundation mobility, isolator characteristics and isolator effectiveness are described elsewhere[2]. In general the effectiveness of isolation decreases if the mobility of the supporting structure is too high.

Size Effects on the Mobility of a Foundation

The mobility of a structure or foundation at a point can be measured. For example one can apply a time varying force with a shaker or an instrumented impact hammer and measure the resulting velocity at the point of application. Mobility at a point can also be theoretically calculated for simple structures.

However, actual connections between a machine and a supporting structure may cover a significant area. For example in ships, engine mounts may be quite large so that force is applied to the supporting structure over a large area (say 300 mm by 150 mm), and not at a point. The aim of the research described here is to explore the significance of area contact as opposed to point contact in determining foundation mobility.

Theoretical Models of Area Contact

A theoretical model[3] was developed where mobility was predicted for a circular region of contact on an infinite plate, Figure 2. Although this is far from a practical situation of a machinery mount in contact with a supporting structure, the model allows one to understand some fundamental characteristics of area contact. The above definition of for mobility under point contact was modified to take into account contact over a surface area[3], and is termed surface mobility.

Figure 3 shows the resulting predictions of surface mobility in non-dimensional terms for the case of a uniform, con-phase force distribution over the circular contact area. On the vertical axis, surface mobility is normalised by dividing by the mobility for point contact on an infinite plate. The horizontal axis is Helmholtz

Number which physically represents π times the ratio of the diameter of the contact region to the governing wavelength of the bending waves in the plate.

$$\text{Helmholtz Number} = ka = \pi\,(2a)/\lambda$$

$$\text{where} \quad \lambda = 2\,\pi\,/k = \text{wavelength}$$

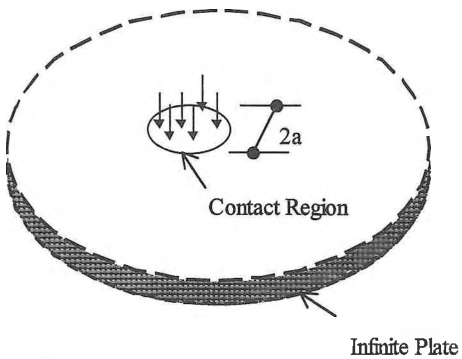


Figure 2 Model of circular surface contact on an infinite plate

The wavelength of bending waves in a plate varies linearly with frequency, long wavelengths at low frequencies and shorter wavelengths at higher frequencies. Hence Figure 3 shows that for this case, mobility depends on the ratio of the size (diameter) of the contact region to the wavelength of the bending waves in the plate. This case also shows a series of minima occurring whenever the diameter of the contact region is approximately equal to an integral number of wavelengths. These minima did not occur when a case of uniform con-phase velocity excitation was studied, however the general downward trend in normalised mobility with Helmholtz number remained.

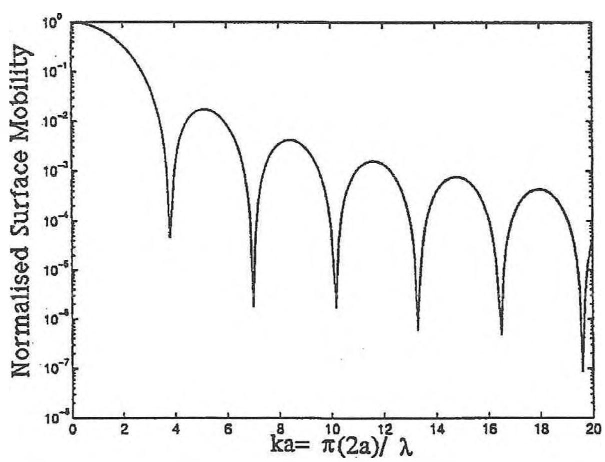


Figure 3 Normalised surface mobility for a circular contact region on an infinite plate subject to a uniform con-phase force distribution

Similar theoretical studies[4] were also carried out for the case of a rectangular contact region on an infinite plate. The results are more involved due to the increased complexity of the geometry, however the general characteristics are quite similar to the circular contact case, that is mobility is strongly affected by the dimensions of the contact area relative to the wavelength of the bending waves in the plate.

Experimental work[3] has verified the theoretical model for circular contact and experimental work is in progress for the verification of cases of rectangular contact. Experimentally, the infinite plate was approximated by a large plate with its edges well damped (buried in sand).

Some work has also been carried out on using finite element techniques to model surface contact. This could lead to a method of predicting surface mobility for more practical support structures.

Discussion

The above theoretical work on surface contact has not yet been developed to the point where it can be quantitatively applied to practical vibration isolation problems. It does however bring out some important points on the nature of the interaction between a large isolator and a supporting structure.

(1) It can be seen that large area contact can result in a mobility which is significantly different to measured or predicted point mobility. In the idealised cases studied, area mobility was less than point mobility, but this need not always be the case. It may be worthwhile developing methods to directly measure surface mobility for critical isolation applications. Such measurements could help to avoid situations where poor isolation occurs in important frequency ranges due to high surface mobility in the foundation.

(2) The research also helps to answer the question, 'How large is large?' In the above models, surface mobility was approximately equal to point mobility when the dimensions of the contact region were small compared to the governing wavelength in the supporting structure. Mobility changed significantly once the contact dimensions were of the order of, or larger than, the governing wavelength in the supporting structure.

Acknowledgements

The authors express their appreciation to the following organisations for supporting this work: Defence Science and Technology Organisation, Australian Research Council and University of New South Wales.

References

1. Tse, F. S., Morse, I. E. and Hinkle, R. T., *Mechanical Vibrations*, 2nd edition, Boston, Allen and Bacon, 1979
2. Williamson, H. M., *Vibration Isolation research at the Australian Defence Force Academy*, 3rd CF/CRAD Meeting on Naval Applications of Materials Technology, Halifax, 1997.
3. Norwood, C., Williamson, H. M. and Zhao, J., *Surface Mobility of a Circular Contact Area on an Infinite Plate*, Journal of Sound and Vibration, vol. 202, no. 1, April 1997.
4. Dai, J., Williamson, H. M. and Li, Y. J., *Investigation of Effective Point Mobility on a Two-dimensional Contact Area*, Proceedings, 6th Western Pacific Regional Acoustics Conference, Hong Kong, pp. 215 – 220, November 1997.

Results of a Survey on Audio and Visual Warning Systems for Military Helicopters

C. I. Mack(1), L. L. Churchill(2), and G. R. Arrabito(1)

(1) Defence and Civil Institute of Environmental Medicine, 1133 Sheppard Avenue West, Toronto, Ontario, M3M 3B9
(2) Daniel Edward Computing Ltd., Ottawa, Ontario.

Introduction

The Defence Research Establishment, Valcartier, (DREV) of the Department of National Defence (DND) is presently developing an Eye-safe Laser-Based Obstacle Warning System (ELBOWS). It is intended to mount this system on one or more Canadian Forces' (CF) helicopters to provide advance warning of obstacles in the flight path of the aircraft. The Defence and Civil Institute of Environmental Medicine (DCIEM), DND's center of excellence in human factors, has been tasked to develop a display for the new warning system.

Aircraft used in the military environment contain numerous warning systems to assist the crew in monitoring the state of their equipment and flight path. For example, the CH-146 Griffon helicopter contains more than 60 warning and indicator lights, and five or six auditory alerts. The auditory alerts provide notification of conditions ranging from routine to critical. For example, one routine alert is a sound similar in nature to the ringing of a telephone, which tells the crew they are being contacted on high-frequency (HF) radio. However, another audio alert indicates that rotor speed has decreased to the point where the helicopter is in immediate danger of losing lift; this critical warning sound is backed up by warning lights in both pilots' primary field of view.

Requirements for the ELBOWS Display

Much is known about warning systems and displays. Basic issues relevant to the ELBOWS system were discussed in a previous report [1]. As a result of this literature review and analysis, it was decided that the system should identify and classify obstacles. Furthermore, the critical nature of avoiding obstacles with little warning time appears to justify an additional audio alert. Other researchers have determined that up to ten audio alarms can be distinctly identified in military helicopters [2]. Laboratory research has implied that the use of spatial, or three-dimensional audio, may provide a reaction time advantage over non-directional audio [3].

However, the literature review left some issues unresolved. For example, what level of tolerance will the crew have for false alarms? Is there any confusion interpreting existing audio alarms? Should the new warning have a cancel feature? Do the air crew want to have a backup to the primary warning? To address these and other questions, it was decided to design and administer an attitude questionnaire on the subject of audio and visual warning systems in military helicopters.

Questionnaire Design

The selected audience for the questionnaire was current regular and reserve force helicopter air crew. The questionnaire was designed to solicit two types of answers: descriptive answers to supply background information and detailed descriptions of warning systems issues; and scaled questions amenable to statistical analysis. The final questionnaire contained 42 items, and was organized into four major sections:

- relevant information about the respondent, to be used for factor analysis;
- general questions on warning systems, to determine the general level of knowledge and attitudes of the respondents towards warning systems;
- specific questions about audio warnings, to investigate specific design issues and operator attitudes toward audio warnings; and
- questions relating to the specific functionality of the proposed obstacle warning system.

In addition to the four main sections of the questionnaire, respondents were asked to indicate whether they had experienced a near miss with wires.

A pilot trial of the questionnaire was performed at a CF helicopter squadron and responses were received from five volunteers. Two of the questionnaires were completed in the presence of the experimenters; three other forms were completed and returned by mail. Several changes were made to the formatting as a result of the pilot trial, and two new questions were added.

Survey Administration

The final version of the questionnaire was administered at two CF helicopter units, - a search and rescue (SAR) squadron and a tactical helicopter squadron. These two operational communities were sampled because of the very different nature of their missions. Additionally, the SAR squadron flies a large twin-rotor helicopter and the tactical helicopter squadron flies the smaller, single-rotor CH-146 Griffon.

Both squadrons were visited by the survey administrators with the intention of getting all surveys completed during the three or four-day visits. The population of helicopter air crew at the SAR squadron was considerably smaller than at the tactical helicopter squadron (less than 20 at the former vs. about 75 at the latter). However, the response rates at the two units were significantly different, resulting in approximately the same number of responses from the SAR and tactical samples.

Results

Thirty completed questionnaires were received. All of the surveys from the SAR air crew were completed and returned to the administrators during the squadron visit; some of the surveys from the tactical squadron were returned by mail.

The personal data section of questionnaire asked for information concerning number and type of flying hours, crew position, years of service and type of service (regular or reserve). However, because of the relatively small sample sizes, the only factor that appeared to be significant in the grouping of responses was whether the respondent was from a tactical or a SAR squadron.

Of all the respondents, one-half answered 'it depends' when asked whether audio or visual warnings are more likely to get their attention. Forty percent replied that audio warnings were more likely to get their attention, and ten percent stated that visual warnings were preferred. Only two of thirty people had experienced confusion while interpreting audio alarms; the response rate was similar for visual alarm confusion.

No one thought that there are too many audio alarms in the two CF helicopters under study. We also asked whether current alarms are too quiet, just right, or too loud, using a word-anchored, seven-point scale. Twenty-four out of twenty-eight replies to this question indicated that the volume of current audio warnings is 'just right'.

In two different questions, the air crew were asked to list the most important warnings and those of secondary importance in the two helicopters. While the individual lists varied somewhat, there was general agreement that the most important warning list includes:

- master caution light (visual)
- radar altimeter (audio)
- fire (visual)
- low rotor speed (audio and visual)

The secondary warning list was much longer, and no item was listed by more than six respondents. Two audio alerts appeared in the secondary list: the emergency locator transmitter alert, and the HF radio notification.

The survey also addressed the cancellation and switching-off of alarms. A cancel function resets the alarm without corrective action being required; a subsequent warning condition will reactivate the alarm. Switching off an alarm means that no warning will be issued. Respondents were asked whether current auditory alerts could be cancelled or switched off. While there was some misinterpretation of this question (some respondents felt that pulling a circuit breaker was a valid way to switch off an alarm), it was possible to determine that some critical auditory alerts can be cancelled and/or turned off. When asked whether the new system should have a cancel function, all respondents answered in the affirmative. Six respondents said the system should not have an on/off function, while the remainder said that it should.

A sensor-based obstacle warning system will generate some false alarms. The tolerance of air crew to false alarms is a subject of great concern in the design of the display: too many false alarms and the air crew will distrust the system. We asked how often current warning systems provide false alarms, and twenty-six of twenty-eight replies indicated rarely, occasionally, or sometimes (one to three on a five-point word-anchored scale.) We asked respondents to rate the degree of acceptability of false alarms on a seven-point, word-anchored scale. One-half of the answers were unacceptable in all or most systems; only one person said false alarms were acceptable in most systems. Air crew were asked whether false alarms could result in incorrect actions, and 27 of 30 agreed to some extent. Most respondents also agreed that multiple false alarms could cause a subsequent valid alarm to be ignored.

The questionnaire provided six possible options for the new warning system display: voice directions, tone only, tone and voice, light only, light and visual direction, and tone and visual direction. The only definite conclusion from the responses is that no one preferred the light only option. Thirty of thirty-nine answers (some respondents indicated multiple choices were acceptable) included audio, and the remaining nine preferred a light and a visual direction.

The last question asked whether the air crew had experienced a near miss. More than one-third of the respondents indicated they had been involved in such a situation, providing strong support for the new warning system.

Conclusions

The results of the survey have provided some valuable and insightful input into the process of display design for the new system. The need for the system is supported by the significant number of near misses reported by the air crew.

Questions related to the use of audio vs. visual alarms tended to favour the use of an audio alert. However, several respondents also indicated visual only preferences. False alarms are a great concern for the operators and the new system will need to be robust against them.

The completion of the survey represents the end of the second of three stages of requirements capture. In the third and final stage, experimental work will be undertaken to verify the nature of the warning display, e.g. tone vs. voice vs. light, etc.

References

1. Mack, C.I., User Requirements for an Obstacle Warning System, DCIEM Report Number 97-R-33, Defence and Civil Institute of Environmental Medicine, 1997.
2. James, S., Proposals for an Integrated Audio Warning Suite for Helicopters, DERA Report DRA/ AS/MMI/ CR9504/1, UK Defence Evaluation and Research Agency, 1995.
3. Calhoun, G.L., Janson, W.P., and Valencia, G., Effectiveness of Three-Dimensional Auditory Directional Cues, Proceedings of the Human Factors Society 32nd Annual Meeting, Santa Monica, CA: Human Factor Society, 1988.

A Multimedia Software for Evaluation of Traffic Noise Reduction Solutions

Alex Boudreau*, André L'Espérance* and Martin Bouchard**
*GAUS, Université de Sherbrooke, Sherbrooke (Québec) J1K 2R1, Canada
**Département de génie électrique, Université d'Ottawa, Ottawa (Ontario) K1N 6N5, Canada

1. INTRODUCTION

The determination of the optimal abatement solution for traffic noise is not easy to carry out. Indeed, sound propagation is a complex phenomenon, and current traffic noise models[1] are still approximate. In fact, most of those models do not consider the multiple reflection effect between two parallel barriers (situation that is often encountered in practice); they neither take into account the type of road surface (very influential factor in practice) nor the spectral content and the absorption properties of the materials as a function of frequency.

Another important deficiency of the existing tools is that they only give a numerical information of the results, that is the sound level and/or the noise reduction in decibels (dB). For most of the people, a level in decibel is however a rather abstract information. What does a reduction of 6, 10 or 12 dB correspond to, in terms of sound perception ? Very few people know it accurately.

The development of *CEEPA* is intended to fill those gaps. CEEPA is a software based on the recent knowledge on sound propagation models and on traffic noise emission. A signal processing algorithm has been integrated to this software in order to provide an audio output to the users. When used with a multimedia functionality computer, this new software thus allows the user to listen to the predicted noise levels. This paper presents the new software developed for a Windows 95 environment.

2. COMPUTATION METHODS

The computation algorithm has for main specificities :

1) Punctual sources model
2) Multiple reflections model
3) Traffic noise emission spectrum
4) Signal processing for audio output

2.1 Punctual Sources Model

Rather than a linear source line model, punctual sources are used to represent the traffic noise. This method allows to use accurate and well known propagation models to evaluate ground effect, diffraction, as well as reflections between barriers. In this model, each punctual source represents a set of vehicles on a unit of road length.

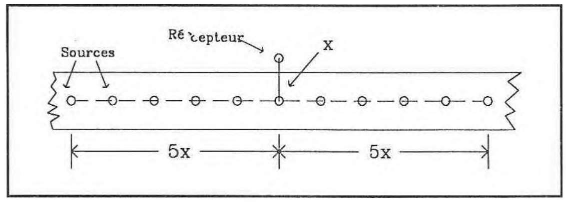


Figure 1 For information on this method, see references 1 to 4.

2.2 Multiple Reflections Model

When there are two parallel barriers, multiple reflections between them occur, which is an aspect that is generally not considered. In the proposed model, this phenomenon is considered using the image source concept (SIM), and the specific reflection coefficient of the material on the barriers is taken into account[1].

2.3 Traffic Noise Emission Spectrum

For a good evaluation of the ground effect, of the absorption effects, and of the diffraction effect, the spectral content of the source sound power should be considered, another aspect generally ignored in common traffic noise models. To overcome this deficiency, the notion of traffic noise emission spectrum[2] has been used. An emission spectrum specifies the contribution of each frequency 1/3 octave band in relation to the spectrum's overall noise level. Since the spectral content and the overall noise level emitted by vehicles depends on the type of road surface, the software offers different particular emission spectra[3]. If required, other particular emission spectra may be evaluated and used[4].

2.4 Signal Processing for Audio Output

The software generates audio files (*.wav)* associated to the receiver locations, for the reference case as well as for the modified case (see below *User Interface and Results* to know the difference between those two cases).

The audio file is carried out by the filtering (in 1/3 octave bands) of pre-recorded traffic noise files (.wav file). The filtering parameters are established following the spectral content of the pre-recorded noise files and of the theoretical spectrum calculated by the software. Figure 2 gives the diagram of this filtering process.

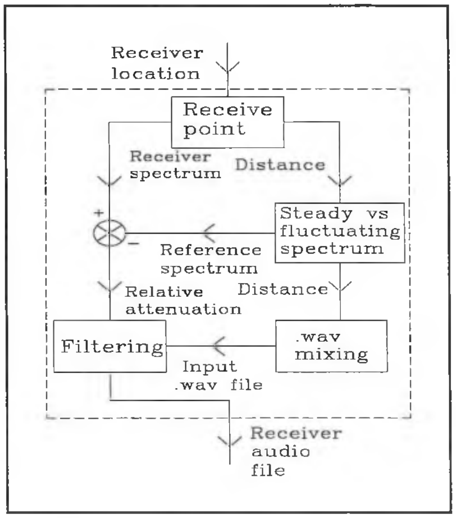


Figure 2 Signal processing of the audio output

3. User interface and results

The interface has been developed following Windows 95 standards.

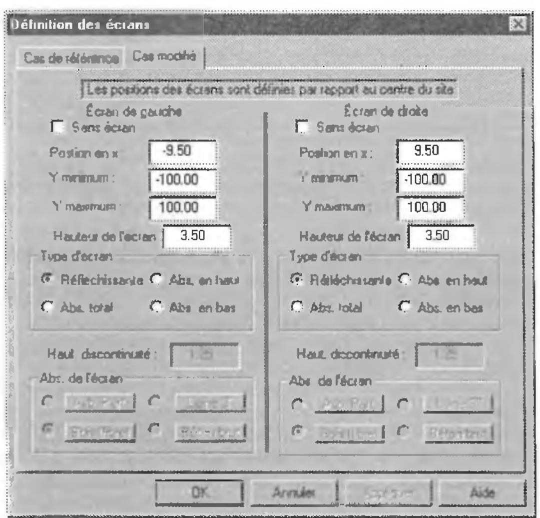


Figure 3 Example of dialog box.

Each analysis is made of two cases called *reference case* and *modified case*. The *reference case* corresponds to the road as it exists before the modifications. The *modified case* is a second configuration where the user gives the various changes desired: addition of barriers, addition of absorbents on the barriers, change of surface, change of traffic parameters, etc.

The noise level computation is done for every 1/3 octave frequency. The computation is automatically done for the reference case and the modified case. This computation is done for a given receiver grid (for example 25x25) distributed uniformly on the studied site (see Figure 4).

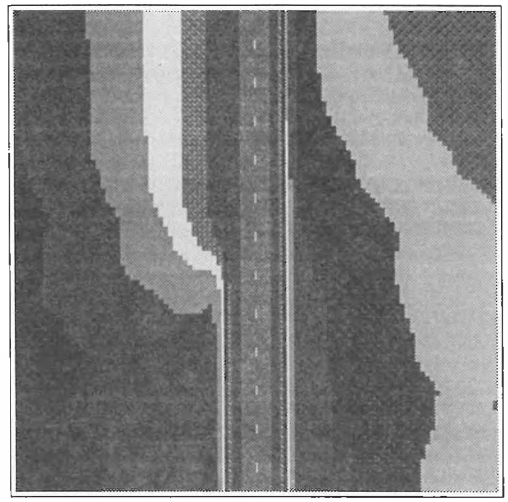


Figure 4 Example of graphical results given by CEEPA.

Using a computer with a sound blaster card, an audio evaluation of the noise abatement strategies can be done. To get this evaluation, the user has to establish the receivers where he wants the analysis done, and CEEPA will generate some audio files (*.wav)* for the reference case and for the modified case. One can thus have a representative impression of the effect of noise abatement strategies for any given location.

4. CONCLUSION

CEEPA is a user-friendly evaluation tool for the traffic noise. Compared to common methods, this software allows to consider ground effect, multiple reflections between barriers and emission spectra of different road surfaces. Besides its user-friendliness and the functionalities of its user interface, CEEPA can generate an audio output, which facilitates the analysis and optimization of the planned acoustical abatement strategies.

References

1-STAMINA model, developed by the FHWA (Federal Highway Administration).

2 Panneton et al. *Development and validation of a model predicting the performance of hard or absorbent parallel noise barriers* , J. Acoust. Soc. Jpn. (E) 14, 4 (1993).

3-Watts, G., *Normalized Traffic Noise Spectra*, Proceeding Inter-Noise 92, Toronto p. 823-827

4 L'Esperance et al. *Méthode d'évaluation de l'efficacité acoustique des revêtements routiers* », Rapport d'étude réalisée pour le ministère des Transports du Québec, Juin 1993, pp. 30.

INSTRUMENTATION SYSTEMS: DEVELOPMENT OF AN INTERNET-BASED COURSE

L.A. Tremblay, D.G. Jamieson, M.F. Cheesman
Hearing Health Care Research Unit, The University of Western Ontario
London, Ontario N6G 1H1 CANADA

This paper outlines the development of a course on Instrumentation Systems, delivered over the Internet. This is a senior undergraduate course for students who plan to enter Western's graduate program in Audiology. Such students come from all regions of Canada, and few have access to such a course at their previous institution. In addition to permitting students from anywhere in the country to access the course, Internet courses can be offered throughout the year, releasing students from the restrictions of a seasonal schedule of classes, and allowing them to take the course at a time and location which is convenient for them. The initial investment in course development is substantial, being many times greater than that for a traditional lecture course. However, course delivery is relatively inexpensive.

Internet-based courses offer certain pedagogical advantages for courses relating to acoustics: 1) they make it easy to incorporate audio demonstrations into course modules; 2) they similarly facilitate the use of computer-based video and illustrations; 3) they permit interactive learning at a pace set by the student; and 4) they permit arbitrary branching by the student, so that course materials can, but need not be, studied in strict linear fashion, as with traditional text- or lecture-format presentation.

Given these advantages, it is not surprising that many institutions are pursuing web-based courses, some in a significant way. As just one example the California Virtual University Foundation is a cooperative effort among 95 State-funded institutions of higher education: The University of California, the California State University and the California Community College, plus Sun Microsystems, Cisco Systems, Pacific Bell and Oracle. The Virtual University already provides 1,600 courses that are available to students at any of these institutions. A similar cooperative effort is underway among the states of the Pacific Northwest.

Our Internet course in Instrumentation Systems is based on an existing course, which is taught in a traditional, 13-week, lecture/lab format to first-year Audiology students at Western. During 1997/98, all students in the class reported that they had both access to and knowledge of the Internet, giving some assurance that the Internet version of the course would have been accessible to them.

The course follows a course in Hearing Science (see Taylor, Jamieson & Cheesman, this volume), which introduces basic concepts in physical acoustics, plus basic anatomy, auditory physiology, and psychoacoustics. The Internet course has been developed as a series of modules, reflecting the topics covered in the current course, during approximately 39 hours of lectures and 26 hours of laboratories.

Each module includes: 1) a statement of learning objectives; 2) explanatory and descriptive hypertext; 3) visuals in the form of diagrams, digital photographs, digital video, and animations; 4) audio demonstrations; 5) worked examples; 6) assignments; 7) a self-test review quiz; 8) references and suggestions for further study, both web-based and traditional; 9) an Adobe-format copy of a required reading; and 10) an optional, advanced-study component.

The course focuses on the following topics:

1) Review of basic concepts introduced in hearing science course
2) Overview of instruments and measurement devices
3) Properties of sound
4) Amplitude and damping
5) Fundamentals of Electricity and circuits
6) Filters and distortion
7) Digital Signals
8) Variables in an Acoustic environment (ambient noise, reverberation, signal-to-noise ratio)
9) Acoustic calibration
10) Introduction to Impedance
11) Introduction to ABR
12) Audiometric equipment (audiometers, hearing aids, tympanometers, OAE systems, and ABR systems)

The module on *Digital Signals* will be discussed as a representative sample of the course materials. This module covers the topics: 1) digital signals vs analog signals; 2) the sampling theorem; 3) quantization; 4) Nyquist rate; 5) aliasing; 6) conversions between analog and digital signals; and 7) digital synthesis of signals. In addition to diagrams and illustrations of concepts, audio demonstrations of the effects of quantization have been included to allow students to play back the signals on demand, in order to hear and appreciate the differences between the signals. These audio examples used in the module were originally prepared for a computer-based teaching module (Jamieson & Cheesman, 1997) using the CSRE software system. The examples provided are of speech signals that have been quantized at 16 bits, 10 bits, 4 bits, and 1 bit.

The student assesses the intelligibility and perceived quality of the differently-quantized speech signals. Worked examples of calculations of the number of amplitude values for each level of quantization are provided and students are then given a series of questions to work on their own. The module includes a self test, with multiple choice questions and definitions of terms. It concludes with a list of links to other web sites (e.g., http://www.comdis.wisc.edu/VCD202/CD202menu.html), providing students with a starting point for further exploration and directing them to relevant on-line demonstrations. Each module also links to the course teaching assistant, to allow personal feedback on individual questions and concerns regarding the course material.

During the 1999 winter term, the web course is being offered as an in-house supplement to the normal lecture course at Western. Revisions based on student evaluations and instructor experiences will occur the following summer. The course is scheduled to be offered as a self-contained Internet course beginning with the Winter, 2000 term.

REFERENCES

Jamieson, D.G. and Cheesman, M.F. (1997) CSRE-based teaching modules for courses in speech and hearing sciences. London: Avaaz Innovations, Inc.

DEVELOPMENT OF AN INTERNET-BASED COURSE IN HEARING SCIENCE

B. Taylor, D.G. Jamieson and M.F. Cheesman
Hearing Health Care Research Unit, The University of Western Ontario
London, Ontario N6G 1H1 CANADA

This paper outlines the development of a course on Hearing Science, delivered over the Internet. This is a senior undergraduate course for students who plan to enter Western's graduate program in Audiology. Such students come from all regions of Canada, and not all have access to such a course at their previous institution. In addition to permitting students from anywhere in the country to access the course, Internet courses can be offered throughout the year, releasing students from the restrictions of a seasonal schedule of classes, and allowing them to take the course at a time and location which is convenient for them. The initial investment in course development is substantial, being many times greater than that for a traditional lecture course. However, course delivery is relatively inexpensive.

Internet-based courses offer certain pedagogical advantages for courses relating to acoustics: 1) they make it easy to incorporate audio demonstrations into course modules; 2) they similarly facilitate the use of computer-based video and illustrations; 3) they permit interactive learning at a pace set by the student; and 4) they permit arbitrary branching by the student, so that course materials can, but need not be, studied in strict linear fashion, as with traditional text- or lecture-format presentation.

Given these advantages, it is not surprising that many institutions are pursuing web-based courses, some in a significant way. As just one example, the California Virtual University Foundation is a cooperative effort among 95 State-funded institutions of higher education: The University of California, the California State University and the California Community College, plus Sun Microsystems, Cisco Systems, Pacific Bell and Oracle. The Virtual University already provides 1,600 courses that are available to students at any of these institutions. A similar cooperative effort is underway among the states of the Pacific Northwest.

Our Internet course in Hearing Science is based on an existing course, which is taught in a traditional, 13-week, lecture/lab format to first-year Audiology students at Western. During 1997/98, all students in the class reported that they had both access to and knowledge of the Internet, giving some assurance that the Internet version of the course would have been accessible to them.

The course is designed to introduce students from a diverse range of backgrounds to basic topics in acoustics, sound generation and transmission, sound characterization for both clinical and scientific purposes, anatomy of the human auditory system, auditory physiology, and auditory sensation and perception. Students will also become familiar with current theories and research on hearing through selected readings. They will learn the influence hearing science has had on audiological practice and understand how the consideration of new information may help advance the field of Audiology.

The Internet course has been developed as a series of modules, reflecting the topics covered in the current course, during approximately 39 hours of lectures and 26 hours of laboratories.

Each module includes 1) a statement of learning objectives; 2) explanatory and descriptive hypertext; 3) visuals in the form of diagrams, digital photographs, digital video, and animations; 4) audio demonstrations; 5) worked examples; 6) assignments; 7) a self-test review quiz; 8) references and suggestions for further study, both web-based and traditional; 9) an Adobe-format copy of a required reading; and 10) an optional, advanced-study component.

The course includes 30 modules covering the following topics: 1) sound transmission in the ear; 2) sine waves; 3) sound fields and resonance; 4) scientific notation and logarithms; 5) decibels; 6) anatomy of the auditory system; 7) auditory physiology; 8) auditory sensitivity; 8) signal detection theory; 9) measurement of hearing; 10) loudness; 11) masking; 12) frequency selectivity; 13) temporal processing; 14) perception of speech; 15) binaural hearing and cross-channel processes; and 16) the effects of hearing impairment on auditory perception.

The modules on *the Anatomy and physiology of the ear* supplement diagrams with animations demonstrating the movement of key anatomical structures during sound transmission and transduction. Students are referred to relevant external sites for advanced study [e.g., 1,2]. Modules concerning sound waves provide animations and video and sound clips to demonstrate how sound generation and trans-mission. Because many of our students have a restricted scientific and mathematical background, modules have been developed to introduce or review such relevant concepts as scientific notation and logarithms; these modules feature worked examples and make use of hypertext links to explain the derivation of equations and provide additional details useful in solving various fundamental problem types. Following these modules, the topics of Decibels and Signal Detection Theory are introduced, using a similar approach. Modules on auditory perception include illustrations, animations, video clips, and audio examples to demonstrate the various psycho-acoustical phenomena and hearing measurement procedures.

Internet-based courses offer certain obvious advantages: they remove course enrolment restrictions based on geography (location of the student), time (both university term and specific course scheduling), and physical space (size of the lecture hall). They facilitate interactive learning and the incorporation of animations, video and audio demonstrations into course materials. The selection of material for study is determined by the student, through pacing, branching and repetition, making it appropriate for use by students with different backgrounds and abilities. Unlike the traditional classroom setting, in which all members of the class necessarily follow the same, largely-predetermined and nearly-linear sequence, students at different levels can move through the course material at a rate and in a sequence that is most appropriate for them. Students can move quickly through the areas with which they are familiar, and spend more time on those topics that are more novel or difficult for them.

Creation of the course has been extremely time consuming, even though it was derived from an existing course, which was highly developed and had many resource materials. In general, more than 80 hours of development time has been required for the initial implementation of each module, with much more time required for those modules where media materials were unavailable.

During the 1998 fall term, the web course in Hearing Sciences is being offered as an in-house supplement to the normal lecture course at Western. Revisions based on student evaluations and instructor experiences will occur over the following winter and summer terms. The course is scheduled to be offered as a self-contained Internet course beginning with the Fall, 1999 term.

REFERENCES

[1] http://ctl.augie.edu/perry/ear/heardis.htm
[2] http://www.comdis.wisc.edu/VCD202/CD202menu.html

Detection of Small Arms Fire.

Bruce Martin
MacDonald Dettwiler & Associates Halifax
Suite 60, Windmill Road, Darmouth, N.S., CANADA, B2B 1L7
E-mail: bruce @iotek.ns.ca; TEL: 902 481-3510 FAX : 902 468-2278
Jacques Bédard
Defence Research Establishment Valcartier
2459, Pie XI Blvd. North, Val-Bélair, Québec, CANADA, G3X 1J5
E-mail: jacques.bedard @drev.dnd.ca; TEL: 418 844-4401 FAX: 418 844-4501

1. Introduction

Since April 1996 MacDonald Dettwiler (MDA) and the Defence Reseach Establishment ValCartier (DREV) have conducted research into the detection, localization, and classification of small arms fire. The result of this research and development is a commercial product, Guardian, that detects, localizes, and classifies small arms fire. The motivation for this development effort was to protect Canadian peacekeepers in locations such as Bosnia. This paper describes small arms fire events, and how the atmosphere affects those events. The detection techniques we investigated are described, along with our conclusions on the most effective detection technique.

2. Small Arms Events

When a small caliber weapon is fired, there are two events that occur. First, the explosion of the gun powder creates a muzzle blast. The intensity is generally strongest in the direction of fire and weakest behind the weapon. The duration of the muzzle blast approximately 2 ms, and has a peak intensity in excess of 155 dB. A typical acoustic event is shown in Figure 1.

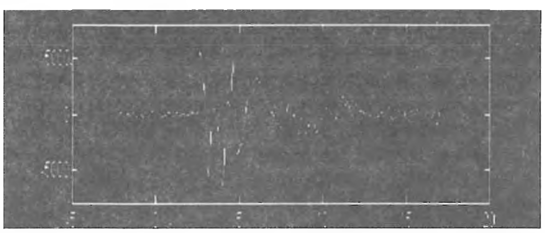


Figure 1. Acoustic Wave from a 7.62 mm rifle fired 200 from the sensor, towards the sensor. (25 ms of data)

The other event that occurs is a shock wave due to the supersonic bullet. The duration of the shock wave is approx 0.5 ms, and it also has a peak intensity in excess of 155 dB. The shock wave travels conically outwards from the line of fire of the bullet, at an angle given by, q = asin (1/M) where M is the bullet Mach number. Typically the initial velocity of the bullet is between Mach 2.5 – 3.0. The shock wave is only detectable in the forward arc of the weapon. Figure 2 shows the shock wave that occurred before the event in figure 1.

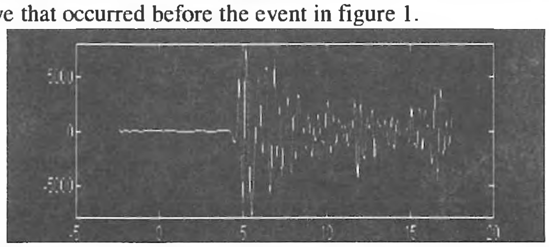


Figure 2.: Shock Wave, 7.62 mm rifle fired 200 meters from the sensor. (25 ms of data)

Because the bullet travels supersonically, the shock wave always arrives at a receiver before the acoustic wave. It also arrives from a different direction. These two properties affect small arms event detection and localization since the acoustic wave can be masked by the shock wave, and because the shock wave does not provide localization information for the source of fire. Therefore, a system like Guardian must be able to detect the weak acoustic events that occur after much stronger shock events in order to localize the source of fire. As the time of arrival difference approaches zero the shock and acoustic waves become mixed. This can cause a system to localize on the shock wave and produce errors that are slightly biased by the shock wave direction. Figure 3 shows the time difference of arrival of shock and acoustic waves as a function of receiver location relative to the shooter at the origin. The direction of fire is along the x-axis (into the page). The red section on the top right indicates that the shock wave does not reach those positions. Figure 4 shows the angular difference of arrival of the events. A system would only have difficulty distinguishing the two events when the time difference is less than 0.05 seconds.

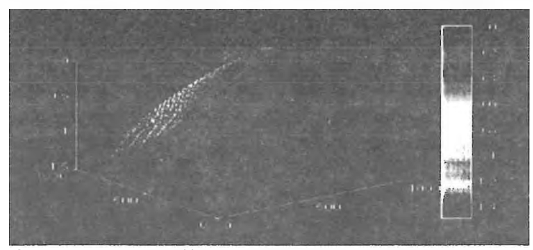


Figure 3.: Time difference of arrival of shock wave before acoustic. X and Y axes units are meters, Z axis is in seconds.

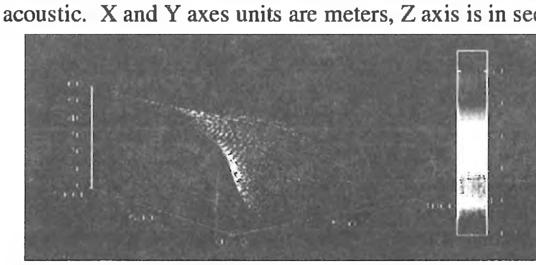


Figure 4: Angular difference of arrival of shock and acoustic waves. X and Y axes units are meters, Z axis is in degrees.

3. Atmospheric Propagation

The atmosphere affects the propagation of the small arms fire events through three primary mechanisms: absorption, refraction, and the geometry of the event. Absorption limits the distance that a high frequency sound can travel, so that frequencies higher than 2 kHz are effectively attenuated after 200 meters[1]. In our application, the atmospheric filter tends to remove some of the frequencies that help distinguish shock waves from acoustic waves.

Refraction occurs when there is a sound speed gradient as a function of height. Typically this occurs when there is a wind or when there is solar ground heating. The effect of refraction is better sound propagation when the receiver is downwind of a source, or when the ground is cooler than the air. This can occur in the evening after a hot day. Conversly, when the receiver is upwind of a source, or on a hot afternoons, very little of the sound will reach the receiver. The energy received in an upward refracting environment can be as much as ten times less than in a downward refracting environment.

The geometry of a small arms fire event has an important impact on the signals at a receiver. Figure 5 shows that the sound measured at the receiver is composed of three contributions: direct wave, reflected wave and ground wave. If there are objects around the receiver, then there are also reflections from the objects. A detector must distinguish these reflections from real shot events. The direct and reflected waves superimpose at the receiver, and can cancel if the source and receiver are close to the ground. The ground wave is the result of the spherical wavefront interacting with a porous ground that is not a perfect reflector. Effectively porous ground acts as a low-frequency filter. Soft ground such as snow, sand or grass will transform the acoustic energy into a long-lasting pulse containing only low frequencies[2].

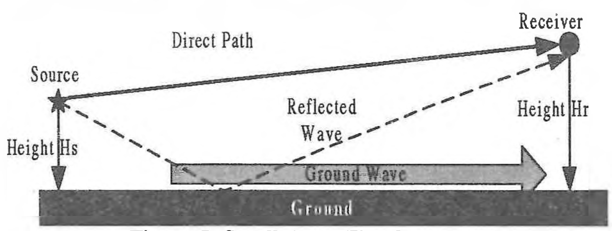


Figure 5: Small Arms Fire Geometry.

The great variety of environmental conditions turns the sound pulses into a continuum of shapes and durations which require compromises during pattern recognition. The acoustic wave in Figure 1 is an example of a situation where the direct and reflected waves do not cancel. The rapid oscillations at the beginning of the figure are the direct and reflected waves, while the wide peak that follows them is due to the ground wave.

4. Detection Techniques

There are several detection methods that are well suited to wide band transients. The simplest of these is an energy detector, which can be improved by filtering the signal into different bands of interest if appropriate. Other detection methods considered were short-time fourier transform (STFT), wavelet transform (WT), and matched filter processing.

As expected, the STFT detector is best suited to narrowband transients since it concentrates energy into narrow frequency bins. The WT detector concentrates energy in 'bins' that correspond to scaled and translated versions of a wavelet basis function [3]. Since many wavelet bases such as Dauchebies4 or 6 have similar shapes to the transients, we expected good performance. Performance was better than the STFT, however, we were unable to find a wavelet basis that significantly improved the probability of detection compared to the energy detector. As well, the wavelet methods are not well suited to real-time processing because they operate on blocks of data, which adds an inherent delay to the detection. A related algorithm, called the adaptive optimum kernel method (AOK)[4] was also investigated because it is a real-time operation and it uses an adaptive kernel that tries to make a best fit to the signals. AOK is computationally very expensive and did not produce significantly improved results over the energy detector.

Matched filter processing involves convolution of the received data with a known signal, and detection of a peak arrival time. When the signal is long compared to the characteristic duration of the noise, this technique can detect wideband signals that are far below the noise. However, for short signals, and in cases where the signal shape correlates well with the noise, there is very little performance improvement compared with the energy detector. Small arms fire events fall into this category.

The conclusion of our analysis was that the energy detector is the most robust, and most effective detector. Because the ground wave is significantly different from the other signals, we are using a two-band detector. One band operates in the ground wave frequencies, the other operates across the full band. The energy detector compares the energy within a short time window to an estimate of the background energy. In estimating the background, it is important to minimize the effects of very high energy signals and to exclude frequency bands that contain more environmental and man-made noise sources than signal.

