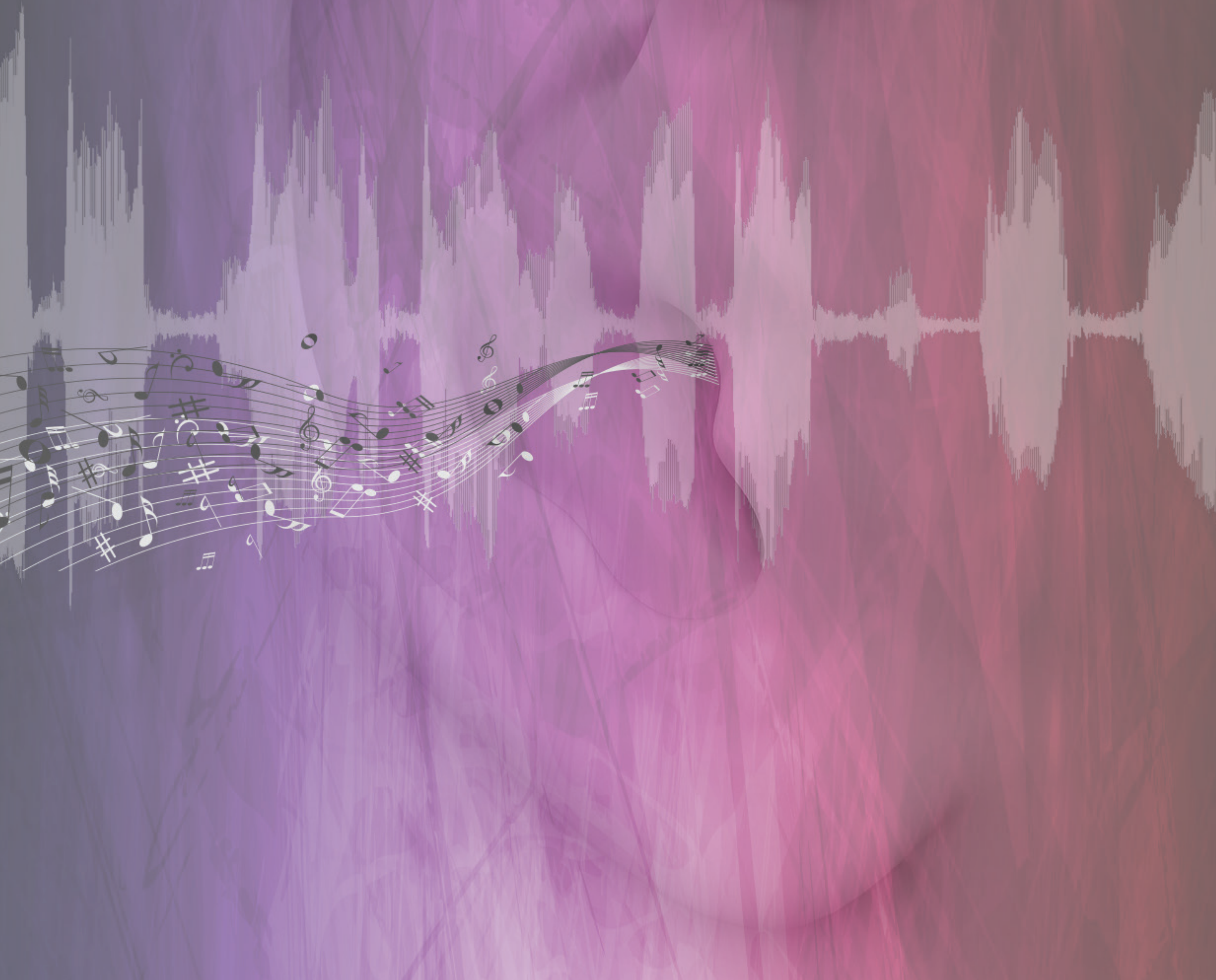
Acoustics Australia

Including
Part 2: Special Issue on Auditory Perception



Australian Acoustical Society   Volume 42 Number 3 December 2014

Acoustics Australia

ACOUSTICS AUSTRALIA EDITORIAL ENQUIRIES
The Editor, Acoustics Australia
AcousticsAustralia@acoustics.asn.au
www.acoustics.asn.au

SUBMISSIONS
All Articles and Technical Notes
www.springer.com/engineering/journal/40857

News, New Products and Books for Review
acousticsaustralia@acoustics.asn.au

SUBSCRIPTION ENQUIRIES
www.springer.com/engineering/journal/40857

Australian Acoustical Society Enquiries see Page 236

Acoustics Australia is published by the Australian Acoustical Society
(A.B.N. 28 000 712 658)
ISSN 0814-6039

Printed by
Cliff Lewis Printing
91-93 Parraweena Rd,
CARINGBAH NSW 2229
Tel (02) 9525 6588
Fax (02) 9524 8712
email: matt@clp.com.au
ISSN 0814-6039

Vol. 42, No. 3 December 2014

PAPERS

Including Special Issue Part 2: AUDITORY PERCEPTION

TECHNICAL NOTES

REGULAR ITEMS

Cover design: Helena Brusic

FROM THE CHIEF EDITOR

This issue of the journal marks the end of an era for the production of Acoustics Australia. It is the last issue that has manual processing of the papers and the publication via the resources of the Australian Acoustical Society. From 2015 we will be using a paper management system and will have Springer as the publisher. There will still be plenty of work for the editorial team but no longer will we be relying on the use of our emails and Dropbox storage.