5. Experiments

The algorithms developed by MDA and DREV have been tested on a database of 10,000+ gun-fire events. There are a large number of false alarms, environmental conditions, and geometries in the database. The data was collected at three locations in CFB ValCartier, and one location near Halifax, NS. Figure 6 shows one of the locations at CFB ValCartier. Overlaid on the figure are results from a Guardian test and evaluation trial conducted by DREV. The vectors on the figure are localizations of shots by Guardian. During the trial Guardian had a very high probability of detection, and localized events within +/-3 degrees. Errors encountered were due to detection of target hits, and to cases where the shock and acoustic waves arrived simultaneously.

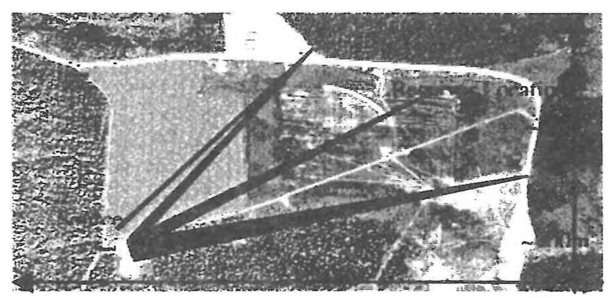


Figure 6: DREV Test Range, with Guardian Results.

6. References

[1] Daigle, G.A. "Atmospheric Acoustics", in *Encyclopedia of Physical Science and Technology, Vol. 2.* Academic Press, 1992. pp 221-235.

[2] Piercy, J.E., Embelton, T.F.W., and Sutherland, L.C. *Review of Noise Propagation in the Atmosphere.* JASA Vol 61, No. 6, June 1977. p1403-1418.

[3] See, for example, del Marco, Stephen, and Wiess, John. *Improved Transient Signal Detection Using a Wavepacket-Based Detector with an Extended Translation-Invariant Wavelet Transform.* IEEE Transactions on Signal Processing, Vol 45, No. 4, April 1997.

[4] Jones, Douglas L. *An Adaptive Optimum Kernel Time-Frequency Representation.* IEEE Transactions on Signal Processing, Vol. 43, No. 10, October 1995.

The Evaluation of Some Temporal Sampling Strategies in Community Noise Measurement

Richard J. Peppin
Scantek, Inc., 916 Gist Ave, Silver Spring, MD 20910
Tel: 495-7738 Fax: 301 495-7739
and
Svein-Arne Nordby and Roberto Borjabad
Norsonic, PO Box 24 , N-3420-Lierskogen, Norway
Tel: 473285 8900 Fax: 47 32 852208

1.0 Introduction

The measurement of community noise, determining outdoor sound pressure level over the course of some time, is probably the most common type of noise measurement. Consultants, environmental officers, police officers, and people who run airports do these measurements. Sometimes, the people who make, or analyze, the measurements understand the complexities of the measurement. But few do. If people are asked to measure the same exact noise at the same exact location, the chances of them getting the same results are slim. The reasons are many: the microphone frequency response may be different, the type, or accuracy, of their sound level meters may be different, the frequency weighting may be different, and how they measure, or what or how sample, may be different. We emphasize this latter category in this paper.

As we all know, the variability of community noise is large. Two people taking measurements for two different time periods will probably not get the same answer. Of course the longer you measure, the better chance the average of two measurements will match. (We have a theory that if we measure the noise for 10,000 years and measure the Leq all reading will coalesce.) Of course the longer averages remove information about detail of measurement. (Perhaps that is why the US EPA used a metric known as a "yearly Ldn." It removed any semblance of reasons for people to complain, any reasonable budget to measure, and any chance loud, brief, disturbing noise to be a problem.)

So how to measure time varying sounds? What metric does one choose? Fortunately, some have no choice. Often the code or regulation might say, "measure the noise for one hour, etc., and take the Maximum Fast reading. If at this setting it is above, say, 55 dBA, the noise producer is dog meat. But unfortunately the problem is worded more vaguely, "Like, I mean, what is the decibel level in these here neighborhoods?"

To quantify an almost random, time-varying signal one could take a statistical approach based on samples of instantaneous (or on latched[1]) maximum (or minimum) fast (or slow) or on samples of an average over the interval period. Hence, given a measurement duration, and a measuring procedure (how to orient the microphone, etc.), the sample method will give values that can statistically be used to characterize the environment[2].

---

[1] Latched is the maximum (or minimum) value found anywhere within the interval.

[2] Keep in mind that, whatever the precision of these measurements in characterizing the sound measured, it bears no relation to the sound just before, or just after measurement period unless other factors are included.

2.0 Discussion

In this study we used the newest real time analyzer from Norsonic. For the same time period, 5 minutes in duration, we sampled the same measured signal over several different intervals to see how they compared. As you will see, and as one would expect, the sample time is critical to the description of the sound. That is, the results will depend on what is sampled and when are samples taken, everything else being equal. The usefulness of any of these results is really more of a political issue than a physical issue, but clearly, the sample time, at least for these common measurements, determines quite a lot about the site characterization.

The measurement was taken over the Oslo - Drammen freeway in Norway, on a footbridge, approximately 8m above the four-lane freeway. The background noise was very low as the measurement point is in the countryside, with no buildings or any other noise source around. A five-minute sample was taken. It could have been longer and we suspect if it were, the data and the conclusions would be quite different. The exactly same sound was sampled for 1s, 2s, ..., up to 5 min and the Leq (over each interval) and Lmax Fast (highest point of each interval during the sample time) was recorded. The number of samples for the period are progressively greater. So there is one 5m sample, five 1-min samples, ten 30s samples, and up to 300 1 s samples.

3.0 Results

We do not present profound results that can be applied to the solution of all community noise measurement problems. Rather discuss the ramifications of the use of different sampling protocols for these specific data.

3.1 Sampling Maximum Levels

If one measures the maximum level occurring in a 5 minute period, and the maximum sound level occurring during many sub intervals, there will always be one sub interval who's maximum equals the overall 5 minute interval. That is, if the maximum recorded A-wtd fast Lp was, say, 80 dB over a five minute period, than sampling every minute and recording the maximum level in each interval will provide at least one interval with a level of 80 dB. This is a much different measurement than one where at the end of every minute sample, the Fast level is recorded. Here it is quite possible to miss the maximum level of the event.

Table 1 presents the sampled Leq and Maxi for three sample intervals. Keeping in mind that the five minute Lmax and Leq were, 83.9 and 76.4 dB, respectively, one can see that recording

Lmax levels at prescribed but not contiguous intervals can lead to strange results.

Table 2 shows the standard deviation (SD) of all the maxima for different time intervals. While we would expect the SD to reduce as the number of samples decreases, the 60 s (five sample) SD was higher than the 30 s (10 sample) value.

If one were to do statistical analysis on the Lmax, as the number of samples increased (interval number decreased) we would expect to see a greater variation. This wasn't what we found. The 1-s sample provided the lowest, but not also the highest fables of the statistical descriptors. See Table 3.

3.2 Sampling Average Levels

The "energy average" or Leq weighs higher levels more than lower levels if one measures the Leq over an interval. For the 30-s intervals in Table 1, had one used Leq at those times for sampling, the results would have varied considerable.

The variation in the values of Leq is higher than the corresponding values of Lmax, perhaps because the Leq is weighted toward the higher levels. Table 2 shows that, as the intervals decrease, the variability of Leq is greater than Lmax for the same data.

Finally, in the statistical descriptors shown in Table 3, the Leq is very consistent in describing highest levels (L1) and least consistent in describing lower levels.

4.0 Conclusions

These data, like most community noise data are unique and probably irreproducible. Nevertheless this study shows that the sample interval and the sample metric is very significant in the interpretation and evaluation of results. It was most useful to be able to measure all different time intervals in the same instrument on the same signal. It helped provide a description and an understanding of the noise environment.

Table 1 A-wtd. Lp recorded at distinct time intervals

Time	30s Leq	10s Leq	1s Leq	30s Max	10s Max	1s Max
0:30	77	76.6	78.8	81.1	79.9	79.4
1:00	78.1	78.7	77.3	82.8	82.1	79.4
1:30	76.3	78.2	76.3	81.7	81.7	77.6
2:00	76.4	74.9	73.2	82	79.6	73.8
2:30	73.7	71.6	70.6	80.2	76	71.4
3:00	72.9	72.3	77.7	80	79.1	79.1
3:30	77.5	74.9	71.7	83.9	82.1	73.1
4:00	77	76.3	74	81.9	79.2	74.8
4:30	75.7	76.9	74.5	83.4	83.4	75.4
5:00	77.2	77.5	75	82.2	80.5	76.4

Table 2 Standard deviation of different samples

Sample Interval	Leq	Lmax
01 s	3.29	3.45
02 s	3.11	3.34
05 s	2.49	2.37
10 s	2.13	1.80
20 s	1.89	1.78
30 s	1.67	1.26
60 s	1.71	1.44

Table 3 Statistical descriptors for different sample times

	60s	30s	20s	10s	5s	2s	1s
Lmax							
L99 approx.	*74.5*	*80.0*	*77.0*	*76.0*	*74.0*	*69.0*	*68.0*
L50 approx.	*77.5*	*82.5*	*82.5*	*81.5*	*80.5*	*79.0*	*78.5*
L10 approx.	*79.5*	*84.5*	*84.5*	*83.5*	*83.5*	*82.0*	*81.5*
L01 approx.	*79.0*	*85.0*	*85.0*	*84.0*	*84.5*	*84.0*	*83.5*
Leq							
L99 approx.	74.0	73.0	73.0	72.5	70.5	67.5	67.5
L50 approx.	77.5	77.5	76.5	77.0	76.5	77.0	76.5
L10 approx.	78.5	79.0	79.0	79.5	79.5	80.0	80.0
L01 approx.	79.0	80.0	80.0	80.5	81.5	81.5	81.5

VERTICAL IMAGING CAPABILITY OF SURROUND SOUND SYSTEMS THROUGH BINAURAL TECHNIQUES

John Klepko

Faculty of Music
McGill University
Montreal, Quebec
H3A 1E3

INTRODUCTION

Recent advances in audio technology have allowed storage and playback of multichannel recordings for music and video applications. With these advances comes a need to research new methods of encoding the soundfield that can take advantage of the enhanced spatial reproduction ability surround-sound has over conventional 2-channel stereophonic systems. The most prevalent multichannel systems in use involve 5 channels of information stored discretely and replayed over 5 separate loudspeakers. The typical loudspeaker layout is shown in Fig. 1.

Most of the recent research involving the imaging capacity of these systems is concerned with the horizontal plane surrounding the listener. Achieving controllable placement of sounds in the 360° surrounding the listener with only 5 channels is a formidable task in itself, but in order to more completely reproduce or simulate a real acoustical space, the vertical dimension should also be represented.

This paper proposes the use of a 5-point microphone array to encode an acoustical event with both horizontal and vertical imaging capabilities. Previous investigations [1] with this system have reported reasonably accurate horizontal imaging so the present study will focus on the vertical imaging.

Complete and accurate vertical imaging is not expected from this system since this is not even possible in natural hearing. Unlike the research in vertical localization, this investigation is based on creating an "illusion" of a sound source position. This is due to the fact that the *intended* vertical positions do not coincide with the *actual* sound sources as there are no loudspeakers above the listener during playback of the recording. As well, it is realized that the difficulty and errors which occur naturally in vertical localization can obscure the test results of this system.

The main objective of this experiment is to ascertain whether this system is able to resolve recorded sound events at these difficult overhead positions. A sub-objective is to discover whether certain sound stimulus types are more often correctly localized overhead than others.

METHOD

The recording apparatus consists of 5 microphones all from the same manufacturer. These were chosen because they use the same transducer element with equal sensitivity (25 mV/p) and noise specifications (10-12 dBA). This way, all microphones have virtually the same dynamic performance which is necessary for consistent level-matching across all 5 channels. Each microphone is assigned to a separate dedicated channel; the left and right are fed from identical super-cardioid pattern microphones, the centre channel from a cardioid microphone. The 2 surround channels are fed from a custom-made dummy-head microphone (with shoulders) incorporating 2 pressure-type omnidirectional microphones fitted into the ear molds. All 5 microphones are amplified by preamps of the same type built into the recording console. All gains are set to the same level (ie. no compensation) and subsequently played back and monitored at matched levels. All 5 channels are recorded to a digital tape-based 8-track at 16-bit word length and 48 kHz sampling rate. The playback setup (Fig. 1) for the experiment consists of 5 loudspeakers of the same type. They are simple 2-way direct radiator loudspeakers. All 5 loudspeakers are placed equidistant from the listening position at 2 metres.

In order to simplify the investigation at this initial stage, the test positions are limited to 5 positions across the median plane as shown in Fig. 2. The sound samples included 5-second samples of; male speech, 2 separate percussion instruments (shaker, sleigh bells) and 1/3-octave band filtered noise played through a small 2-way loudspeaker. The center frequencies were; 0.25, 0.5, 1, 2.5, 4, 6, 8, 10, 12.5 and 16 kHz. All 5-second samples were divided into 3 short, repeated phrases (or bursts). They were recorded with the 5-point microphone array from a distance of 1.5 metres. In an attempt to simulate a practical application of this system, the recording was performed on a large concert hall stage. These sound sample types and positions were randomly reorganized and re-recorded onto a test tape. The total number of sample points was 65 and the duration of the test was about 20 minutes. 5 subjects (with no reported hearing defects) were involved in the test. They were instructed to note down the perceived position (along the median plane) of the sound event. The subjects were seated but not restricted from moving.

FIGURE 1

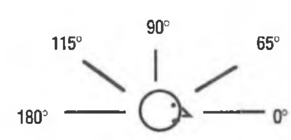


FIGURE 2

RESULTS

Figure 3 compares the 2 general sound types simplified into 3 positions (with above = 65°, 90° and 115°). Figure 4 compares the percentage of correct responses for the 5 different positions for all sound types. Figures 5 and 6 compare the percentage of correct responses for the natural sounds and narrow-band noise respectively.

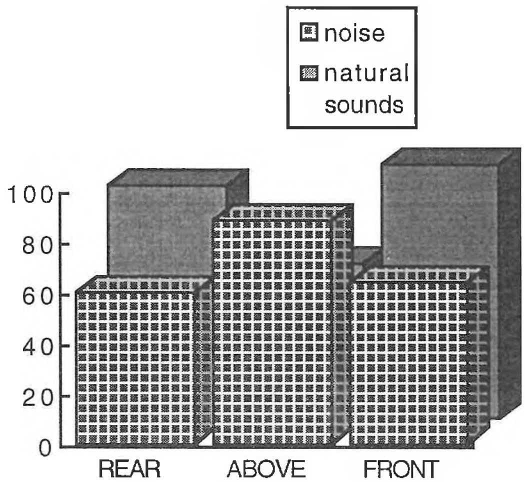


Figure 3

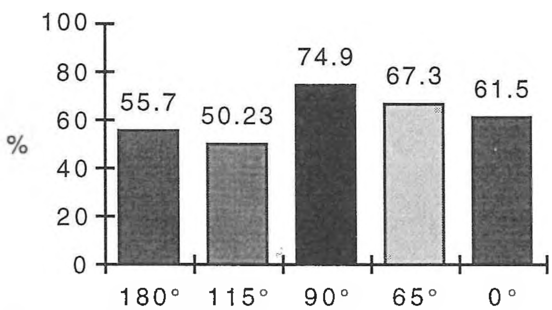


Figure 4

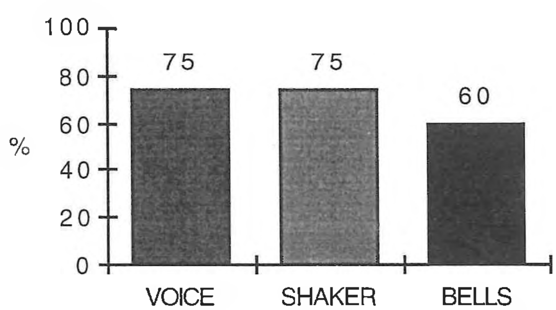


Figure 5

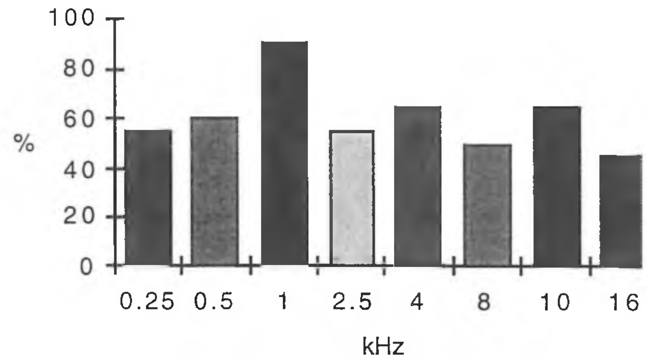


Figure 6

DISCUSSION

From these results it can be shown that it is possible to perceive overhead sounds despite there not being a sound source (loudspeaker) above the listener. The high value for the direct overhead position (Fig. 4) was due to the noise samples which were more often localized above. The natural sounds (i.e. voice, shaker, bells) did much worse for the overhead sounds, but scored excellent for the front and rear positions. The explanation for these results may be that the natural sounds are more familiar and are "expected" to be on ground level rather than above. The narrow-band noises are synthetic sounds which are unfamiliar. These intangible sounds carry no expectations possibly allowing the listener to believe the illusion of an overhead position. But it should also be noted that the natural sound types scored the highest overall percentages over the noise types which is due to the good localization for the front and back positions.

As expected, the low frequency noise bands were difficult to localize, especially overhead. But part of this can be explained by the natural hearing system's difficulty with such sounds. However, there was no trend towards better localization with increasing center frequency as reported by Gardner [2]. Perhaps the sound samples were too short. (Informal listening experiments with longer samples that were not randomized provided significantly better results).

CONCLUSION

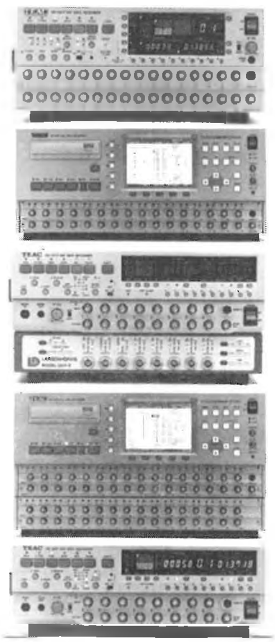
A new method of encoding the soundfield in its total spatial dimension has been presented with particular attention to overhead imaging capability. Although not very precise, this investigation shows that the proposed system has the ability to represent overhead sounds to some controllable degree. More investigations of different sound types are planned.

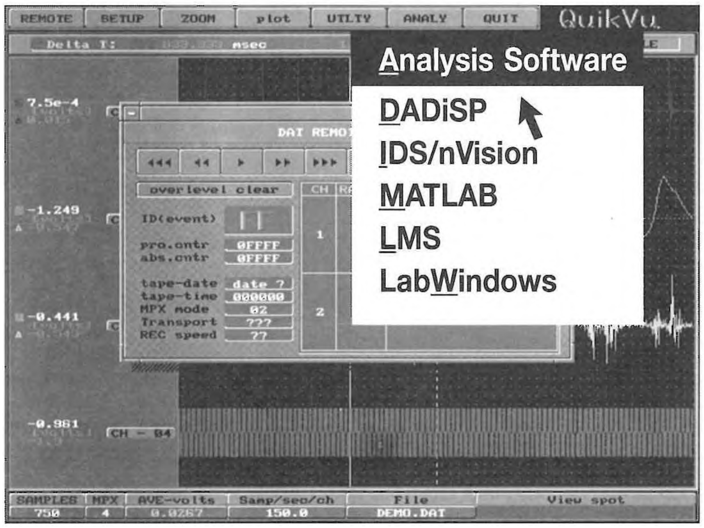
REFERENCES

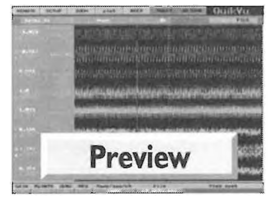
[1] KLEPKO, J., "5-Channel Microphone Array with Binaural Head for Multichannel Reproduction", AES Convention, New York, 1997, preprint # 4541.

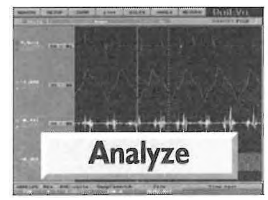
[2] GARDNER, M., "Some Monaural and Binaural Facets of Median Plane Localization", JASA, vol. 54, no. 6, 1973.

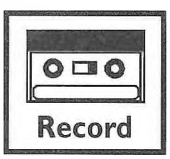
Choice. Convenience. Cost.
No one else compares.











Accurately recording test data is a critical issue for every test engineer. You only get one shot, because repeats are expensive, if not impossible. That's a fact we understand. That's why TEAC offers more combinations of affordable recorders, data analysis programs and downloading options than anyone else. For digital recording we have seven different compact, portable DAT models with extended dynamic range and high signal-to-noise ratios.

For Hi-8mm recording, our RX series offers a very low cost-per-channel, in either a 16 or 32 channel model. Using multiple units, you can synchronously record up to 128 channels. These units feature S/N ratios greater than 80dB, a 20kHz bandwidth and up to 60 minutes of record time. And using on-board, menu-driven programming, you can easily record up to 7GB of data.

And you'll appreciate TEAC QuikVu®, a real-time PC or PCMCIA software program that lets you preview test data before you record. It allows you to monitor data in real time as well as archive or transfer data to other systems. And speaking of software, TEAC offers another major advantage.

We have relationships with all the major data analysis software companies, so our recorders are compatible with their programs. That makes your data analysis much easier ... much faster. Just pick your favorite analytical program and go to work.

Choice. Convenience. Cost.

We're the only one who can give you a total solution for virtually any application. In fact, if we don't have it, you probably don't need it.

TEAC®

INSTRUMENTATION DATA RECORDERS

Dalimar Instruments Inc. *At your service since 1986*

193, Joseph Carrier, Vaudreuil-Dorion, Quebec, Canada J7V 5V5 Tel.: (450) 424-0033 Fax: (450) 424-0030
90 "C" Centurian Drive, Suite 206, Markham, Ontario, Canada L3R 8C5 Tel.: (905) 948-8345 Fax: (905) 948-8344
E-mail: info@dalimar.ca Website: www.dalimar.ca

SYSTEM 824 SLM/RTA

Five sophisticated acoustical instruments in One!

1 **Integrating Sound Level Meter** meeting Type 1 Standards with simultaneous measurement of sound pressure levels using fast, slow, and impulse detector, and simultaneous A, C, and flat weighting. It measures 48 sound pressure parameters at once! All this with a 105 dB linearity range!

Simple Sound Analyzer with simultaneous sound pressure level measurement and real-time 1/3 octave frequency analysis.

Logging Sound Level Meter permits data gathering of broadband sound pressure levels and frequency spectra over user-defined time intervals.

Real Time Frequency Analyzer with 1/1 and 1/3 octave analysis over a 20kHz frequency range and optimized for characterizing steady-state or high speed transient events.

5 **Fast Fourier Transform Analyzer** with 100, 200, and 400 lines resolution and 20kHz range for specific frequency investigations.

For use in a wide variety of applications

Research and Development

- Building Acoustics
- Sound Power Determination
- Vibration measurements
- Statistics
- Simple Point Shoot
- Transient Capture

Environmental

- Aircraft Noise
- Industrial Noise
- General Surveys
- Transportation Noise
- Community Noise

Worker Safety

- Noise Exposure Measurements
- Work Place Surveys
- Machinery Noise
- Audiometric Calibration
- Simultaneous C minus A Measurements

Listen with Larson•Davis

Dalimar Instruments Inc.

At your service since 1986

193, Joseph Carrier, Vaudreuil-Dorion, Quebec, Canada J7V 5V5 Tel.: (450) 424-0033 Fax: (450) 424-0030
90 "C" Centurian Drive, Suite 206, Markham, Ontario, Canada L3R 8C5 Tel.: (905) 948-8345 Fax: (905) 948-8344
E-mail: info@dalimar.ca Website: www.dalimar.ca

Total and Relative Duration as Cues to Surface Structure in Music

Michael E. Lantz & Lola L. Cuddy
Department of Psychology
Queen's University, Kingston, Ontario, Canada

1. INTRODUCTION

The purpose of our research is to understand perceived *pitch structure* in music. Listening to a melody establishes a pitch structure, generally known as a 'sense of key', and we can easily detect a tone that sounds sour or does not fit into that structure.

Krumhansl and Shepard (1979) developed a technique, the probe-tone technique, to quantify perceived pitch structure. A sequence of tones is played to a listener, followed by a probe tone, one of the 12 chromatic divisions of the octave. The listener rates how well the probe tone fits the sequence just heard. The sequence is repeated with a different probe tone at the end, until all 12 chromatic tones have been presented and rated. The resulting profile of 12 ratings is then examined for evidence of perceived pitch structure. Krumhansl and Kessler (1982) showed that key-defining sequences yielded a hierarchically structured probe-tone profile—ratings were highest for the key tone, or tonic; followed by two other tones forming a major triad with the tonic; followed by the other tones in the key; followed finally by tones not in the key.

The hierarchy of probe-tone ratings is consistent with descriptions of pitch structure provided by Western music theory. The ratings correspond to music-theoretic proposals concerning the relative salience and function of tones within a key. An important question arises concerning the psychological processes underlying this correspondence: Is the hierarchical structure acquired through prolonged exposure to a particular musical system, or can the structure be readily perceived through cues available at the musical surface?

Two possible candidates for surface cues can be considered. The first candidate is the frequency of occurrence of tones within a composition. Knopoff and Hutchinson (1983) and Youngblood (1958) (tabulated by Krumhansl, 1990) provided counts of the number of occurrences of each tone in many compositions across many eras of Western tonal music. For purposes of the counts, all compositions were transposed to one key. It was found that the distribution of frequencies with which tones were sounded was very similar to the probe-tone profile for that key. The tones that occurred most often were the most important tones in the probe-tone hierarchy and the tones that occurred least often were the least important. The second candidate is the duration of the tones. Unpublished results from our laboratory showed high correlations between the duration of tones and their frequency of occurrence across several Western compositions across several centuries. Listeners may be able to rely on both frequency of occurrence and duration to help them perceive pitch structure.

To test the idea that pitch structure could be perceived through cues at the musical surface, we created sequences that did not conform to the conventional Western tonal context. Certain tones within each sequence were designated to be consistently longer in duration than the other (shorter) tones, while the shorter tones would consistently occur more often. With the probe-tone technique, we have previously shown that the longer tones in similar contexts received systematically higher ratings (Lantz & Cuddy, 1996). This previous work, however, did not control the *total* duration for which the longer and shorter tones were presented. In the present experiment, we manipulated the total duration of the shorter, more frequent, tones to determine if the cumulative durations of the shorter tones would overcome the perceptual advantage of the longer tones.

2. METHOD

2.1 Participants

Forty-eight listeners (mean age = 19.6 yrs, range 17-30 yrs) from an Introductory Psychology class took part for course credit. Half the listeners (22 women, 2 men) had at least Royal Conservatory (Toronto) Grade VIII music training or equivalent. The other listeners (16 women, 8 men) had much less music training, ranging from no training to five years of high school music classes.

2.2 Sequence construction

Sequences were generated from six-tone tonesets. Each toneset contained the tones of the tonic major triad from two maximally distant keys. Thus, the tonesets did not correspond to any scale within the conventional Western tonal system. The three tones of one major triad were always longer in duration (Triad Du) than the tones of the other triad. Each tone in Triad Du occurred three times in each sequence at 1000 ms per occurrence. The three tones of the other major triad were shorter but occurred more frequently (Triad Fr). The total duration of Triad Fr was systematically increased across three levels from half to one and a half times the total duration of Triad Du. Table 1 shows that the increase was accomplished by either (a) increasing frequency of occurrence, holding duration constant (125 ms); or (b) lengthening the duration of each occurrence (from 125 ms to 375 ms), holding frequency of occurrence constant. Tone order in each sequence was quasi-random. The same tone did not occur twice in succession and successively adjacent tones were at least two semitone steps apart.

Table 1. *Six sequence conditions were created by holding the duration and frequency of occurrence constant for Triad Du tones, while increasing the total duration of Triad Fr tones. Total duration of Triad Fr tones increased by* ***(a)*** *increasing the number of occurrences, and* ***(b)*** *increasing the duration of each occurrence of Triad Fr tones.*

Durations are 𝅗𝅥 = 1000 ms, ♪. = 375 ms, ♪ = 250 ms, 𝅘𝅥𝅯 = 125 ms.

Condition		Triad Du		Triad Fr	Total Duration
(a)	Level 1	𝅗𝅥	vs	𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯	Du > Fr
	Level 2	𝅗𝅥	vs	𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯 𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯	Du = Fr
	Level 3	𝅗𝅥	vs	𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯 𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯	Du < Fr
(b)	Level 1	𝅗𝅥	vs	𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯	Du > Fr
	Level 2	𝅗𝅥	vs	♪♪♪♪	Du = Fr
	Level 3	𝅗𝅥	vs	♪. ♪. ♪. ♪.	Du < Fr

2.3 Apparatus

Sequences were Musical Instrument Digital Interface (MIDI) files. Presentation of sequences and collection of responses were controlled by a Zenith Z-200 computer running 'Midiplay' software (Ergincan, 1993). All tones were pure tones in the range of C4 to B4, and were generated by a Yamaha TX802 FM tone generator with MIDI input from the Z-200. Sequences were delivered through Sennheiser 480 headphones in a sound-attenuated chamber.

2.4 Procedure

Listeners were tested individually with the probe-tone technique. Each of the six sequences—one for each of the six sequence conditions in Table 1—was heard 14 times in succession. The first two trials were practice trials with the probe tone selected randomly. The 12 experimental trials were each followed by one of the 12 chromatic tones in random order. Listeners rated how well each of the probe tones 'fit' the context of the sequence on a scale of '1 - fits very poorly' to '7 - fits very well'. Listeners then filled in a form about their music training.

3. RESULTS AND DISCUSSION

No differences between the two levels of music training were found. Probe-tone ratings for each listener for each condition were entered into the data file in the order (1) three tones of Triad Du, (2) three tones of Triad Fr, and (3) six tones that did not occur in the sequence. Two 12 X 3 (Probe-Tone Ratings X Levels of Condition) ANOVAs with orthogonal contrasts were applied to the data. The first analysed the results of the three levels of condition (a), increasing frequency of occurrence of Triad Fr; and the second, the results of the three levels of condition (b), increasing duration of Triad Fr. In both cases, the contrasts of interest were between Triad Du ratings and Triad Fr ratings.

For condition (a), shown in Figure 1, there was an overall preference for the longer tones of Triad Du, $F(1,47) = 14.43$, $p < .001$. The preference for the long tones was evident even when the total duration of Triad Fr was greater than the total duration of Triad Du. The difference between Triad Du and Triad Fr ratings did not change significantly as the total duration of Triad Fr increased. Increasing the total duration of the shorter tones by increasing their frequency of occurrence did not increase their perceptual salience.

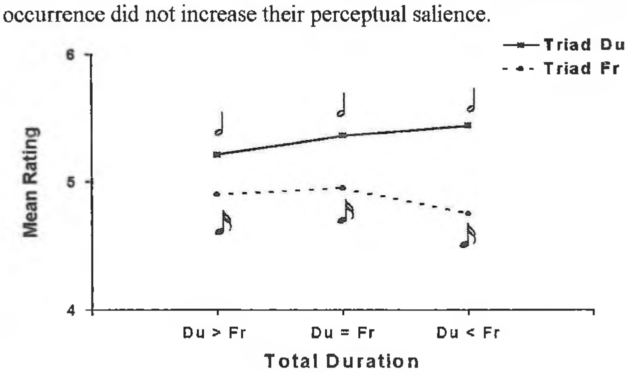


Figure 1. Mean ratings for Triad Du and Triad Fr as the total duration of Triad Fr is increased by increasing the frequency of occurrence of Triad Fr tones.

For condition (b), shown in Figure 2, the preference for the longer tones was evident again at level 1, $F(1,47) = 30.01$, $p < .0001$. However, the preference disappeared when the duration of the shorter tones approached that of the longer tones ($p > .20$ at both levels). The differences between Triad Du and Triad Fr ratings were significantly smaller at levels 2 and 3 (when the total duration of Triad Fr increased) than at level 1, $F(1,47) = 30.90$, $p < .0001$. The salience of the shorter tones appeared to increase as the duration of each individual tone increased, relative to the longer tones.

The two findings together suggest that longer tones in a sequence are preferred to shorter tones except when the shorter tones approach the duration of the longer tones. In the sequences in which the long 1000 ms tones were opposed to the short 125 ms tones, total duration had little effect on the salience of the short tones. Only when the shorter tones increased in duration to 250 and 375 ms were Triad Fr ratings influenced, approximating the ratings for the longer tones.

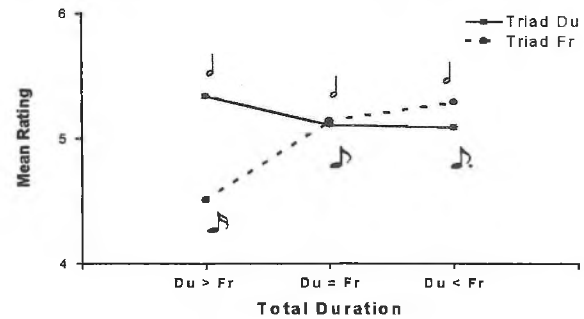


Figure 2. Mean ratings for Triad Du and Triad Fr as the total duration of Triad Fr is increased by increasing the duration of each Triad Fr tone.

A further conclusion is that perceived pitch structure can be derived from surface cues in the sequences. Listeners could not have used highly-learned schema developed through long-term exposure to a music system because the sequences in this study were unconventional. Sequences did not comply with any Western scale or key. Nevertheless, there were systematic preferences for certain tones over others within the probe-tone paradigm. Absolute duration of single tones and relative duration each appeared to be important cues to perceived pitch structure.

4. REFERENCES

Ergincan, F. (1993). *Midiplay.* Unpublished computer program. Queen's University, Kingston, Ontario.

Knopoff, L., & Hutchinson,W. (1983). Entropy as a measure of style: The influence of sample length. *Journal of Music Theory, 25,* 75-97.

Krumhansl, C.L. (1990). *Cognitive foundations of musical pitch.* New York: Oxford University Press.

Krumhansl, C.L., & Kessler, E.J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89,* 334-368.

Krumhansl, C.L., & Shepard, R.N. (1979). Quantification of the hierarchy of tonal function within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 579-594.

Lantz, M.E., & Cuddy, L.L. (1996). The effects of surface cues in the perception of pitch structure: Frequency of occurrence and duration. *Proceedings of the Fourth International Conference for Music Perception and Cognition* (pp. 281-286). Montreal, Canada: McGill University.

Youngblood, J.E. (1958). Style as information. *Journal of Music Theory, 2,* 24-35.

This study was supported by a research grant to the second author from the Natural Sciences and Engineering Council of Canada.

JUDGED DIRECTION OF PAIRS OF OCTAVE-RELATED COMPLEXES (SHEPARD TONES): TRAINING EFFECTS AND REVERSAL

Annabel J. Cohen[1], Michael Lamoureux[2], Richard Fleming[3] Debora Dunphy[4]
[1]Department of Psychology, University of Prince Edward Island, Charlottetown, P.E. C1A 4P3,
[2]Department of Mathematics and Statistics, University of Calgary, Calgary
[3]Department of Psychology, University of Prince Edward Island, Charlottetown, P.E.
[4]Enigma Animations Productions, Vancouver

1. Introduction

Shepard (1964) constructed a set of octave-related sinusoids whose amplitudes were scaled by a bell-shaped envelope which removed the usual cues to pitch height but preserved chroma (e.g., A, A#). For simplicity, he divided an octave into 10 rather than 12 equal units to produce a "chromatic" scale of 10 complex tones. When pairs of these tones were presented, listeners, who had previously demonstrated good frequency directional sensitivity, judged the direction of the second tone to be higher than the first when the chroma of the second tone was 1or 2 semitones higher than that of the first tone. Conversely, when the chroma of the second tone was 8 or 9 semitones higher than that of the first, it was judged as lower. The results were consistent with a proximity principle. Judgements by this group of listeners were consistent for the tones separated by intermediate distances. A group of listeners who had poorer pitch directional judgment ability in the pretest produced less consistency for the smaller and larger intervals as well.

Although typically the direction of octave-complex intervals around 5 to 7 semitones is inconsistent, we noted in our laboratory, that presentation of a particular series of intervals overrode this inconsistency. In such a series, each interval began on the same note and the second note of successive intervals increased or decreased by one semitone. For example, an increasing series would be C-C#, C-D, C-D#, C-E, C-F, C-F#, C-G... For this increasing pattern, we typically observed consistent ascending judgments for intervals of 7 and 8 semitones, although violation to the proximity principle. The descending series led to systematically descending judgments. A similar phenomenon was reported by Tuller and Giangrande (1991). This phenomenon resembles biases induced with the psychophysical method of limits (Fechner, 1860) when the ascending series produces the opposite bias in judgment from the descending series. To explain the phenomena, we considered the possibility that hearing an interval in one direction establishes a propensity to hear the interval in the same direction. Such a proposal has also been made in the context of music perception by Narmour (1990), however only for small intervals. Large intervals, he claims, bias expectations in the other direction.

In the present study, examples of each of the intervals from 1 to 11 semitones were presented to two groups of listeners, differing in level of musical training, who were asked to judge the direction of a second tone of each interval relative to the first. Percentage of judgments in each direction was measured for each interval. In order to determine whether interval direction biassed successive directional judgments, the second order statistics of directional responses were investigated.

2. Method

Stimuli. All 12 tones consisted of 10 sinusoidal components that were related by octaves with amplitudes governed by a fixed, bell-shaped spectral envelope (cf., Shepard, 1964). The lowest component of the lowest tone was 20 Hz. Successive fundamentals were separated by one equal-tempered semitone. Thus, the fundamental of the 12^{th} tone of the set was separated by 11 semitones from the fundamental of the first tone. Each tone was .5 sec in duration, and a .5 sec silence preceded the second tone of the interval.

Procedure. Testing was performed in a quiet room. On each trial the subject indicated whether the two tones ascended or descended by positioning a cursor on a 1 inch square up or down button displayed on a 17" computer screen. The up button was above the down button.

There were 132 different tone pairs resulting from the combination of all possible pairs of the 12 different tones with the exception of repetitions. Each pair was selected in a random order without replacement. This resulted in 12 examples of each of 11 interval sizes of from 1 to 11 semitones.

Apparatus. Tones were generated on a NeXT computer system at a sampling rate of 44.1 kHz and 16-bit resolution (13-bits linearity, Lamoureux & Cohen, 1993). Playback was accomplished through ADS loudspeakers on either side of the listener.

Subjects. There were 34 students, 19 of whom had more than 5 years of musical training.

3. Results

The percentage of ascending judgments as a function of interval size was calculated separately for the two groups of subjects. The pattern of responses was similar for the two groups, but more differentiated for the more musically trained listeners. The second tone was always judged higher when it was from 1 to 3 steps clockwise from the first tone and almost always judged lower when it was 9 to 11 steps clockwise (i.e., 1 to 3 steps counterclockwise). In the vicinity of 6 steps, the judgments were less consistent and the second tone was judged higher about as often as it was judged lower. For the untrained listeners, there also was some inconsistency for the smallest and largest intervals.

Second order statistics. The influence of the direction of a prior interval judgment was investigated by dividing the responses into two classes: those following a preceding "up" response and those following a preceding "down" response. For musically trained listeners, for every interval size, there was a greater tendency to respond up when the preceding interval was judged down, than when the preceding interval was judged up. This reversal effect was largest over the intervals 4 to 6 semitones. The size of the phenomenon at its maximum was over 10%. This tendency was not apparent in the data of the less trained listeners.

The effect was explored in other ways all leading to the same basic result. In one case, only the intervals (1 to 3 and 9 to 11 semitones) which were unambiguous were considered. In another, only the intervals which were unambiguous and also judged "correctly" were considered.

4. Conclusions

The results replicate the original demonstration of Shepard's (1964) circular pitch phenomena. Shepard's conditions differed slightly from those of the present study. In Shepard's study, each tone was presented for only 125 ms followed by 875 ms silence. The results extend to longer tones and shorter ITI's. Shepard also compared different groups of listeners according to a pretest of discrimination of direction of brief sine tone stimuli. In the present study, classifying subjects on the basis of musical experience led to the same pattern of data which had differentiated the extreme groups discussed by Shepard. In the present study, it is likely that the less trained listeners had difficulty in resolving the tones of the smallest intervals and therefore had difficulty in determining their relative direction. In other work, we have shown that scales constructed of adjacent semitone intervals lead to poorer performance of a task of serial order recognition than do scales with wider spacing (Cohen & Frankland, 1990; Parncutt & Cohen, 1995) and that subjects with less musical training perform more poorly.

Finally, our study was motivated by the question, can ambiguity of circular pitch phenomena be overridden by cognitive predispositions to hear successive intervals in the same direction. Observations in our lab, and those reported by Tuller and Giangrande (1991) suggested the possibility that an interval would be judged as rising if the preceding interval was judged as rising. Our results indicate that listeners do not get "caught up in this momentum of the direction of the previous interval". While this might be an adaptation phenomenon, further analyses with third order statistics suggest that it is not, however, too few data points of the type necessary for the analysis prevent drawing any strong conclusions.

The reversal phenomenon demonstrated in this study cannot easily be explained but it rules out the possibility that a rising interval predisposes listeners to hear subsequent intervals as rising. We believe, however, that it should be possible to induce such tendencies through presentation of contour patterns, or macrocontours (Cohen, Trehub & Thorpe, 1989).

Contrary to the notion that a prior direction will prime the same direction, and contrary to Narmour's notion of registral direction, that small intervals imply succeeding intervals in the same direction, the present study showed a bias to hear intervals in the opposite direction to that of the prior interval judged. In relating Narmour's theory to the present study, it should be recognized that Narmour's predictions are for the direction of the interval created by adding a new note N3 to the original interval N1-N2. Instead, the present study looked at the interval between N3 and N4. We might nevertheless have expected the implication of N1-N2 to carry through to N3-N4.

5. References

Cohen. A. J. & Frankland, B. F. (1990). Scale and serial order information in melodic perception. *Can. Acoustics*, *18*, 2-10.

Cohen, A.J., Trehub, S.E. & Thorpe, L.E. (1989). Effects of uncertainy on melodic information processing. *Perception & Psychophysics*., *8*,21-29.

Fechner, G.T. (1860). *Elemente der Psychophysik*. H.E. Adler et al. (trans). N.Y.: Holt, Rinehart

Lamoureux, M. & Cohen, A.J. (1993). Evaluating the NeXT for psychoacoustic and music perception research. *Computer Music J.*, *17*, 78-84.

Narmour, E. (1992). *Analysis and cognition of basic melodic structures*. Chicago: U. Chicago

Parncutt, R. & Cohen, A. J. (1995). Identification of microtonal melodies. *Perc. & Psychoph.*, *57*, 835-6.

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *J.A.S.A.*, *36*, 2346-2353.,

Tuller, B., & Giangrande, J. (1991). Hysteresis in the perception of directional pitch. *J.A.S.A.* abs. 8MU7, *89*, 1988.

6. Acknowledgment and Note

The research was supported by equipment and operating grants from the Natural Sciences and Engineering Research Council made to A. J. Cohen. Correspondence may be directed to acohen@upei.ca.

Level, spectral and temporal cues in children's detection of masked signals

Prudence Allen, Rhiannon Jones, Pamela Slaney & Louise Korpela[1]
School of Communication Sciences and Disorders,
University of Western Ontario, London, Ontario N6G 1H1.

Abstract: Preschool-aged children and adults detected masked signals in four conditions evaluating the role of level, spectral and temporal cues on performance. Performance was similar in fixed and roving level conditions for both age groups suggesting use of level-invariant cues. When the signal was moved to the spectral edge of the masker the performance of the adults improved but that of the children did not suggesting that the children did not benefit from cues provided by the off-center signal. Children's performance worsened when the signal was a narrow band noise rather than a pure tone but the adults' did not, suggesting children's reliance on temporal changes in the masker with the introduction of the signal. Analyses of the stimuli suggested that the children's thresholds corresponded to signal-to-noise ratios at which multiple cues were present at magnitudes that were large enough to be discriminable. These findings were in agreement with post hoc analyses of tonal signals presented in flat- and notched-spectrum maskers used to examine frequency resolving ability in an earlier study (Allen, Wightman, Kistler & Dolan, 1989, J. Speech Hear. Res., 32, 317-324). In that study young children's thresholds were obtained at levels for which level, frequency and periodicity values differed between the signal and no signal stimuli at magnitudes that would be discriminable, independent of the spectral qualities of the masker. [Work supported by NSERC].

Introduction

Children's masked detection thresholds are higher than those of adults and do not mature until children reach school-age. To date, no satisfactory explanation for these observations has been found. Explanations based on immature attention have been useful but are often based upon the assumption that children listen for the signals in a frequency selective manner and that their attention bands are either too broad or the child is unable to focus attention consistently at the signal frequency. However, it is possible that children may not listen in a frequency selective manner. Children may attend to non frequency specific cues that are available primarily at higher signal-to-noise ratios.

The first study reported here examined several such potential cues including average frequency (or overall pitch), periodicity and overall level. These cues were evaluated through comparisons between selected combinations of data obtained in four experimental conditions. The second study was retrospective and consisted of acoustic measurements of level, average frequency, and periodicity in stimuli used to evaluate children's frequency resolving abilities in a previously published notched-noise masking study (Allen et al., 1989). The goal was to determine what acoustic cues may have been available to the children and adults at their measured threshold levels.

Experiment 1: Signal in noise detection cues.

Preschool aged children and adults participated in four conditions in a repeated measures design. Performance was evaluated in pairs of conditions designed to estimate the role of level, frequency/pitch and periodicity.

Method. Signals were 328 msec in duration, gated simultaneously with a noise masker (800-1200 Hz bandwidth). The condition 1 signal was a 1000 Hz tone, spectrally centered in a fixed level (30 dB N_0) noise masker. Condition 2 was identical to condition 1 but level was roved $\pm$ 5 dB, re: 30 dB N_0. The signal for Condition 3 was a 1175 Hz tone which was therefore spectrally off-center in the roving level masker. The Condition 4 signal was a noise band, 50 Hz wide, centered at 1000 Hz. Masker level was roved.

Comparison of performance in conditions 1 and 2 enabled evaluation of the role of absolute level information by comparing detection of a 1000 Hz tonal signal in fixed and roving level conditions. Comparison of conditions 2 and 3 assessed the role of spectral (pitch) information by comparing detection of signals that were spectrally centered or off-center in the roving level masker. Conditions 2 and 4 were combined to evaluate periodicity cues by comparing detection of a pure tone and a noise signal, both spectrally centered in a roving level noise masker. As well, an acoustic analysis of all signal plus masker combinations was performed to determine changes in average frequency, periodicity and level as signal to noise ratio was varied.

Listeners were 11 preschool-aged children (3.7 to 5.3 years, Mean = 4.6) and 6 adults (24.7 to 25.8 years, Mean = 25.2 years). Of the 11 children, 6 produced data to which fitted psychometric functions (see Allen & Wightman, 1994, J. Speech Hear. Res., 37, 205-215, for details) showed improvements with increasing signal to noise ratio and 75% thresholds within the range levels tested. The data from these children and the adults were submitted to a Multivariate Analysis of Variance. The data from the remaining 5 children were examined independently. Inclusion of their data in individual conditions did not produce a change in the pattern of results

Results and Discussion. Children's thresholds were consistently higher than those of the adults, $F(1,10) = 193.88$, $p < .0001$, and the slopes of their psychometric functions were shallower, $F(1,10) = 58.8$, $p < .0001$. The age by condition interaction was significant only for thresholds, $F(3,30) = 3.5$, $p = .021$. Threshold data were therefore subjected to the three planned comparisons noted above but slope estimates were not. Mean thresholds (dB) and standard deviations are shown below.

Condition:	1	2	3	4
Children	33.76 (2.8)	29.21 (1.6)	27.29 (1.6)	33.87 (2.1)
Adults	19.14 (3.5)	18.83 (3.2)	11.06 (3.0)	18.35 (1.9)

Level cues. Neither the adults nor the children showed a significant change in threshold when the level was fixed (1) Vs roved (2) suggesting that listeners in both age groups were able to use level invariant cues for detection.

Spectral cues. Comparisons of performance when the signal was spectrally centered (2) or placed at the edge of the masker (3) showed no significant differences for the children. Adult listeners showed a slight but significant improvement in thresholds when the signal was placed at the spectral edge of the masker. Acoustic analyses of the signal plus masker combinations showed that the overall pitch of the noise masker changed when the signal was added at signal to noise ratios above 20 dB, higher than where most adult thresholds were measured. This argues against their use of

[1] Experiment 1 is based on Allen, P., Jones, R., & Slaney, P., The role of level, spectral and temporal cues in children's detection of masked signals, J. Acoust. Soc. Am., in press; Experiment 2 is from an unpublished master's thesis by L. Korpela (1998).

pitch information to detect the signals. It is more likely that the adult listeners may have been able to shift their attention to the spectral edge of the masker where off-frequency listening could have improved the signal-to-noise ratios. The children, who may have reduced inability to selectively attend to specific frequency regions, were likely not able to use this additional cue.

Periodicity. Children's thresholds were approximately 5 dB higher when the signal was a narrow band of noise (4) compared to when it was a pure tone (2). The performance of the adults was unchanged. This suggests that the children may have been more sensitive to the changes in the temporal fine structure of the masker produced with the addition of the signals. Acoustic analyses of the stimuli showed that children's thresholds in all 4 conditions corresponded to the signal-to-noise ratios at which the masker alone and signal plus masker stimuli differed in overall periodicity.

Conclusions. Previous studies (Allen & Wightman, 1992, J. Speech Hear. Res., 35, 222-233) have shown that children require greater spectral shape contrast to detect and discriminate between complex signals. The overall higher thresholds for the children observed in this further argue that children's thresholds may have been elevated because they required greater cross channel differences (higher s/n ratios) to discriminate between a noise alone and signal plus noise stimulus. Also, the results suggested that periodicity may provide useful information to the children.

Experiment 2: Post-hoc analysis of signals in notched-noise masking study (Allen et al., 1989).

An acoustic analysis of signals used in a previous study of frequency resolution in which flat- and notched-spectrum masked thresholds were obtained, was conducted to determine what acoustic cues may have been present at the threshold signal-to-noise ratios measured.

Method. The signal plus masker combinations at which preschool-aged children and adults masked thresholds were obtained (Allen et al., 1989) were analyzed for changes in the acoustic features noted in Experiment 1. Two conditions were evaluated: Signal plus flat and notched spectrum noise maskers. Notch width was 40% of the signal frequency. Spectrum level was 40 dB. Signal to noise ratios corresponding to detection thresholds for adults and preschool-aged children were evaluated as was an intermediate signal-to-noise ratio. Three frequencies were analyzed, 500, 2000, and 4000 Hz.

Signals were analyzed as a whole unit, without filtering, and at several filtered conditions meant to simulate frequency selective listening. Filtering was set at 1000, 500, and 250 Hz bandwidths to evaluate the potential impact of varying degrees of frequency selective attention.

Results and Discussion.

4000 Hz.

Flat spectrum masker. At the level at which children's thresholds were observed for the 4000 Hz signal in the flat spectrum masker (79 dB) there were large level (6dB) and periodicity (.23) differences in the noise alone and signal plus noise stimulus without any degree of frequency selective listening (filtering). Frequency differences were small (69 Hz) but may have been discriminable. Had even nominal filtered listening been in place (1000 Hz wide), level differences of 6 dB were present at signal levels much below that where the children's thresholds were measured. At the adult detection threshold a potentially discriminable level difference (4.8 dB) was present but the periodicity and frequency differences were quite small. As filtering was employed, the level cue increased but the magnitude of the other cues remained the same or decreased. This argues that the adults as well as the children could have discriminated between the noise alone and signal plus noise stimulus purely on the basis of level without any frequency selective listening.

Notched spectrum masker: Children's thresholds had been observed at the same levels as with the flat spectrum masker (79 dB) in spite of the large spectral notch present at the signal frequency. At these levels the cues were large (6 dB level, .37 periodicity difference, and 300 Hz frequency). The magnitude of all three cues decreased with signal level such that at the adult threshold levels (59 dB) there were no frequency or periodicity differences and the level cue was reduced to 2.5 dB. Had either group of listeners used even a nominal filtering (1000 Hz), the magnitude of the level cue would have been equal to the signal level as most of the noise would have been removed and performance should have been observed at much lower levels. This argues that the children, and possibly the adults, detected the signals in these notched maskers with little or no frequency selective listening and a reliance on level differences.

2000 Hz data. Flat masker: At the levels where the children's thresholds were measured (79 dB), there were large level (6dB), frequency (228 Hz) and periodicity (.42) differences. Had 1000 Hz filtering been employed discriminable differences would have been present at much lower intensities of the signal and thresholds would have been lower. At the adult's thresholds (67 dB), the level difference was 2.7 dB and the periodicity and frequency differences were likely not discriminable.

Notched spectrum masker. At the children's thresholds (68 dB) there were likely no discriminable cues without the need for at least some filtering. With 1000 Hz filtering an 8 dB level cue was available as well as a periodicity cue of .23. Thus, at least some degree of frequency selective listening must have been employed. The adult thresholds were measured at a level (56 dB) where a 2.8 dB intensity cue was available without filtering suggesting that frequency selective listening was not necessary.

500 Hz data:

Flat spectrum masker. At the children's thresholds (76 dB) there was a large frequency cue (631 Hz) in an unfiltered condition. With only minimal filtering (1000 Hz) large level (9 dB) and periodicity (.45) cues were available but the frequency difference would have been reduced. At slightly lower levels the level cue would have been small and likely not detectable. For the adults, there was a 2.7 dB level cue available without filtering. The magnitude of this cue would not have changed significantly with filtering.

Notched spectrum masker. At the level of children's thresholds (68 dB) a frequency difference (158 Hz) was available without filtering but the periodicity and level cues would have been below threshold. Only with filtering of at least 500 Hz bandpass would a large level difference (6dB) have been present. At wider filtering (1000 Hz) only a periodicity cue would have been present. For the adult threshold levels (51 dB) a 2.3 dB level cue would have been available with no filtering. Cue values at that level did not change with filtering.

General Discussion

In spite of the large differences in threshold values measured from the children with different frequency and masker combinations, there were consistently large level cues available that would have likely been discriminable. Frequency and periodicity cues were also sometimes available at these levels. Frequency selective listening would only have been necessary to achieve these discriminable level cues at the very low frequencies. For adults level differences of 2.0 dB or greater were consistently present regardless of frequency or masker. Thus, level differences in the signal plus masker and masker alone stimuli could have been responsible for both children's and adult's detection performance.

Notched-noise and quiet detection thresholds as early indicators of auditory system damage in school-aged children.

Prudence Allen & Cindy Lee Marples
School of Communication Sciences and Disorders,
University of Western Ontario, London, Ontario N6G 1H1.

Abstract: Children may be exposed to harmful levels of environmental and recreational noise, the negative effects of which may be cumulative over the child's lifetime. Early detection is an important component of hearing health promotion programs but audiometric screening methods may be relatively insensitive to early stages of damage. This paper reports on a preliminary study that examined notched-noise detection thresholds and audiometric thresholds at 1000 and 4000 Hz in 30 elementary- and 30 high-school- aged students as indicators of auditory system damage. Damage due to noise exposure, if present, was expected to be revealed by an elevation in 4000 Hz thresholds for the older listeners, especially when the notched-noise thresholds were examined. Consistent with expectations, the notched-noise detection thresholds obtained from the high school-aged students at 4000 Hz largely fell at least one standard deviation above those obtained from the younger children at that frequency. No such differences were found for the notched-noise thresholds measured at 1000 Hz, or for the audiometric thresholds measured at either frequency. These results suggest that auditory system damage may be appearing relatively early and that this damage may be detectable when non-traditional audiometric measures are used.

Introduction

Children are repeatedly exposed to environmental factors that may injure their hearing. Damage from noise and other toxins is cumulative over the individuals lifetime and, while posing no immediate threat to functioning in the early years, may produce significant impairments of communication later in life. But through early awareness and hearing conservation programs many of the ill effects of hearing loss can be significantly reduced. One important component of a conservation program is early identification.

Damage due to noise exposure is most often observed as a destruction of outer hair cells in the 4000 Hz region of the cochlea. Hair cell loss may be substantial before an individuals ability to detect a sound in that region becomes impaired beyond clinically acceptable levels. Our current methods of early identification include audiometric screening procedures in which individuals are asked to detect 20 dB HL tonal signals at frequencies from 1000 to 4000 Hz (American Speech-Language and Hearing Association, *Guidelines for screening for hearing impairment and middle ear disorders,* 1989). This protocol is useful for identifying individuals with substantial hearing losses who may be need of immediate intervention, but it may fail to identify individuals who are only beginning to show the damaging effects of noise and for whom the hair cell loss may not be at levels that would cause large threshold elevations. The current study examined the potential use of notched-noise detection thresholds as an early indicator of cochlear damage. This method was compared to audiometric threshold assessment, a more descriptive procedure than that used in typical screening programs.

Method

Two groups of children, 30 young school-aged, 7-8 years of age (Mean = 8 years, 1 month), and 30 high school aged children, 15-19 years of age (Mean = 16 years, 9 months) participated. Three protocols were administered.

Audiometric thresholds were obtained at 2 frequencies, 1000 and 4000 Hz thus providing more detailed information regarding hearing status than a screening at 20 dB HL. These thresholds were obtained using standard clinical procedures.

Masked thresholds for a 1000 and 4000 Hz tonal signal were measured in the presence of a 30 dB spectrum level notched noise masker centered at the signal frequency, using a computer driven 2-alternative-forced choice procedure with visual feedback. Signals and maskers were 200 msec in duration. Notch width was 0.4*Cf. The flanking noise bands were 0-800 Hz and 1200-3200 Hz for the 1000 Hz signal, and 1200-3200 Hz and 4800-6800 Hz for the 4000 Hz signal.

Children also completed a brief Hearing conservation questionnaire that examined knowledge and practices in hearing conservation.

Results

Audiometric thresholds showed no significant differences by age or frequency. As there was no significant effect of ear, mean thresholds shown in table 1 were averaged across ear. Standard deviations are also shown. Had a criterion of thresholds at 20 dB or less been used, 2 of the high school and 2 of the elementary aged- children would have failed showing no age-related differences.

Table 1. Mean audiometric thresholds and standard deviations.

	1000 Hz	4000 Hz
Younger	11.0 (5.4)	9.0 (6.70
Older	7.8 (4.7)	6.0 (6.3)

Notched-noise thresholds are shown in Table 2. Results are averaged across ear. Results were similar for both age groups at 1000 Hz but there was a trend for slightly higher 4000 Hz thresholds in the older age group when compared to the younger.

Table 2. Mean notched-noise thresholds and standard deviations.

	1000 Hz	4000 Hz
Younger	44.2 (4.4)	55.59 (7.2)
Older	44.6 (4.4)	59.31 (6.9)

Because greater damage due to noise exposure was expected to reveal itself in slightly higher thresholds for the older listeners, the data from these listeners was compared to that obtained from the younger listeners by normalizing all scores relative to the distribution obtained for the younger listeners at each frequency. In this way, the mean and standard deviation of the younger listeners would fall to 0 and 1, respectively, and the data from the older listeners could be evaluated as deviations from the baseline, or younger children's data. When this was done it was seen that the normalized scores for the older listeners fell at or near the mean of obtained from the younger children at 1000 Hz, ± 1 standard deviation, but at 4000 Hz, the data from the older children tended to fall 1 to 2 standard deviations above the mean. This suggested a trend for increased damage in the older children's hearing at 4000 Hz, but not at 1000 Hz, consistent with expectations of greater damage due to noise exposure.

Evaluation of the hearing conservation questionnaire showed that only 36% of the teenage participants reported their intention to use hearing protection when in noisy environments. This was in contrast to a positive response form 67% of the younger children. Similarly, 80% of the younger children reported that they would take a break from high music levels but only 62% of the teenagers reported this. However, 86% of the younger children and 83% of the teenagers reported that they would turn down the volume of the television if it was too loud. In general, this survey suggests that older school-aged children may be less likely to engage in good hearing health care practices while many of the younger children reported that they would, or at least that at that time of their life, they knew the proper way to respond.

Discussion

The results of this study suggested that while audiometric thresholds failed to show any age-related changes at 1000 or 4000 Hz, the notched-noise thresholds did suggest a potential for greater damage at 4000 Hz in the older listeners. There was no such trend at 1000 Hz further arguing that the increased thresholds at 4000 Hz may have truly reflected greater damage from noise exposure. The trend for elevated notched-noise thresholds at 4000 Hz was also significant in light of the observation that audiometric thresholds were well within the range of clinically acceptable levels.

Over an individual's life time the effects of noise and other toxic agents on hearing are cumulative. High levels of environmental as well as recreational noise can contribute to much of the hearing problems commonly observed in our aging population. It is interesting that if sensitive measures are used, some of this damage may be detectable much earlier than once thought. The responses to the hearing conservation questionnaire from the older listeners, suggest that greater emphasis may be needed in the area of hearing health care. Perhaps if augmented with early detection procedures that are more sensitive to subtle signs of damage, many of the hearing losses commonly encountered in older listeners can be minimized.

The Impact of Signal Bandwidth and Generic Head-Related Transfer Functions for Localization on the Horizontal Plane in Virtual Acoustic Space

G. Robert Arrabito, R. Brian Crabtree, Sharon M. McFadden and Sherry Van Blyderveen
DCIEM, 1133 Sheppard Ave. W., P.O. Box 2000, Toronto, ON M3M 3B9

INTRODUCTION

Auditory localization refers to the ability to determine the location of a source of sound solely on the presence of acoustic information. The localization of sound sources depends on the encoding of interaural differences in the level and time of arrival of the sound at the two ears in addition to spectral and temporal information contributed by the pinnae, head, and upper torso. A common phenomenon in auditory localization is for the listener to make front/back reversals, i.e., perceiving the mirror image of the sound source. In order to reduce front/back reversals, auditory signals must be broadband and contain spectral energy above approximately 3 kHz when presented in the free-field.

The use of directional auditory cuing via a three-dimensional (3-d) audio display to pilots in military aircraft could potentially increase pilot performance by off-loading the visual modality. In order to synthesize the location of a sound in virtual auditory space over headphones, digital filters are used to implement head-related transfer functions (HRTFs) which have been measured from humans or acoustic mannequins for many sound source positions in the free-field. The physical differences between individuals cause HRTFs to differ from person to person. Upon playback of a spatial signal over headphones, the HRTFs could be either those measured from the listener's head ("individualized") or those measured from another head ("generic").

Due to the limited bandwidth of aircraft communication systems, one would expect that front/back reversals would be increased. Such confusions could have potentially serious consequences in practical applications. An experiment in progress is examining auditory localization on the horizontal plane in a 3-d audio display using different signal bandwidths and generic HRTFs. A subsequent auditory localization experiment in the free-field will assist to psychoacoustically validate the results obtained in virtual auditory space.

METHOD

Participants

To date, three women and two men have been tested in a pilot study. The mean age was 33 years. A Békésy audiometric test confirmed that all participants had no more than a 20 dB bilateral hearing loss at any frequency between 125 Hz and 8 kHz.

Stimuli and Apparatus

The stimuli were low pass 3 kHz, high pass 3 kHz, and low pass 14 kHz white noise. These were chosen to allow an assessment of the effectiveness of binaural and spectral cues. Seven different generic HRTFs were used to filter the stimuli in order to synthesize the location of the sound in virtual auditory space.

Participants were tested individually in an IAC sound-attenuating listening booth. The Tucker-Davis Technologies (TDT) PD1 was used to spatialize the stimulus in real-time in conjunction with custom software. The TDT system is a suite of audio equipment (e.g., filters, analog-to-digital converters, digital-to-analogue converters, attenuators, convolvers), under computer control, which is responsible for creating and presenting sound to observers, and for collecting and storing their responses. The stimulus was presented over Stax electrostatic headphones (model SR-Λ Signature) at a sound level of approximately 71 dB(A). A localization judgement was made by pressing a button on a response box whose buttons were arranged in the same configuration as the virtual speaker array.

Procedure

The participant's task was to identify the perceived location of the stimulus that had a 300 ms duration. The stimulus was presented at one of eight static virtual positions on the horizontal plane at 45 degree intervals with the 0 degree azimuth position located in front of the participant. Each block consisted of 8 practice trials followed by 40 experimental trials, with each position used once in the practice trials and five times in the experimental trials. Each trial began by flashing a 0.5 second warning light on the wall in front of the participant followed by a 0.5 second delay prior to the presentation of the stimulus. A maximum of ten seconds was given to make a response corresponding to the virtual speaker that had emitted the stimulus. Each block was comprised of one of the three stimuli and one of seven HRTF conditions. The order of trials and blocks was presented according to a balanced Latin square design. A session contained 21 blocks. Participants completed four sessions each on separate days. No feedback was given to the participants as to the correctness of the localization judgements.

RESULTS AND DISCUSSION

A preliminary analysis of the pilot data suggests that localization performance measured by percent correct, type of reversal (e.g., front/back, left/right, and diagonal), and response time was not significantly affected by the choice of HRTF or stimulus. These findings were not expected. One might expect that performance would be significantly affected by the process of measuring the HRTFs, as measurement techniques differ across laboratories and are motivated by the different goals of the investigators (see references 2, 3, 5-27 of (Moller, Sorensen, Hammershoi & Jensen, 1995)). Some of the parameters that vary significantly in the measurement of HRTFs are type of test stimulus (e.g., sinusoidal tones or noise bursts), the point in relation to the ear canal where the measurement is made (e.g., at the blocked ear canal or a point somewhere along the ear canal), and the number of source positions. For example, Wightman and Kistler (1989) measured 144 source positions while Bronkhorst (1995) measured 967 source positions. A smaller number of measured sound source positions requires that the between sources interpolation procedure be more precise. The seven HRTFs used in this study were each measured differently. As for the choice of stimulus, one would expect that the low pass 3 kHz and 14 kHz stimulus would be significantly different from the high pass 3 kHz stimulus due to the dominant role of interaural time differences (ITD) (Wightman & Kistler, 1992). In that study Wightman and Kistler (1992) reported that when lower frequency ITD cues are present in the stimulus they override interaural level difference (ILD) and spectral shape cues present in low and high frequencies.

Similar results to those reported in this study were obtained by Bronkhorst (1995) who examined auditory localization in the free-field and under virtual conditions as a function of the cutoff frequency of the stimulus. He reported that front/back reversals under virtual presentation on the horizontal plane did not vary significantly as a function of the cutoff frequency regardless of the choice of HRTFs, (individualized or generic). However, under free-field presentation on the horizontal plane, Bronkhorst (1995) found that the number of front/back reversals was statistically affected by the cutoff frequency. This latter finding was also reported by King and Oldfield (1997). The implications of our findings on auditory localization performance in the free-field will be reported as further data become available.

REFERENCES

Bronkhorst, A.W. (1995). Localization of real and virtual sound sources. J Acoust Soc Am, 98(5), 2542–53.

King, R.B. and Oldfield, S.R. (1997). The impact of signal bandwidth on auditory localization: Implications for the design of 3-dimensional audio displays. Hum Fact, 39(2), 287–95.

Moller, H., Sorensen, M.F., Hammershoi, D. and Jensen, C.B. (1995). Head-related transfer functions of human subjects J Aud Eng Soc,43(5), 300–21.

Wightman, F.L. and Kistler, D.J. (1992). The dominant role of low-frequency interaural time differences in sound localization. J Acoust Soc Am, 91(3), 1648–61.

Wightman, F.L. and Kistler, D.J. (1989). Headphone simulation of free-field listening part I: Stimulus synthesis. J Acoust Soc Am, 85(2), 858–67.

THE INTEGRATION OF ACTIVE NOISE REDUCTION AND BINAURAL TECHNOLOGIES IN THE DESIGN OF COMMUNICATION HEADSETS

Christian Giguère*, Sharon M. Abel and G. Robert Arrabito*****

*** Programme d'audiologie et d'orthophonie, Université d'Ottawa, Ottawa (ON), K1N 6N5**
**** The Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto (ON), M5G 1X5**
***** Defence and Civil Institute of Environmental Medicine, Toronto (ON), M3M 3B9**

1.0 INTRODUCTION

This study reviewed the fundamental basis and the technical aspects involved in integrating two relatively new technologies in the design of communication headsets. The first technology, known as active noise reduction (ANR), can improve speech intelligibility by reducing the amount of interfering noise from the environment. The second technology, known as binaural technology, allows the simulation of 3D auditory displays, which can enhance speech intelligibility and situational awareness over monophonic listening. A typical use of such a device would be inside an aircraft cockpit environment [1,2].

2.0 ANR TECHNOLOGY

Communication headsets with sound attenuation capabilities are often used in situations where an individual must be in contact with others at a remote location while operating in a noisy environment. The most common design is based on a passive circumaural hearing protective device (HPD) fitted with earphones inside the earcups and a boom microphone in front of the mouth. The reduction in environmental noise depends on the attenuation characteristics of the earcups. Passive circumaural HPDs tend to be most effective above 1000 Hz [3], and often do not provide much more than 10-15 dB of attenuation at low frequencies, even in the best designs and under controlled laboratory conditions.

In situations of intense low-frequency environmental noise, the performance of passive communication headsets can be enhanced using ANR technology. A miniature microphone picks up the noise inside the earcup and produces a phase-inverted copy for sound wave cancellation. The active attenuation achieved in this way is currently limited to frequencies below 500-1000 Hz, where it adds to the passive attenuation provided by the earcup. Maximum active low-frequency attenuation in the order of 10-20 dB has been measured over the passive mode [1,4].

The additional low-frequency reduction in environmental noise achieved with ANR headsets points to improvements in auditory perception for signals transmitted through the communication channels. However, this has not been consistently demonstrated in practice [5]. Firstly, ANR devices often show a negative amount of active attenuation of about 3-6 dB in the region of 1000-3000 Hz [3], which can have adverse effects on auditory perception, particularly speech intelligibility. Secondly, the frequency response of the communication channels and the effect of the ANR circuitry on the speech transmission quality are also important determinants of intelligibility. Table I lists additional factors affecting auditory perception with communication headsets.

3.0 BINAURAL HEARING AND TECHNOLOGY

The detection, discrimination and recognition of a signal in the presence of other competing signals and/or interfering noise can sometimes be markedly improved when listening binaurally rather than monaurally. In particular, binaural speech intelligibility in noise has been investigated extensively over the past decades in headphone, sound field and simulated sound-field studies [1,2]. Comparison of binaural listening to monaural listening through the ear with the best S/N ratio showed a constant binaural advantage equivalent to a reduction in speech reception threshold of about 3 dB across a range of experimental conditions [6]. This relatively small decibel advantage could give rise to a substantial speech intelligibility improvement given the very steep slope of the intelligibility function near the 50% intelligibility level, typically 15% per dB for sentence material.

Communication headsets with 3D auditory display capabilities are needed to take full advantage of the binaural system [7]. In addition to improved speech intelligibility in noise, virtual auditory environments and 3D models of the listening space can be created, which can greatly facilitate the monitoring and interpretation of the various acoustic sources of information. The binaural technology procedures [8] necessary to create 3D auditory displays require knowledge of the head-related transfer functions of the listener, and proper equalization of the sound delivery system, typically a set of headphones.

4.0 HEADSET DESIGN

Table II lists the principal technical features to consider in the selection of a suitable ANR headset for binaural technology applications. In a prototype system, the binaural signal processing unit should likely be external to the ANR headset itself. This unit would store the head-related transfer functions of the listener and execute all the operations required to transform each acoustic source into left and right ear signals, given the direction of (virtual) sound incidence. These signals would be presented to the listener via the communication channels of the ANR headset. Depending on the application, the binaural unit would also store a more or less complex model of the 3D auditory environment to be created, and execute all angular adaptations related to the head motion of the listener.

ANR communication headsets based on analog electronics have been commercially available for the past 10-15 years. Table III lists the devices surveyed. All devices support stereophonic listening, a pre-condition for binaural technology applications. However, they differ in several other respects such as the amount of attenuation provided, and the method of processing and adjusting the level of the communication signals [1,2]. The most likely candidates are the Peltor, Sennheiser and

TechnoFirst devices. However, many important technical characteristics including the amount of cross-talk attenuation, the earphone linearity, the electronic noise floor and the degree of interaural earphone matching are unspecified by the manufacturers, and would require extensive testing.

Finally, prototype ANR devices based on digital technology have been tested in research laboratories in the past few years [9], and the first commercial digital ANR headsets have been recently introduced. Since binaural technology is also based on digital signal processing, digital ANR headsets could lead to more completely integrated and compact ANR-binaural systems than analog ANR headsets.

[Work conducted under a contract from the Defence and Civil Institute of Environmental Medicine]

REFERENCES

[1] Abel, S.M. and Giguère, C. (1997). *A review of the effect of hearing protective devices on auditory perception: The integration of active noise reduction and binaural technologies.* Final Report for Contract No. W7711-6-7316/001 SRV, DCIEM, 50 pp.

[2] Giguère, C., Abel, S.M. and Arrabito, G.R. (1998). "Binaural technology for application with active noise reduction communication headsets: Design considerations." Submitted to Acta Acustica.

[3] Casali, J.G. and Berger, E.H. (1996). "Technological advancements in hearing protection circa 1995: Active noise reduction, frequency/amplitude-sensitivity, and uniform attenuation." Am. Ind. Hyg. Assoc. J. 57, 175-185.

[4] McKinley, R.L., Steuver, J.W. and Nixon, C.W. (1996). "Estimated reductions in noise-induced hearing loss by application of ANR headsets." In: *Scientific Basis of Noise-Induced Hearing Loss,* Edited by A. Axelsson et al. (Thieme, New York), pp. 347-360.

[5] Gower, D.W.Jr. and Casali, J.G. (1994). "Speech intelligibility and protective effectiveness of selected active noise reduction and conventional communication headsets." Human Factors 36, 350-367.

[6] Bronkhorst, A.W. and Plomp, R. (1992). "Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing." J. Acoust. Soc. Am. 92, 3132-3139.

[7] McKinley, R.L., Erikson, M.A.. and D'Angelo, W.R. (1994). "3-Dimensional auditory displays: development, applications, and performance." Aviation, Space, and Environmental Medicine, May, A31-A38.

[8] Moller, H. (1992). "Fundamentals of binaural hearing." Appl. Acoust. 36, 171-218.

[9] Pan, G.J., Brammer, A.J., Zera, J. and Goubran, R. (1995). "Application of adaptive feed-forward active noise control to a circumaural hearing protector." Proc. of Active 95 (Newport Beach, CA), 1319-1326.

Table I: General factors affecting auditory perception with communication headsets

Task:	Detection Discrimination Word Recognition (Virtual) Sound localization
Listener:	Age Fluency Hearing status
Background:	Quiet or Noisy Type, Spectrum and Level of noise S/N ratio
Device:	Attenuation characteristics Passive or Active Communication channel characteristics

Table II: Technical factors affecting the reproduction of virtual audio signals with communication headsets

Listening mode
Volume control
Cross-talk attenuation
Earphone linearity
Earphone frequency response
Electronic noise floor
Coupling to the ear
Interaural matching

Table III: Analog ANR communication headsets surveyed

Peltor ANR Aviation Headset
Sennheiser NoiseGard
Bose Aviation Headset
Bose Aviation Series II
David Clark DCNC Headset
David Clark H1013X
Telex ANR Headset System
Telex ANR 4000
TechnoFirst NoiseMaster

THE EFFECT OF AGING ON SPATIAL HEARING

Sharon M. Abel*, Angela Consoli*, Christian Giguère and Blake C. Papsin***

Dept. of Otolaryngology, University of Toronto*
and Prog. d'audiologie et d'orthophonie, Université d'Ottawa**

1.0 INTRODUCTION

The ability to localize sounds is available at birth.[1] It is known that auditory spatial acuity improves rapidly during the first few years of life but begins to deteriorate by middle age.[2,3] Later deficits may reflect age-related degeneration of the peripheral and central auditory pathways.[4] Such changes could interfere with the encoding of binaural and spectral cues that are essential for accurate sound localization.[5] To test this hypothesis, an experiment was conducted to measure the utilization of these cues from adolescence to old age.

2.0 METHODS AND MATERIALS

2.1 Subjects

Seven groups of sixteen subjects, aged 10-19, 20-29, 30-39, 40-49, 50-59, 60-69 and 70-79 years, respectively, participated. All were screened for hearing loss in the region of 0.5-4 kHz.

2.2 Apparatus

Subjects were tested individually in a semi-reverberant sound proof chamber that modelled real-world listening in a small office. The specifications of the chamber and stimulus generating and loudspeaker systems have been described previously.[6] Subjects responded using a laptop response box with a set of microswitches in the same configuration as the loudspeaker array.

2.3 Procedure

The subject's task was to identify the direction of a 75 dB SPL 300-ms sound (1/3 octave noise band centred at 0.5 or 4 kHz or broadband noise), randomly emanating from a set of four or eight loudspeakers, surrounding her/him at a distance of 1 metre. For the 4-speaker array, speakers were placed either close to the midline or the interaural axis, in each quadrant. For the 8-speaker array, the separation between pairs of speakers, placed within each quadrant, was varied (15, 30, 45 or 60 deg).

One block of sixteen random presentations of the stimulus through each speaker in the array was given for each of the 18 listening conditions. A trial began with a 0.5-s warning light on the response box, followed by a 0.5-s delay, and then the presentation of the stimulus. The warning light was the subject's cue to keep the head steady and fixate a straight-ahead visual target attached to the wall of the booth. A maximum of 7 s was given for the choice of response key corresponding to the speaker that had emitted the stimulus. No feedback was given about the correctness of the judgements.

3.0 RESULTS AND DISCUSSION

Within each condition, accuracy in speaker identification was highest with broadband noise and lowest with the 1/3 octave noise band centred at 0.5 kHz. The greater the number of speakers in the array (4 vs 8), the lower the percent correct. The mean percent correct achieved by the 16 subjects in the 20-29, 40-49 and 60-69 year old groups is shown in Fig. 1 for the 8-speaker arrays. There was an overall trend toward poorer localization with aging.

Table I shows the effect of aging on correct identification of spatial quadrant for the 8-speaker arrays. Low-frequency left frontal (LF) superiority displayed by the 20 and 40 year olds was absent in the elderly. At the higher frequency, a progressive inability to perceive that sounds had come from behind (B) was apparent by middle age. In front (F), accuracy in quadrant identification was higher on the left (L) than the right (R) side of space.

Sound source identification was best when both binaural and spectral cues were available (broadband noise). Inter-aural level differences with 4 kHz were more effective than time-of-arrival differences with 0.5 kHz. With aging, there was a decline in the use of binaural cues, and greater difficulty with front/back discrimination. Left frontal superiority may reflect a right hemisphere advantage either for spatial acuity or for the processing of complex spectral information.[7]

Acknowledgements

Supported by The Physicians' Services Inc. Foundation.

References

1. Morrongiello BA, Fenwick KD, et al. (1994). "Sound localization in newborn human infants." Dev Psychobiol 28: 519-538.
2. Ashmead DH, Clifton RK, Perris (1987). "Precision of auditory localization in human infants." Dev Psychol 23: 641-647.
3. Abel SM, Hay VH (1996). "Sound localization: The interaction of HPDs, aging and hearing loss." Scand Audiol. 25: 3-12.
4. Casey MA, Feldman ML (1982). "Aging in the rat medial nucleus of the trapezoid body. I. Light microscopy." Neurobiol Ag 3: 187-195.
5. Blauert J (1997). Spatial Hearing: The Psychophysics of Human Sound Localization, rev. (MIT Press, Cambridge Mass).
6. Giguère C, Abel SM (1993). "Sound localization: Effects of room reverberation, speaker array, stimulus frequency and stimulus rise/decay." J Acoust Soc Am 94: 769-776.
7. Abel SM, Giguère C, Consoli A, Papsin BC. (1998). "Mirror image reversal errors and left/right asymmetry in sound localization." Submitted to Acta Acustica.

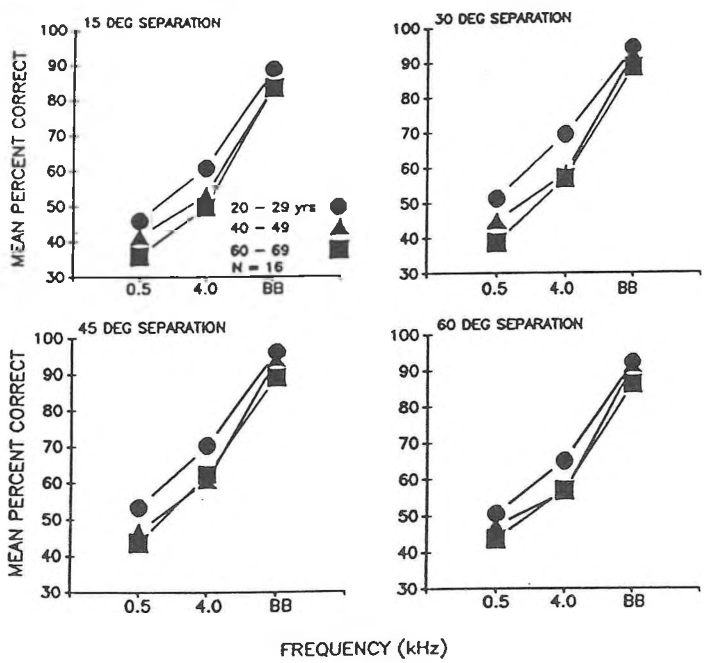


Fig.1 *The effect of age on sound source identification for the 8-speaker arrays.*

Table I. *The effect of age on quadrant accuracy for the 8-speaker arrays.*

Age (yrs)	Quad	1/3 Octave Bands 0.5 kHz L	R		4 kHz L	R		Broadband Noise L	R	
20-29	F	67*	54	*61*	80	72	*76*	99	99	*99*
	B	57	54	*56*	86	84	*85*	97	97	*97*
	Avg	*62*	*54*		*83*	*78*		*98*	*98*	
40-49	F	71	63	*67*	83	72	*78*	98	99	*99*
	B	37	43	*40*	60	60	*60*	96	97	*97*
	Avg	*54*	*53*		*72*	*66*		*97*	*98*	
60-69	F	48	48	*48*	84	77	*81*	97	99	*98*
	B	53	52	*53*	61	62	*62*	92	86	*89*
	Avg	*51*	*50*		*72*	*70*		*95*	*93*	

* Percent correct/quadrant

HIGH-FREQUENCY BONE CONDUCTION AUDIOMETRY USING A PIEZOELECTRIC TRANSDUCER

Marie-Claude Gallichan*, Sofia Ravenna*, Christian Giguère*, Tony Leroux*, Lixue Wu** and George S.K. Wong**

* Programme d'audiologie et d'orthophonie, Université d'Ottawa, Ottawa (ON), K1N 6N5
** Institute for National Measurement Standards, National Research Council of Canada, Ottawa (ON), K1A 0R6

1.0 INTRODUCTION

The bone vibrator used with most clinical audiometers, the Radioear B-71, is limited to threshold measurements not exceeding 4 or 6 kHz. A bone vibrator operating at higher frequencies would allow to assess more accurately the exact nature of a hearing loss, conductive or sensorineural, and would provide an alternative to air-conduction transducers for high-frequency audiometry applications. Unfortunately, the frequency response of the B-71 shows steep resonance peaks and sharp drops at high frequencies [1]. The vibration output of electromagnetic bone vibrator devices, like the B-71, also tends to decrease with frequency above 6-8 kHz. This study evaluated a new bone vibrator device based on the piezoelectric effect, developed at the National Research Council of Canada, and designed to overcome these limitations [2].