Under the new system, once a paper has been processed, gone through the review stage and is accepted, a pre-publication version will be immediately available online. All the records will be archived and the articles will be fully indexed. Despite these changes, the benefits of the AAS membership of free access to the journal and the papers within it will continue.

So at this point, I must take the opportunity to thank the past editorial teams, all the authors and reviewers upon whose hard work over the years we have built upon. Their diligence has led to improvements in the journal and subsequent increase in quality submissions. Without their attention to the task we would not have the journal that you see today.

It is fitting that this is a composite issue. Some papers follow on from the August special issue on Auditory Perception and the others are general submissions.

The three articles in Part 2 of the special issue on auditory perception review contemporary research on: intensity dynamics and loudness change (Olsen); methods for, and issues in, measuring the quality of transmitted speech in telecommunication contexts (Köster, Möller, Antons, Arndt, Guse & Weiss); and electrophysiological investigation of language experience-dependent effects in pitch processing of tonal language (Krishnan & Gandour). Our thanks to the reviewers of these papers: Robert Eklund, Neville Fletcher, Nina Kraus, Frederic Marmel, Peter Počta, and Robert Teghtsoonian.

The first issue in 2015 will be a special issue on Room Acoustics with Sir Harold Marshall as the guest editor. As well as papers on this topic we welcome contributed papers and technical notes on all aspects of acoustics.

Marion Burgess
Chief Editor
and Kate Stevens
Guest Editor

FROM THE PRESIDENT

Season's Greetings everyone,

It's been another busy year - hard to believe that this is the December issue of Acoustics Australia. InterNoise 2014 was a fantastic way to draw the year to a close. Congratulations to the Victorian Division, especially Norm Broner in making InterNoise the spectacular success that it was. There was certainly something for everyone - excellent plenary speakers, a HUGE array of papers across the days, enormous exhibition hall, coffee to keep us all going and a fabulous App to keep us organised. The enormous effort put in certainly showed in what appeared to be a seamless event that was enjoyed by the many I have spoken to since.

This is my first message in Acoustics Australia as President. I am looking forward to working with you over the next 2 years to keep up the great work of those I follow. On that note I would like to thank Norm again for his efforts as President over the last two years. I'd also like to take this opportunity to thank all the Federal Councillors for their work over 2014. In particular I would like to make note of the work put in by Terry McMinn and our General Secretary Richard Booker in progressing the website. We are looking forward to our new look website early in 2015.

It is with much pleasure that I now advise you of the elevation of Norm Broner (VIC) and John Macpherson (WA) to Fellows of the Society. As the conference dinner this year was an InterNoise event we were unable to make any AAS announcements. The announcements were made at the AGM. Congratulations to you both on a well-deserved recognition. The Victorian Division's Gerald Riley Award of $2500 for the best AAS member paper at Internoise 2014 was also presented to Sebastian Oberst at the AGM for his paper entitled 'An innovative signal processing technique for the extraction of ants' walking signals'. Congratulations Sebastian.

Lastly, this will be the final edition of Acoustics Australia in this format. From next year the journal will be published by the international publishing company Springer. This change brings with it many opportunities for the journal to reach a broader audience. The journal will maintain a similar appearance. Members will receive a copy of the journal by email or can download from the web page (http://www.springer.com/engineering/journal/40857). Another prompt to log on to the AAS website and make sure your records are up to date.

I wish everyone the very best for the Season. Enjoy the break and come back refreshed and ready in 2015.

Tracy Gowen
President AAS

INTENSITY DYNAMICS AND LOUDNESS CHANGE: A REVIEW OF METHODS AND PERCEPTUAL PROCESSES

Kirk N. Olsen
The MARCS Institute, University of Western Sydney, Australia
k.olsen@uws.edu.au

In real-world listening domains such as speech and music, acoustic intensity and perceived loudness are dynamic and continuously changing through time. The percept of loudness *change* in response to continuous increases (up-ramps) and decreases (down-ramps) of intensity has received ongoing empirical and theoretical interest, the result of which has led to conflicting findings from a range of key paradigms. Therefore, the aim of this brief review is to: (a) describe key paradigms used to measure changes in loudness in response to continuous intensity change; (b) identify methodological issues associated with each paradigm; and (c) discuss the mechanisms proposed to explain differences in loudness change when methodological constraints and response biases are controlled. It is concluded that direct and indirect measures of loudness change reflect two distinct aspects of auditory perception. Specifically, magnitude estimation and continuous loudness paradigms reflect changes in perception associated with a ramp's direction and magnitude of intensity change, and empirical evidence supports the conclusion that greater loudness change in response to down-ramps relative to up-ramps is the real-time perceptual outcome. On the other hand, retrospective global judgements of loudness change are disproportionally weighted on end-level intensity rather than magnitude of intensity change. However, an up-ramp-specific effect of duration on global loudness change is evident when end-level response bias is controlled, and this may be associated with end-point time-of-arrival responses to real and apparent looming auditory motion.

INTRODUCTION

The association between acoustic signals and their perception is a fundamental issue in auditory psychophysics. In psychological terms, the subjective percept of loudness is closely related to a sound's physical intensity and is broadly defined as the magnitude of auditory sensation [1, 2]. However, the relationship between acoustic intensity and loudness is not straightforward. Additional acoustic parameters such as frequency play a significant role in loudness perception, as evident by frequency-dependent equal-loudness contours [3, 4]. Mapping loudness across the frequency spectrum has been made possible by the use of psychoacoustic steady-state stimuli. These are stimuli with intensity profiles that do not vary through time and thus offer the experimenter a high degree of stimulus control. However, almost all real-world sounds are dynamic, with continuous changes in acoustic and perceptual parameters such as increases and decreases of intensity and loudness.

For the purpose of the present paper, a continuous rise of intensity is defined as an 'up-ramp'; a ramp of intensity (in experiments, most often linear) that continuously rises from relatively low intensity to relatively high intensity. Conversely, a continuous decrease of intensity is defined as a 'down-ramp'. Paradigms measuring *changes* in loudness both directly and indirectly have been used to investigate the mechanisms underlying the perception of time-varying, temporally dynamic intensity stimuli. This is an important line of research in auditory psychophysics, as intensity and loudness are dynamic aspects of real-world listening in domains such as speech and music. Furthermore, fundamental psychophysical research on dynamic intensity and loudness change has far-reaching implications for and applications to fields such as ecological psychoacoustics, music perception, composition, and performance, sound engineering, and the design of informative auditory warnings [5-9]. However, paradigms used to investigate changes of intensity and loudness have led to a range of conflicting results; results that can begin to be reconciled with a systematic analysis of methodological similarities, differences, benefits, and shortcomings.

Therefore, the overarching aim of the present paper is to organise and briefly review research investigating the dynamic percept of loudness change in response to continuous acoustic intensity change. Specifically, the paper will: (a) describe the key paradigms used to measure changes in loudness in response to continuous intensity change; (b) identify methodological benefits and shortcomings associated with each paradigm; and (c) discuss the mechanisms proposed to explain differences in loudness change when methodological constraints and response biases are controlled.

DOWN-RAMP DECRUITMENT AND LOUDNESS MAGNITUDE ESTIMATION

Early studies investigating changes in loudness as a function of continuous increases and decreases of intensity used traditional psychophysical measurements such as magnitude estimation [10]. Developed from the seminal work of S. S. Stevens [11], magnitude estimates of loudness in the current context require listeners to make discrete numerical estimations of loudness in response to the intensity at stimulus onset, offset, and sometimes intermittently throughout each dynamic intensity sweep. In its simplest form, loudness change is 'indirectly' calculated as the ratio between the two discrete onset and offset loudness estimates. For pure-tone up-ramp and down-ramp stimuli presented at durations from ~10 s to 180 s, the ratio of loudness change for a down-ramp is greater than a corresponding up-ramp stimulus matched on parameters such as frequency, duration, range, and region of intensity change [10, 12-15]. The greater magnitude of loudness change in response to down-ramp tonal stimuli has been termed 'decruitment'. As can be seen in Figure 1, decruitment is due to the observation that loudness falls more rapidly as the continuous 65-20dB sound pressure level (SPL) decrease of intensity falls below ~40dB SPL [10]. This, in turn, leads to the relatively low 20dB SPL end-level (offset) of the down-ramp to be perceived as 'softer' in loudness than the equivalent 20dB SPL onset of the 20-65dB SPL up-ramp. The reciprocal phenomenon of 'upcruitment' does not elicit such pronounced effects on up-ramp perception, resulting in a smaller ratio of loudness magnitude estimates between up-ramp onset and offset levels. Furthermore, when intermittent tones are presented with decreasing levels, decruitment is not observed [10]. These data show that the specific direction and continuity of intensity change over time are significant factors for down-ramp decruitment.