2.0 METHODS AND MATERIALS

2.1 Subjects

Eight subjects, aged from 18 to 30 years (mean = 24 years), participated. This corresponds to the age range for establishing the reference equivalent thresholds for clinical audiometry under ANSI S3.6-1996 [3]. Subjects had no previous medical history of hearing problems. They all had normal hearing thresholds in the region of 0.25-12 kHz, and normal middle ear function.

2.2 Instrumentation and calibration

Two bone vibrators were under study, the new piezoelectric device from 1 to 12 kHz, and the conventional Radioear B-71 from 1 to 4 kHz for comparison purposes. The stimuli were generated by an Audioscreen Essilor C76FX3 portable extended high-frequency audiometer. The B-71 was connected to the bone vibrator port at the back of the audiometer. In order to supply stimuli up to 12 kHz, the piezoelectric device was connected to one of the two earphone ports at the back of the audiometer. Preliminary tests showed that this arrangement would provide sufficient vibration output to measure bone-conduction (BC) thresholds for normal subjects.

The two bone vibrators were calibrated using a Brüel & Kjær Type 4930 Artificial Mastoid in compliance to the general test procedures from ANSI S3.6-1996. The force sensitivity level of this artificial mastoid is specified up to 10 kHz by the manufacturer. For the calibration of the piezoelectric device at 12 kHz, the force sensitivity curve was extrapolated from 10 to 12 kHz. In a previous study [1], the force sensitivity level of the Brüel & Kjær Type 4930 Artificial Mastoid was shown to increase smoothly from 8 to 16 kHz.

During BC threshold measurements, narrow-band masking noise was applied to the non-test ear of the subjects using a GSI-10 clinical audiometer and a Telephonics TDH-50P earphone. All calibration and threshold measurements took place in an audiometric room at the University of Ottawa. The ambient noise met the requirements of ANSI S3.1-1991 [4].

2.3 Procedures

All BC threshold measurements were obtained using the Hughson-Westlake ascending/descending method with a step size of 3 dB. Only the right ear of the subjects was tested using a mastoid placement of the bone vibrators. The B-71 vibrator was tested at 1, 2, 3 and 4 kHz. In addition to these frequencies, the piezoelectric vibrator was also tested at 6, 8, 10 and 12 kHz. For each vibrator, thresholds were obtained with and without an earplug occluding the right ear. These two conditions permitted to test whether acoustic radiation from the shell of the vibrators could interfere with BC threshold measurements. In all conditions, the non-test left ear was masked at a level of 30 dB HL. Finally, for two of the subjects, BC threshold measurements were repeated 16 times for each of the two vibrators under the condition of occluded test ear. The bone vibrator was refitted each time on the subject's mastoid. This permitted to assess the intra-subject threshold variability due to vibrator placement and behavioural response.

3.0 RESULTS AND DISCUSSION

Based on Student t-test statistical analyses of the unoccluded BC threshold measurements, there was no significant difference ($p > 0.05$) between the two vibrators at 1, 2 and 3 kHz. At 4 kHz, a significant difference ($p = 0.004$) was observed between the two vibrators, the B-71 yielding a threshold value well above (15.8 dB) the reference equivalent threshold force level (RETFL) from ANSI S3.6-1996. In contrast, the piezoelectric vibrator yielded a threshold value only slightly lower (-3.7 dB) that the RETFL.

There was no statistical difference in threshold measurements between the occluded and unoccluded test ear conditions for the B-71 vibrator at 2, 3 and 4 kHz. At 1 kHz, occluded thresholds were 10.8 dB lower on average than unoccluded thresholds and the difference was significant ($p = 0.001$). For the piezoelectric vibrator, there was no statistical difference at any frequency from 1 to 12 kHz. However, at 1 kHz the piezoelectric device showed the same tendency as the B-71 for lower occluded thresholds. The difference was 6.0 dB and it almost reached statistical significance ($p = 0.07$). The threshold differences at 1 kHz are consistent with the reported occlusion effect of 7.6 dB for earplugs under an average insertion in the ear canal [5]. Altogether, these results confirm that acoustical radiations are negligible for the two bone vibrators and that both can be used for BC threshold measurements with the test ear unoccluded.

Figure 1 compares the mean unoccluded BC threshold values obtained in this study for the Radioear B-71 and the piezoelectric device to the RETFLs from ANSI S3.6-1996, and to the median unoccluded BC threshold values reported by Hallmo et al. [6] for males and females in the age group 18-24 years using the Prăcitronic KH70 electromagnetic transducer. With the exception of the B-71 at 4 kHz, the reported thresholds are similar. Over the 1-4 kHz region, the piezoelectric device provided a better match to the RETFLs from ANSI S3.6-1996 than the B-71. Above 4 kHz, the piezoelectric device yielded slightly lower thresholds than the RETFLs from ANSI or the data from Hallmo et al. by about 3 to 8 dB.

Tables I and II list the standard deviation for the 16 repetitions of BC measurements carried out by two subjects. In the 1-4 kHz region, the variability in threshold measurements is very similar for the two bone vibrators. The main exceptions are the larger standard deviation with the B-71 at 3 kHz for subject 1, and the larger standard deviation with the piezoelectric device at 1 kHz for subject 2. For both vibrators, the intra-subject variability does not exceed the commonly accepted clinical test-retest threshold criterion of ± 5 dB.

In summary, the new piezoelectric transducer is suitable to measure unoccluded BC thresholds from 1 kHz up to at least 12 kHz for normal young adults. Further tests are needed with older subjects and hearing-impaired individuals to establish the maximum hearing levels at which the piezoelectric transducer can be used.

REFERENCES

[1] Richter, U. and Frank, T. (1985). "Calibration of bone vibrators at high frequencies."Audiol. Acoust. 24, 52-62.

[2] Wong, G.S.K., Wu, L., Kunov, H. and Madsen, P.B. (1996). "A wide-band piezoelectric transducer for otoacoustic emission applications." Proc. of the Int. Congress on Noise Control Engineering (Liverpool, UK), Vol. 5, 2565-2569.

[3] ANSI S3.6-1996. *American National Standard Specification for Audiometers* (American National Standards Institute, New York).

[4] ANSI S3.1-1991. *American National Standard Maximum Permissible Ambient Noise Levels for Audiometric rooms* (American National Standards Institute, New York).

[5] Schroeter, J. and Poesselt, C. (1986). "The use of acoustical test fixtures for the measurement of hearing protector attenuation. Part II." J. Acoust. Soc. Am. 80, 505-526.

[6] Hallmo, P., Sundby, A. and Mair, I.W.S. (1994). "Extended high-frequency audiometry." Scand. Audiol. 23(3), 165-170.

Table I: Standard deviation (dB) for 16 BC threshold measurements under occluded test ear condition (Subject 1)

Frequency (Hz)	Radioear B-71	Piezoelectric device
1	3.2	4.0
2	2.1	2.0
3	5.1	2.1
4	2.1	3.2
6	-	2.2
8	-	2.9
10	-	3.4
12	-	2.8

Table II: Standard deviation (dB) for 16 BC threshold measurements under occluded test ear condition (Subject 2)

Frequency (Hz)	Radioear B-71	Piezoelectric device
1	2.5	5.0
2	3.0	3.7
3	3.4	2.8
4	2.7	2.9
6	-	3.2
8	-	4.0
10	-	3.5
12	-	3.2

[Work conducted under a collaborative research agreement between the University of Ottawa and the National Research Council of Canada]

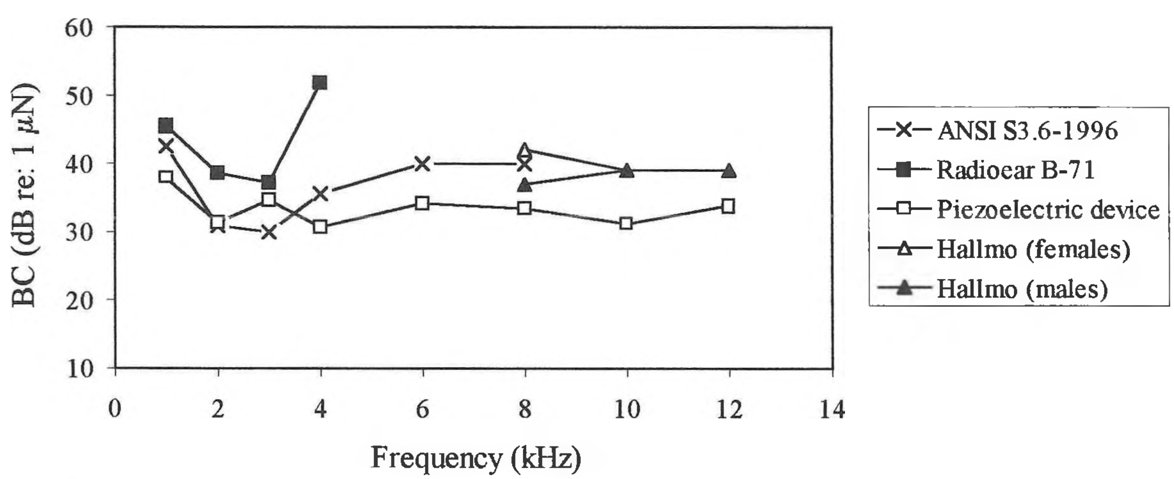


Figure 1: Comparison of BC thresholds from different sources and devices: ANSI S3.6-1996 (RETFLs), Radioear B-71 (n=8), Piezoelectric device (n=8), and Pracitronic KH70 device from Hallmo et al. (1994).

Classification of simple sensory events in younger and older adults: Are they different.

Dana R. Murphy & Bruce A. Schneider
Department of Psychology
University of Toronto at Mississauga
Mississauga, ON L5L 1C6

In a number of listening tasks, including speech, listeners have to be able to identify and keep track of a number of different stimuli (the absolute identification problem). The following research was undertaken in an attempt to determine if younger and older adults differ in terms of their ability to identify and keep track of simple auditory stimuli as the number of stimuli increases. In the first experiment, younger and older adults were tested for their ability to identify pure tones varying only in intensity. The capacity to keep a large number of tones in mind was evaluated by increasing the number of intensities to be identified from 2 to 8 intensities. In the second experiment, the ability of younger and older adults to identify 4 tones varying in intensity was tested. After determining how well younger and older adults could identify these four tones, we added a fifth tone that differed in intensity from the original 4 by as much as 50 dB. Some theorists (Hasher & Zacks, 1988) claim that older and younger adults differ in their ability to inhibit unwanted information from entering working memory and interfering with task relevant processing. When the added tone differs significantly in intensity from the original four tones, it is likely to have a very distinct sensory representation. Therefore, any tendency for this tone to interfere with the identification of the original four could be attributed to higher-order, non-sensory processes. Adding such a tone to the set should disrupt the processing of the older adults more than that of the younger adults. Thus, in the second experiment, we expected to see the older adults' identification of the original 4 tones be more disrupted by the addition of the 5th tone than the younger adults' identification.

Experiment 1
Method

Participants

Ten younger adults (mean age = 22.30) and 10 older adults (mean age = 69.00) participated in this study. All participants conformed to the following audiometric profile: thresholds ≤ 25 dB HL for all frequencies ≤ 2 kHz (allowances were made if a single frequency was 30 dB HL); ≤ 35 dB HL at 3 kHz; and ≤ 45 dB HL at 4 kHz.

Stimuli

Eight intensities (52, 58, 64, 70, 76, 82, 88 and 94 dB SPL) of a 1-kHz tone (5 ms rise/fall time) were used in this experiment. Each tone was generated on a microcomputer and presented via headphones (TDH 49) through the Tucker/Davis System 3 sound system. The tones were presented binaurally to each participant.

Procedure

Participants completed the 2 (70, 76 dB SPL) and 4 tone (64, 70, 76, 82 dB SPL) portions of the experiment in their first session. The 6 (58, 64, 70, 76, 82, 88 dB SPL) and 8 tone conditions were completed in later independent 1 hour sessions. Listeners were informed that they would be hearing a series of tones and that they were to identify each tone with a button press. The softest tone was always identified by the left most button on the array, while the loudest tone was always identified with the right most button on the array with array size varying with number of stimuli. Intermediate tones were identified with intermediate buttons with intensity increasing from left to right.

Participants first completed a practice session consisting of 20 trials using a set size of two stimuli. Following the practice, the participant then identified tones in 50 trial blocks with the number of blocks being equivalent to the number of tones (2, 4, 6, or 8).

Feedback was provided by a 200 ms light which indicated the correct response following the participant's button press.

Results and Discussion

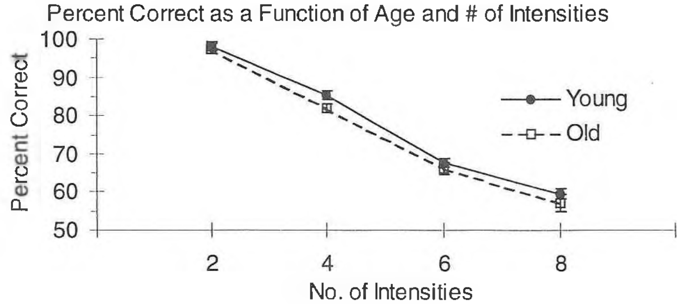
Percentage of correct responses for each tone were tabulated and an overall percentage correct score was calculated. Figure 1 graphically depicts the accuracy results of this study as a function of number of tones and age group. As can be seen in this figure, the performance of the older adults was almost identical to the performance of the younger adults. This was confirmed by a 2 age (younger vs. older) x 4 tone condition (2 vs. 4. vs. 6 vs. 8 tones) ANOVA with percent correct as the dependent measure. In this analysis, neither the main effect for age, nor the interaction between age and tone condition were significant ($F(1,18)=2.26$, $p>.15$, and $F(1,18)<1.0$, respectively).

Despite the hearing differences between younger and older adults, older adults were able to identify and keep track of a large number of tones differing only in intensity. Furthermore, while accuracy declined as the number of tones increased, younger and older adults demonstrated a similar pattern of decline. The older adults were no more influenced by the increased number of tones than were the younger adults.

Experiment 2

In Experiment 2, we first evaluated the performance of younger and older adults on a set of four 1 kHz tones (25, 30, 35, and 40 dB SPL). We then added a fifth tone. The sound pressure level of this fifth tone varied across sessions from 45 to 90 dB SPL. When the fifth tone was added to the original four, we expected performance on the original four to decrease. As the intensity separation between the added tone and the original four increases, the source of the performance decrement should become more cognitive in nature. In the first session, individuals were required to identify 4 tones. Over the course of seven more sessions, participants were required to identify 5 tones, four of which were the exact same intensities they identified in the first session and one tone was of the same frequency and duration, but differed from the original 4 in intensity. The intensity of the 5th tone ranged from 45 dB (5 dB above the highest of the original 4 tones) to 90 dB, 50 dB above the highest of the original 4 tones.

Figure 1
Percent Correct as a Function of Age and # of Intensities



Method

Participants

Ten younger adults (mean age = 21.6) and 10 older adults (mean age = 71.8) participated in this study. All participants met the same hearing criteria as was used in Experiment 1.

Stimuli

In every session, four 1-kHz tones were presented. These tones differed in intensity with the four intensities being 25, 30, 35, and 40 dB SPL. In addition, in sessions 2-7 a fifth tone was added. This tone was also a 1-kHz tone ranging in intensity from 45 to 90 dB. The exact intensities included for the 5th tone were: 45, 50, 55, 60, 70, 80, 90.

Procedure

All participants were tested in the 4 tones alone condition first. In the first session, participants completed a practice session of 40 trials and then completed 4 blocks of 50 trials each. In sessions 2-7, each participant completed a practice sequence of 50 trials followed by 5 blocks of 50 trials each.

On every trial, the computer randomly selected a tone from the set of 4 tones (5 tones in sessions 2-7) and presented that tone to the participant. The participants then pressed the button representing the tone they thought they heard. Upon the button press indicating each participant's response, a light above the correct button would flash for 200 ms.

Results and Discussion

Figure 2 presents the mean percent correct responses for younger and older adults. These data represent the percent correct on the original 4 tones only as a function of the intensity of the added 5th tone. As can be seen in this figure, adding a 5th tone disrupts the accuracy for the original 4 tones. This is true for both younger and older adults. However, there is no difference between the accuracy of the younger and the older adults. A 2 age (younger vs. older) x 8 condition (base 4 alone vs. add 45 dB vs. add 50 vs. add 55 vs. add 60 vs. add 70, vs. add 80 vs. add 90) ANOVA confirmed this finding. While adding the fifth tone did make it more difficult to identify the original four tones as indicated by a significant condition effect, $\underline{F}(7,126)=34.13$, MSE=24.73, $\underline{p}<.0001$, there was no overall age effect, $\underline{F}(1,18)=.12$, $\underline{MSE}=50.37$, ns, nor was there an interaction between age and condition, $\underline{F}(7,126)=.55$, $\underline{MSE}=24.73$, ns. Thus there is no indications that older adults are more disrupted by the addition of the fifth tone, independent of its intensity.

General Discussion

Over the course of the two experiments reported here, younger and older adults were tested for their ability to identify and keep track of tones varying only in intensity. In the first experiment, we tested the ability to identify up to 8 intensities. There was no difference between the younger and older adults in their ability to classify this number of intensities. In the second experiment, we tested the influence of added a 5th intensity on the ability to identify the same 4 tones. Adding a 5th intensity made it more difficult to identify the original 4 tones. In addition, as the intensity of the added tone increased, identification of the original 4 tones became more difficult. However, as was the case in the first experiment, younger and older adults did not differ in terms of their ability to identify the 4 tones with an added 5th tone. This was contrary to some of the theories of cognitive aging. In particular, to the extent that performance was based on cognitive factors, the inhibition hypothesis predicted that adding stimuli to the identification set would disrupt the processing of the older adults more than the processing of the younger adults, when the added stimulus was significantly different from the original four. Given these results, there is little evidence to support the notion that identification of simple auditory events changes with age. Thus, it is unlikely that the speech processing difficulties of the older adults (CHABA, 1988) are related to their ability to identify and keep track of very simple auditory events.

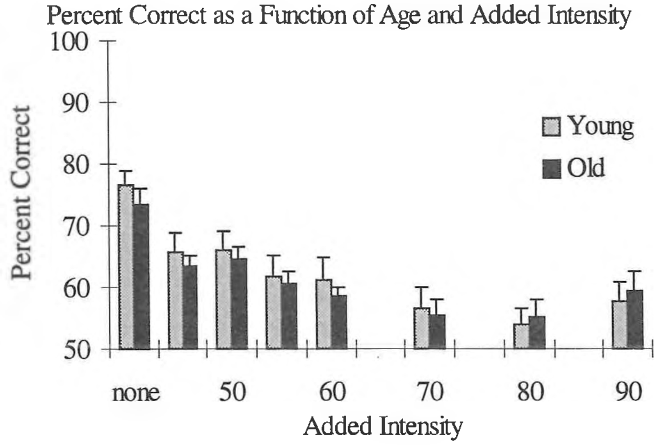
References

CHABA (Committee on Hearing, Bioacoustics, and Biomechanics). (1988). Speech understanding and aging. Journal of the Acoustical Society of America, 83(3), 859-895.

Hasher, L. & Zack, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. The Psychology of Learning and Motivation, 22, 193-225.

Figure 2

Percent Correct as a Function of Age and Added Intensity



Sound Levels from Headphone/Portable Compact Disc Player Systems

Stephen Bly, Stephen Keith and Robert Hussey
Radiation Protection Bureau, 775 Brookfield Road,
Ottawa, Ontario, Canada K1A 1C1, Postal Locator 6301B

Introduction

Concerns have arisen over the potential for noise induced hearing loss in users of personal cassette and portable compact disc players. The results of several studies in other countries [1-3] suggested that, of the children and young adults using personal cassette players, approximately 10 - 15% received exposures exceeding the Canadian Federal occupational noise exposure limit. Compared to cassette tape recordings, compact discs can provide a wider dynamic range and higher quality sound at loud levels. As a result, the recent growth in the popularity of portable compact disc (CD) players has added to the concerns about noise exposure.

Different approaches can be taken to manage the risk of noise induced hearing loss from use of both types of players. For example, France has recently adopted legislation limiting the acoustic output of these devices to 100 decibels A-weighted (dB(A)). To help examine options for Canada, sound level measurements were made on a variety of headphone/portable CD player combinations. This paper presents the findings of that measurement survey.

The results are discussed with regard to the risk of noise induced hearing loss from prolonged use at maximum volume settings.

Method and Apparatus

The maximum sound levels were measured from combinations of eight recent model portable compact disc (CD) players and 18 headphones sold in Canada. The eight CD players were identified as CD1. . . CD8. Only two of these eight models were from the same manufacturer (CD2 and CD8).

Sound levels were measured for 24 combinations of the headphones and CD players. Eight combinations consisted of the CD player and headphone with which it was packaged at purchase. Ten additional combinations were obtained using separately purchased headphones with CD1. This player was used to assess the relative sensitivities of these headphones because its output power varied by less than 0.8 dB over the range of headphone impedances (16 to 32 ohms). The most sensitive of these 10 headphones, identified as P6, was also combined with CD3 to CD8.

Most of the CD players were supplied with supra-aural headphones with porous ear pads. The exceptions were CD1 and CD5. These were provided with intra-concha headphones, which had no headband. Instead, the earpiece fit into the concha, with the diaphragm facing the ear canal.

Seven of the 10 additional headphone models studied were intra-concha and one was supra-aural with a porous ear pad. For the other two models, the earpiece fit into the concha and resembled that of the intra-concha headphone. However, the earpieces were mounted on a headband with the diaphragm facing forwards, instead of into the ear canal. The diaphragm orientation prevented this other type of headphone from making a seal around the ear canal.

Two recorded sounds were measured: (i)a "heavy metal" CD music track (Mackaye et al., "Filler/I don't want to hear it" in: "Undisputed Attitude," Slayer, American CDW 43072) and (ii)a filtered pink noise track conforming to the IEC requirements for a simulated programme signal [4], with a 13 dB crest factor and the peak recorded at the maximum level on the CD. The durations of the two tracks were 148 and 128 seconds, respectively.

Measurements were made at the internal microphone of a Bruel & Kjaer type 4128 Head and Torso Simulator (HATS) using a Hewlett Packard 35670A analyser to provide real time 1/12 octave band sound pressure levels. For each measurement, the CD player's volume control was set to maximum and the player's "bass boost" was switched on. The data was obtained by starting the CD player, allowing 7 seconds for the analyser filters to settle and averaging over 135 and 64 seconds for the "heavy metal" and IEC tracks, respectively. Before and after each set of measurements, the internal microphone sensitivity of the HATS was checked at 1 kHz using a Bruel & Kjaer type 4230 piezo calibrator. During all tests, the output from the HATS was monitored by an external speaker to ensure acceptable operation of the headphones. The background noise level was at least 30 dB(A) lower than the measurements.

For measurements, headbands were positioned vertically on the HATS. The earpiece was aligned visually on the HATS's anatomically realistic pinna for optimum coverage of the ear canal and cavum. For the intra-concha headphones, the tragus was pulled forward and then the headphone was inserted in the cavum. After the initial fitting, the headphones were checked for a secure fit, and realigned if necessary.

The sound levels of interest were the A-weighted, equivalent continuous sound pressure levels (Leq) in the absence of a subject, that would produce the sound pressure levels measured at the HATS's internal microphone. To obtain the Leq values, the appropriate frequency response of the HATS, as supplied by the manufacturer, was subtracted from the measured 1/12 octave band data to yield the sound pressure levels in a diffuse field. The A-weighting was then applied and the band levels summed. Values for Leq were also calculated for sound incident normal to the forehead. For all but one combination, it was found that the difference from the corresponding diffuse field level was less than 1 dB. The reproducibility of the Leq values was estimated as ±5 dB for intra-concha headphones and ±2 dB for the two other types.

Results and Discussion

The results in Table 1 showed that, for "heavy metal" music, a number of headphone/CD player combinations could create exposures exceeding the Canadian Federal occupational noise limit. For all the systems that were tested as sold, the measured short duration Leq values for the "heavy metal" track ranged from 90 dB(A) to 105 dB(A). The occupational noise limit was an equivalent continuous sound pressure level of 87 dB(A) for 8 hours per day, with a 3 dB(A) exchange rate [5]. Therefore, the daily listening durations needed to exceed the occupational noise limit varied from about 8 minutes to 4 hours. Five of the eight systems required daily listening times of no more than 1 hour. For the most sensitive headphone, P6, using the IEC signal, the limit was exceeded for listening times ranging from 45 seconds to about 40 minutes for the various CD players. Similar results could be inferred for the "heavy metal" track because of the similarities in the Leq values when the original headphones were used with both recordings.

The listening times required to exceed the occupational noise limit were within the ranges found in surveys of listening habits of users of personal cassette players [1-3, 6]. In those studies, typical listening times ranged from about 30 to 90 minutes per day, but a few percent of the population surveyed listened as much as 4 hours per day.

To assess the potential health impact further, an estimate was also made of the predicted permanent noise induced hearing loss [7] if users listened to the sound levels in Table 1 for 1 hour per day for 5 years, the estimated typical exposure duration [1-3,6]. Calculations were made of the noise induced permanent threshold shifts, averaged over the pure tone (speech) frequencies of 500, 1000, 2000 and 3000 Hz for the 5% most susceptible ears in a screened population of 18 year olds. This pure-tone-average is often used in the evaluation of hearing impairment for compensation purposes [8]. The calculated hearing losses ranged from 0 to 13 dB for the original headphone combinations. Use of the most sensitive headphone, P6, increased the calculated hearing loss to between 4 and 39 dB.

The results presented in Table 1 also showed a wide range in Leq values, 28 dB, depending on the headphone/portable CD player combination. This large range was due to the differences in headphone sensitivity and the different maximum output powers of the CD players. For example, the headphone sensitivity changes 17 dB when the original headphone for CD6 is replaced by P6. The effect of the CD player output is shown in the last line of Table 1, the Leq varies by 17 dB when the same headphone, P6, was used on each unit.

In addition, the results showed that intra-concha headphones tended to have a higher sensitivity than the other two types of headphones. For the measurements on CD1, with the ten separately purchased headphones, the seven intra-concha headphones yielded Leq values between 102 and 108 dB(A). By contrast, the other two types only produced 92 to 96 dB(A). Comparison of the first and third rows of Table 1 also supports this conclusion. The intra-concha headphone P6 was 11-17 dB more sensitive than all but one of the supra-aural headphones originally supplied with the CD players. By contrast, P6 was only 2-6 dB more sensitive than the intra-concha headphones supplied with CD1 and CD5.

Although detailed comparisons were not possible, the largest values found in other studies of maximum sound levels from portable CD and personal cassette players [2-3, 9] were within the broad range of values found in this study.

Conclusions

The information provided in this work on the potential for noise induced hearing loss can be used to warn of the consequences of prolonged listening to "heavy metal" music through headphones at maximum volume settings.

For this type of music, all the headphone/CD player combinations packaged together at purchase could provide exposures exceeding Canadian Federal occupational noise exposure limits. However, for typical exposure durations, even if all users listened at maximum volume settings, estimated noise-induced hearing losses varied strongly, from 0 - 13 dB. These variations were greatly increased, from 0 - 39 dB, when separately purchased headphones were combined with the CD players.

Potential hearing losses depended significantly on both the headphone and the CD player. In particular, for a given CD player, there was a tendency for sound levels to be significantly greater if intra-concha headphones were used. This suggests that the extent to which headphones are purchased separately may be an important consideration in evaluating options for reducing the risk of noise-induced hearing loss from use of these devices.

References

[1] Rice, C.G., Rossi, G., and Olina, M., "Damage risk from personal cassette players," British Journal of Audiology **21**, 279-288 (1987).

[2] Ising, H., Hanel, J., Pilgramm, M., et al., "Risk of hearing loss caused by listening to music via headphones," HNO **42**, 764-768 (1994).

[3] Airo, E., Pekkarinen, J. and Olkinuora, P., "Listening to music with earphones: An assessment of noise exposure," Acustica **82**, 885-894 (1996).

[4] IEC 268-1 (1985), "Sound System Equipment Part 1: General."

[5] Canada Occupational Safety and Health Regulations - Amendment "Part VII Levels of Sound" Canada Gazette Part II, **125**:16, 2370-2379.

[6] Aono, Sh., Ohta, T., Yamanaka, N., et al., "Listening level of music from the personal stereo player and the effect on hearing in terms of TTS," J. Acoust. Soc. Japan **53**, 440-447 (1997).

[7] ISO1999 (1990) "Acoustics-Determination of occupational noise exposure and estimation of noise-induced hearing impairment."

[8] Alberti, P.W., "Occupational hearing loss in Canada," in *Occupational hearing loss*, Edited by Sataloff, R.T., Sataloff, J., Marcel Dekker, Inc., New York, 713-721 (1993).

[9] Loth, D., Avan P., Menguy C., et al., "Auditory hazard in relation to listening to portable digital compact-disc players," Bull. Acad. Natl. Med. **176**, 1245-1252 (1992).

TABLE 1. Equivalent continuous, diffuse field, A-weighted sound pressure level, short duration Leq, obtained using the IEC simulated programme signal and the "heavy metal" track. The eight CD players were combined with either the original headphones which accompanied them at purchase or with P6, the most sensitive headphone purchased separately.

Signal	Head phone	Portable CD Player							
		CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8
IEC	Orig.	102	98	99	91	102	87	87	93
Metal	Orig.	105	101	101	94	100	90	90	96
IEC	P6	108	-	115	104	104	104	98	98

OPTICAL IMAGING OF INTRINSIC SIGNALS IN THE AUDITORY CORTEX

[1]Noam Harel, [2]Richard J. Mount and [1,2]Robert V. Harrison.
Auditory Science Laboratory, [1]Department of Physiology, University of Toronto, Toronto, Ontario, Canada M5S 1A8
[2]Department of Otolaryngology, The Hospital for Sick Children, 555 University Avenue
Toronto, Ontario, Canada M5G 1X8

Introduction

The central representation (topographic mapping) of the sensory epithelium is a fundamental feature of visual, somatosensory and auditory systems. The auditory system is organized with orderly projections from the cochlea resulting in cochleotopic maps at the brainstem, midbrain, thalamus and cortex. Such projections were originally elucidated using neuro-anatomical techniques (1) and subsequently confirmed by detailed electrophysiological mapping studies using sound stimuli (2). At the level of auditory cortex, tonotopic maps derived from electrophysiological studies have been revealed in numerous mammalian species for example: cat (3), ferret (4), chinchilla (5) and human subjects (6). In addition, a number of other methods (such as voltage sensitive dyes [7], scalp recorded evoked potentials [8] and neuromagnetic measurements [9]) have been refined for studying cortical patterns of neural activity and have revealed the tonotopic axis in auditory cortex. Refinements in the spatial resolution of medical imaging techniques (e.g. positron emission tomography [PET] and functional magnetic resonance imaging [fMRI]) have allowed for tonotopic mapping in human and non human subjects.

Another method for exploring patterns of central auditory activity is based on the detection of small changes to the optical properties of active neural tissue; optical imaging of intrinsic signals. The earliest investigations of such signals started decades ago (10) but only recently, with modern imaging technology, has useful application been possible (11, 12, 13).

There are a number of processes which accompany the activation of neurons and which can alter the optical properties of neural tissue, i.e. increasing the absorption of certain wavelengths of light. These processes include the transition of oxyhemoglobin to deoxyhemoglobin, local changes in blood volume and flow and physical changes to cell membranes and cell volume which accompany neural activation (11). These small (<1%) changes to the properties of neural tissue can be detected optically.

The vast majority of reports on the use of optical imaging of intrinsic signals are coming from studies in the visual cortex (e.g. 11, 14, 15). Somatosensory cortex has also been imaged successfully (16). Only a few preliminary reports have been made on the optical imaging of intrinsic signals from auditory cortex (18 - 22). We report here on acoustically evoked intrinsic signals arising from cortex.

Materials and Methods

Cortical imaging was carried out on anesthetized adult chinchillas (*Chinchilla laniger*) weighing between 500 - 700 g. All procedures were carried out within the standards of care of the local animal care committee and following guidelines of the Canadian Council on Animal Care.

The bone over the temporal lobe was removed (10 mm diameter) exposing the auditory cortex. The dura mater was kept intact. A well of Vaseline was build around the craniotomy, filled with silicon oil and covered with a glass cover-slip (see Fig. 1 for set up).

Intrinsic signals were recorded and analyzed using the IMAGER 2001 video acquisition system (Optical Imaging, Germantown, NY). Detailed descriptions of this system have been given elsewhere (11, 14 - 17). The essential operation of the system is that a reference (no stimulation) image is subtracted from video images of stimulated cortex and the resulting signal is amplified.

The cortical surface was illuminated with a monochromatic (green; 540 nm wavelength) light. Images of cortex were acquired with a CCD video camera. The camera was positioned and focused with its optical axis perpendicular to the cortical surface.

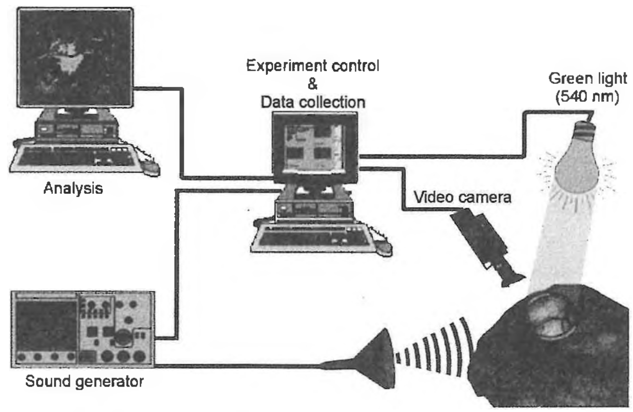


Fig. 1 Experimental setup for recording intrinsic signals

Images were collected for a 7.5 s period and stored as 15x500 ms data frames (repeated 8 - 16 times). Acoustic stimulation (10 ms rise/fall; 50 ms plateau; 10/s; frequencies 0.5, 1, 2, 4, 8, or 16 kHz and at a level between 0 - 80 dB SPL) was presented only during 4s of the data collection period (from 0.5 - 4.5 s; the initial 500 ms period provides a "no-stimulus" data frame). To allow for the relaxation of activity-dependent metabolic/vascular changes, each period of data collection was followed by a 12 s no-stimulus interval.

Results

Intrinsic signals in response to acoustic stimulation were found in a cortical area previously defined as primary auditory cortex in the chinchilla (5). The time course of the intrinsic signal is typically as shown in figure 2. The stimulus evoked intrinsic signal has an onset 0.5 - 1 s after stimulus onset and reaches a maximum after 3 - 4 s. After stimulation, the intrinsic signal decays at a slower rate. The graph quantitatively shows the changes in (540 nm) light absorption in three regions of the optically monitored area: an area (10 x 10 pixels) within the activated auditory cortex (A; solid line), and activity in two control areas: a non auditory cortical area (B; dashed line), and the bone on the edge of the craniotomy (C; dotted line).

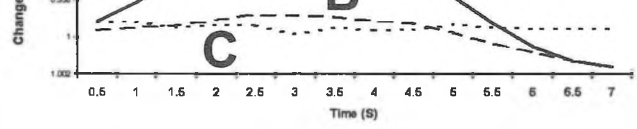


Fig. 2 Time course of intrinsic signal, see text for details

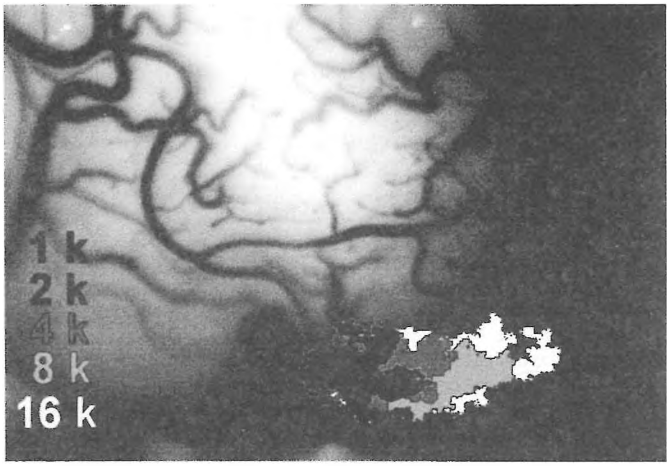


Fig. 3 A frequency (tonotopic) map of auditory cortex in the chinchilla derived by using optical imaging of intrinsic signals.

Tonotopic maps of auditory cortex were derived using pure tone stimuli (from 1 - 16 kHz; octave intervals) presented at 80 dB SPL (Fig 3). In randomly ordered experiments, areas of intrinsic signal activity were derived for each stimulus frequency. These areas are shown superimposed against the full grayscale reference image of the cortical surface. A clear shift in the region of activity in response to different stimulus frequency can be noted. The progression of low frequency anteriorly to high frequency posteriorly is fully consistent with the tonotopic axis in chinchilla AI cortex derived by single unit electrophysiological studies (5).

Discussion

In the present study we show that frequency specific suprathreshold stimuli produce intrinsic signal activity in different regions of auditory cortex such that tonotopic maps can be derived. Acoustically evoked intrinsic signals can be identified quite reliably using the time course (Fig. 2) of the signal as the main criterion. We have also found that in the absence of any acoustic stimulus one can often record apparently spontaneous waves of cortical activity.

We have found that the region of intrinsic signal origin coincides with the known position of primary auditory cortex in the chinchilla (5, 22). In addition, in one animal of the present study we have verified the position of AI cortex using surface recorded (ball electrode) auditory evoked cortical responses prior to optical imaging and found an accurate superimposition.

The potential of this method as a tool for exploring development and plasticity of auditory cortex lies in its relative non-invasiveness and speed. Under ideal conditions a tonotopic map can be derived in 30 minutes.

Acknowledgments

This research was supported by MRC (Canada) and by the Masonic Foundation of Ontario.

References

1. Woolsey CN, Walzl EM: Topical projection of nerve fibers from local regions of the cochlea to cerebral cortex. Bull Johns Hopkins Hosp 1942;71:315-344.
2. Merzenich MM, Brugge JF: Representation of the cochlear partition on the superior temporal plane of the macaque monkey. Brain Res 1973;50:275-296.
3. Merzenich MM, Knight PL, Roth GL: Representation of cochlea within primary auditory cortex in the cat. J Neurophys 1975;38:231-249.
4. Philips DP, Judge PW, Kelly JB: Primary auditory cortex in the ferret (Mustela putorius): neural response properties and topographic organization. Brain Res 1988;443:281-294.
5. Harrison RV, Kakigi A, Hirakawa H, Harel N, Mount RJ: Tonotopic mapping in auditory cortex of the chinchilla. Hear Res 1996;100:157-163.
6. Howard MA 3rd, Volkov IO, Abbas PJ, Damasio H, Ollendieck MC, Granner MA: A chronic microelectrode investigation of the tonotopic organization of human auditory cortex. Brain Res 1996;724:260-264.
7. Taniguchi IJ, Horikawa T, Moriyama M, Nasu M: Spatio-temporal pattern of frequency representation in the auditory cortex of guinea pigs. Neurosci Let 1992;146:37-40.
8. Pantev C, Bertrand O, Eulitz C, Verkindt C, Hamson S, Scuirerer G, Elbert T: Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. Electroencephalog Clin Neurophysiol 1995;94:26-40.
9. Pantev C, Hoke M, Lutkenhoner B, Lehnertz K: Tonotopic organization of auditory cortex: pitch versus frequency representation. Science 1989;246:486-488.
10. Cohen LB: Changes in neuron structure during action potential propagation and synaptic transmission. Physiol Rev 1973;53:373-418.
11. Grinvald A, Frostig RD, Lieke E, Hildesheim R: Optical imaging of neuronal activity. Physiol Rev 1988;68:1285-1366.
12. Orbach HS, Cohen LB, Grinvald A: Optical mapping of electrical activity in rat somatosensory and visual cortex. J Neurosci 1985;5:1886-1895.
13. MacVicar BA, Hochman D: Imaging of synaptically evoked intrinsic optical signals in hippocampal slices. J Neurosci 1991;11:1458-1469.
14. Ts'o DY, Frostig RD, Lieke EE, Grinvald A: Functional architecture of primate visual cortex revealed by high resolution optical imaging. Science 1990;249:417-420.
15. Bonhoeffer T, Grinvald A: The layout of orientation domains in cat area 18 revealed by intrinsic optical imaging. J Neurosci 1993;13:4157-4180.
16. Frostig RD, Lieke EE, Ts'o DY, Grinvald A: Cortical functional architecture and local coupling between neuronal activity and the microcirculation revealed by in vivo high resolution optical imaging of intrinsic signals. Proc Natl Acad Sci 1990;87:6082-6086.
17. Bonhoeffer T, Grinvald A: Optical imaging based on intrinsic signals: the methodology; in Toga AW, Mazziotta JC (eds): Brain mapping: the methods. San Diego, Academic, 1996, p 55.
18. Bakin JS, Kwon MC, Masino SA, Weinberger NM, Frostig RD: Suprathreshold auditory cortex activation visualized by intrinsic signal optical imaging. Cerebral Cortex 1996;6:120-130.
19. Hess A, Scheich H: Optical and FDG mapping of frequency specific activity in auditory cortex. Neuroreport 1996;7:2643-2647.
20. Dinse HR, Godde B, Hilger T, Reuter G, Cords SM, Lenarz T, von Seelen W: Optical imaging of cat auditory cortex: cochleotopic selectivity evoked by acute electrical stimulation of a multi-channel cochlear implant. Eur J Neurosci 1997;9:113-119.
21. Harrison RV, Harel N, Raveh E, Kakigi A, Mount RJ: Representation of sound frequency and intensity in auditory cortex revealed by intrinsic signals optical imaging. Soc Neurosci Abst 1997;23:1033
22. Harrison RV, Harel N, Raveh E, Kakigi A, Mount RJ: Optical imaging of intrinsic signals in chinchilla auditory cortex. Audiol Neurootol. 1998; 3(2-3): 214-223.

Effects of cochlear hypoxia on otoacoustic emissions compared with auditory evoked potentials

Shoichi Sawada, Claudine Gysin , Richard Mount, Robert V. Harrison

Auditory Science Laboratory, Department of Otolaryngology,
University of Toronto and the Hospital for Sick Children, Toronto, Canada

PURPOSE

The purpose of this study was to investigate the effects of mild systemic hypoxia on inner haircell (IHC) and outer haircell (OHC) function. A number of studies have explored the effects of anoxia/hypoxia and ischemia of the inner ear [e.g. 1,2,3,4,5] but have not addressed whether there is a dissociation between IHC and OHC function during such an insult. To monitor OHC activity we recorded otoacoustic emissions (OAEs) which are known to originate at OHC level. IHC/cochlear afferent nerve function was assessed indirectly from auditory brainstem evoked responses (ABR). We hypothesize that in chronic cochlear hypoxia, IHC (and cochlear afferent nerve) activity deteriorates before that of OHCs. Our study represents a model for effects of perinatal hypoxia or anoxia on the inner ears of high risk neonates which may be a contributing factor to the hearing deficits seen in auditory neuropathy [6,7]. We show that changes in IHC and OHC function during mild hypoxia occur with different time courses.

METHODS

All experiments were carried out in a sound-attenuating booth. Five adult chinchillas weighing from 400 to 600g, free from ear disease, were used. All animals were anesthetized with ketamine (15mg/kg, IM) and xylazine (25mg/kg, IM). To prevent airway obstruction by secretions, atropine (0.1mg/kg, IM) was injected. Half doses of ketamine and xylazine were given for maintenance for anesthesia. All animals were tracheotomized and a small tube was inserted into trachea. To produce systemic hypoxia, a dead space was added to the tidal volume by connecting a syringe with a small hole at the distal end to the tracheal tube. A 20ml volume of dead space was sufficient to change ABR thresholds [3]

Otoacoustic emissions measurement

For transient evoked otoacoustic emissions (TEOAE) and distortion product otoacoustic emissions (DPAOE) recording, we used ILO88/92 systems (Otodynamics). TEOAE and DPOAE were measured at 10 minute intervals before and during the period of experimental hypoxia. The parameters for TEOAE were the default protocol for human measurements in the non-linear mode. Stimuli were 80 μs clicks at 80 dB peSPL. The responses were the average of 260 sweeps. For DPOAE measurement, test stimuli of f_2 was at 8kHz, from 35 to 80dB SPL. The emissions were generated by presentation of the same level (f_1=f_2) at a separation ratio of f_2/f_1=1.22. To plot the time course (see figure), we selected amplitudes at 50 dB SPL. (We have also measured emission input/output functions to verify that thresholds were not changing independently of these amplitude values.)

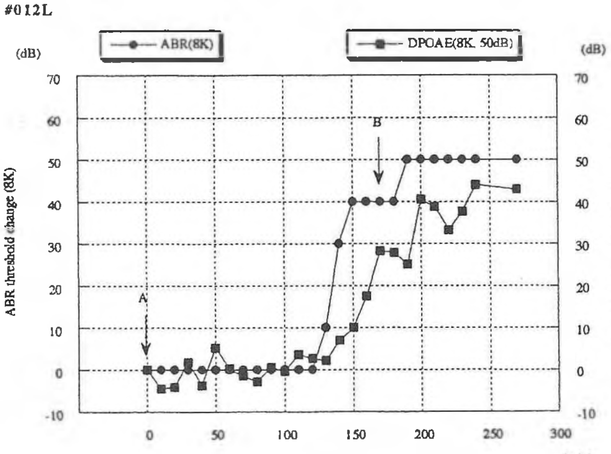
Auditory brainstem evoked response measurement

ABR responses were recorded with needle electrodes in standard vertex - bulla (mastoid) configuration. Stimuli were from 0 to 80 dB SPL tone bursts, presented through a calibrated sound system. The responses were amplified and filtered conventionally, and waveforms were averaged (BioSig; Tucker-Davis Technologies). Before and after hypoxia, we measured six frequencies (0.5, 1, 2, 4, 8, 16 kHz). During hypoxia, thresholds at 8 kHz were measured at 10 minute intervals.

RESULTS

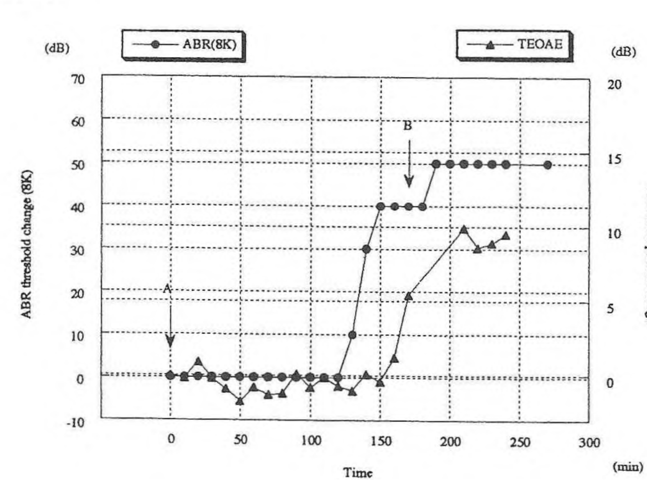
The figures show the time courses of ABR thresholds and of and DPOAE (left panels) TEOAE amplitudes (right panels) during many hours of sustained hypoxia (start and finish of hypoxic period indicated by arrows A and B). In all animals, ABR thresholds were significantly increased within 90-150 minutes, whilst otoacoustic emission amplitudes showed either little deterioration (e.g. #013R, #014L) or lagged behind ABR changes (#012L). It is of importance to note the general correspondence between TEOAE and DPOAE amplitude changes measured from the same cochlea.

#012L

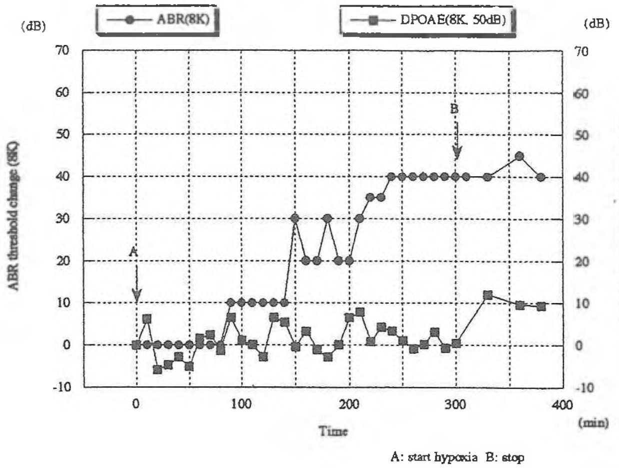


A: start hypoxia B: stop

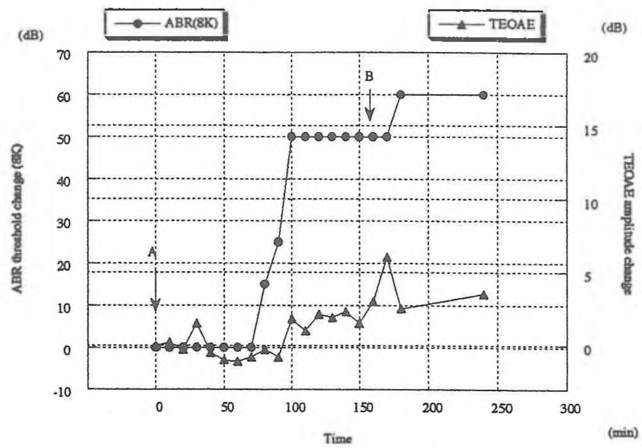
#012L



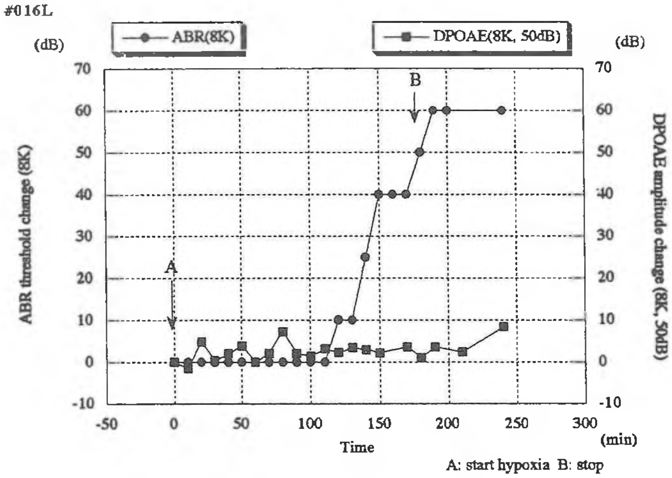
#013R



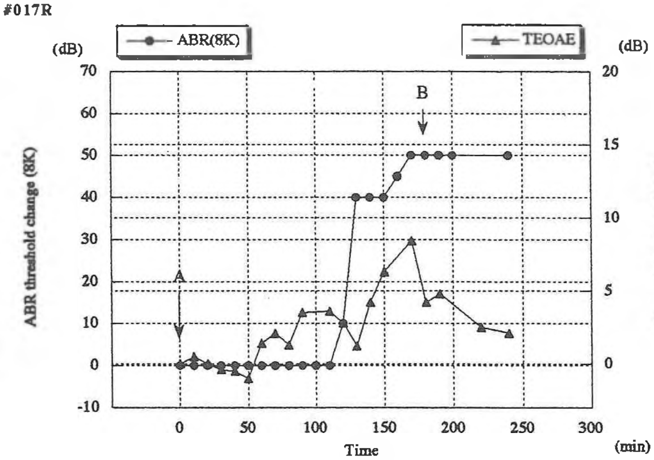
#014L



#016L



#017R



DISCUSSION

We show that IHC/cochlear afferent function as reflected in neural evoked responses are more sensitive to prolonged, mild cochlear hypoxia than the activity of OHCs as revealed in otoacoustic emission measures. Our focus on cochlear hypoxia relates to the clinical issue of possible IHC/cochlear afferent damage as a result of hypoxia in high risk infants, either resulting from difficult birth, or circulatory insufficiency in utero. Such IHC damage could result in auditory neuropathy [6], a hearing disorder which appears to result from a sub-total depletion or desychronization of cochlear afferent neurons, but in which OHC function (as shown by otoacoustic emission or cochlear microphonic recording) is much less damaged. The present study which shows that hypoxia can have an initial effect on the IHC/cochlear afferents supports our theory [7] that hypoxia is an important etiological factor in some types of auditory neuropathy.

REFERENCES

[1] Konishi T et al. (1961): Effect of anoxia on cochlear potentials. J Acoust Soc Am 33: 349-356.
[2] Kusakari J et al. (1981): The effect of transient anoxia upon the cochlear potentials. Auris Nasus Larynx 8: 55-64.
[3] Shirane M and Harrison RV (1987): The effects of hypoxia on sensory cells of the cochlea in chinchilla. Scanning Microsc 1, 1175-1183.
[4] Shirane M and Harrison RV (1987): The effects of deferoxamine mesylate and hypoxia on the cochlea. Acta Otolaryngol 104, 99-107.
[5] Billett TE et al. (1989): The nature and progression of injury in the organ of Corti during ischemia. Hear Res 41:189-198.
[6] Starr A et al. (1996): Auditory neuropathy. Brain 119: 741-753.
[7] Harrison RV (1998): An animal model of auditory neuropathy. Ear and Hearing (in press)

EFFECTS OF LONG-TERM ANAESTHESIA ON AUDITORY EVOKED POTENTIAL AMPLITUDE AND LATENCY

Naoki Mori, Shoichi Sawada, Richard J. Mount and Robert V. Harrison
The Auditory Science Laboratory, Department of Otolaryngology,
The Hospital for Sick Children, University of Toronto

INTRODUCTION

In auditory science, a number of experimental techniques involve the use of anaesthetized animal models. In some studies anaesthesia is maintained for many hours, for example during electrophysiological mapping or optical imaging experiments. We, and others, have often noted a deterioration in auditory function during such long-term experiments. Here we explore systematically these effects with the particular interest in whether changes observed result from effects on central auditory pathways, or are caused by deterioration in cochlear function. We report here the effect of long-term anaesthesia, using a ketamine-xylazine combination, on auditory evoked potentials. We have monitored auditory brainstem evoked responses (ABRs) and middle latency responses (MLRs) from the chinchilla during 12 hours of anaesthesia, and we describe the deterioration in auditory function which results over this period.

MATERIAL AND METHODS

Adult chinchillas weighing between 460 g and 700 g and free from ear disease were used in this experiment. All procedures were carried out within the guidelines of the Canadian Council on Animal Care.

For induction of anaesthesia, atropine 0.04 mg/kg, xylazine 2.4 mg/kg and ketamine 15 mg/kg were injected i.m. Supplemental xylazine 1.2 mg/kg and ketamine 7.5 mg/kg were given at approximately one hour intervals to maintain level of anaesthesia during the experiment.

ABR and MLR potentials were recorded using skin needle electrodes in a mastoid-vertex configuration. The ABR was elicited using a range of tone pip frequencies between 2 and 24 kHz. The responses were amplified and filtered by convention methods. Responses were realized by averaging 300 sweeps with a 20 ms time window (Bio-Sig; Tucker-Davis Technologies). The MLR was elicited using 8 kHz tone bursts at 80 dB SPL. Responses were average of

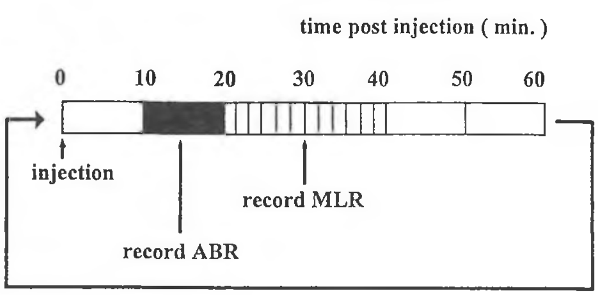


Figure 1. The protocol for recording ABR and MLR.

1,000 sweeps using a 50 ms time window. ABR responses were recorded ten minutes after the initial injection of anaesthetic, ten minutes later, MLR were recorded. This protocol was repeated for 12 hours (Figure 1). Then the anaesthesia was stopped. Twenty four hours later, ABR and MLR were again measured.

RESULTS

Figure 2 shows typical results from one subject. The waveforms of MLR, recorded at hourly intervals, are plotted. Early and late peaks are identified (small arrow symbols). These peak latencies gradually increased as the period of anaesthesia was prolonged. Whilst all peaks in the waveform

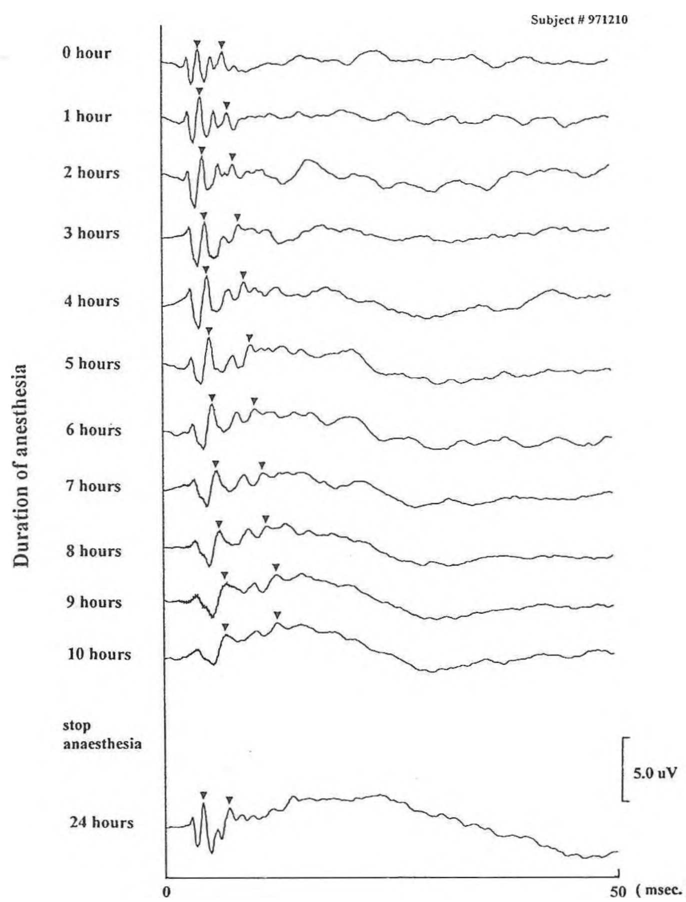


Figure 2. Auditory evoked potentials during long-term ketamine anaesthesia.

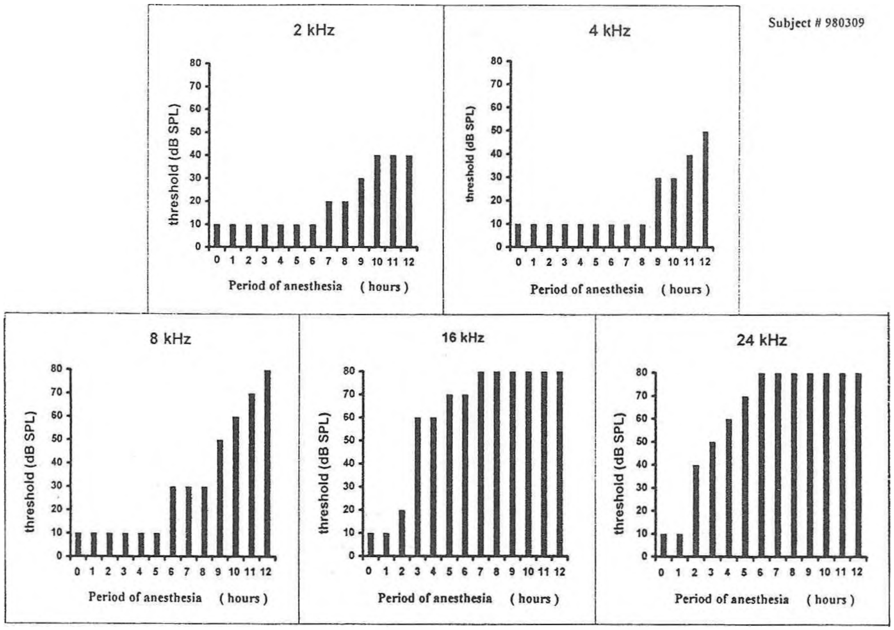


Figure 3. Effects of long-term ketamine anaesthesia on ABR thresholds.

showed increased latency over time, the change in the later peaks was most obvious. One day after the end of the anaesthetic period the waveforms appear normal as indicated in the lower trace of fig. 2. This indicates the reversible nature of the anaesthetic effects.

Figure 3 shows graphs of the ABR thresholds to 2, 4, 8, 16 and 24 kHz tone stimuli, as a function of time during anaesthesia. The threshold increase during anaesthesia was greater on ABR thresholds evoked by high frequency stimuli than by lower frequencies. The threshold shift at 24 kHz and 16 kHz occurred after two hours of anaesthesia. The thresholds increased rapidly reaching 80 dB SPL after six or seven hours. At 8 kHz, threshold changes started after six hours of anaesthesia and gradually increased over c.7 hours to reach 80 dB SPL. At lower frequencies, 4 kHz and 2 kHz, the changes occurred later, between seven and nine hours. The increase in threshold occurred less rapidly, and reached only 40 - 50 dB SPL.

DISCUSSION

The changes reported here agree with related work reported on by others. Sohmer et al. [1] have shown that experimental hypotension depressed ABR in cats. The ABR loss began with the later waves and progressed to the earlier waves. He suggested that hypotension induced cerebral ischaemia and decreased the oxygen supply and that the oxygen supply was the important factor for ABR.

Sanford and Colby [2] reported that in the rabbit, a ketamine-xylazine combination produces a drop in blood pressure and respiratory rate and a reduction of heart rate which recovers between four and six hours following injection. These conditions could influence cochlear function and therefore auditory evoked potentials. In contrast, in a human study it has been reported that ketamine does not suppress auditory evoked potentials in either peak latency or amplitude [3].

Billett et al. [4] reported damage in the cochlea during ischaemia. The damage appeared first in the cells of the basal turn of the cochlea and gradually progressed towards the apex with increasing period of ischaemia. Our results match these findings in that the ABR threshold change appeared first in the responses evoked by higher frequencies and only later progressed to affect responses evoked by lower frequencies.

We interpret our results as indicating that the changes originate at the cochlear level, because all parts of the ABR including the earliest peaks of the waveform show changes. We believe that the changes shown here occur because long-term anaesthesia can result in metabolic or ionic changes which affect haircells and/or nerve cells and their synapses. Functionally the cells in the basal turn of cochlea appear to deteriorate earlier than those apically. Our interest in this study is mainly practical. We are currently exploring a number of different anaesthetic techniques (e.g. barbiturate, halothane) to find the most effective agent for long term studies in chinchilla. There is also a clinical issue. Does long-term surgery in human subjects cause cochlear changes, and are these changes temporary (reversible) or not?

ACKNOWLEDGMENTS

This work was funded by The Masonic Foundation of Ontario and MRC Canada.

REFERENCES

[1] Sohmer H., Freeman S. et al. (1986) The depression of the auditory nerve - brain-stem evoked response in hypoxaemia - mechanism and site of effect. Electroencephalography and clinical Neurophysiology 64 : 334-338.

[2] Sanford T.D. and Colby E.D. (1980) Effect of xylazine and ketamine on blood pressure, heart rate and respiratory rate in rabbits. Laboratory Animal Science 30 : 519-523.

[3] Schwender D., Klasing S. et al. (1993) Mid-latency auditory evoked potentials during ketamine anaesthesia in humans. British J Anaesthesia 71 : 629-632.

[4] Billett T.E., Thorne P.R. et al. (1989) The nature and progression of injury in the organ of Corti during ischemia. Hearing Research 41 : 189-198.

Denture Influence on Resonance Balance in Speech

Jana M. Scarsellone, M.S., CCC-SLP, Anne P. Rochet, Ph.D., CCC-SLP

3-48 Corbett Hall, University of Alberta, Edmonton, Alberta T6G 2G4

and

John F. Wolfaardt, BDS, MDent, Ph.D.

Misericordia Hospital, 16940 – 87 Avenue, Edmonton, Alberta T5R 4H5

Introduction

Normal resonance of the voice begins when the sound waves produced by vocal fold vibration interact with resonating tubes and chambers in the throat, mouth, nose and cranium. Normally, this vibrating airstream is selectively resonated and filtered within the vocal tract and, as it moves headward, is directed primarily through the oral cavity for non-nasal sounds and through the nasal cavity for nasal sounds. Disordered resonance, revealing itself as either hyponasality or hypernasality, is detrimental to the acceptability and intelligibility of speech and thus constitutes a major clinical problem in speech-language pathology.

In the past, clinical assessment and management of disordered resonance have depended on perceptual judgments which are subjective in nature and are influenced by many different speaking variables. However, oral/nasal resonance balance now can be measured instrumentally with the Nasometer (Kay Elemetrics Corporation, Lincoln Park, New Jersey, USA), a computer-based tool that calculates nasalance based on acoustical information received from both oral and nasal microphones. Nasalance, the ratio of nasal acoustic energy (N) to nasal-plus-oral acoustic energy (N+O) expressed as a percentage [(N)/(N+O) x 100], is a correlate of the human perception of resonance balance (Hardin, Van Demark, Morris, & Payne, 1992; Dalston, Warren, & Dalston, 1991a, 1991b). The Nasometer is useful, therefore, in complementing clinical judgments of resonance disorders (Dalston, Neiman, & Gonzalez-Landa, 1993; Dalston, Warren, & Dalston, 1991a, 1991b; Hardin et al., 1992).

As normal reference data for nasalance have accumulated, small but reliable differences in nasalance values have emerged across age groups; that is, older speakers tend to obtain higher nasalance scores than younger speakers. Explanations for the trend toward increased nasalance with age include naturally-occurring physiological and structural changes of the speech mechanism. In addition, with advancing age, individuals meet with a greater probability of requiring dentures. Conceivably, dentures have the potential to impose artificial structural changes that could influence nasalance scores. Therefore, the possibility exists that the presence of an upper denture may act as an intervening variable in the measurement of nasalance.

The passive acoustic effects on oral/nasal resonance balance related to the presence of dentures are not understood. In theory, the presence of an upper denture may increase or decrease nasalance according to one of the two following hypotheses. The first hypothesis suggests that the baseplate of the upper denture, which covers the entire hard palate and extends onto the soft palate, could impede the passive transmission of sound waves through the palate and into the nasal cavities. Thus, a reduction in the overall nasal energy in the nasalance ratio would result in decreased nasalance. Alternatively, it may be argued that a denture would decrease the volume of the oral cavity and increase oral impedance thereby reducing the oral energy component of the nasalance equation, resulting in increased nasalance. Thus, the presence of an upper denture appears to have the potential to alter resonance, however the passive acoustic effects on the oral/nasal resonance balance related to the denture are not fully understood.

The purpose of this study was to investigate the effects of upper dentures on nasalance values in normal elders. Specifically, the experimental question posed was "Is there a difference between nasalance obtained with and without full upper dentures for the same speaker reading the same passage aloud?"

Method

Participants

Twenty women between the ages of 61 and 81 years (mean = 71 years) participated in this study. All had been residents of western Canada for at least 45 years and were fluent in the English language. All participants had worn upper or upper and lower dentures for at least 1 year prior to data collection. In addition, all participants were satisfied with the speech function of their dentures. Histories of respiratory, neurological, laryngeal, velopharyngeal, or craniofacial disorders were criteria for exclusion of potential participants. In addition, all participants were free of upper respiratory infections at the time of recording.

Instrumentation and Materials

A Nasometer (model 6200) was used to collect nasalance data. Reading stimuli included three English language passages used routinely in clinical practice and in the collection of nasalance data in North America. Use of these passages ensured that the speech samples obtained would vary with respect to the presence of nasal consonants.

Procedure

Denture-speaking status was randomly counterbalanced across participants' performances: half of the women read all passages first with their upper denture in place and then with it removed. This order was reversed for the remaining participants. The three passages also were presented in random order, and each participant read each passage three times. After reading all passages in the first condition, the upper denture was either removed or inserted, depending upon the initial condition which had been assigned. The participant then read each passage three times again, in random order.

Data Analysis

The nasalance data were submitted to a three-factor (2 x 3 x 3) (denture condition; passage; trial) within-subjects factorial Analysis of Variance (ANOVA). An alpha level of .05 was used for all statistical tests.

Results

Significant main effects were revealed for denture condition, $F(1,19) = 8.18$, $p<.01$, and passage, $F(2,38) = 5.21$, $p<.01$, but not for trial, $F(2,38) = 3.25$, $p>.05$. A significant interaction was found between denture status and trial, $F(2,38) = 3.25$, $p<.05$. All remaining two-way and three-way interactions were not significant. With respect to the main effect for denture status, participants displayed significantly lower nasalance scores without their upper dentures than with them. This finding was consistent across the group means for all three reading passages, $F(3,76) = 4.04$, $p<.01$.

Discussion

The purpose of this study was to determine if the presence of an upper denture affects nasalance values. Two hypotheses regarding the effects of dentures on nasalance values were offered. The first hypothesis proposed that the baseplate of an upper denture may impede the passive transmission of sound waves into the nasal cavities thereby reducing nasalance values. This hypothesis was not supported in this study. Nasalance values were higher with an upper denture in place than without one. The second hypothesis proposed that an upper denture would change the dimensions and the relative acoustical impedance of the oral cavity in a manner that would result in an increase in nasalance values. This hypothesis was supported by the results of the present study. An explanation for this finding may be found by considering it in relation to the acoustical theory of vowel production (Fant, 1960) and in the context of the equation used to compute nasalance (Fletcher et al., 1989).

When considering both the oral and nasal channels of the vocal tract, the one with less acoustical impedance will transmit a greater proportion of sound energy through it. Specific to this experiment, an upper denture may decrease the volume and length of the oral cavity relative to those same dimensions with no denture in place. Additionally, an increase in oral impedance may result from increased contact of the tongue with the palate that has been found when speaking with dentures versus without (Ylppo and Sovijarvi, 1962; Wictorin and Agnello, 1970). The result of a decrease in oral resonating volume and an increase in oral acoustical impedance with dentures in could result in a decrease in oral energy received by the Nasometer's oral microphone compared to oral energy values with dentures out. Likewise, with dentures removed: oral volume would increase; oral impedance would decrease and oral resonance received by the Nasometer's oral microphone would increase. Because the oral component of the nasalance ratio is increased, the nasal component would be relatively weaker thereby decreasing the nasalance value.

This hypothesis could be tested through first and second formant frequency analyses and second formant (F2) loci comparisons across denture conditions. The frequency of all formants will decrease as vocal tract volume increases (Lindblom & Sundberg, 1971; Murry & Bone, 1989). With respect to dentures, one would expect that speech produced without a denture (increased vocal tract length and volume) would be characterized by oropharyngeal formants which were lower in frequency than those obtained for speech produced with a denture (decreased vocal tract length and volume). Because the behavior of F2 responds to changes of dimension within the anterior oral cavity, F2 loci comparisons might be sensitive to differences in oral cavity constriction and increased impedance associated with the presence of a denture. F2 is higher in frequency when the anterior oral space becomes constricted and lower in frequency when the oral cavity is elongated or more open (Lindblom & Sundberg, 1971). One would expect that speech produced with dentures would be characterized by higher F2 values than speech produced without dentures.

Conclusion

The nasalance values obtained from the speakers in this study differed significantly between denture and edentulous conditions. The small yet significant differences were explained by considering them as a function of the nasalance ratio in the context of the acoustical theory of vowel production. That is, higher nasalance values obtained with dentures in place were attributed to a decrease in the dimensions of the oral cavity and an increase in oral acoustical impedance resulting in a reduction in the amount of acoustical energy transmitted to the Nasometer's oral microphone.

References

Dalston, R., Neiman, G., & Gonzalez-Landa, G. (1993). Nasometric sensitivity and specificity: a cross-dialect and cross-culture study. Cleft Palate-Craniofacial Journal, 30, 285-291.

Dalston, R., Warren, D., & Dalston, E. (1991a). Use of nasometry as a diagnostic tool for identifying patients with velopharyngeal impairment. Cleft Palate-Craniofacial Journal, 28, 184-189.

Dalston, R., Warren, D., & Dalston, E. (1991b). A preliminary investigation concerning the use of nasometry in identifying patients with hyponasality and/or nasal airway impairment. Journal of Speech and Hearing Research, 34, 11-18.

Fant, CGM. (1960). Acoustic theory of speech production. Gravenhage, Mouton: The Hague.

Hardin, M.A., Van Demark, D.R., Morris, H.L., & Payne, M.M. (1992). Correspondence between nasalance scores and listener judgments of hypernasality and hyponasality. Cleft Palate-Craniofacial Journal, 29, 346-351.

Lindblom, B.E. & Sundberg, J.E. (1971). Acoustical consequences of lip, tongue, jaw, and larynx movement. The Journal of the Acoustical Society of America, 50, 1166-1179.

Murry, T. & Bone, R.C. (1989). Acoustic characteristics of speech following uvulopalatopharyngoplasty. Laryngoscope, 99, 1217-1219.

Classifying Subgroups of Individuals with ALS: Acoustic and Aerodynamic Characteristics

H.A. Leeper, Ph.D., School of Communication Sciences and Disorders, Elborn College, University of Western Ontario, Michael Strong, M.D., Department of Clinical Neurological Sciences, London Health Sciences Centre, London, Ontario.

Amyotrophic Lateral Sclerosis (ALS) is a progressive degenerative neuromuscular disease that involves the upper and lower motor neurons. Damage to the lower motor neurons is manifested by muscle weakness, fatigue, muscle atrophy, and fasciculations, whereas upper motor neuron damage is manifested by spasticity, muscle weakness, cramping, increased tone, and hyperactive deep-tendon reflexes. Upper and/or lower motor involvement may occur during the course of the disease, but eventually, both systems are involved.

At initial presentation, individuals are described primarily as being either bulbar or nonbulbar, as determined by presence or absence of neuromotor symptoms. Usually, bulbar signs are typified by rapid deterioration, while nonbulbar signs point to slightly slower deterioration, especially of the cranial nerves affecting oral-pharyngeal-laryngeal motor coordination and speech production (Darley, Aronson & Brown, 1975).

Progressive deterioration of the oral, velopharyngeal, and laryngeal subsystems serving speech leads to speech production difficulties and decreased speech intelligibility in individuals with ALS. Major characteristics of speech difficulties include imprecise articulation, hypernasality and nasal air emission, strained-strangled, harsh, breathy, low-pitch and low intensity voice production, and slow speaking rate. These characteristics may occur in different individuals at various times throughout the course of the disease (Kent, Sufit, Rosenbek, Weismer, Martin, & Brooks, 1991).

Acoustical and aerodynamic assessment of voice and speech production are useful in early detection, differential categorization, and the monitoring of deterioration of speech in individuals with ALS. Further, differential speech subsystem assessment of subgroups of individuals with bulbar and nonbulbar signs may lead to better management strategies of these groups of individuals relative to "quality of life" issues. This series of experiments characterizes acoustical and aerodynamic changes in speech deterioration of individuals with bulbar and nonbulbar signs which may allow for specific management strategies over time.

Method

Subjects:

The subjects selected for presentation included 95 men and women with diagnosed ALS. Subgroupings led to classifications with 44 persons with bulbar symptoms and 51 with nonbulbar symptoms upon initial classification by two neurologists (AH;MS). The subjects ranged from 40 years to 78 years of age. The length of illness ranged from 4 months to 12 years, with a mean duration of illness of approximately 2 years. The subjects were assessed twice, approximately 6 months apart.

Stimulus Material:

The stimuli were chosen to investigate subtle changes in the laryngeal, velopharyngeal and oral articulatory subsystems underlying speech production. Respiratory-laryngeal function was assessed using a maximum phonation time technique employing production of the isolated vowel /a/. Vocal fold diadochokinetic function of vocal fold activity was sampled using the syllable glottal fricative /h/ + vowel /a/ (/ha/). Velopharyngeal function was sampled employing the /mp/ blend in the word "hamper", sustained vowel /i/, syllables /ipi/ and /isi/, a nasal sentence, and an all "oral" element passage ("Zoo Passage"). Oral articulatory function was sampled employing 12 CVC words consisting of 6 word initial plosives (/p,t,k,b,d,g/) combined with the vowel (/i/).

Recording Procedures:

Each subject was seated comfortably in a quiet room with an electret condenser microphone held at a constant mouth-to-microphone distance of 15 cm. Speech materials were recorded on a research quality cassette tape recorder and/or directly to one channel of a computer with A-D/D-A converters and a commercial acoustical recording/analyses package (CSpeech, Milenkovic, 1987).

Acoustical information for nasalization was acquired using a Nasometer (Kay Elemetrics) with a dual oral-nasal microphone system and the resultant signal sent to a computer and accompanying nasometer software package for data display and analyses.

Aerodynamic assessment of speech was accomplished with a commercial pressure-flow device (RC Electronics - Computer Scope) employing separate differential pressure transducers for oral-nasal pressure measures and a pneumotachograph-differential pressure transducer for nasal airflow measurements. Measures of peak and durational aspects of the oral-nasal aerodynamic features of speech were analyzed using the accompanying software package on the same computer used for other analyses.

Results

Both inferential and descriptive statistics were employed for each of the data sets to describe differences in the subgroupings of bulbar and nonbulbar individuals with ALS over two time periods (approximately 6 months apart).

Respiratory-Laryngeal Subsystem:

An analysis of variance procedure revealed significant differences ($p = .008$) in Maximum Phonation Time (MPT) over time for the vowel /a/ production for the ALS subjects, with the bulbar group having numerically lower MPTs then the nonbulbar at each assessment period. That is, the bulbar group decreased from 16.7 seconds to 11.3 seconds over time, while the nonbulbar decreased from 21 to 16.7 seconds over the same time period.

The bulbar subjects also demonstrated significantly ($p = .003$) slower vocal fold diadochokinetic rates for /ha/ production (3.46 s/s) than the nonbulbar subjects (4.88 s/s). In addition, the last segment of the repetition phase was significantly different for the groups from Time 1 to Time 2, with the bulbar subjects showing the poorer performance.

Microanalytical acoustical measures of frequency and amplitude perturbation from the sustained vowel /a/ productions showed a numerically poorer performance for the bulbar subjects for modal fundamental frequency, increased jitter and shimmer measures, and decreased signal-to-noise ratios. In addition, there were statistically significant reductions ($p = .001$) for vocal frequency range, intensity range, and maximum and minimum intensity range, with the bulbar individuals showing more restricted performance within and across assessment times than the nonbulbar individuals.

Velopharyngeal Subsystem:

Results of the acoustical assessment of the velopharyngeal (VP) subsystem indicated that the bulbar individuals had significantly ($p = .04$) greater nasalance scores of the vowel /i/, syllables /ipi/ and /isi/ and the "Zoo Passage" than the nonbulbar subjects.

Aerodynamic assessment of the VP port area demonstrated a significant difference ($p = .003$) in nasal airflow duration for the /mp/ blend at both Time 1 and Time 2 between the bulbar and nonbulbar individuals. In addition, significant differences ($p = .02$) in timing of closure between the /m/ and the /p/ elements was determined, with the bulbar group demonstrating the shortest time (i.e., most overlap between the nasal and non-nasal elements) when compared with the nonbulbar subjects.

Oral-Articulatory Subsystem:

Analyses of the acoustical voice-onset-time (VOT) data for the two groups over time indicated a significant difference ($p = .001$) for voicing contrasts (with voiceless phonemes being significantly longer in VOT than voiced consonants). Further, at assessment Time 2, the bulbar individuals showed a significant ($p = .05$) difference in VOT when compared to the nonbulbar subjects. A significant ($p = .05$) Voicing by Position (i.e., anterior versus posterior articulator position) was determined showing the bulbar group having more deviant articulator activity at Time 1 and 2. In addition, the bulbar individuals were more variable within and across time than the nonbulbar subjects.

Discussion:

The series of investigations of the respiro-laryngeal, velopharyngeal, and oral-articulatory subsystems described in this presentation provides acoustical and physiological information concerning the course of "short term" deterioration of the subsystems serving speech in individuals with ALS. These data support the approach of differentiating ALS subjects into bulbar and nonbulbar groupings based upon speed and variability of deterioration within the various subsystems (Leeper, Millard, Bandur, Hudson, 1996; Renout, Leeper, Bandur, Hudson, 1995; Delorey, Leeper, Bandur, Hudson, 1993; Herbert, Leeper, Bandur, Hudson, 1996). The present data are also generally consistent with acoustical and perceptual information from previous investigations of ALS subjects (Kent, et al., 1991; Langmore and Lehman, 1994; Caruso and Burton, 1987).

Given the differences in respiratory-laryngeal support, the loss of aerodynamic integrity at the VP port, and the differential reduction in labial, tongue tip, and tongue blade control between the two subgroups of ALS subjects studied, it would seem important to continue such descriptive categories for management purposes. There is, however, as Strong (1995) has suggested, no current medical therapies that can stop or even slow down the progress of ALS. However, management related to "quality of life" issues such as breathing and communication are important to the well being of the individual and his/her supporting family. A variety of augmentative and assistive communication devices may be used to improve quality of life within the suggested life span. Monitoring of changes in communication skills via non-invasive acoustical and aerodynamic assessment methods will allow for more precise allocation of these techniques for individuals with one of the subgroupings of ALS.