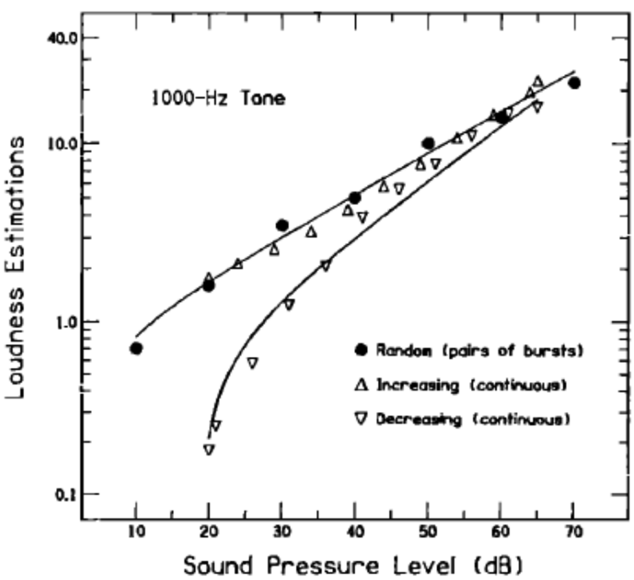


Figure 1. Geometric mean loudness magnitude estimates of a 180 s 1kHz pure tone plotted as a function of level for three modes of signal presentation: upward triangles represent responses to a continuously increasing (up-ramp) tone from 20-65dB SPL; downward triangles represent responses to a continuously decreasing (down-ramp) tone from 65-20dB SPL; and circles represent responses to pairs of intermittent tones presented at seven levels in random order [10]. The loudness curve rapidly steepens as the down-ramp continuously decreases below 40dB SPL, whereas the loudness curve in response to the up-ramp increases linearly on a log scale across the entire range of intensity change. (Reprinted with permission from [10]. Copyright 1990, Acoustical Society of America).

The candidate mechanism underlying down-ramp decruitment is still not completely understood. However, sensory adaptation has been proposed [10, 14, 16]. If the early and relatively high intensity portions of the down-ramp were to adapt neurons responsible for coding the sound, the latter portions of the down-ramp would become less audible. By contrast, the early and relatively low intensity portions of the up-ramp may not cause substantial adaptation for latter, higher-intensity portions. In this scenario, the offset of the down-ramp is likely to be perceived 'softer' than the equivalent intensity onset of the up-ramp. Indeed, this is the case in observations of decruitment. It is unlikely, however, that a sensory mechanism such as adaptation can completely explain down-ramp decruitment [12]. The relative influence of cognition has been investigated using a dual-task paradigm, with results suggesting that the magnitude of decruitment depends on whether listeners actively attend to the stimulus [14]. Computational models that can account for decruitment do not currently exist, and ongoing work aims to develop a model that explains behavioural data and the cognitive and sensory mechanisms underlying this phenomenon.

GLOBAL LOUDNESS CHANGE AND A 'PERCEPTUAL BIAS' FOR RISING INTENSITY

One potential shortcoming associated with measures of loudness change from discrete magnitude judgments is that perceived change is not measured directly. Rather, it is inferred from differences in static loudness responses that reflect *overall* loudness at specific points in time [17]. One alternative to measuring static judgements of loudness across stimulus presentation is to ask listeners to directly judge the magnitude of loudness change after stimulus presentation. A retrospective discrete post-stimulus judgement of perceived change in loudness is termed 'global loudness change'. Neuhoff [18] used this method to investigate loudness change in response to 1.8 s up-ramps and down-ramps presented as 1kHz pure-tone, white-noise, or vowel stimuli (/ə/ - sounds like the 'a' in 'about'). The range of each ramp was 15dB and participants made a global judgement of loudness change in response to single-ramp trials (a single-stimulus paradigm) using a computer-based visual analogue scale. From this post-stimulus response, pure-tone and vowel up-ramps were perceived to change significantly more in loudness than down-ramps; a finding opposite to those of decruitment studies. No significant differences were observed for white noise. These results have now been replicated and extended using 3.6 s vowel stimuli in a comparable single-stimulus paradigm [19], as well as a paired-stimulus paradigm where pairs of 1.8 s 30dB up-ramps and down-ramps were presented in each trial [19, 20]. In a paired-stimulus paradigm, participants are required to (a) indicate which item in a pair changed more in loudness; and (b) rate the magnitude of this difference when one was perceived.

Greater perceived loudness change for pure-tone and vowel up-ramps in this paradigm has been hypothesised as evidence of an evolved, adaptive perceptual bias to looming auditory motion in the environment [18, 20]. Continuous increases of acoustic intensity are a vital cue for looming (or approaching) auditory motion [21]. An overestimation of loudness change for up-ramp 'looming' stimuli may function as a survival response that provides a selective advantage for organisms able to underestimate the arrival of a looming object, effectively allowing extra time to 'err on the side of caution' when taking appropriate action (e.g., avoidance or retreat) [18, 20, 22]. Specific neural processes have been identified for auditory looming in the human brain [23-25], and evidence suggests that the hypothesized adaptive bias from judgements of global loudness change is influenced by sex differences [26]. Perceived time-to-contact and time-of-arrival studies in auditory and visual domains support the notion that real and apparent looming motion is perceived to arrive at a point in space significantly sooner than actual source arrival [27, 28]. The observation in [18] that white noise up-ramps and down-ramps are perceived similarly is explained by first assuming that white noise commonly represents multiple sound sources in the environment (e.g., ocean, rain, wind through trees) [18, 20]. According to Neuhoff, multiple sound sources should not necessarily demand equivalent behavioural priority when compared to simple (pure-tone) and complex (vowel) tonal stimuli, which are arguably more closely associated with a single sound source. However, the suggestion that looming and potentially threatening single sound sources in the environment are characterised by spectral properties akin to a pure-tone or vowel stimulus is yet to be completely substantiated.