References:

1) Darley, F., Aronson, A. & Brown, J. Motor Speech Disorders. Philadelphia, W.B. Saunders, 1975.

2) Kent, R., Sufit, R., Rosenbek, J., Kent, J., Weismer, G., Martin, R., & Brooks, B. Speech deterioration in amyotrophic lateral sclerosis: A case study. Journal of Speech and Hearing Research. 34, 1269-1275, 1991.

3) Leeper, H., Millard, K., Bandur, D., & Hudson, A. An investigation of deterioration of vocal function in subgroups of individuals with ALS. Journal of Medical Speech-Language Pathology. 4, 163-181, 1996.

4) Renout, K., Leeper, H., Bandur, D., & Hudson, A. Vocal fold diadochokinetic function in individuals with amyotrophic lateral sclerosis. American Journal of Speech-language Pathology. 4, 73-80, 1995.

5) Herbert, K. & Leeper, H. Effects of disease progression on speech in individuals with ALS. Paper presented at the American Speech-Language-Hearing Association Meeting, December, 1995.

6) Langmore, S. & Lehman, M. Physiologic deficits in the orofacial system underlying dysarthria in amyotrophic lateral sclerosis. Journal of Speech and Hearing Research. 37, 28-37, 1994.

7) Caruso, A. & Burton, E. Temporal acoustic measures of dysarthria associated with amyotrophic lateral sclerosis. Journal of Speech and Hearing Research. 30, 80-87, 1987.

8) Strong, M. ALS: Update on a tragic illness. Medicine North America. September, 780-788, 1995.

ACOUSTIC CORRELATES OF MONOTONE SPEECH IN PARKINSON'S DISEASE

Scott Adams, Kim Reyno-Briscoe
and Laura Hutchinson
Communication Sciences and Disorders
University of Western Ontario

INTRODUCTION

One of the most frequently perceived symptoms in Parkinson's disease (PD) is monotone speech[1,2]. Speech in PD is often described as monotonous or lacking in emotional expressiveness. Terms such as monopitch (flatness in vocal pitch) and monoloudness (flatness in vocal loudness) are also used to describe the monotone speech of PD[1]. Acoustic studies of PD speech have suggested that monotone speech may be associated with a reduction in frequency variation and intensity variation during spoken sentences[3,4]. However, only one report has attempted to correlate these acoustic measures with perceptual ratings of monotone speech[5]. Unfortunately, this study[5] failed to observe a significant correlation between the acoustic and perceptual measures of monotone speech. Thus, the results of these previous studies indicate that the acoustic correlates of monotone speech remain to be determined. The purpose of the present study was to continue to investigate the acoustic correlates of monotone speech in PD using a variety of novel acoustic measures. One such measure that has not previously been examined in PD speech is fundamental frequency (F0) declination. In normal speech production, F0 declination is typically seen as a gradual decrease in pitch (F0) from the beginning to the end of a sentence. F0 declination has received attention in several previous studies of normal speech[6,7,8]. However, the relationship between F0 declination and the perception of monotone PD speech has not been previously examined.

METHODS

Perceptual ratings and acoustic measures of speech were obtained from 3 groups of subjects: 10 PD patients with mild monotone speech, 10 PD patients with severe monotone speech, and 10 age-matched normals. All subjects read aloud a paragraph (the grandfather passage)[1]. The second sentence from this paragraph was used to obtain both the perceptual and acoustic measures. A direct magnitude estimation procedure[9], involving 5 listeners, was used to obtain perceptual measures of monotone speech. In particular, perceptual ratings were obtained for two dimensions: monopitch and monoloudness.

Several acoustic measures were obtained from the fundamental frequency (F0) and intensity contours of the subjects' spoken utterances. All acoustic measures were obtained using the Computer Speech Lab (CSL) program (Kay Elemetrics, Inc.). One set of measures examined the overall variability of the F0 and intensity contours. These included the standard deviation, the range, and the coefficient of variability of F0 and intensity contours. A second set of measures examined the declination pattern of the F0 and intensity contours. These included estimates of the slope of the F0 and the slope of the intensity contours across various segments of the utterance.

RESULTS

The severe PD patient group was perceived to have significantly higher monopitch and monoloudness scores than the mild PD patients and the normals (Table 1). The severe PD patient group was also found to have significantly lower F0 slope values than the other two groups (Table 1.)

The acoustic measures of the declination pattern were found to have the highest correlations with the perceptual measures of monotone speech (Table 2). In particular, the slope of the F0, estimated across an entire phrase, was found to have the highest correlation ($r=.70$) with the perceptual ratings of monopitch and monoloudness.

DISCUSSION

These results indicate that PD patients with severe monotone speech show a significantly different F0 declination pattern than less severe PD patients and normals. In particular, these monotone PD patients showed a much more gradual F0 declination (almost a flat F0) than was seen in the normal and mild PD subjects (see Figure 1). The finding of a fairly high correlation between F0 declination and monotone PD speech has not been previously reported. In the one previous report that examined the acoustic correlates of monotone speech in PD, the results were disappointing[5]. In particular, Ludlow and Bassich[5] failed to find a significant correlation between several measures of F0 change in sentences and perceived monopitch ($r= -.10$ to $.32$). The present results suggest that a potentially important psychoacoustic relationship may exist between F0 declination and perceived monotone speech in PD. Previous clinical efforts to treat monotone PD speech have not focused on the F0 declination patterns[10,11]. The present findings suggest that the systematic modification of F0 declination needs to be evaluated in future studies of monotone PD patients.

REFERENCES

1. Darley FL, Aronson AE, Brown JR: **Motor Speech Disorders**,
Philadelphia, W. B. Saunders Company, 1975, pp 192-297.

2. Adams, S.G. "Hypokinetic dysarthria in Parkinson's disease". In M. R. McNeil (Ed.), **Clinical Management of Sensorimotor Speech Disorders**. New York: Thieme Pub., 1997.

3. Canter GJ: Speech characteristics of patients with Parkinson's disease. **J Speech & Hearing** 1963; 28: 221-229.

4. Metter EJ, Hanson WR: Clinical and acoustic variability in hypokinetic dysarthria. **Communication Disorders** 1991; 19: 347-366.

5. Ludlow CL, Bassich CJ: Relationships between Perceptual Ratings and Acoustic Measures of Hypokinetic Speech, in McNeil MR, Rosenbek JC, Aronson AE (eds): **The**

Dysarthrias: Physiology, Acoustics, Perception, Management. San Diego, College-Hill Press, 1984, pp 163-192.
6. Cohen A, Collier R, 'tHart J: Declination: Construct or intrinsic feature of pitch? **Phonetica** 1982; 39: 254-273.
7. Gelfer CE, Harris KS, Collier R: Is declination actively controlled? In I. Titse (Ed), Vocal fold physiology. Iowa City: Iowa University Press.
8. Pierrehumbert J: The perception of fundamental frequency declination. **J. Acoustical Society of America** 1979; 66: 363-369.
9. Schiavetti N, Metz D, Sitler R. Construct validity of direct magnitude estimation and interval scaling of speech intelligibility: Evidence from a study of the hearing impaired. **Journal of Speech and Hearing Research** 1981; 24: 441-445.
10. Scott S, Caird FI: Speech therapy for patients with Parkinson's disease. **British Medical Journal** 1981; 283:1088.
11. Scott S, Caird FI: Speech therapy for Parkinson's disease. **Journal of Neurology, Neurosurgery, and Psychiatry** 1983; 46:140-144.

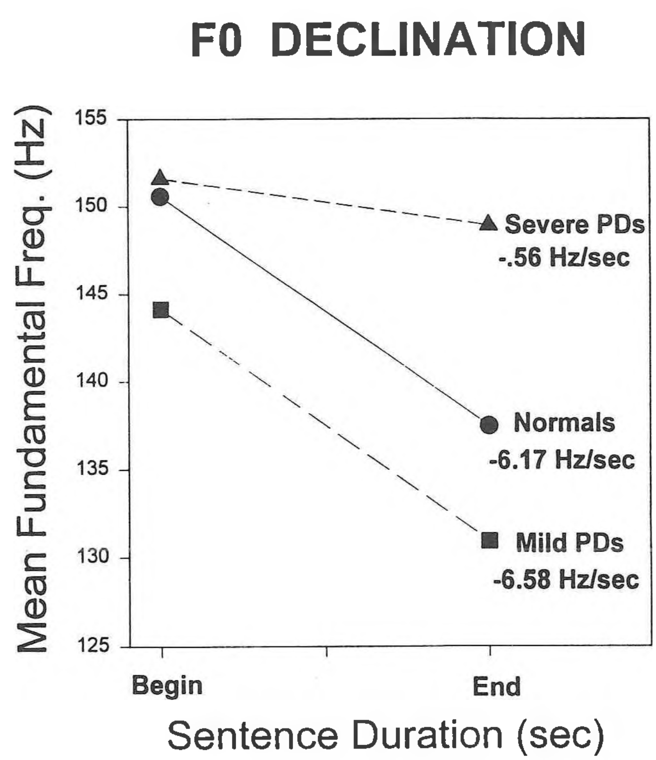
Table 2. Correlations between the perceptual and acoustic measures of monotone speech (Pearson r values obtained using data from all 30 subjects).

	Monopitch	Monoloud
Monopitch	--	--
Monoloud	.92	--
Slope F0	-.69	-.59
Slope Intensity	-.26	-.21
Stand. Dev. F0	-.47	-.36
Range F0	-.14	.08
Stand. Dev. Intensity	-.22	-.24
Range Intensity	.10	.13

Table 1. Mean perceptual and acoustic measures for the normal and PD groups (* = significant t-test, p<.05).

	Normals	Mild PDs	Severe PDs
Monopitch	66.5 (20.2)	53.0 (15.1)	183.5 (33.7)*
Monoloud	65.0 (17.2)	66.5 (19.5)	182.0 (19.8)*
Slope F0	-6.17 (1.33)	-6.58 (1.54)	-0.56 (1.72)*
Slope Intensity	-1.11 (1.71)	-0.72 (1.01)	-0.47 (0.60)
Stand. Dev. F0	16.0 (3.78)	16.6 (6.51)	12.5 (5.50)
Range F0	88.6 (24.5)	82.1 (30.5)	87.4 (53.7)
Stand. Dev. Intensity	3.59 (0.92)	3.53 (0.64)	3.57 (0.62)
Range Intensity	19.3 (2.91)	19.3 (3.24)	19.5 (2.68)

Figure 1. Estimates of F0 declination and the mean F0 slope values for the PD and normal subject groups.



Acoustic Structure of Stops Produced by Tracheoesophageal Speakers

Philip C. Doyle, Ph.D., School of Communication Sciences and Disorders, Voice Production Laboratory, The University of Western Ontario, London, Ontario.

Overview

Use of the tracheoesophageal (TE) voice prosthesis (Singer & Blom, 1980) for postlaryngectomy voice restoration has become standard practice across North America. TE voice production involves use of pharyngoesophageal muscular tissue as a vibratory source following removal of the larynx. The prosthesis permits pulmonary air to act as an aerodynamic driving source to this vicarious voicing mechanism. Tracheoesophageal (TE) speech has provided an added rehabilitation option to individuals who undergo laryngectomy . TE speech is supplied by pulmonary air, thus, distinguishing it aerodynamically from esophageal speech. Use of pulmonary air has been shown to favorably affect acoustic aspects of TE voice (Robbins, 1984). Yet concerns about the relative impact of a pulmonary air source on temporal features of TE speech have been raised (Doyle, Danhauer, & Reed, 1988; Weinberg, Horii, Blom, & Singer, 1982). It has been suggested that increased access to pulmonary air may allow the PE segment to initiate and terminate vibration more rapidly (Doyle et al., 1990; Robbins, Christensen, & Kempster, 1986), and hence, may potentially result in unique perceptual confusions (Doyle et al., 1988).

Perceptually, data have shown that TE speakers exhibit some unusual voicing patterns for cognate phonemes (Doyle et al., 1988; Doyle et al., 1990, Doyle & Haaf, 1989; Gomyo & Doyle, 1989) with a tendency toward voiceless-for-voiced cognate errors. Further, simple patterns of voice onset time do not appear to correlate well with the perceived phoneme in symmetrical CVC constructions. Thus, the purpose of this investigation was to identify and describe the temporal acoustic structure of stop production within an intervocalic stimulus context. These acoustic measures were obtained from a small group of excellent TE speakers.

Method

Three adult male TE speakers who were selected from a larger pool of 31 speakers judged to be "excellent" speakers. The larger group of speakers were initially identified and referred by experienced SLPs as being among the best TE speakers they had encountered. These 3 speakers were consistently identified by at least 2 independent judges as "superior" speakers within the group.

Speech Stimuli

The stimuli under investigation in the present work were comprised of the six English stops (/p,t,k,b,d,g/). Each stop was produced within a nonsense CVCVC construction. The stimuli were all initiated with the nasal /m/, followed by one of the vowels; the mid-consonant (one of the target stops) was then followed by the same vowel and then terminated with the nasal (Doyle et al., 1988). The nasal /m/ was chosen because of its demonstrated ease for alaryngeal speakers, and the two vowels were used because they have been shown to exhibit the greatest amplitude for esophageal vowels. Six samples of each stop with each vowel were obtained from each speaker. Stimuli were produced in the phrase " _______ is a word".

Procedure and Data Analysis

Speakers were recorded in a sound suite on research quality equipment with the microphone at a fixed distance. All speaker samples were perceptual evaluated using an open-set response paradigm. Stimuli also were acoustically analyzed using the Canadian Speech Research Environment (CSRE) software (Jamieson & Nearey, 1988). CSRE provided broadband spectrograms with concommitant amplitude displays for analysis. The temporal acoustic measure of voice onset time (VOT) was measured according to the procedures outlined by Lisker and Abramson (1967) and used by Robbins et al. (1986).

Perceptual Assessment

The entire pool of speaker stimuli were randomized and submitted to perceptual evaluations. Listeners were 10 naive listeners who had no prior exposure or experience with alaryngeal speech. Listeners were requested to transcribe their identification of the middle consonant in the stimuli. These data were then collated and placed in a confusion matrix for further evaluation of intelligibility.

Results

Perceptual Evaluation

Based on the confusion matrices generated, it was determined that 47% of the voiceless stops were correctly identified by listeners. In contrast, 79% of voice stops were correctly identified. Thus, a clear advantage in production and perception of voiced targets when compared to their voiceless cognates was observed.. This finding is inconsistent with earlier data (Doyle et al., 1988; Doyle & Haaf, 1989; Gomyo & Doyle, 1989).

Speaker 1

Speaker 1 showed consistent increases in VOT (Table 1) as loci for stops moved from front-to-back for both V+ and V- stops. This was most apparent with the vowel /i/. Relating to the voiced-voiceless distinction, Speaker 1 tended to exhibit relatively rapid VOTs (<25 msec) for the more anterior voiceless targets (/p/ and /t/).

Table 1

Means, standard deviations and ranges of VOT (in msec): Speaker 1

Stop	Vowel	M VOT	SD	Range
/p/	/i/ /u/	10.37 10.40	5.31 2.61	5.2-15.8 8.7-13.4
/t/	/i/ /u/	25.30 22.63	6.54 4.54	17.8-29.8 17.8-26.8
/k/	/i/ /u/	46.67 18.43	6.61 9.26	40.3-53.5 8.9-27.4
/b/	/i/ /u/	8.03 14.00	6.78 4.76	3.3-15.8 9.1-18.[illegible]
/d/	/i/ /u/	31.80 17.73	3.55 6.73	27.7-33.9 13.7-25.5
/g/	/i/ /u/	46.43 47.67	9.26 6.37	37.0-55.5 41.9-54.5

Speaker 2

Speaker 2 exhibited similar mean VOT (Table 2) patterns across stimuli, with values ranging from +30 to +60 msecs. Data also indicate that Speaker 2 increased VOT based on articulatory loci for V+ targets (front to back), but was inconsistent for V- stops. Overall, Speaker 2 exhibited VOTs >25 msec for all stops.

Table 2

Means, standard deviations and ranges of VOT (in msec): Speaker 2

Stop	Vowel	M VOT	SD	Range
/p/	/i/ /u/	52.13 39.73	14.46 5.33	43.0-68.8 35.8-45.8
/t/	/i/ /u/	33.43 50.60	3.76 6.66	30.6-37.7 45.8-58.2
/k/	/i/ /u/	60.70 51.80	8.82 13.09	54.5-70.8 42.7-66.8
/b/	/i/ /u/	31.10 31.83	5.95 5.09	25.1-37.0 28.6-37.7
/d/	/i/ /u/	35.13 33.70	9.75 3.38	26.9-45.9 29.8-35.7
/g/	/i/ /u/	51.80 41.17	10.28 1.69	40.2-59.8 39.3-42.6

Speaker 3

Speaker 3 exhibited consistent decreases in mean VOT (Table 3) for V+ stops, regardless of vowel. While fairly consistent patterns of VOT were noted for bilabial and velar cognate pairs regardless of vowel, relatively greater VOT differences were noted for alveolars..

Table 3

Means, standard deviations and ranges of VOT (in msec): Speaker 3

Stop	Vowel	M VOT	SD	Range
/p/	/i/ /u/	17.23 19.83	4.16 6.36	12.6-22.8 12.5-23.8
/t/	/i/ /u/	28.80 46.17	5.98 7.40	25.1-35.7 38.7-53.5
/k/	/i/ /u/	29.33 25.40	6.15 10.50	23.3-35.6 13.7-34.0
/b/	/i/ /u/	10.33 14.57	2.66 7.91	8.7-13.4 9.9-23.7
/d/	/i/ /u/	23.97 34.23	2.87 5.66	22.3-27.3 29.1-40.3
/g/	/i/ /u/	24.67 15.50	8.04 5.98	17.4-33.3 9.9-21.8

VOT by Temporal Cluster

In order to provide a comparative index of each speaker's productive performance, VOT data were clustered into arbitrarily selected time categories for further evaluation (0-25 msec, 26-50 msec, and 51-75 msec). The data reveal that all three speakers exhibited +VOT values (i.e., post-stop release). These values were highly individual to speakers. Speaker 1 produced a majority of stops with VOTs between 0-50 msec, Speaker 2 within the range of 26-75 msec, and Speaker 3 within the range of 0-50 msec.

Summary and Conclusions

Normal English speakers have been shown to systematically vary VOT by increasing durations as the stop moves from labial to velar loci (Lisker & Abramson, 1967). This variation has been noted with excellent esophageal speakers. Speaker 1 did exhibit this pattern for all V+ stops in both intervowel contexts. However, for V- stops this pattern was only noted with the /u/ vowel. Speaker 3 exhibited this pattern for both V+ and V- in the /i/ vowel context, but was inconsistent with /u/.

When the temporal cluster data were inspected, these speakers exhibited idiosyncratic patterns which are likely related to unique postsurgical productive systems. It should be noted that all speakers produced V+ stops with positive VOTs. This is likely due to the effects of a powerful aerodynamic system on the PE segment. Further, these TE speakers produced unique temporal clusters in relation to VOT. Again, this may be due to aerodynamic influence, postsurgical anatomy, or both. Although limited, the present data suggest that multiple acoustic cues likely signal stop perception for TE speakers. Additional research in our laboratory suggests that proficient TE speakers are able to effectively signal voicelessness, even when acoustic data reveal that no break in voicing exists. This raises questions regarding the use of air in the vocal tract and upper airway turbulence as a compensatory mechanism in these speakers. This finding and the present data suggest that the proficiency of TE speech may be related how the speaker is able to utilize the air source once it trangresses the pharyngoesophageal segment (the muscular tissue which serves as an alaryngeal voice source). Time varied features of stops are worthy of further inquiry (Kewley-Port, 1993).

Previous research with normal speakers has shown that they vary VOT to distinguish prevocalic V- stops from V+ cognates. Speaker 1 did not effect this distinction. In most cases, his VOTs for stops within cognate pairs were similar. This suggests a restricted VOT range. Although Speaker 2's VOTs were produced within a rather narrow range and were always >25 msec, all V- stops except one were produced with VOTs greater than those noted for V+ cognates. It appears that Speaker 2 attempted to make this distinction, but as a result of physiologic constraints, could only do so within a restricted VOT range. Subject 3 produced V- stops /t/ and /k/, but not /p/ with overall VOTs >25msec. Thus, inconsistency in relation to articulatory loci was again noted. This finding further confirms the idiosyncratic patterns that may characterize many TE speakers.

References

Doyle, P.C., Danhauer, J.L., & Lucks, L.E. (1990). SINDSCAL analysis of perceptual characteristics of consonants produced by esophageal tracheoesophageal talkers. Journal of Speech and Hearing Disorders. 55. 756-760.

Doyle, P.C., Danhauer, J.L., & Reed, C.G. (1988). Listeners' perceptions of consonants produced by esophageal and tracheoesophageal talkers. Journal of Speech and Hearing Disorders. 53. 400-407.

Doyle, P.C., & Haaf, R.G. (1989). Pre-and post-vocalic consonant intelligibility in tracheoesophageal talkers. Journal of Otolaryngology. 18. 350-353.

Doyle, P.C., Swift, E.R., & Haaf, R.G. (1989). Effects of listener sophistication on judgments of tracheoesophageal talker intelligibility. Journal of Communication Disorders. 22. 105-113.

Gomyo, Y., & Doyle, P.C. (1989). Perception of stop consonants produced by esophageal and tracheoesophageal speakers. Journal of Otolaryngology. 18. 184-188.

Kewley-Port, D. (1983). Time-varying features as correlates of place of articulation in stop consonants. Journal of the Acoustical Society of America. 72. 322-335.

Lisker, L., & Abramson, A.S. (1967). Some effects of context on voice onset time in English stops. Language and Speech. 10. 1-28.

Robbins, J. (1984). Acoustic differentiation of laryngeal, esophageal, and tracheoesophageal speech. Journal of Speech and Hearing Disorders. 27. 577-585.

Robbins, J., Christensen, J., & Kempster, G. (1986). Characteristics of speech production after tracheoesophageal puncture: Voice onset time and vowel duration. Journal of Speech and Hearing Research. 29. 499-504.

Singer, M.I., & Blom, E.D. (1980). An endoscopic technique for restoration of voice after laryngectomy. Annals of Otology. Rhinology. and Laryngology. 89. 529-533.

TEMPORAL CUES SUPPORT SYNTACTIC IDENTIFICATION

Michelle D. Wiley and M. Kathleen Pichora-Fuller
School of Audiology & Speech Sciences, University of British Columbia
5804 Fairview Ave., Vancouver, BC, V6T 1Z3

INTRODUCTION

Previously, we demonstrated that normal-hearing young listeners' comprehension of complex syntactic structures declined more than comprehension of simpler syntactic structures as the signal-to-noise ratio became more adverse, with these declines not being attributable simply to declines in word recognition (Dillon, 1995; Pichora-Fuller, Lloyd, Dillon, & Kirson, 1998). The present study was conducted to examine the roles of temporal and spectral cues in syntactic processing.

METHOD

Participants. Fifteen normal-hearing, female, native English-speakers between the ages of 21 and 29 were paid for their participation in Experiment 1. An additional 5 listeners participated in Experiment 2. Their reading working memory spans (Daneman & Carpenter, 1980) ranged from average (2.7) to excellent (5.7). All were university graduates with sufficient knowledge of linguistics to ensure familiarity with the syntactic classifications used in the experiment.

Materials. Sentence materials were prepared previously (Dillon, 1995) based on the sentences used in a comprehension test developed for use with aphasics (Caplan, Baker, & Dehaut, 1985). Dillon's materials consisted of three equivalent lists of 45 sentences; each list had 5 exemplars of each of 9 different syntactic types (Table 1).

Table 1: Examples of Syntactic Types

Code	Type	Example
A	Active	The duck chased the mouse.
P	Passive	The owl was tapped by the pig.
CS	Cleft Subject	It was the fox that touched the mouse.
CO	Cleft Object	It was the duck that the owl bumped.
D	Dative	The dog smacked the frog to the goose.
DP	Dative Passive	The pig was kicked to the duck by the fox.
C	Coordinated	The owl tripped the dog and grabbed the pig.
SO	Subject-Object	The mouse that the duck grabbed bumped the owl.
OS	Object-Subject	The pig kissed the owl that bumped the mouse.

For the present study, three new versions of the sentence lists were prepared. One list of sentences was spectrally-inverted. For another list, words read in citation form were recorded and then concatenated with a 50 msec silence between each word to form sentences with the same lexical (spectral) content but without normal sentence prosody. For a third list, the concatenated sentences without normal sentence prosody were also spectrally-inverted. All test materials were digitized at a sampling rate of 32 kHz and stored in soundfiles.

Following the technique described by Blesser (1972), spectral inversion was accomplished by band-pass filtering (.2-4 kHz, 48 dB/octave) an audio tape-recording of the intact sentences, applying spectral pre-emphasis of 24 dB/octave between .5 and 3.5 kHz, and then spectrally inverting (around 2.1 kHz) using customized in-house hardware (Benguerel, 1998). The consequence of the spectral inversion manipulation was to severely disrupt the spectral properties of speech while preserving its temporal envelope and some pitch information.

In a pilot experiment with 10 other participants, it was found that words and phrases excised from the spectrally-inverted sentences were almost never recognized when an open-set response format was used, whereas the corresponding segments excised from the intact version of the sentences were recognized nearly perfectly. However, following familiarization with the corresponding intact materials and using a close-choice test format with 34 alternatives, re-testing with the spectrally-inverted materials yielded scores close to 30% correct. The increase in scores seemed to be due mostly to the ability of the listeners to correctly identify the few multi-syllabic alternatives; of the 34 alternatives, 5 had two syllables and 2 had three syllables.

Conditions. In Experiment 1, three blocked conditions were presented in the following fixed order: first, an "intact condition" using the list of sentences with normal spectral and temporal cues; second, a "spectrally-inverted condition" using the list of sentences with reduced spectral cues; third, a "concatenated condition" using the list of sentences with reduced temporal cues. In Experiment 2, we tested a different group of listeners in a "concatenated-inverted condition" using the list of sentences with both spectral and temporal cues reduced.

Procedures. Sentences were presented monaurally at 70 dB SPL in quiet over TDH 39 earphones to listeners in a double-walled IAC booth. CSRE software and TDT hardware were used to control the experiment. An example of each of the nine possible syntactic types was displayed adjacent to the numbers 1 to 9 on a computer screen. Following presentation of each sentence, the listener used the computer mouse to select the number of the sentence structure they thought they heard. All subjects received a practice session prior to the test conditions.

RESULTS

Performance was virtually perfect in the intact and near-perfect (98.9% correct) in the concatenated condition. Even in the spectral inversion condition, most sentence structures remained highly recognizable (81.4% correct). In the spectral inversion condition, fewer errors were made on the simpler syntactic types than on the more complex types (Figure 1). Most errors involved confusions between structures with the same number of syllables (Tables 2 and 3). In contrast, performance was only around chance (12.4% correct) for the sentences with both spectral and temporal cues reduced. Furthermore, there was no obvious pattern to the errors when both the spectral and temporal properties of the sentences were reduced.

Figure 1: Mean number of sentences of each syntactic type identified correctly in the four conditions (intact, inverted, concatenated, concatenated and inverted).

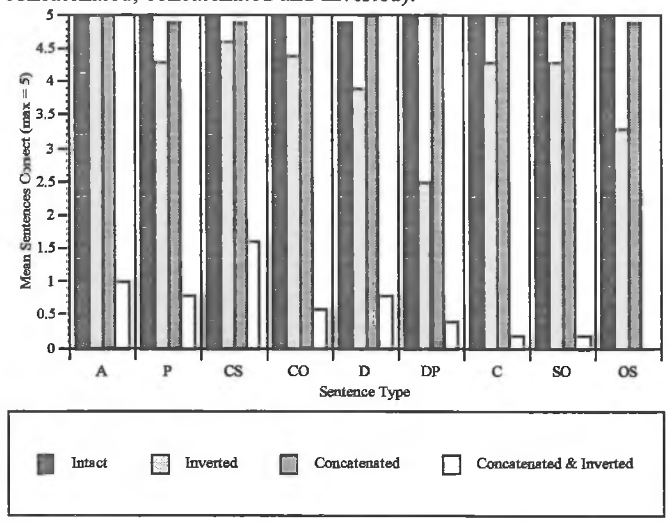


Table 2: Confusion matrix for target-response patterns in the spectral inversion condition by sentence type. Number of responses are shown (max. = 75).

	Target Sentence Type								
Response	A	CS	CO	C	P	SO	D	OS	DP
A	75		1		8				
CS		69	6	1		2		2	
CO		3	66			1		1	2
C				65			3	16	14
P		3	1		64	2	8		5
SO						64	1	1	4
D				3	2	4	59	3	12
OS				6		1		50	1
DP			1		1	1	4	2	37

Table 3: Confusion matrix for target-response patterns according to number of syllables. Sentence type A has 5 syllables; sentence type P has 7 syllables; sentence types CS, CO, and D have 8 syllables; sentence types C, SO, and OS have 9 syllables; sentence type DP has 10 syllables.

	Number of Syllables in Target Sentence				
Number of Syllables in Response	5	7	8	9	10
5	100%	11%	1%		
7		85%	5%	1%	7%
8		3%	90%	8%	19%
9			2%	90%	25%
10		1%	2%	1%	49%

Also of interest is the finding that subjects with low working memory spans were less able to identify spectrally inverted sentences than were those with high working memory spans. Correlations between working memory span and sentence identification scores were highly significant ($p<.05$) for all syntactic types except OS and A, with the highest correlations being observed for the sentence types C ($r = .83$) and SO ($r = .80$) both of which are two-verb structures. Note that it was not possible to observe meaningful correlations with scores for sentence type A since no errors were made on this sentence type and that the majority (64%) of errors on type OS, the third two-verb structure, involved the choice of one other sentence type with the same syllabic pattern (C).

DISCUSSION

The results of this experiment demonstrate that listeners are able to correctly identify most common syntactic structures when spectral cues are severely reduced but prosodic cues are largely preserved or when normal sentence-level prosodic cues are altered but spectral cues are preserved, whereas they are unable to perform better than chance when both types of cues are reduced.

An examination of the pattern of errors in the spectral inversion condition suggests that listeners are guided by syllabicity cues insofar as they usually choose sentence types with the same (or slightly fewer) syllables than the target sentence type. For the two sentence types (OS and DP) that were most poorly recognized, the most frequent response was a more common sentence type with a shallower syntactic tree structure (C). These findings are compatible with the recent arguments presented by Greenberg (1998) regarding he usefulness of syllabic information for speech intelligibility in everyday listening conditions.

The high correlation between working memory span and accuracy in identifying the sentence types, especially two of the three more complex two-verb types with a high number of syllables, is consistent with the idea that subjects are storing information about the number and relative timing of syllables and using this information to select one of the alternative sentence types.

REFERENCES

Benguerel, A.-P. (1998). Stress-timing vs syllable-timing vs. mora-timing: The perception of speech rhythm by native speakers of different languages. Etudes et Travaux, Institute des Langues Vivantes et de Phonétique, Université Libre de Bruxelles.

Blesser, B. (1972). Speech perception under conditions of spectral transformations, I: Phonetic characteristics. Journal of Speech & Hearing Research, 15, 5-41.

Caplan, D., Baker, C., & Dehaut, F. (1985). Syntactic determinants of sentence comprehension in aphasia. Cognition, 21, 117-175.

CSRE (4.2) (1995). Computer Speech Research Environment. London, Ontario: AVAAZ Innovations Inc.

Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behaviour, 19, 450-466.

Dillon, L.M. (1995). The effect of noise and syntactic complexity on listening comprehension. Master of Science Thesis, University of British Columbia.

Greenberg, S. (1998). Auditory processing of speech. Invited tutorial, NATO Advanced Study Institute on Computational Hearing, Il Ciocco, Italy.

Pichora-Fuller, M.K., Lloyd, V., Dillon, L., & Kirson, R. Effects of sentence type and background noise on comprehension by listeners with low- or high-span working memory. Contributed poster, Seventh Cognitive Aging Conference, Atlanta, Georgia, April 1998.

Mismatch Negativity Measures of Gap Detection

Renée Desjardins and Laurel Trainor
Department of Psychology, McMaster University, Hamilton, ON, L8S 4K1

Introduction

This paper examines the feasibility of using Event Related Potentials to measure temporal resolution in adults. Measures of temporal resolution have recently become interesting to language acquisition researchers because of claims that certain kinds of language impairment may result from an underlying lack of temporal resolution (e.g., Lowe & Campbell, 1965; Tallal et al., 1985; Tallal et al., 1993). Many of the features distinguishing speech sounds, such as voice onset time, rely on timing differences of a few milliseconds. A child diagnosed with poor temporal resolution also presumably had poor temporal resolution in infancy. Currently, children with language impairment are not usually diagnosed until at least 3 or 4 years of age. Thus, if temporal resolution could be measured in prelinguistic infants, it could be used to identify infants at risk for later language impairment.

Gap detection is the most common way of measuring temporal resolution. We used Gaussian-enveloped sine wave tones rather than broad band noise as markers for the silent gap as we are interested in gap detection at particular frequencies in the absence of adaptation effects. Gaussian envelopes were used as they minimize the spectral splatter that occurs when a sound is turned on or off and the degree of spectral splatter is independent of the size of the gap (see Schneider et al., 1994). No-gap standard stimuli were constructed as in Schneider et al. (1994) to match the gap stimuli in duration and energy, and to approximate them in spectral content.

The purpose of the present study is to find a measure that does not require attention or a behavioural response. The mismatch negativity (MMN) appears to be ideally suited to this purpose. When an infrequent stimulus (e.g., a tone with an embedded silent gap) occurs in a series of frequent stimuli (e.g., a tone with no gap), the electrical activity recorded at the scalp is more negative for the infrequent than for the frequent stimuli between about 100-300 ms after the onset of the stimulus. MMN appears to be a rather pure measure of sensory processing, as it is affected very little by attention (Näätänen, 1992).

In the present study, we attempted to measure adults' gap detection thresholds using MMN.

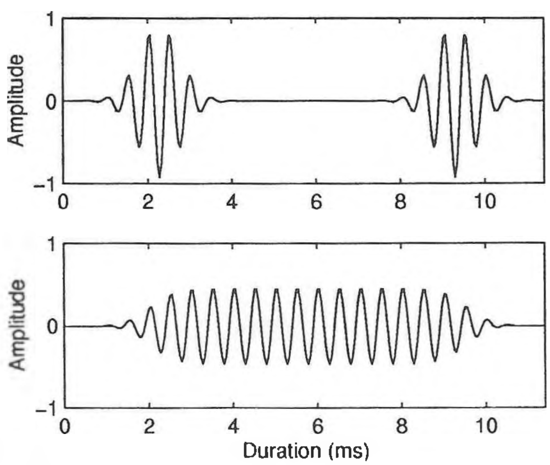


Figure 1. The gap (upper panel) and no-gap (lower panel) stimuli for gap size 7.

Method

Participants

Eight adults (age range = 21 to 24 years; 4 female, 4 male) were tested with all gap sizes. The data for one condition (gap size 8 ms, see *Stimuli*) was unusable for one participant.

Stimuli

In each of 5 conditions, gap stimuli were constructed with two 2000 Hz Gaussian-enveloped tone pip markers (standard deviations of .5 ms) whose peak amplitudes were separated by 4, 5, 6, 7, or 8 ms (see Figure 1). The matching no-gap stimuli were created as in Schneider et al. (1994) to match the gap stimuli in duration and energy, and roughly in spectral content (see Figure 1).

Apparatus

The sounds were generated with inhouse software running on a Comptech pentium computer with a sound card. They were presented with a Denon PMA 480R amplifier and a Grason Stadler speaker. The EEG was recorded with NeuroScan 4.0 software, using 32-channel Synamps, and electrocaps with Ag/AgCL electrodes in a shielded room.

Procedure

A target-nontarget oddball methodology was used. In each gap condition (4, 5, 6, 7, 8 ms) 400 trials were presented, of which 80% were no-gap trials and 20% were gap trials. In order to mimic the infant procedure, adults were tested in a passive listening paradigm, whereby they were simply instructed to watch a screen saver.

Recordings

Recordings were made from the following 27 sites (see Figure 3): Fpz, Fp1, Fp2, Fz, F3, F4, F7, F8, FC1, FC2, FC5, FC6, Cz, C3, C4, T3, T4, T5, T6, PC5, PC6, Pz, P3, P4, Oz, O1, O2. The sampling rate was 500 Hz, and the bandpass was set between .15 and 30 Hz. Impedance levels were maintained below 5 kOhms. Cz was used as a reference during recording, although a common average reference was used for the analyses.

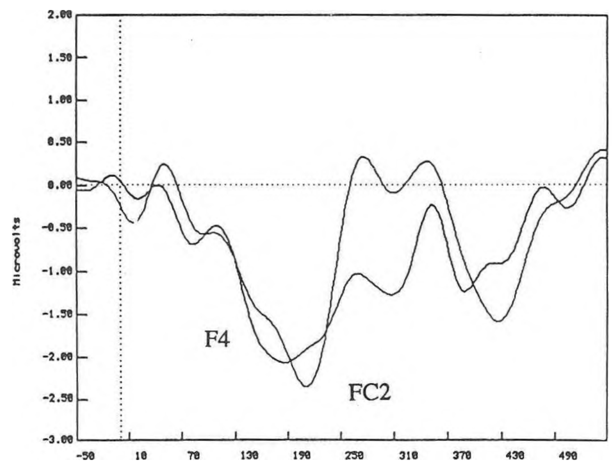


Figure 2. Difference waves (oddball minus standard) for gap size 7 at F4 (upper panel) and FC2 (lower panel).

Data Analysis

The recordings were low pass filtered at 18 Hz. Baseline was defined as the mean amplitude for the 50 ms preceding the onset of the stimulus. Epochs were defined as 550 ms beginning from the onset of the stimulus. FP1, FP2, F7, and F8 were used to rejected trials contaminated by eye movement artifact.

For each participant in each condition, the waveforms on the standard (no-gap) trial epochs were averaged together, as were the waveforms on the oddball (gap) trial epochs. Then the averaged standard waveform was subtracted from the averaged oddball waveform to create a difference wave. T-tests were employed to determine the portions of the difference wave between 100 and 300 ms that were significantly less than 0 across participants.

Results

The difference wave was significantly below 0 ($p < .05$) at a number of sites in the region expected for MMN for gaps of 8, 7, 6, and 5 ms. At 4 ms, there was significance only for very brief portions of the waveform at only 3 sites. From this, we can conclude that the gap threshold with these stimuli as measured by MMN is in the neighourhood of 4 ms. Figure 2 shows difference waves for a gap of 7 ms.

In detail, for 8 ms gaps, significance was found at FC2 (186-234 ms), Cz (170-194 ms), C4 (194-248 ms), and Pz (144-192 ms). For 7 ms gaps, significance was found at Fz (152-256 ms), F3 (178-216 ms), F4 (146-252 ms), FC2 (138-244 ms), FC6 (174-208 ms), C4 (176-222 ms), and CP6 (216-242 ms). For 6 ms gaps, significance was found at F4 (160-258 ms), Cz (166-232 ms), C4 (210-218 ms), and CP6 (194-206 ms). For 5 ms gaps, significance was found at Fz (166-196 and 218-236 ms), FC1 (152-240 ms), FC2 (126-248 ms), FC6 (224-264 ms), Cz (148-244 ms), C4 (210-250 ms), T4 (236-292 ms), CP6 (206-274 ms), and T6 (248-300 ms). For 4 ms gaps, significance was found at F8 (180-196 ms), FC6 (174-192 ms), and T4 (162-186 ms).

Figure 3 shows the distribution of sites showing significant MMN. It can be seen that the locus of the effect is predominantly right frontal.

Discussion

The threshold measured by MMN of around 4 ms is in agreement with the behavioural literature. Schneider et al. (1994) found that practiced young adults had thresholds in the neighbourhood of 2 to 3 ms with these stimuli. Philips et al. (1998) found that gap thresholds with broadband markers with similar leading-marker durations to those of the present study were between 4 and 6 ms for inexperienced listeners. Thus the MMN appears to be a good tool for measuring gap detection thresholds.

The MMN in the current study had a right frontal locus. According to Näätänen (1992), the MMN has two underlying generators, one located bilaterally in auditory cortex and one located in the right frontal hemisphere. As the stimuli were presented in the sound field (i.e., binaurally), it appears that we are primarily measuring the effects of the latter generator.

Although our task was a passive listening one, we did not follow the usual practice of having adults read a book or engage in problem solving tasks in order to eliminate all attention to the auditory channel. Our data actually shows evidence of a P3a following the MMN, perhaps indicating inadvertent capture of attention. However, even under these conditions, we could measure a clear MMN, which suggests that this methodology may be a good one for measuring thresholds in nonlinguistic infants. A study is currently underway to determine gap thresholds in infants using MMN.

References

Lowe, A. D., & Campbell, R. A. (1965). Temporal discrimination in aphasoid and normal children. *Journal of Speech and Hearing Research, 8,* 313-314.