Indeed, the 'perceptual bias for rising intensities' hypothesis and the use of global loudness change as a sensitive perceptual measure of real-time changes of intensity have been challenged [e.g., 15, 29, 30, 31]. For example, the continuous increase of intensity change that characterizes up-ramps and auditory looming is not absolutely necessary to elicit differences in global loudness change predicted by the 'perceptual bias' hypothesis [30, 32]. Furthermore, the global loudness change measure relies on retrospective post-stimulus judgements and is therefore susceptible to cognitive-based response biases that will now be addressed.

End-level bias and recency in memory

Empirical evidence shows that direct ratings of global loudness change are influenced by a 'bias for end level' [15, 19, 31]. For example, as up-ramp end-level increases in dB, so does the magnitude of global loudness change, even when the magnitude of intensity change in up-ramps is held constant [15]. Specifically, with every 15dB increase in end-level from an up-ramp with a fixed magnitude of intensity change, perceived loudness change approximately doubles. This is evidence that post-stimulus retrospective global judgements of loudness change are weighted towards the most recent portion of the up-ramp – the end-level – and not the entire magnitude of intensity change. Susini et al. [33, 34] explain an end-level bias with reference to memory-based recency. Simply, recency is defined as a memory recall bias for the last item presented in a sequence of stimuli [35]. In the context of global loudness change, the most recent portion of an up-ramp or down-ramp is its final intensity level. Recency may bias judgements of global loudness change toward the final level of intensity (end-level) of an up-ramp or down-ramp. Take the case where global loudness change in response to a 60-80dB SPL up-ramp is compared with an 80-60dB SPL down-ramp. If a cognitive-based recency mechanism were to impact this response, it is not surprising that perceived change is greater for up-ramps because, in the example above, the up-ramp ends on a level 20dB greater than that of the down-ramp. This can be described as an end-level recency mechanism in global judgements of loudness change.

To investigate the hypothesis of an end-level recency mechanism, Olsen et al. [19] balanced end-level differences in an analysis comparing 50-70dB SPL up-ramps with 90-70dB SPL down-ramps using Neuhoff's [18] /ə/ vowel stimulus at 1.8 s and 3.6 s ramp durations. In this comparison, up-ramps and down-ramps have an equivalent intensity of 70dB SPL at the end of the ramp (in other words, equivalent or 'balanced' end-levels). Participants were presented with a single-stimulus paradigm and were required to judge global loudness change retrospectively using a visual analogue scale. Results from this analysis do not provide evidence of down-ramp decruitment, and when end-level differences between 1.8 s up-ramps and down-ramps are removed, the original Neuhoffian [18] 'bias for rising intensities' is not recovered. This suggests that an end-level recency mechanism can explain the greater magnitude of global loudness change in response to 1.8 s up-ramps when end-level intensity differs between up-ramps and down-ramps [e.g., 18]. However, at the 3.6 s duration, global loudness change was significantly greater for up-ramps relative to down-ramps, even when end-level recency was controlled in the balanced end-level analysis. These data provide evidence of an up-ramp-specific effect of duration under balanced end-level conditions: global loudness change increases as a function of duration for up-ramps only, while down-ramp global loudness change is not affected by stimulus duration.

An earlier experiment using direct and unconstrained magnitude estimations of global loudness change has also investigated up-ramps and down-ramps with balanced end-levels [15]. When such a measure of global loudness change replaces the visual analogue scale used in [19], greater perceived changes in loudness in response to up-ramps were not observed for 1.8 s 1kHz pure-tone stimuli. In fact, loudness change was numerically greater in response to 1.8 s down-ramps, and this difference increased as the sweep size of each ramp doubled from 15dB to 30dB (no inferential statistics were conducted on these specific comparisons [15]).

Reasons for the somewhat varied results between experiments that controlled up-ramp and down-ramp end-level differences are not clear. Differences in scaling methods may be a factor: the experiment in [15] used an unconstrained magnitude estimation procedure to measure global loudness change, whereas the predefined visual analogue loudness scale used in [19] constrains listeners' responses. Future research directly comparing these two procedures with a design comprising multiple regions and ranges of intensity change (cf. [15]) will shed further light on how global judgements of loudness change relate to the sweep size and end-level of up-ramps and down-ramps.