Näätänen, R. (1992). *Attention and brain function.* Hillsdale: LEA.

Phillips, D. P., Hall, S. E., Harrington, I. A., & Taylor, T. L. (1998). "Central" auditory gap detection: A spatial case. JASA, 103, 2064-2068.

Schneider, B. A., Pichora-Fuller, M. K., Kowalchuk, D., & Lamb, M. (1994). Gap detection and the precedence effect in young and old adults. *JASA, 95,* 980-991.

Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiological basis of speech: A case for the preeminence of temporal processing. In Tallal, P., Galaburda, A. M., Llinas, R. R., & von Euler, C. (Eds.) *Temporal information processing in the nervous system: Special reference to dyslexia.* New York: New York Academy of Sciences.

Tallal, P., Stark, R. E., & Mellits, E. D. (1985). Identification of language-impaired children on the basis of rapid perception and production skills. *Brain and Language, 25,* 314-322.

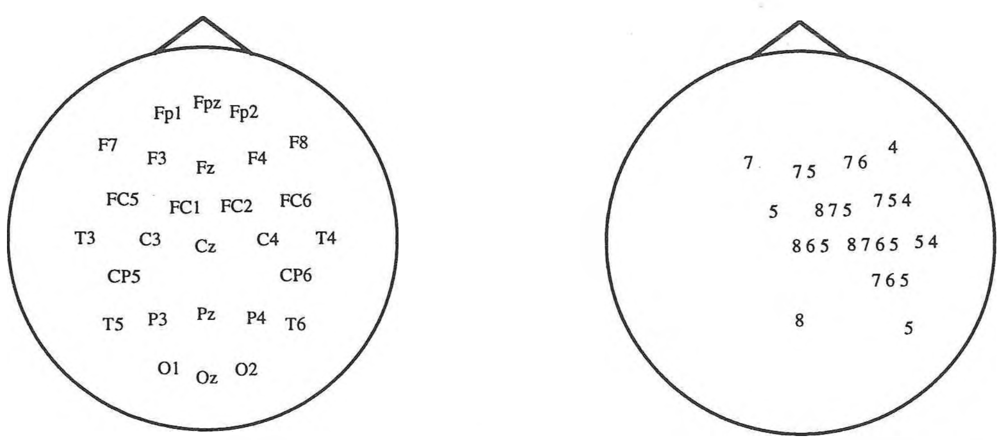


Figure 3. Left panel: The 27 recording sites. Right panel: Numbers represent where significant MMN was found for that gap size.

The influence of a secondary task on the understanding of continuous discourse by younger and older adults.

Dana R. Murphy, Meredyth Daneman, and Bruce A. Schneider
Department of Psychology
University of Toronto at Mississauga
Mississauga, ON L5L 1C6

Schneider, Daneman, Murphy, and Kwong See (1998) showed that when sound levels were adjusted to compensate for individual differences in hearing, the ability of both younger and older adults to answer questions based on connected discourse was nearly equivalent at all levels of noise. In the absence of such adjustments, older adults answered fewer questions correctly than younger adults. These two results suggest that poorer hearing, rather than a decline in cognitive processing, is the reason why older adults are less able to recall information from connected discourse. However, in the previous study, the participants were able to focus their attention on the connected discourse. In everyday listening situations, we often have to divide our attention between two different tasks. Thus, it is possible that age-related difference in performance would emerge in a divided attention situation even after stimulus levels had been adjusted for individual differences in hearing. In two experiments we tested younger and older adults' ability to extract and remember information from connected discourse in the presence of a distracting secondary task.

Experiment 1

Method

Participants

Twelve younger adults (mean age of 21.5) and twelve older adults (mean age of 71.5) participated in this study. All participants had normal hearing (thresholds at or below 25 dB HL up to 2000 Hz). Thresholds for low-predictability, sentence final words in the modified SPIN test were also determined for each participant.

Materials

Passages. Digital recordings were made of a male actor reading six passages. In the noise condition these digitized passages were added to a background consisting of a 12-speaker babble and presented to the listener over the right earphone.

Secondary Task. Throughout the experiment, a circle appeared on the computer monitor sitting in front of the participant inside the sound booth. At random points around its perimeter a small square would occasionally appear. Participants were required to use the ball on top of a stationary mouse to move the curser from the center of the circle to the square target on its perimeter as quickly as possible.

Procedure

Passages were presented either in quiet (Q) or in a moderate level of noise (N) under one of three distraction conditions. In the no distraction (ND) condition, participants only had to listen to the passage. In the low distraction (LD) condition, the circle appeared every 6 seconds and participants were required to move the cursor to the circle while listening to the passage; in the high distraction (HD) condition the circle appeared every 3 seconds. In all cases, participants were instructed to make listening to the story their primary responsibility. Immediately after each passage, the participants answered a series of multiple choice questions regarding the material they had just heard. One half of these questions concerned specific details mentions in the story (detail questions), while the other half of the questions required participants to synthesize material (integrative questions).

In all conditions, the story was presented monaurally to the right ear 50 dB above the level of each participant's right-ear babble threshold. In the noise condition, the level of noise was adjusted relative to each individual's low-context SPIN threshold. Those with a higher tolerance for noise (i.e. a low SPIN threshold) received a higher level of noise, while those with a low tolerance for noise received less noise. The ratio of discourse to babble for a person whose SPIN threshold was 0 dB was -12 dB. The SN ratio for other participants was obtained by adding -12 to each individual's low-context SPIN threshold.

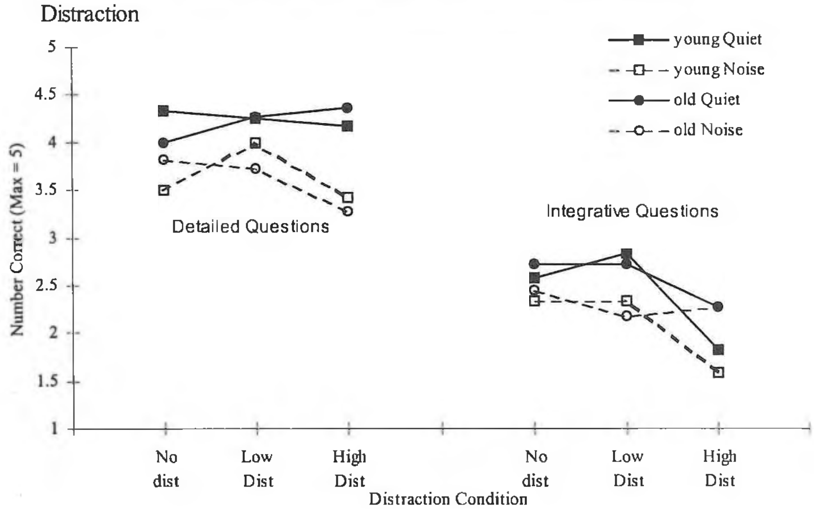
All tests were administered in a double-walled sound attenuating chamber.

Results

Figure 1 presents the number of questions answered correctly as a function of age, question type, noise, and distraction. A 2 Age (younger vs. Older) x 2 Question Type (Detail vs. Integrative) x 2 Noise Condition (Quiet vs. Noise) x 3 Distraction (No Distraction vs. Low Distraction vs. High Distraction) MANOVA was conducted on the number of questions answered correctly. This analysis revealed a significant main effect for question type, $F(1,21)=184.61$, $p<.0001$. Therefore, the data in Figure 1 are split according to this main effect, with the lines on the left hand side of the figure representing performance on detail questions and the lines on the right hand side of Figure 1 representing the performance on integrative questions. Interestingly, the younger and older adults performed similarly across the two types of questions (i.e. there was no age x question type interaction).

Noise had a significant influence on the ability to remember details from the discourse. This noise effect was similar for the two age groups. However, there was no effect of level of distraction, and no interactions of degree of distraction with any of the other variables including age. Thus, at least when the secondary task was visual, younger and older adults were equally good at completing the two tasks simultaneously. These

Figure 1
Mean number correct as a function of Question Type, Age, noise, and Distraction



results differ from what might be expected given the literature on the cognitive abilities of older adults. In particular, younger and older adults are often found to have more difficulty completing a secondary task at the same time that they complete a primary task (McDowd & Craik, 1988). However, in most research of this nature, the primary and secondary tasks tend to be very similar and tap the same cognitive resources. In Experiment 1, the primary task was a verbal task (listening to discourse), while the secondary task was non-verbal. Perhaps differences could be found between the age groups if a verbal distracting task were used. Thus, in Experiment 2, we had younger and older adults monitor the computer screen and answer true/false questions that appeared at regular intervals throughout the experiment. These questions were in the form of single sentences. Individuals were instructed to answer the questions as quickly and as accurately as possible while at the same time maintaining the task of listening to the story as their primary responsibility. Given that the participants were now required to focus on verbal information in both the primary and secondary task, we expected them to be more influenced by the introduction of the distracting secondary task. In addition, given the greater difficulty that older adults are know to have in selective and divided attention tasks, we expected the secondary distraction task to influence the performance of the older adults more than that of the younger adults.

Experiment 2

Method

Participants

A second group of 12 younger (mean age = 20.08) and 12 older adults (mean age = 73) who met the same screening criteria as in Experiment 1, served as participants for this experiment.

Materials

The discourse passages were the same as in Experiment 1. The distraction task consisted of the presentation of a sentence to which the participant was to respond "True" or "False".

Procedure

The identical two noise conditions were crossed with three levels of distraction: no distraction (ND); low distraction (LD, sentences presented every 20 seconds); and high distraction (HD, sentences presented every 10 seconds). All other testing procedures were the same as in Experiment 1.

Results and Discussion

Figure 2 presents the mean number of questions answered correctly as a function of question type, distraction condition, age group, and noise condition. The data were subjected to a 2 age (younger vs. older) x 2 question type (detail vs. integrative) x 2 noise (quiet vs. noise) x 3 distraction (no distraction vs. l ow distraction vs. high distraction) MANOVA. As can be seen in Figure 2, there was a significant Question Type effect with participants being most accurate with detail questions. Thus, the data in Figure 2 were split according to question type, with the lines on the left side of the figure representing the performance on detail questions and the lines on the right side representing the integrative questions.

Noise again reduced the accuracy with which questions were answered by the same amount for both younger and older adults. Interestingly, this influence of noise was only noticeable in the detail questions.

Unlike the visual secondary task used in Experiment 1, the T/F sentence secondary task used in this experiment proved to make answering questions more difficult. Again, the influence of distraction was similar between younger and older adults as the age x distraction interaction was not significant.

There was also a significant question type x noise interaction as well as a question type by distraction interaction. However, separate ANOVAs completed on the two different types of questions revealed that these interactions were driven by the fact that very little was happening for Integrative questions. That is, while there was a significant noise and distraction main effect among the detail question data, there was no effects for these two variables among the Integrative Question data.

General Discussion

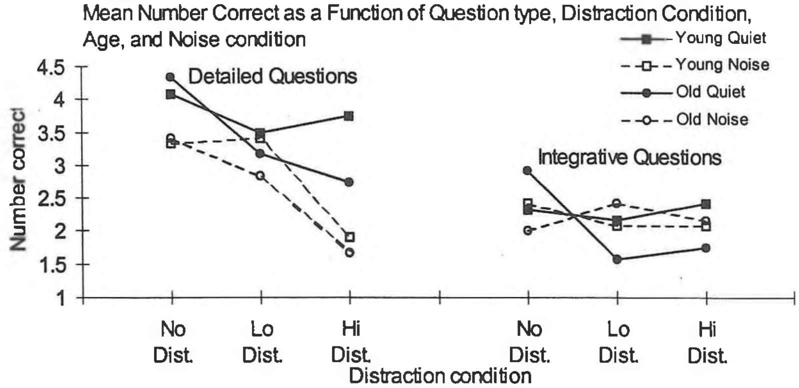
Younger and older adults with relatively good hearing were tested for their ability to hear and remember continuous discourse in background noise while simultaneously engaged in a second, lower-priority task. All participants were tested at signal and noise levels which compensated for differences in hearing (babble) threshold and differences in their ability to recognize individual words in noise. In two experiments, noise had identical effects on the ability of younger and older adults to remember continuous discourse. In Experiment 1, neither the younger nor the older adults were influenced by the requirement to perform a secondary visual task. However, in Experiment 2, the participants' ability to answer detail questions about the discourse dropped significantly when they were required to simultaneously answer a number of T/F questions. This indicates that in order for the secondary task to interfere with listening to the passage, it may have to engage some of the same processes that are being utilized to process the passage. Interestingly, the younger and older adults were equally affected by this secondary task. Thus, it appears that older adults without significant hearing loss can understand continuous discourse as well as younger adults if their individual thresholds for speech in noise are taken into consideration and that these older adults may also be able to deal with distracting and secondary events as efficiently as younger adults.

References

McDowd, J. M. & Craik, F. I. M. (1988). Effects of aging and task difficulty on divided attention performance. Journal of Experimental Psychology: Human Perception and Performance, 14, 267-280.

Schneider, B. A., Daneman, M., Murphy, D. R., & Kwong See, S. T. (1998). Discourse comprehension in noise: The effects of aging. Manuscript submitted for publication.

Figure 2
Mean Number Correct as a Function of Question type, Distraction Condition, Age, and Noise condition



Fundamental Frequency Influences Vowel Discrimination in 6-Month-Old Infants

Renée N. Desjardins and Laurel J. Trainor

Department of Psychology, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1

Introduction

The ability to discriminate different vowel tokens found in one's native language is a fundamental skill necessary for the acquisition of language. Previous research suggests that young infants are able to discriminate readily between vowel categories (e.g., Polka & Werker, 1994; Swoboda, Morse & Leavitt, 1976; Swoboda, Kass, Morse & Leavitt, 1978) as young as 8-weeks of age.

However, previous research has relied primarily on the use of tokens from a male speaker, possibly because these are easier to analyze using commercial speech analysis programs and to produce synthetically. Moreover, vowel discrimination studies have typically employed tokens in which the fundamental frequency (F0) remains constant over the token. One exception, is an experiment by Kuhl and Miller (1982) in which they used both monotone and falling pitch contours with a male speaker. Infants were able to ignore variations in pitch while attending to changes in the vowel category.

Unlike in the vowel perception experiments, speech to infants varies significantly in F0 (e.g., Fernald & Simon, 1984; Papousek, Papousek & Haekel, 1987) These variations in F0 as well as an overall increase in F0 are notable characteristics of infant-directed speech. Infants show a strong preference, both attentional and affective, for infant-directed speech over adult-directed speech even when the passages are in a non-native language (e.g., Werker, Pegg & McLeod, 1994).

This experiment was designed to study the effects of a female F0 on the discrimination of two vowel tokens, /i/ and /I/, both found in the English language and both known to be discriminable to very young infants (e.g., Swoboda et al, 1976; 1978). We chose to examine three different fundamental frequencies: the low steady F0 of a typical female voice in adult-directed speech, 250 Hz; the F0 more typical of speech directed to infants, 350 Hz; and a falling F0 reminiscent of infant-directed pitch contours, 350 Hz to 250 Hz.

Method

Subjects. Seventy-two infants (36 male and 36 female) between 6 months 0 days and 6 months 30 days were tested. All infants had no reported hearing problems, were healthy at the time of testing and heard English spoken at least 90% of the time. Twenty-four infants were tested at each of three fundamental frequencies; 12 infants in each condition had /i/ as the background token and /I/ as the change token and 12 infants had /I/ as the background and /i/ as the change token.

Stimuli. Synthetic tokens of two English vowels /i/ (as in "heed") and /I/ (as in "hid") were produced using SenSyn (Sensimetrics Corporation) for the Macintosh on a Macintosh IIci computer. Values for the first 3 formants (F1, F2 & F3) were taken from the average values for women's voices as measured by Peterson and Barney (1952). Values for F4 were determined using Syrdal's (1985) formula. Bandwidths were taken from Dunn (1961).

Both vowel tokens were synthesized at three different fundamental frequencies. In the falling condition the F0 began at 350 Hz and fell linearly to 250 Hz over the duration of the vowel (500 ms). In the high steady condition, the F0 remained at 350 Hz for the duration of the vowel and in the low steady condition, the F0 remained at 250 Hz for the duration of the vowel.

Tokens were equalized for loudness by 3 adult listeners and presented at an average of 62 dB (A) in a quiet environment.

Apparatus. Infants were tested in a sound-attenuating chamber (Industrial Acoustics Co.). A Macintosh IIci computer with an audiomedia card presented the stimuli. A Strawberry Tree I/O card connected to a custom interface box was used to operate a button box, lights, and mechanical toys located in the soundbooth. Stimuli were presented via a Dennon amplifier (PMA 480) to a single GSI loudspeaker located 90° to the infant's left. A box with a smoked Plexiglass front was located underneath the speaker. During reinforcement (see below) lights and a toy in one of the four compartments in the box became visible; otherwise the contents were hidden.

Procedure. Infants were tested individually in a go/no-go conditioned head turn procedure. The infant sat on his/her parent's lap across from an experimenter who sat behind a small table. Both the parent and the experimenter listened to masking music through headphones for the duration of the experiment. When the infant was attentive and facing forward (towards the experimenter), the experimenter pressed at button to indicate that a trial was to begin. During both training and testing phases the background vowel was presented continuously repeating every 2 s.

During the training phase, the experimenter familiarized the infant to the contingency between a head turn towards the speaker when a change in the vowel occurred and the animated toy reinforcer. Only change trials occurred during training and the vowel was presented 5 dB(A) louder than the background. Thus, both the vowel token and loudness were cues that indicated that a change had occurred. Infants were required to make four consecutive correct responses within 20 trials in order to proceed to the testing phase. A head turn was considered correct if it occurred within 2 s of the onset of the change and if the infant turned at least 45° to the left to face the speaker.

During the testing phase, a total of 24 trials were presented (12 change and 12 no change control trials, in random order). Only correct head turns were reinforced; the experimenter, who was unaware of what the infant was hearing, recorded by a button press whenever a head turn occurred.

Results

For each infant an A' score was computed. (A' is a nonparametric equivalent of d' and may be better suited for infant studies in which only a small number of trials is used (Grier, 1971).) A' scores were submitted to an Analysis of Variance with 2 between-groups factors: F0 (falling, high steady, low steady) and Direction of vowel change (/I/ to /i/, and /i/ to /I/).

A significant main effect for F0 was obtained, $F(2, 12) = 3.76$, $p < .03$. Protected t-tests were performed in order to determine which conditions differed. Infants in the low steady and falling F0 conditions obtained higher A' scores on average than did infants in the high steady conditions, $p < .02$, $p < .03$, respectively. These results are summarized in Figure 1.

The main effect of Direction of change was not significant. The interaction between F0 and Direction of Change was also not significant.

T-tests were also computed in order to determine whether infants' performance in each of the three F0 conditions differed significantly from chance (A' of .50) in their A' scores.

In the falling fundamental condition, infants who had /I/ as the background token performed significantly above chance, $t(11) = 2.949$, $p < .007$. However, infants who had /i/ as the background token performed at chance, $t(11) = 1.072$, $p > .15$. A' scores for these two groups are plotted in Figure 2.

In the low steady F0 condition, infants obtained A' scores significantly above chance in both background vowel conditions: /I/ background, $t(11) = 3.197$, $p < .005$, and /i/ background, $t(11) = 2.874$, $p < .008$. A' scores for the low steady groups are plotted in Figure 3.

In the high steady F0 condition, performance did not exceed chance for either of the background vowel conditions: /I/

background, both *ps* > .15. The A' scores for the high steady groups are plotted in Figure 4.

Discussion

Taken together the results of this experiment suggest that discrimination of /i/ and /I/ vowel tokens for 6-month-old infants is possible when the F0 is low but steady (250 Hz), and when the F0 is falling (from 350 Hz to 250 Hz)—but only when /I/ is the background token. However, with a high steady F0 (350 Hz), infants perform at chance suggesting that they are unable to discriminate the two vowel tokens.

These findings are surprising given that the /i/-/I/ distinction is one that infants as young as 8 weeks of age can discriminate (e.g. Swoboda et al, 1976; 1978). However, these studies employed only steady F0 tokens in the range of a typical male speaker (i.e., approximately 120 Hz). The infants in the lowest F0 condition in our experiment —250 Hz—were also able to easily discriminate the two tokens. The 250 Hz fundamental is within the range of a woman's F0 in normal adult conversation.

The falling fundamental and high steady F0 conditions correspond more closely to the pitch of a female speaker's voice when directed to an infant. Maternal speech to infants can range in pitch up to at least 600 Hz at times (Fernald & Simon, 1984; Papousek, et al, 1987), yet our results suggest that a high steady F0 makes it difficult for infants to discriminate two vowel tokens.

The wide pitch excursions found in infant-directed speech correspond most closely to the falling F0 condition in this experiment. Infants were able to discriminate the falling fundamental tokens of /I/ and /i/, but when /I/ was the background token. Perhaps, then, the wide pitch excursions of infant-directed speech compensate for the difficulty of discrimination of vowels posed by the use of high fundamental frequencies.

In sum, these findings highlight the necessity for examining more closely the effects of F0 on vowel discrimination in infants.

The results suggest that high F0 has a negative effect on 6-month-old infants' ability to discriminate two vowel tokens, /I/ and /i/. Further research is necessary in order to determine whether features such as jitter (frequency variation from cycle to cycle of the wave form) normally occurring in natural tokens, would make discrimination of high steady tokens less difficult.

References

Fernald, A. & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. Developmental Psychology, 20, 104-113.

Grier, J.B. (1971). Nonparametric Indexes for sensitivity and bias: Computing formulas. Psychological Bulletin, 75, 424-429.

Kuhl, P.K. & Miller, J.D. (1982). Discrimination of auditory target dimensions in the presence or absence of variation in a second dimension by infants. Perception & Psychophysics, 31, 279-292.

Pasousek, M., Papousek, H. & Haekel, M. (1987). Didactic adjustments in fathers' and mothers' speech to their 3-month-old infants. Journal of Psycholinguistic Research, 16, 491-516.

Polka, L & Werker, J.F. (1994). Developmental changes in perception of nonnative vowel contrasts. Journal of Experimental Psychology: Human Perception and Performance, 20, 421-435.

Swoboda, P.J., Kass, J., Morse, P.A., & Leavitt, L.A. (1978). Memory factors in vowel discrimination of normal and at-risk infants. Child Development, 49, 332-339.

Swoboda, P., Morse, P. & Leavitt, L. (1976). Continuous vowel discrimination in normal and at-risk infants. Child Development, 47, 459-465.

Werker, J.F., Pegg, J.E. & McLeod, P.J. (1994). A cross-language investigation of infant preference for infant-directed communication. Infant Behavior & Development, 17, 323-333

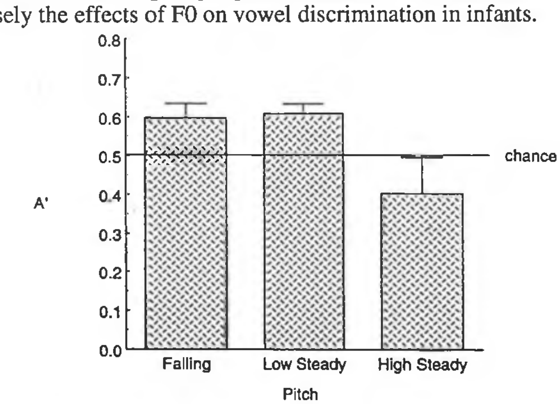


Figure 1. A' scores by pitch condition

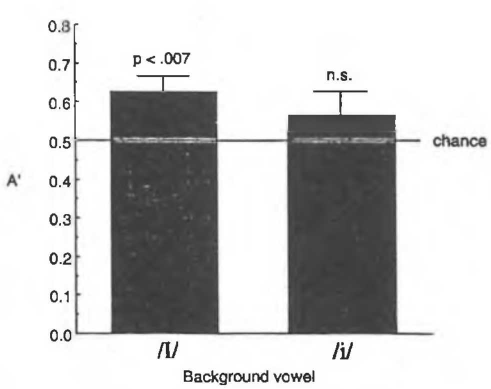


Figure 2. A' scores in the falling pitch condition

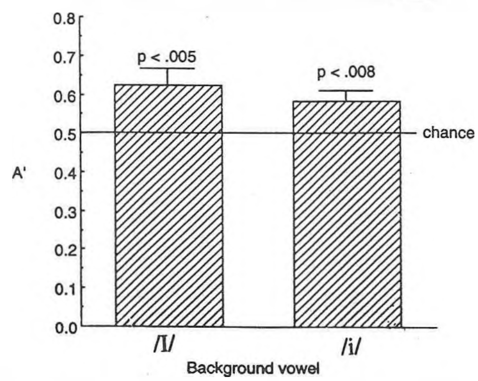


Figure 3. A' scores in the low steady pitch condition.

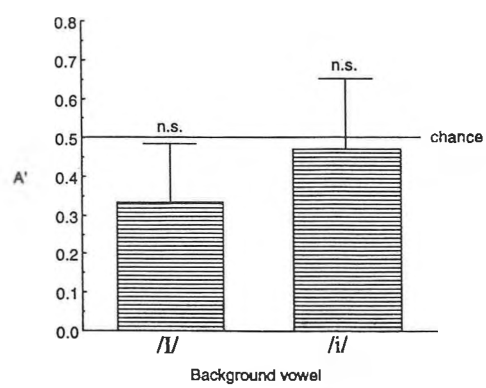


Figure 4. A' scores in the high steady pitch condition

INTELLIGIBILITY IN CLASSROOM NOISE FOR YOUNG SCHOOL AGED CHILDREN*

Garry Kranjc and Donald G. Jamieson
Hearing Health Care Research Unit, School of Communication Sciences and Disorders
The University of Western Ontario, London, Ontario, Canada N6G 1H1

INTRODUCTION

Favourable classroom acoustics contribute to children's level of academic performance [1, 2]. Inappropriate classroom noise levels reduce speech intelligibility and compromise psycho-educational and psychosocial development [2, 3]. The low frequency background noise found in a many classrooms can mask some speech sounds through upward spread of masking [4]. Other components of Classroom noise share certain spectral and temporal characteristics with target (e.g., teacher) speech. The extent to which noise masks speech depends on the long term spectrum of the noise, fluctuations in noise intensity over time, and the intensity of noise relative to the intensity of speech [2].

Such background noise reduces the intelligibility of speech by masking or distorting acoustic cues in the speech signal. Background noise in the classroom thus increases the attentional demands on students. This load is potentially greatest for the youngest children, which is a particular problem as this group creates the most background noise [5].

Background. Studies of speech perception in noise have focussed adults, but some studies have examined the performance of children, including hearing-impaired children [2, 3]. Previous studies have used adult multi-talker babble as background noise, not noise that is representative of the background noise a child is exposed to in the classroom (i.e., voices of other children, furniture noise, ventilation and other equipment noise). The present study used noise recorded from an occupied classroom to obtain a more representative sample.

Studies of speech perception in noise for young children often require the child to respond verbally or in writing. Oral response are open to errors, as young children frequently display poor articulation [6]. Written responses may tax the young child=s abilities, delaying testing and reducing prospects for the child to complete the study. Computer-based testing using pictures reduces opportunities for such errors and increases the quality of the data.

The present study sought to: (a) measure the speech perception abilities of young children using real classroom noise; (b) examine how identification accuracy varied for children of different ages, in the various listening conditions; and (c) confirm the age-appropriateness of the word list and task devised for this study.

METHOD AND PROCEDURE

Subjects were 40 students, consisting of ten subjects in each of Kindergarten (aged 5), grade 1 (aged 6), grade 2 (aged 7) and grade 3 (aged 8). Inclusion criteria were: 1) normal appearance of the ear canal and tympanic membrane; 2) pure tone air conduction thresholds no worse than 15 dB HL at 1, 2 and 4 kHz in either ear; 3) acoustic immittance measures for compliance between 0.3 and 1.6 ml; 4) middle ear pressure between 50 and -150 daPa bilaterally; 5) native speakers of English.

Experimental Speech Stimuli. Speech stimuli were 60 words (24 monosyllables, 12 spondees, 12 trochees, 12 trisyllables identified as being within the vocabulary of young children and able to be represented pictorially. Each word was spoken by one adult female speaker. The words were sampled as .WAV files at 22.5 kHz using the Time Frequency Response (TFR) software [7] and processed to equalize the RMS value of the vowels for syllables in the monosyllabic, spondees and trisyllabic words. The RMS of the entire syllable was adjusted to achieve 1.75 mV RMS for a 100 ms window centered on the peak of the vowel. The secondary syllable in the trochees was edited to have an RMS value equalling approximately 0.88 mV for a 100 ms sample of the vowel (i.e., half the RMS of the primary syllable), in order to distinguish trochees from spondees. The RMS values of individual final consonants were adjusted as required to ensure the item sounded natural. Each word was mixed with the sample of classroom noise at a signal-to-noise ratio (SNR) of 0 dB, -6 dB, and -12 dB, using a mixer program [7]. This mixer computed the RMS level of the signal, then computed the RMS level of the noise. The RMS value was then equalized to the RMS value of the original signal. Noise was also appended to this composite file, providing a 200 ms buffer of classroom noise at the beginning and end of the composite signal.

All aspects of the experiment were controlled by a PC computer, running the ECoS/Win experiment control software [8]. All test signals were replayed over a 16-bit Sound Blaster SB16 sound card and presented bilaterally over Telephonics TDH 49 supra aural headphones. Words were played at 65 dB SPL over the headphones to simulate the vocal intensity normally used by school teachers in the classroom, speaking at a distance of 1 meter. Subjects were tested individually in a portable classroom (ambient background noise at the test location from 54.9 to 59.3 dBC), while seated in front of a computer monitor that displayed the response alternatives. On each trial, the listener pointed to one of the 12 pictures on the monitor that best represented the word presented over the headphones.

Kindergarten and Grade 1 subjects participated in two sessions of 15 minutes each, with the quiet and -6 dB SNR conditions in the first session, followed by the 0 dB SNR and the -12 dB SNR in the second condition, within the same day. Grade 2 and 3 children participated for one 30-minute session, with the quiet, 0 dB SNR condition, -6 dB SNR condition, and -12 dB SNR condition presented. Word type (monosyllables-front to mid vowels; monosyllables - mid to back vowels; spondees; trochees; and trisyllables) was blocked within SNR conditions for all children.

RESULTS AND DISCUSSION

Older children tended to perform better than younger children in all listening conditions. Trisyllables were more understandable than bisyllables (spondees and trochees), which were generally more understandable than monosyllables. A pronounced effect was found with MF words, as SNR decreased to -6 dB: Kindergarten and Grade 1 children performed much worse than the older children (Figure 1). Kindergarten subjects also had lower scores than older children for spondees in the - 6 dB SNR condition.

Although direct comparisons between this study and previous studies is not possible due to stimuli and procedural differences, some generalizations are possible. First, in the quiet conditions, our children performed at a level comparable to that reported in other studies. In the 0 dB SNR and -6 dB SNR conditions, children in this study performed better than those in previous studies. We conclude that the word lists and testing protocols devised for this study were at an age-appropriate level.

Monosyllables with initial consonants having low intensity were particularly susceptible to masking, leading to frequent confusions with similar sounding words (e.g., confusions of *hair* with *chair* and *bear)*. Classroom noise appeared to interact with target speech to create the perception of new phonemes, not present in the target stimuli. The lower intensity of the second syllable in the trochees was readily masked by background noise, causing overall scores to be lower for two-syllable trochees than for two-syllable spondees. Such errors for key individual contact words may well cause younger children to lose the context of the teacher=s message so that they may not be able to follow along effectively in classroom activities.

GENERAL DISCUSSION

In general, our children performed at least as well as children tested under similar conditions in previous studies. We conclude that the word lists used here were age-appropriate and within the vocabulary of our children, and that the task was both understood and within of the ability of our subjects. MANOVAs showed that Kindergarten and Grade 1 children are especially affected by noise,

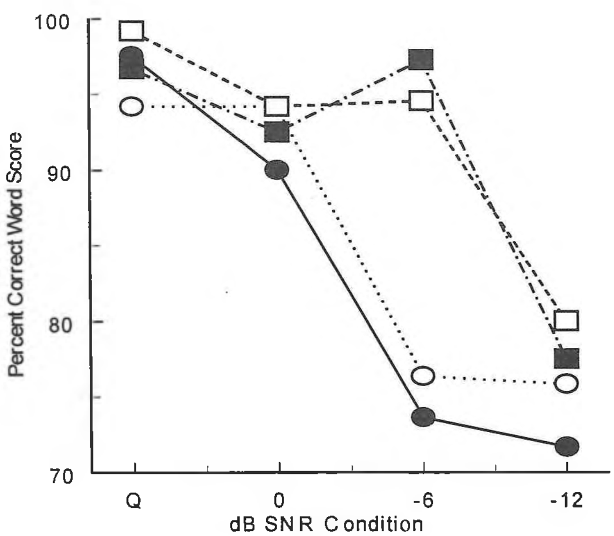


Figure 1. Mean percent correct word identification for monosyllable: front-mid vowel stimuli at each SNR condition for each group of listeners. (Symbols as for Figure 2.)

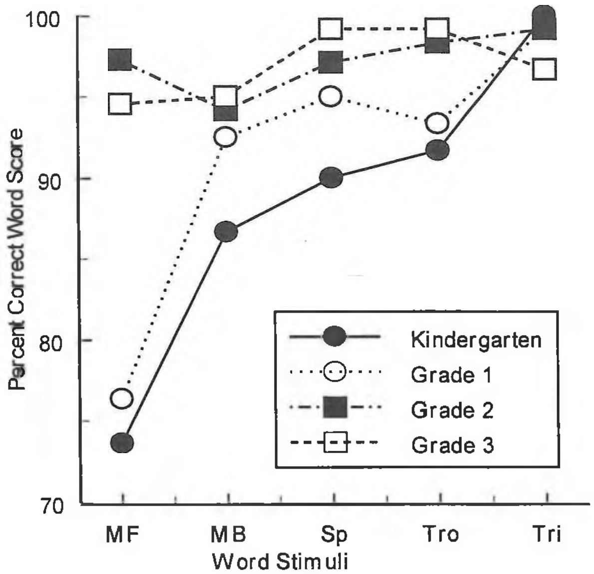


Figure 2. Mean percent correct word identification in the -6 dB SNR condition for each type of word stimulus for the four groups of listeners.

particularly for monosyllable: front-mid vowel words. Post hoc analyses indicated that in such conditions the performance of Kindergarten and Grade 1 subjects in this study decreased significantly when the SNR reached -6 dB, while the performance of Grade 2 and 3 subjects significant declined only when the SNR reached -12 dB. Kindergarten subjects also performed more poorly for trisyllables in the -12 dB SNR condition compared with children in other grades. In the -6 dB SNR condition, Kindergarten subjects also had significantly lower mean scores than Grade 3 subjects for spondees. These are not atypical listening conditions: studies of classroom acoustics have reported classroom noise levels consistent with or exceeding -6 dB SNR. The present study shows that in such circumstances, younger children may have particular difficulty understanding speech, especially monosyllable words containing low intensity initial consonants.

Currently, Canada lacks federal or provincial standards for classroom noise levels. The American Speech-Language and Hearing Association has recommended that classrooms should be constructed such that the SNR at the child=s ear is greater than +15 dB SNR, and the noise level in an unoccupied classroom does not exceed 35 dB(A) or NCC = 20 dB [9]. Studies of ambient classroom noise show that most classrooms do not meet these guidelines [10].

There is mounting evidence that Atypical@ classroom noise levels have significant, negative effects on the performance of young children. The present study adds further weight to this body of evidence, particularly for the youngest, school-aged children. Perhaps the evidence for degraded speech perception in presence of excessive classroom noise will be useful in persuading legislators and educators of the need for improved acoustic conditions for classrooms. This would include actions to a) reduce noise levels in unoccupied classrooms through reductions in ventilation and other noise sources; b) construction specifications to better control sound transmission and reverberation; and the availability of sound-field FM amplification systems to improve signal to noise ratio throughout the classroom environment, for all students.

REFERENCES

[1] Nabelek, A.K. & Nabelek, I.V. (1994). Room acoustics and speech perception. In J. Katz (Eds.), *Handbook of Clinical Audiology, Fourth Edition.* Baltimore, MD: Williams & Wilkins.

[2] Finitzo-Hieber, T. & Tillman, T.W. (1978). Room acoustics effects on monosyllabic word discrimination ability for normal and hearing-impaired children. *Journal of Speech and Hearing Research, 21*, 440-458

[3] Crandell, C.C. (1993). Speech recognition in noise by children with minimal degrees of sensorineural hearing loss. *Ear & Hearing, 14*, 210-216.

[4] Crandell, C.C. & Smaldino, J.J. (1995). Speech perception in the classroom. In C.C. Crandell, J.J. Smaldino, & C. Flexor (Eds.), *Sound-field FM Amplification: Theory & Practical Applications.* San Diego, CA: Singular

[5] Picard, M. & Bradley, J.S. (1997). Revisiting speech interference by noise in classrooms and remedial solutions. In Review.

[6] Ross, M. & Lerman, J. (1970). A picture identification test for hearing-impaired children. *Journal of Speech and Hearing Research, 13*, 44-53.

[7] Avaaz Innovations Inc. (1997). *TFR: The Avaaz Signal Analysis Package: User=s Guide.* London, ON: Avaaz Innovations Inc.

[8] Avaaz Innovations Inc. (1997). *Avaaz ECos/Win: Experiment Generator & Controller: User=s Guide.* London, ON: Avaaz Innovations Inc.

[9] American Speech-Language-Hearing Association (1995). Position statement and guidelines for acoustics in educational settings. *Asha, 37 (Suppl. 14)*, 15-19.

[10] Edwards, C. (1997). Today=s lesson: Noise in the classroom. *Vibes. The Canadian Hearing Soc. Quarterly Magazine, July*, 19-20.

* Work supported by NSERC and ORTC.

Please direct communications to jamieson @audio.hhcru.uwo.ca.

NEWS / INFORMATIONS

CONFERENCES

The following list of conferences was mainly provided by the Acoustical Society of America. If you have any news to share with us, send them by mail or fax to the News Editor (see address on the inside cover), or via electronic mail to desharnais@drea.dnd.ca

1998

14-18 September: 35th International Conference on Ultrasonics and Acoustic Emission, Chateau of Treste, Czech Republic. Contact: H. Kotschova, Geophysical Institute, AS Bocni II/401, 14131 Prague 4, Czech Republic; Fax: +42 2 761 549; Email: hko@ig.cas.cs; Web: www.ig.cas.cz

17-18 September: Annual Meeting of INCE/Japan, Tokyo, Japan. Fax: +81 423 27 3847.

17-18 September: 5th Mexican Congress on Acoustics, Queretaro, Qro., Mexico. Contact: S. Beristain, P.O. Box 75805, Lindavista 07300 Mexico, D.F. Mexico; Fax: +52 5 523 4742; Email: sberista@maxwell.esimez.ipn.mex

21-25 September: 3rd International CARIS conference - CARIS and the year of the Ocean, De Ruwenberg, The Netherlands. Contact: Universal Systems Ltd., c/o CARIS and the Year of the Ocean, 264 Rookwood Avenue, Fredericton, New Brunswick, Canada, E3B 2M2; Tel.: 506-458-8533; FAX: 506-459-3849.

4-7 October: Euro-noise 98, Munich, Germany. Contact: CSM Congress @ Seminar Management, Industriestr. 35, 82194 Gobenzell, Germany; Fax: +49 8142 5 47 35; Email: csm_congress@compuserve.com

12-16 October: 136th meeting of the Acoustical Society of America, Norfolk, VA. Contact: ASA, 500 Sunnyside Blvd., Woodbury, NY 11797, Tel.: 516-576-2360; FAX: 516-576-2377; Email: asa@aip.org; WWW: http://asa.aip.org

12-16 October: International Conference on Signal Processing (ICSP 98), Beijing, China. Contact: Yuan Baozong, Inst. on Information Science, Northern Jiaotong Univ., Beijing 100044, China; Fax: +86 10 6828 3458; Email: yuanbz@sun.ihep.ac.cn; Web: cie-china.org

14-16 October: 16th Yugoslav Conference on Noise and Vibration, Nis. Yugoslavia. Email: noise@znrfak.znrfak.ni .ac.yu

31 October - 2 November: AES International Conference, Audio, Acoustics and Small Spaces, Copenhagen, Denmark. Contact: J. Voetmann, DELTA Acoustics & Vibration, Building 356, Akademivey, 2800 Lyngby, Denmark; Fax: +45 45 93 19 90; Email: jv@delta.dk

CONFÉRENCES

La liste de conférences ci-jointe a été offerte en majeure partie par l'Acoustical Society of America. Si vous avez des nouvelles à nous communiquer, envoyez-les par courrier ou fax (coordonnées incluses à l'envers de la page couverture), ou par courrier électronique à desharnais @drea.dnd.ca

1998

14-18 septembre: 35e Conférence internationale sur les ultrasons et les émissions acoustiques, Château de Treste, République Tchèque. Info: H. Kotschova, Geophysical Institute, AS Bocni II/401, 14131 Prague 4, Czech Republic; Fax: +42 2 761 549; Email: hko@ig.cas.cs; Web: www.ig.cas.cz

17-18 septembre: Rencontre annuelle de l'INCE/Japon, Tokyo, Japon. Fax: +81 423 27 3847.