Post-stimulus sensory persistence of excitation

Post-stimulus persistence of neural excitation is one candidate sensory mechanism proposed to explain the 'residual' differences in global loudness change when cognitive mechanisms such as recency are controlled under balanced end-level conditions [36, 37]. The rationale behind the persistence of excitation hypothesis is that the auditory system continues to respond to a sensory stimulus after it ceases to be presented. A longer post-stimulus *sensory* response of greater magnitude may result in a subjectively larger *perception* of change for that stimulus. This hypothesis was investigated using psychophysical forward masking [36], defined as "an elevation of hearing threshold for a target signal presented after another stimulus event: the masker. The difference between masked signal threshold and signal threshold in quiet is an indicator of masking magnitude, a measure of the auditory system's response to a sensory stimulus beyond its physical presence at a particular point in time" (p. 596). Greater masking magnitude from up-ramps relative to down-ramps under balanced end-level conditions would provide evidence of an underlying sensory mechanism most likely occurring at peripheral stages of auditory processing [but see, 38]. As displayed in Figure 2, results from a forward masking paradigm using 3.6 s vowel up-ramp and down-ramp maskers show that differences in mean masked thresholds between 40-60dB SPL up-ramps and 80-60dB SPL down-ramps were below 1.34dB SPL at masker-offset to signal-offset delays of 10 ms to 170 ms (the signal was a 10 ms 1.5kHz pure tone). These differences were not significant. After ~180-200 ms, masked thresholds returned to baseline thresholds in quiet. These results subsequently rule out post-stimulus persistence of excitation as an explanatory mechanism for differences in global loudness change between 3.6 s up-ramps and down-ramps with balanced end-levels.

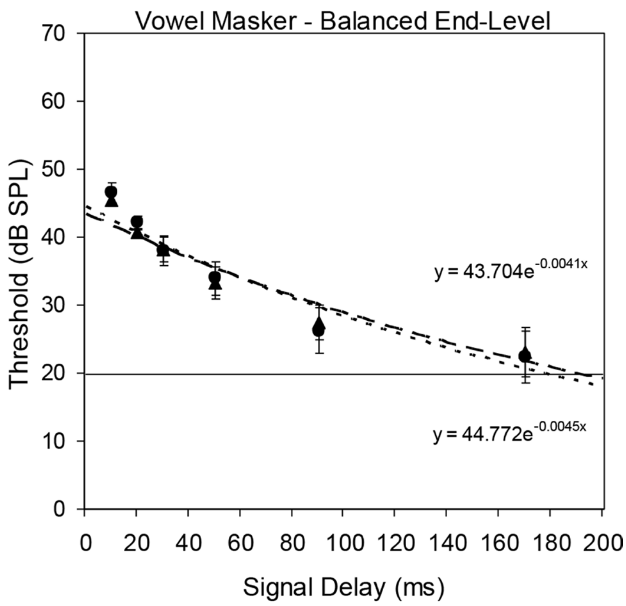


Figure 2. Mean forward-masking patterns ($N = 3$) from 3.6 s up-ramp (solid triangle) and down-ramp (solid circle) vowel maskers with masker-offset to signal-offset delays of 10, 20, 30, 50, 90, and 170 ms [36]. The signal was a 10 ms 1.5kHz pure tone. Balanced end-level comparisons (40–60dB SPL up-ramps versus 80–60dB SPL down-ramps) are displayed. Exponential curves are shown for up-ramps (dashed line, top equation) and down-ramps (dotted line, bottom equation), and error bars represent standard error of the mean. The solid horizontal line represent mean signal threshold in quiet. (Reprinted with permission from [36], Copyright 2012, Pion Ltd. www.pion.co.uk; www.envplan.com).

CONTINUOUS LOUDNESS MEASUREMENT: FURTHER EVIDENCE FOR DOWN-RAMP DECRUITMENT

An underlying similarity between direct loudness change measured from retrospective global judgements and indirect loudness change measured from static 'snapshots' of loudness magnitude is that neither are completely sensitive to changes in loudness on a continuous moment-to-moment basis. Therefore, the third key loudness change paradigm reviewed here is continuous loudness measurement. One of the benefits of a continuous measure of loudness is that end-level recency is inherently removed because a retrospective judgement of loudness change is not required. In auditory perception research, continuous responses have been used to investigate relationships between acoustic properties such as intensity, spectral flatness (a global parameter of timbre), the perception of affect (e.g., emotional arousal and valence/pleasantness), and loudness. Such studies have been undertaken in contexts ranging from traffic noise [39] to music from classical [40-42] and electroacoustic [43, 44] genres.

Only a handful of experiments, however, have systematically manipulated increases and decreases of acoustic intensity when measuring loudness continuously. Susini and colleagues [33, 34] developed an 'analogical/categorical' (A/C) scaling device for continuous loudness measurement. The response tool comprised a physical box with a slider that listeners used to continuously rate loudness on a scale containing seven categorical labels, from 'very, very loud' to 'very, very soft', with 'mid' serving as the mid-point of scale. Loudness change was calculated as the difference between loudness values recorded at the beginning and end of each ramp, analogous to magnitude estimates of loudness but in the context of continuous perceptual measurement. Using a paired-stimulus paradigm and 1kHz pure tone up-ramps (60-80dB SPL) and down-ramps (80-60dB SPL) presented at durations of 2, 5, 10, and 20 s per item, an omnibus ANOVA with N=15 did not result in a significant main effect of intensity ramp [34]. This suggests that both the 'bias for rising intensity' and decruitment effects disappear when 'indirect' loudness change is measured from a continuous response and calculated similarly to the magnitude estimates used in decruitment studies [10, 12-15].

However, in [34], sample size and statistical power was moderate at best, and no specific contrasts between up-ramps and down-ramps at each duration were analysed. From close inspection of the results of Experiment 1 in [34], loudness change was numerically greater for down-ramps relative to up-ramps across all durations. Furthermore, the mean loudness values at down-ramp offset (60dB SPL) were lower than the mean loudness values in response to the equivalent 60dB SPL up-ramp onset at stimulus durations of 5, 10, and 20 s. Taken together, these two trends in results provide some support for the main observations of decruitment: (1) that down-ramps are perceived to change more in loudness than up-ramps when calculated as the difference between ramp onset and offset loudness ratings; and (2) that greater loudness change in response to down-ramps is due to a 'softer' loudness response to down-ramp offset intensity, relative to the equivalent up-ramp onset intensity. As previously discussed, decruitment may be explained by a down-ramp adaptation mechanism, where early and relatively high intensity portions of a down-ramp adapt neurons responsible for coding the sound, resulting in latter portions of the down-ramp to become less audible. These results using continuous loudness as a tool to investigate perceptual change support this hypothesis, but here in a higher region of intensity change than those used in earlier decruitment studies [e.g., 10].

The significant role of down-ramp end-level loudness 'softening' in this decruitment-like effect has recently received further support in a musical context [45]. Using a similar continuous loudness paradigm to [34] but with a computer-based visual analogue loudness scale, 29 participants continuously rated loudness in response to a range of 6.4 s monophonic melodies constructed with up-ramp and down-ramp intensity profiles and ascending and descending melodic contours [45]. The range of each intensity sweep was either 15dB (70-85dB SPL) or 30dB (55-85dB SPL) and loudness change was calculated as the difference between loudness values at the beginning of the continuous response and the offset of each melody. Linearity of the continuous loudness responses was also investigated. Overall, musical down-ramps were perceived to change significantly more in loudness than musical up-ramps. As can been seen in Figure 3, this is explained by the significantly 'softer' loudness response to down-ramp offset intensity, relative to loudness in response to the equivalent up-ramp onset intensity. Furthermore, continuous loudness responses to down-ramps in the region between 55-85dB SPL were essentially linear. This is in contrast with the non-linearity in loudness change that characterises decruitment as down-ramps continuously decrease below 40dB SPL [10]. This difference in the loudness curve is likely due to the region of intensity change used in [45], which remained at an overall higher intensity region than decruitment studies that include levels as low as 20dB SPL.

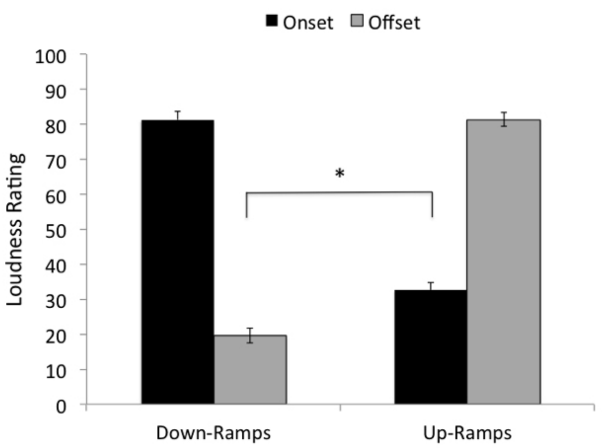


Figure 3. Mean loudness ratings at the onset of the continuous loudness response and the offset of each stimulus for 6.4 s down-ramp and up-ramp melodies presented in [45] (Experiment 1). On the y-axis, zero represents a 'soft' loudness response and 100 represents a 'loud' loudness response. No significant difference in loudness in response to the onset of the down-ramp and the offset of the up-ramp was observed. However, loudness was significantly lower in response to the offset of a down-ramp, relative to the onset of the up-ramp. Error bars report standard error of the mean; *$p < .001$. (Reprinted from [45], Copyright 2014, with permission from Elsevier).

CONCLUDING REMARKS

The present paper aimed to provide a brief review of research investigating the dynamic percept of loudness change in response to continuous acoustic intensity change. Description and evaluation of the three key paradigms used in this field of research was presented, and cognitive and sensory mechanisms that may underpin differences in perceived loudness change were discussed. It is clear from this review that conflicting results due to direct and indirect measures of loudness change reflect two distinct aspects of auditory perception.