17-18 septembre: 5e congrès mexicain d'acoustique, Queretaro, Qro., Mexique. Info: S. Beristain, P.O. Box 75805, Lindavista 07300 Mexico, D.F. Mexico; Fax: +52 5 523 4742; Email: sberista@maxwell.esimez.ipn.mex

21-25 septembre: 3e conférence internationale CARIS - CARIS et l'Année de l'océan, De Ruwenberg, Pays-Bas. Info: Universal Systems Ltd., a/s CARIS et l'Année de l'océan, 264 Rookwood Avenue, Frédéricton, Nouveau Brunswick, Canada, E3B 2M2; Tél.: 506-458-8533; Fax: 506-459-3849.

4-7 octobre: Euro-noise 98, Munich, Allemagne. Info: CSM Congress @ Seminar Management, Industriestr. 35, 82194 Gobenzell, Germany; Fax: +49 8142 5 47 35; Email: csm_congress@compuserve.com

12-16 octobre: 136e rencontre de l'Acoustical Society of America, Norfolk, VA. Info: ASA, 500 Sunnyside Blvd., Woodbury, NY 11797, Tél.: 516-576-2360; FAX: 516-576-2377; Email: asa@aip.org; WWW: http:// asa.aip.org

12-16 octobre: Conférence internationale sur le traitement de signal (ICSP 98), Beijing, Chine. Info: Yuan Baozong, Inst. on Information Science, Northern Jiaotong Univ., Beijing 100044, China; Fax: +86 10 6828 3458; Email: yuanbz@sun.ihep.ac.cn; Web: cie-china.org

14-16 octobre: 16e conférence yougoslave sur le bruit et les vibrations, Nis. Yougoslavie. Email: noise@znrfak.znrfak.ni .ac.yu

31 octobre - 2 novembre: Conférence internationale AES, audio, acoustique et espaces restraints, Copenhague, Danemark. Info: J. Voetmann, DELTA Acoustics & Vibration, Building 356, Akademivey, 2800 Lyngby, Denmark; Fax: +45 45 93 19 90; Email: jv@delta.dk

11-13 November: 1st Asia Pacific Conference on Acoustics and Vibration (APAV 98), Singapore. Contact: APAV 98 Secretariat, 1 Selegie Road #09-01, Paradiz Centre, Singapore 188308, Singapore; Fax: +65 334 7891; Email: apavcon@singnet.com.sg

12-15 November: Institute of Acoustics (UK) Autumn Conference: Speech and Hearing, Windemere, UK. Contact: Institute of Acoustics, Agriculture House, 5 Holywell Hill, St. Albans, Herts. AL1 1EU, UK; Fax: +44 1727 850 355; E-mail: acoustics@clus1.ulcc.ac.uk

16-18 November: Inter-Noise 98, Christchurch, New Zealand. Contact: New Zealand Acoustical Society, P.O. Box 1181, Auckland, New Zealand.

20 November: Recreational Noise, Queenstown, New Zealand. Contact: P. Dickenson, NZ Ministry Health, PO Box 5013, Wellington, New Zealand; Fax: +64 4 496 2340; Email: philip.dickenson@mohwn.synet.net.nz

22-26 November: Noise Effects 98, Sydney, Australia. Contact: GPO Box 128, Sydney NSW 2001; Tel: +61 2 9262 2277; Fax: +61 2 9262 3135; Email: noise98@tourhosts.com .au; WWW: www.acay.com.au/~dstuckey /noise-effects98

30 November - 4 December: 5th International Conference on Spoken Language Processing, Sydney, Australia. Contact: ICSLP Secretariat, Tour Hosts, GPO Box 128, Sydney, NSW 2001, Australia; Fax: +61 2 9262 3135; Email: tourhosts @tourhosts.com.au; WWW: http://cslab.anu.edu.au/icslp98

15-16 December: Sonar Signal Processing, Loughborough, UK. Contact: Institute of Acoustics, Agriculture House, 5 Holywell Hill, St. Albans, Herts AL1 1EU, UK; Fax: +44 1727 850 533; Email: acoustics@clus1.ulcc.ac.uk

1999

15-19 March: Joint Meeting of Acoustical Society of America/European Acoustics Association, Berlin, Germany. Contact: ASA, 500 Sunnyside Blvd., Woodbury, NY 11797; Tel: 516-576-2360; Fax: 516-576-2377; Email: asa@aip.org; WWW: asa.aip.org

10-14 May: 4th International Conference on Theoretical and Computational Acoustics, Trieste, Italy. Contact: A. Marchetto, ICTCA,99, Osservatorio Geofisico Sperimentale, P.O. Box 2011-Opicina, 34016 Trieste, Italy; Fax: +39 40 327040; Email: ictca99@ogs.trieste.it

17-20 May: Society of Automotive Engineers (SAE) and Noise and Vibration Conference & Exposition meeting, Traverse City, MI. Contact: M.J. Asensio, SAE/Troy, 3001 W Big Beaver Rd, Troy, MI, USA. Tel: 248-649-4920, ext. 3106.

11-13 novembre: 1e conférence Asie-Pacifique sur l'acoustique et les vibrations (APAV 98), Singapour. Info: APAV 98 Secretariat, 1 Selegie Road #09-01, Paradiz Centre, Singapore 188308, Singapore; Fax: +65 334 7891; Email: apavcon@singnet.com.sg

12-15 novembre: Conférence d'automne de l'institut d'acoustique (Royaume-Uni): Parole et audition, Windemere, Royaume-Uni. Info: Institute of Acoustics, Agriculture House, 5 Holywell Hill, St. Albans, Herts. AL1 1EU, UK; Fax: +44 1727 850 355; E-mail: acoustics@clus1.ulcc.ac.uk

16-18 novembre: Inter-Noise 98, Christchurch, Nouvelle-Zélande. Info: New Zealand Acoustical Society, P.O. Box 1181, Auckland, New Zealand.

20 novembre: Bruit récréatif, Queenstown, Nouvelle-Zélande. Info: P. Dickenson, NZ Ministry Health, PO Box 5013, Wellington, New Zealand; Fax: +64 4 496 2340; Email: philip.dickenson@mohwn.synet.net.nz

22-26 novembre: Effets du bruit 98, Sydney, Australie. Info: GPO Box 128, Sydney NSW 2001; Tél: +61 2 9262 2277; Fax: +61 2 9262 3135; Email: noise98@tourhosts.com .au; Web: www.acay.com.au/~dstuckey /noise-effects98

30 novembre- 4 décembre: 5e conférence internationale sur le traitement de la langue parlée, Sydney, Australie. Info: ICSLP Secretariat, Tour Hosts, GPO Box 128, Sydney, NSW 2001, Australia; Fax: +61 2 9262 3135; Email: tourhosts @tourhosts.com.au; WWW: http://cslab.anu.edu.au/icslp98

15-16 décembre: Traitement de signal sonar, Loughborough, Royaume-Uni. Info: Institute of Acoustics, Agriculture House, 5 Holywell Hill, St. Albans, Herts AL1 1EU, UK; Fax: +44 1727 850 533; Email: acoustics@clus1.ulcc.ac.uk

1999

15-19 mars: Rencontre conjointe de l'Acoustical Society of America et de l'Association d'acoustique européenne, Berlin, Allemagne. Info: ASA, 500 Sunnyside Blvd., Woodbury, NY 11797; Tél: 516-576-2360; Fax: 516-576-2377; Email: asa@aip.org; WWW: asa.aip.org

10-14 mai: 4e conférence internationale sur l'acoustique théorique et informatisée, Trieste, Italie. Info: A. Marchetto, ICTCA,99, Osservatorio Geofisico Sperimentale, P.O. Box 2011-Opicina, 34016 Trieste, Italy; Fax: +39 40 327040; Email: ictca99@ogs.trieste.it

17-20 mai: Conférence et exposition de la Société des Ingénieurs d'autos (SAE) et conférence Bruit et Vibrations, Traverse City, MI. Info: M.J. Asensio, SAE/Troy, 3001 W Big Beaver Rd, Troy, MI, USA. Tél: 248-649-4920, poste 3106.

27-30 June: ASME Mechanics and Materials Conference, Blacksburg, VA. Contact: Mrs. Norma Guynn, Dept. of Engineering Science and Mechanics, Virginia Tech, Blacksburg, VA 24061-0219; Fax: 540-231-4574; Email: nguynn@vt.edu; WWW: http://www.esm.vt.edu/mmconf/

28-30 June: 1st International Congress of the East European Acoustical Association, St. Petersburg, Russia. Contact: EEAA, Moskovskoe Shosse 44, St. Petersburg 196158, Russia; Fax: +7 812 127 9323; Email: krylspb @sovam.com

28 June - 1 July: Joint Conference of Ultrasonics International '99 and World Congress on Ultrasonics '99 (UI99/WCU99), Lyngby, Denmark. Contact: L. Bjorno, Department of Industrial Physics, Technical University, Building 425, 2800 Lyngby, Denmark; Fax: +45 45 93 01 90; E-mail: lb@ipt.dtu.dk; WWW: www.msc.cornell.edu/~ui99/

4-9 July: 10th British Academic Conference in Otolaryngology, London, UK. Contact: BOA-HNS, The Royal College of Surgeons, 35-43 Lincoln's Inn Field, London WC2A 3PN, UK; Fax: +44 171 404 4200.

1-4 September: 15th International Symposium on Nonlinear Acoustics (ISNA-15), Gottingen, Germany. Contact: W. Lauterborn, Drittes Physikalisches Institut, Universitat Gottingen, Burgerstr. 42-44, 37073 Gottingen, Germany; Fax: +49 551 39 7720; Email: lb@physik3.gwdg.de

3-5 October: WESPRAC VII, Kumamoto, Japan. Contact: Computer Science Dept., Kumamoto Univ., 2-39-1 Kurokami, Kumamoto, Japan 860-0862; Fax: +81 96 342 3630; Email: wesprac7@cogni.eecs.kumamoto-u.ac.jp

CAREER OPPORTUNITIES

JOB SOUGHT: Acoustician from Russia, Ph. D., seeks a research opportunity abroad. 25 years research experience in musical acoustics and sound quality problems including about one year research in Sweden and two years of NATO Science Fellowship in Canada. Please contact Alexander Galembo: galembo@pavlov.psyc.queensu.ca, fax 1-613-5452499 or tel. 1-613-545600 ext 5754.

OPPORTUNITÉS DE CARRIÈRE

RECHERCHE TRAVAIL: Acousticien de Russie, Ph. D., cherche travail de recherche à l'étranger. 25 ans d'expérience de recherche en acoustique musicale et problèmes de qualité sonore, incluant une année de recherche en Suède, et deux ans au Canada avec bourse de recherche en sciences de l'OTAN. Contactez Alexander Galembo: galembo@pavlov.psyc.queensu.ca, fax 1-613-5452499, tél. 1-613-545600 poste 5754.

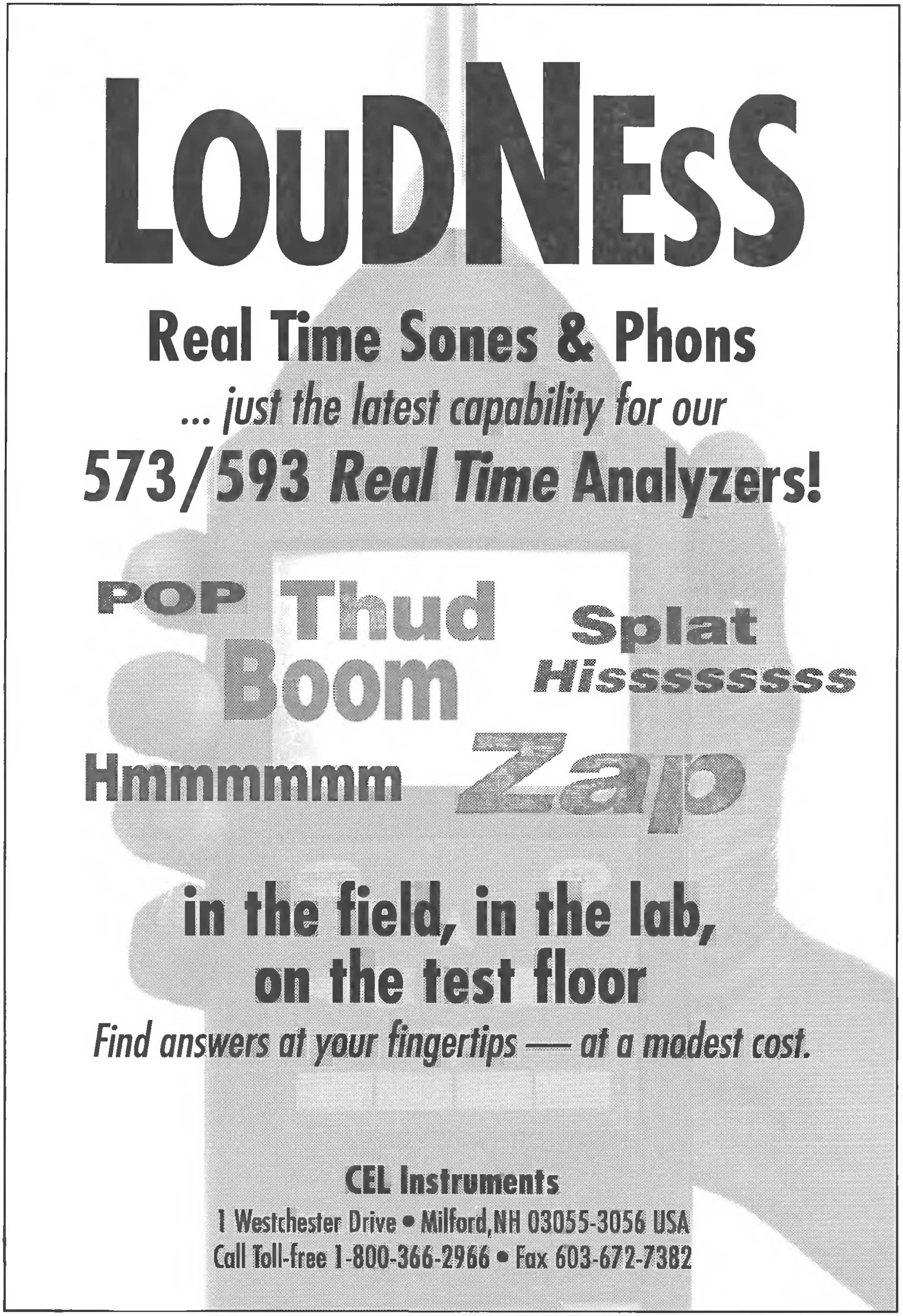
27-30 juin: Conférence ASME sur la mécanique et les matériaux, Blacksburg, VA. Info: Mrs. Norma Guynn, Dept. of Engineering Science and Mechanics, Virginia Tech, Blacksburg, VA 24061-0219; Fax: 540-231-4574; Email: nguynn@vt.edu; WWW: http://www.esm.vt.edu/mmconf/

28-30 juin: 1er Congrès international de l'Association d'acoustique de l'Europe de l'Est, St. Petersburg, Russie. Info: EEAA, Moskovskoe Shosse 44, St. Petersburg 196158, Russia; Fax: +7 812 127 9323; Email: krylspb @sovam.com

28 juin - 1 juillet: Conférence conjointe de "Ultrason International '99" et "Congrès mondial '99 sur les ultrasons" (UI99/WCU99), Lyngby, Danemark. Info: L. Bjorno, Department of Industrial Physics, Technical University, Building 425, 2800 Lyngby, Denmark; Fax: +45 45 93 01 90; E-mail: lb@ipt.dtu.dk; WWW: www.msc.cornell.edu/~ui99/

4-9 juillet: 10e Conférence académique britannique sur l'otolaryngologie, Londres, Royaume-Uni. Info: BOA-HNS, - 4 septembre: 15e Symposium international sur l'acoustique non-linéaire (ISNA-15), Gottingen, Allemagne. Info: W. Lauterborn, Drittes Physikalisches Institut, Universitat Gottingen, Burgerstr. 42-44, 37073 Gottingen, Germany; Fax: +49 551 39 7720; Email: lb@physik3 .gwdg.de

1-4 septembre: 15e Symposium international sur l'acoustique non-linéaire (ISNA-15), Gottingen, Allemagne. Info: W. Lauterborn, Drittes Physikalisches Institut, Universitat Gottingen, Burgerstr. 42-44, 37073 Gottingen, Germany; Fax: +49 551 39 7720; Email: lb@physik3 .gwdg.de

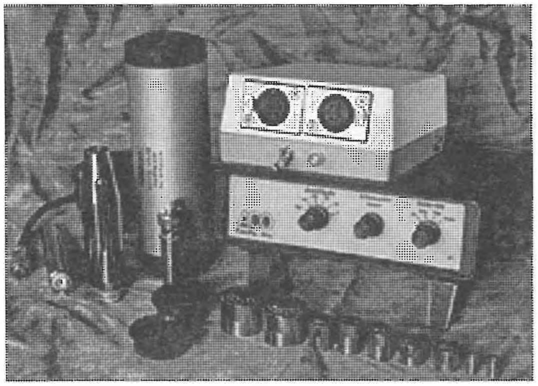

LOUDNESS
Real Time Sones & Phons
... just the latest capability for our
573/593 Real Time Analyzers!
POP
Thud
Splat
Boom
Hisssssss
Hmmmmmm
Zap
in the field, in the lab,
on the test floor
Find answers at your fingertips — at a modest cost.
CEL Instruments
1 Westchester Drive • Milford,NH 03055-3056 USA
Call Toll-free 1-800-366-2966 • Fax 603-672-7382

PRECISION MICROPHONES and MICROPHONE SYSTEMS

ACO Pacific, Inc.

2604 Read Ave.
Belmont, CA 94002 U.S.A.
Tel:650-595-8588 FAX:650-591-2891
e-mail: acopac@acopacific.com

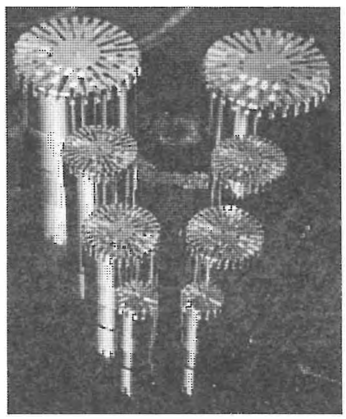
Established in 1978 ACO Pacific, Inc. is chartered to serve the needs of both End-users and Equipment Manufacturers. Headquartered in the San Francisco Bay Area, ACO Pacific, Inc. has manufacturing facilities and suppliers both in California and internationally.



Our OEM relationships include Hewlett Packard, Audio Precision, SPS, DRA Labs and many other large and small instrumentation and process control manufacturers. Our End-user customers include: manufacturers of loudspeakers; computers and peripherals; heavy equipment; automobiles and automotive parts - tires, brakes, engines; universities; aerospace; and an alphabet soup of government agencies US and foreign.

ACO Pacific's Product Family

Measurement Microphones
Microphone Preamplifiers
Microphone Power Supplies
SPL Calibrators
Simple Intensity™ Sound Intensity Systems
Very Random™ Noise Generators



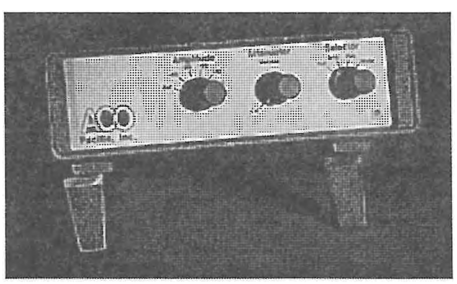
The "Alternative"

Type 1 Measurement Microphones
1, 1/2 and 1/4 Inch Models
Stainless Steel
and
Titanium Diaphragms
Quartz Insulators
Frequency Responses to 120 kHz
Noise Floors below 10 dBA
Meets or Exceeds IEC and ANSI Standards

Model 3024

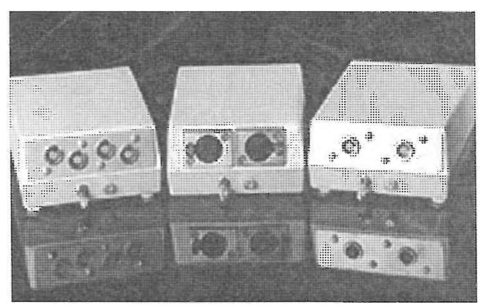
Very Random™Noise Generator

Pink and White Noise, 1kHz Sine Outputs
1.6 Hz to 39 kHz (-3dB)
Portable - Battery and AC Power



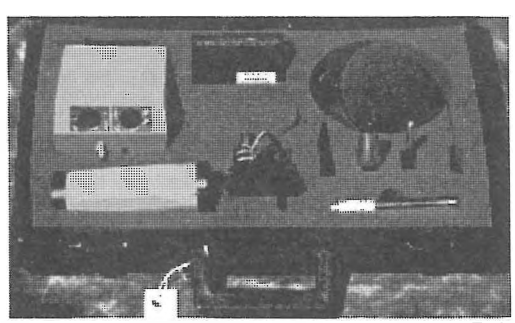
ACOustical Interface™

Precision Microphone Power Supplies
2 and 4 Channels - Gain Available
XLR and Lemo™ Connectors



PS9200KIT

Includes: PS9200 (XLR) Power Supply
AC Adaptor
4012 Preamplifier w/CA4012-5 Cable
Selection of 1/2 Inch Type 1 Microphone
WS1 - 3 inch Windscreen
SC1 Die Cut Storage Case (SC2 optional)
Options: 511E SPL Calibrator (shown)
"G" Gain Stage



ACOustics Begins With ACO™

ANNUAL STUDENT PRESENTATION AWARDS

The Canadian Acoustical Association makes awards to students whose papers are presented at the CAA Annual Symposium. Students contemplating presenting papers at the Symposium should apply for these awards with the submission of their abstract.

RULES

1. These awards are presented annually to authors of outstanding student papers that are presented during the technical sessions at Acoustics Week in Canada.
2. In total, three awards of $500.00 are presented.
3. Presentations are judged on the following merits:
   i) The way the subject is presented;
   ii) The explanation of the relevance of the subject;
   iii) The explanation of the methodology/theory;
   iv) The presentation and analysis of results;
   v) The consistency of the conclusions with theory and results.
4. Each presentation is judged independently by at least three judges.
5. The applicant must be:
   i) a full-time graduate student at the time of application;
   ii) the first author of the paper;
   iii) a member of the CAA;
   iv) registered at the meeting.
6. To apply for the award, the student must send this application simultaneously with the abstract. Multiple authors are permitted, but only the first author may receive an award.

PRIX ANNUELS RELATIFS AUX COMMUNICATIONS ETUDIANTES

L'Association Canadienne d'Acoustique décerne des prix aux étudiant(e)s qui présenteront une communication au congrès annuel de l'ACA. Les étudiant(e)s qui considèrent présenter un papier doivent s'inscrire à ce concours au moment où ils (elles) soummettent leur résumé.

REGLEMENTS

1. Ces prix sont décernés annuellement aux auteurs de communications exceptionelles presentées par des étudiants lors des sessions techniques de la Semaine Canadienne d'Acoustique.
2. Au total, trois prix de 500$ sont remis.
3. Les présentations sont jugées selon les critères suivants:
   i) La façon dont le sujet est présenté;
   ii) Les explications relatives à l'importance du sujet;
   iii) L'explication de la méthodologie;
   iv) La présentation et l'analyse des résultats;
   v) La consistence des conclusions avec la théorie et les résultats.
4. Chaque présentation est evaluée séparément par au moins trois juges.
5. Le candidat doit être:
   i) un étudiant à temps plein de niveau gradué au moment de l'inscription;
   ii) le premier auteur du papier;
   iii) un membre de l'ACA;
   iv) un participant au congrès.
6. Afin de s'inscrire au concours, l'étudiant doit envoyer ce formulaire d'inscription en même temps que son résumé. Plusieurs auteurs sont permis, mais seul le premier auteur peut recevoir le prix.

APPLICATION FOR STUDENT PRESENTATION AWARD AT ACOUSTICS WEEK IN CANADA

FORMULAIRE D'INSCRIPTION POUR LES PRIX DECERNES AUX ETUDIANTS LORS DE LA SEMAINE CANADIENNE D'ACOUSTIQUE

NAME OF THE STUDENT:____________________NOM DE L'ETUDIANT

SOCIAL INSURANCE NUMBER:____________________NUMERO D'ASSURANCE SOCIALE

TITLE OF PAPER:____________________TITRE DU PAPIER

UNIVERSITY/COLLEGE:____________________UNIVERSITE/COLLEGE

NAME, TITLE OF SUPERVISOR:____________________NOM ET TITRE DU SUPERVISEUR

STATEMENT BY THE SUPERVISOR: The undersigned affirms that the above-named student is a full-time student and the paper to be presented is the student's original work.

Signature:____________________

DECLARATION DU SUPERVISEUR: Le sous-signé affirme que l'étudiant(e) mentionné(e) ci-haut est inscrit(e) à temps plein et que la communication qu'il (elle) présentera est le fruit de son propre travail.

Date:____________________

APPLICATION FOR STUDENT TRAVEL SUBSIDY TO ACOUSTICS WEEK IN CANADA

Travel subsidies are available to students presenting papers at Acoustics Week in Canada if they live at least 150 km from the conference venue, if the subsidy is needed, if supporting receipts are submitted, and if they publish a summary of their paper in the proceedings issue of *Canadian Acoustics*.

I wish to apply for a CAA Travel Subsidy: _____yes _____no.

STATEMENT BY THE SUPERVISOR: The undersigned affirms that the CAA Travel Subsidy, combined with other travel funds that the above-named student may receive to attend the meeting will not exceed his/her travel costs.

Signature:____________________

FORMULAIRE DE DEMANDE DE REMBOURSEMENT POUR FRAIS DE DEPLACEMENT A LA SEMAINE CANADIENNE D'ACOUSTIQUE

Un remboursement de frais de déplacement est offert aux étudiants qui présentent une communication lors de la Semaine Canadienne d'Acoustique, s'ils demeurent à plus de 150 km du site du congrès, si le remboursement est nécessaire, si les reçus à l'appui sont soumis et s'ils publient un résumé dans les Actes du Congrès.

Je désire demander un remboursement: _______oui _______non.

DECLARATION DU SUPERVISEUR: Le sous-signé affirme que le remboursement, jumelé à d'autres fonds que l'étudiant(e) ci-haut mentionné(e) peut recevoir ne dépasseront pas ses coûts réels de voyage.

Date:____________________

INSTRUCTIONS TO AUTHORS FOR THE PREPARATION OF MANUSCRIPTS

Submissions: The original manuscript and two copies should be sent to the Editor-in-Chief.

General Presentation: Papers should be submitted in camera-ready format. Paper size 8.5" x 11". If you have access to a word processor, copy as closely as possible the format of the articles in Canadian Acoustics 18(4) 1990. All text in Times-Roman 10 pt font, with single (12 pt) spacing. Main body of text in two columns separated by 0.25". One line space between paragraphs.

Margins: Top - title page: 1.25"; other pages, 0.75"; bottom, 1" minimum; sides, 0.75".

Title: Bold, 14 pt with 14 pt spacing, upper case, centered.

Authors/addresses: Names and full mailing addresses, 10 pt with single (12 pt) spacing, upper and lower case, centered. Names in bold text.

Abstracts: English and French versions. Headings, 12 pt bold, upper case, centered. Indent text 0.5" on both sides.

Headings: Headings to be in 12 pt bold, Times-Roman font. Number at the left margin and indent text 0.5". Main headings, numbered as 1, 2, 3, ... to be in upper case. Sub-headings numbered as 1.1, 1.2, 1.3, ... in upper and lower case. Sub-sub-headings not numbered, in upper and lower case, underlined.

Equations: Minimize. Place in text if short. Numbered.

Figures/Tables: Keep small. Insert in text at top or bottom of page. Name as "Figure 1, 2, ..." Caption in 9 pt with single (12 pt) spacing. Leave 0.5" between text.

Photographs: Submit original glossy, black and white photograph.

References: Cite in text and list at end in any consistent format, 9 pt with single (12 pt) spacing.

Page numbers: In light pencil at the bottom of each page.

Reprints: Can be ordered at time of acceptance of paper.

DIRECTIVES A L'INTENTION DES AUTEURS PREPARATION DES MANUSCRITS

Soumissions: Le manuscrit original ainsi que deux copies doivent être soumis au rédacteur-en-chef.

Présentation générale: Le manuscript doit comprendre le collage. Dimensions des pages, 8.5" x 11". Si vous avez accès à un système de traitement de texte, dans la mesure du possible, suivre le format des articles dans l'Acoustique Canadienne 18(4) 1990. Tout le texte doit être en caractères Times-Roman, 10 pt et à simple (12 pt) interligne. Le texte principal doit être en deux colonnes séparées d'un espace de 0.25". Les paragraphes sont séparés d'un espace d'une ligne.

Marges: Dans le haut - page titre, 1.25"; autres pages, 0.75"; dans le bas, 1" minimum; latérales, 0.75".

Titre du manuscrit: 14 pt à 14 pt interligne, lettres majuscules, caractères gras. Centré.

Auteurs/adresses: Noms et adresses postales. Lettres majuscules et minuscules, 10 pt à simple (12 pt) interligne. Centré. Les noms doivent être en caractères gras.

Sommaire: En versions anglaise et française. Titre en 12 pt, lettres majuscules, caractères gras, centré. Paragraphe 0.5" en alinéa de la marge, des 2 cotés.

Titres des sections: Tous en caractères gras, 12 pt, Times-Roman. Premiers titres: numéroter 1, 2, 3, ..., en lettres majuscules; sous-titres: numéroter 1.1, 1.2, 1.3, ..., en lettres majuscules et minuscules; sous-sous-titres: ne pas numéroter, en lettres majuscules et minuscules et soulignés.

Equations: Les minimiser. Les insérer dans le texte si elles sont courtes. Les numéroter.

Figures/Tableaux: De petites tailles. Les insérer dans le texte dans le haut ou dans le bas de la page. Les nommer "Figure 1, 2, 3,..." Légende en 9 pt à simple (12 pt) interligne. Laisser un espace de 0.5" entre le texte.

Photographies: Soumettre la photographie originale sur papier glacé, noir et blanc.

Références: Les citer dans le texte et en faire la liste à la fin du document, en format uniforme, 9 pt à simple (12 pt) interligne.

Pagination: Au crayon pâle, au bas de chaque page.

Tirés-à-part: Ils peuvent être commandés au moment de l'acceptation du manuscrit.

The Canadian Acoustical Association | l'Association Canadienne d'Acoustique

SUBSCRIPTION INVOICE

Subscription for the current calendar year is due January 31. New subscriptions received before July 1 will be applied to the current year and include that year's back issues of *Canadian Acoustics*, if available. Subscriptions received from July 1 will be applied to the next year.

FACTURE D'ABONNEMENT

L'abonnement pour la présente année est dû le 31 janvier. Les nouveaux abonnements reçus avant le 1 juillet s'appliquent à l'année courante et incluent les anciens numéros (non-épuisés) de *l'Acoustique Canadienne* de cette année. Les abonnements reçus après le 1 juillet s'appliquent à l'année suivante.

Check ONE Item Only: / Cocher la case appropriée :

CAA Membership	___	$ 50	___	Membre individuel
CAA Student membership	___	$ 10	___	Membre étudiant(e)
Institutional Subscription	___	$ 50	___	Membre institutionnel
Sustaining Subscription	___	$ 150	___	Abonnement de soutien

Total Remitted $_______ Versement total

INFORMATION FOR MEMBERSHIP DIRECTORY / RENSEIGNEMENT POUR L'ANNUAIRE DES MEMBRES

Check areas of interest (max 3): / Cocher vos champs d'intérêt (max. 3):

Architectural Acoustics	1. _____	1.	Acoustique architecturale
Engineering Acoustics / Noise Control	2. _____	2.	Génie acoustique / Contrôle du bruit
Physical Acoustics / Ultrasound	3. _____	3.	Acoustique physique / Ultrasons
Musical Acoustics / Electroacoustics	4. _____	4.	Acoustique musicale / Electroacoustique
Psychological / Physiological Acoustics	5. _____	5.	Physio/psycho-acoustique
Shock and Vibration	6. _____	6.	Chocs et vibrations
Hearing Sciences	7. _____	7.	Audition
Speech Sciences	8. _____	8.	Parole
Underwater Acoustics	9. _____	9.	Acoustique sous-marine
Signal Processing / Numerical Methods	10. _____	10.	Traitement des signaux / Méthodes numériques
Other	11. _____	11.	Autre

Business telephone number (_____) ______________ Numéro de téléphone au bureau
Business facsimile number (_____) ______________ Numéro de télécopieur au bureau
Business E-Mail number ______________________________ Numéro de courier électronique au bureau

PLEASE TYPE NAME AND ADDRESS BELOW:

VEUILLEZ ECRIRE VOTRE NOM ET VOTRE ADRESSE CI-DESSOUS:

Faites parvenir ce formulaire à l'adresse suivante en prenant soin d'y joindre un chèque fait au nom de L'ASSOCIATION CANADIENNE D'ACOUSTIQUE:

Make cheques payable to THE CANADIAN ACOUSTICAL ASSOCIATION. Mail this form with payment to:

Trevor Nightingale
Secretary, Canadian Acoustical Association
P. O. Box 74068
Ottawa, Ontario K1M 2H9

The Canadian Acoustical Association
l'Association Canadienne d'Acoustique

PRESIDENT **PRÉSIDENT**	**John Bradley** IRC, NRCC Ottawa, Ontario K1A 0R6	(613) 993-9747 john.bradley@nrc.ca
PAST PRESIDENT **PRÉSIDENT SORTANT**	**John Hemingway** 2410 Old Pheasant Road Mississauga, Ontario L5A 2S1	(416) 798-0522 jrh@mail.globalserve.net
SECRETARY **SECRÉTAIRE**	**Trevor Nightingale** P. O. Box 74068 Ottawa, Ontario K1M 2H9	(613) 993-0102 trevor.nightingale@nrc.ca
TREASURER **TRÉSORIER**	**John Hemingway** 2410 Old Pheasant Road Mississauga, Ontario L5A 2S1	(416) 798-0522 jrh@mail.globalserve.net
MEMBERSHIP **RECRUTEMENT**	**Don Jamieson** Hearing Health Care Res. Unit Elborn College University of Western Ontario London, Ontario N6G 1H1	(519) 661-3901 jamieson@audio.hhcru.uwo.ca
EDITOR-IN-CHIEF **RÉDACTEUR EN CHEF**	**Ramani Ramakrishnan** Aiolos Engineering 51 Constellation Court Suite 200 Toronto, Ontariol M9W 1K4	(416) 674-3017 ramani@aiolos.com
DIRECTORS **DIRECTEURS**	**Li Cheng** **Annabel Cohen** **David DeGagne** **Stan Dosso** **Dalila Giusti** **Jean Nicolas** **Winston Sydenborgh** **David Quirt**	(418) 656-7920 (902) 628-4331 (403) 297-3200 (604) 363-2877 (905) 660-2444 (819) 821-7157 (416) 844-7113 (613) 993-9746
WORLD WIDE WEB HOME PAGE:	**http://www.uwo.ca/hhcru/caa/**	

SUSTAINING SUBSCRIBERS / ABONNES DE SOUTIEN

The Canadian Acoustical Association gratefully acknowledges the financial assistance of the Sustaining Subscribers listed below. Annual donations (of $150.00 or more) enable the journal to be distributed to all at a reasonable cost. Sustaining Subscribers receive the journal free of charge. Please address donation (made payable to the Canadian Acoustical Association) to the Secretary of the Association.

L'Association Canadienne d'Acoustique tient à témoigner sa reconnaissance à l'égard de ses Abonnés de Soutien en publiant ci-dessous leur nom et leur adresse. En amortissant les coûts de publication et de distribution, les dons annuels (de $150.00 et plus) rendent le journal accessible à tous nos membres. Les Abonnés de Soutien reçoivent le journal gratuitement. Pour devenir un Abonné de Soutien, faites parvenir vos dons (chèque ou mandat-poste fait au nom de l'Association Canadienne d'Acoustique) au secrétaire de l'Association.

ACO Pacific
2604 Read Ave.
Belmont, CA, USA 94002
(650) 595-8588 FAX: (650) 591-2891

Acoustec Inc.
Attn: Dr. J.G. Migneron
1381 rue Galilée, Suite 103
Québec, Québec G1P 4G4
(418) 682-2331 FAX: (418) 682-1472

Aercoustics Engineering Limited
Barman & Associates
50 Ronson Drive, Suite 127
Rexdale, Ontario M9W 1B3
(416) 249-3361 FAX: (416) 249-3613

Atlantic Acoustical Associates
P.O. Box 96, Station M
Halifax, Nova Scotia B3J 2L4
(902) 425-3096

H. L. Blachford Ltd.
Annt: Mr. D.E. Watson
2323 Royal Windsor Dr.
Mississauga, Ontario L5J 1K5
(905) 823-3200 FAX: (905) 823-9290

Bruel & Kjaer Canada Ltd.
90 Leacock Road
Pointe Claire, Québec H9R 1H1
(514) 695-8225 FAX: (514) 695-4808

J. E. Coulter Associates Ltd.
Suite 507
1200 Sheppard Ave. E
Willowdale, Ontario M2K 2S5
(416) 502-8598 FAX: (416) 502-3473

Dalimar Instruments Inc.
193, Joseph Carrier
Vaudreuil-Dorion, Québec J7V 5V5
(514) 424-0033 FAX: (514) 424-0030

Eckel Industries of Canada Ltd.
Attn: Mr. Blake Noon
P.O. Box 776
Morrisburg, Ontario K0C 1X0
(613) 543-2967 FAX: (613) 543-4173

Environmental Acoustics Inc.
Attn: Mr. H.J. Doedens
#13 - 5155 Spectrum Way
Mississauga, Ontario L4W 5A1
(905) 238-1077 FAX: (905) 238-9079

Hatch Associates Ltd.
Attn: Tim Kelsall
2800 Speakman Dr.
Mississauga, Ontario L5K 2R7
(905) 855-7600 FAX: (905) 855-8270

HGC Engineering
Plaza One, Suite 203
2000 Argentia Road
Mississauga, Ontario L5N 1P7
(905) 826-4044 FAX: (905) 826-4940

Hydro-Quebec
Vice-presidence Environnement
75 Rene Levesque ouest, 16e etage
Montreal, Québec H2Z 1A4

Industrial metal Fabricators (Chatham) Ltd.
Industrial Noise Control
Attn: Mr. Frank Van Oirschot
P.O. Box 834, 288 Inshes Ave.
Chatham, Ontario N7M 5L1
(519) 354-4270 FAX: (519) 354-4193

Integral DX Engineering Ltd.
907 Admiral Ave.
Ottawa, Ontario K1Z 6L6
(613) 761-1565 FAX: (613) 729-4337

Jade Acoustics Inc.
545 North Rivermede Road, Suite 203
Concord, Ontario L4K 4H1
(905) 660-2444 FAX: (905) 660-4110

John Swallow Associates
Attn: John C. Swallow
250 Galaxy Boulevard
Etobicoke, Ontario M9W 5R8
(416) 798-0522

Larson Davis Laboratories
1681 West 820 North
Provo, Utah, USA 84601
(801) 375-0177

MJM Conseillers en Acoustique Inc.
Attn: M. Michel Morin
6555 Cote des Neiges, Suite 400
Montréal, Québec H3S 2A6
(514) 737-9811 FAX: (514) 737-9816

Nelson Industries Inc.
Corporate Research Dept.
P.O. Box 600
Stoughton, WI, USA 53589-0600
(608) 873-4370

OZA Inspections Ltd.
P.O. Box 271
Grimsby, Ontario L3M 4G5
(416) 945-5471 FAX: (416) 945-3942

Peutz & Associes
Attn: Marc Asselineau
103 boul. Magenta
F-75010 Paris, France
+33 1 42858485 FAX: +33 1 42821057

J. L. Richards & Assoc. Ltd.
Attn: Fernando Ribas
864 Lady Ellen Place
Ottawa, Ontario K1Z 5M2
(613) 728-3571 FAX: (613) 728-6012

Scantek Inc.
916 Gist Ave.
Silver Spring, MD, USA 20910
(301) 495-7738 FAX: (301) 495-7739

SNC/Lavalin Environment Inc.
2 Felix Martin Place
Montréal, Québec H2Z 1Z3
(514) 393-1000

Spaarg Engineering Limited
Noise and Vibration Analysis
822 Lounsborough St.
Windsor, Ontario N9G 1G3
(519) 972-0677 FAX: (519) 972-0677

State of the Art Acoustik Inc.
Attn. Dr. C. Fortier
Unit 43, 1010 Polytek St.
Ottawa, Ontario, K1J 9J3

Tacet Engineering Ltd.
Attn: Dr. M.P. Sacks
111 Ava Road
Toronto, Ontario M6C 1W2
(416) 782-0298 FAX: (416) 785-9880

University of Alberta
MEANU, Dept. of Mech. Eng.
6720 - 30 St.
Edmonton, Alberta T6P 1J6
(403) 466-6465 FAX: (403) 466-6465

Valcoustics Canada Ltd.
30 Wertheim Court, Unit 25
Richmond Hill, Ontario L4B 1B9
(905) 764-5223 FAX: (905) 764-6813

Wilrep Ltd.
1515 Matheson Blvd. E, Unit C 10
Mississauga, Ontario L4W 2P5
(905) 625-8944 FAX: (905) 625-7142