Indirect loudness change derived from magnitude estimation and continuous loudness paradigms reflect, at the least, changes in perception associated with a ramp's direction and magnitude of intensity change. In these paradigms, loudness is measured throughout the entire dynamic stimulus; statically in the case of magnitude estimation, and continuously in the case of the continuous response. Therefore, they are the most sensitive tools for understanding real-time perception of intensity change and the mechanisms that underpin those perceptions as they unfold though time. Continued empirical evidence supports the conclusion that greater loudness change in response to down-ramps is the real-time perceptual outcome. Greater loudness change in response to down-ramps relative to up-ramps is characterised by: (1) a steeper linear loudness curve in response to down-ramps presented at intensity regions above 40dB SPL; and (2) a further rapid non-linear steepening of the loudness curve in response to down-ramps that continuously decrease below 40dB SPL. The extent to which these results are underpinned by peripheral and central mechanisms is a question that requires further behavioural evidence and computational modeling. For example, output from models of various stages of auditory processing could be used to compare with behavioural data to identify the location(s) of sensory adaptation [46-49].

Any method that purports to be a valid measure of loudness change must be supported by evidence that it is indeed sensitive to the magnitude of intensity change. It is clear from this review that global loudness change is disproportionally weighted on end-level intensity perception, rather than the magnitude of intensity change within each dynamic sweep. However, when differences between up-ramp and down-ramp end-levels are controlled in experimental design and analysis, an up-ramp-specific effect of duration remains: global loudness change in response to up-ramps increases as a function of duration, even when the range of physical intensity change remains constant and end-level recency is controlled. Neuhoff [17] argued that a direct global judgement of loudness change in response to an up-ramp looming stimulus in the environment is more useful for localizing a moving sound source than snapshot judgements of loudness at discrete points in time. The results of global loudness change reviewed here are most consistent with auditory and visual research that investigates anticipatory or predictive perceptions to real and apparent looming motion; perceptions that are usually measured after motion offset. In these studies, looming stimuli are perceived to arrive at a point in space sooner than what would be expected from the physical velocity of the approaching stimulus [27, 28].

The magnitude of underestimation for a looming stimulus' time-of-arrival increases as the velocity of a visual looming stimulus becomes slower [50, 51]. This is analogous to the up-ramp-specific effect of duration reported above, where longer durations of up-ramp 'looming' stimuli contain slower rates of intensity change over time, but elicit a greater magnitude of global loudness change than stimuli with shorter durations and faster rates of change. One may speculate, therefore, that global loudness change reflects – at least in part – the effects of duration and rate (but not magnitude) of intensity change for end-point time-of-arrival responses to real and apparent looming auditory motion.

ACKNOWLEDGEMENTS

Thanks to Robert Teghtsoonian, one anonymous reviewer, Kate Stevens, and the MARCS Institute Music Cognition and Action research group for helpful comments on an earlier draft.

REFERENCES

[1] Moore, B. C. J., *An introduction to the psychology of hearing.* 5th ed., Boston: Academic Press 2003

[2] Olsen, K. N., Loudness and intensity, in *Music in the Social and Behavioral Sciences: An Encyclopedia,* W.F. Thompson, Editor. 2014, Sage Publications: New York.654-657.

[3] Fletcher, H. and Munson, W. A., "Loudness, its definition, measurement and calculation", *Journal of the Acoustical Society of America* **5**, 82-108 (1933)

[4] Suzuki, Y. and Takeshima, H., "Equal-loudness-level contours for pure tones", *Journal of the Acoustical Society of America* **116**, 918-933 (2004)

[5] Olsen, K. N. and Stevens, C. J., "Psychophysiological response to acoustic intensity change in a musical chord.", *Journal of Psychophysiology* **27**, 16-26 (2013)

[6] Thompson, W. F., Peter, V., Olsen, K. N., and Stevens, C. J., "The effect of intensity on relative pitch", *Quarterly Journal of Experimental Psychology* **65**, 2054-2072 (2012)

[7] Dean, R. T., Olsen, K. N., and Bailes, F., "Is there a 'rise-fall temporal archetype' of intensity in the music of Joseph Haydn? The role of the performer", *Journal of Music Performance Research* **6**, 39-67 (2013)

[8] Neuhoff, J. G., *Ecological psychoacoustics*. Amsterdam: Elsevier Academic Press 2004

[9] Croghan, N. B., Arehart, K. H., and Kates, J. M., "Quality and loudness judgments for music subjected to compression limiting", *Journal of the Acoustical Society of America* **132**, 1177-1188 (2012)

[10] Canévet, G. and Scharf, B., "The loudness of sounds that increase and decrease continuously in level", *Journal of the Acoustical Society of America* **88**, 2136-2142 (1990)

[11] Stevens, S. S., "The direct estimation of sensory magnitudes: Loudness", *The American Journal of Psychology* **69**, 1-25 (1956)

[12] Teghtsoonian, R., Teghtsoonian, M., and Canévet, G., "The perception of waning signals: Decruitment in loudness and perceived size", *Perception & Psychophysics* **62**, 637-646 (2000)

[13] Canévet, G., Teghtsoonian, R., and Teghtsoonian, M., "A comparison of loudness change in signals that continuously rise or fall in amplitude", *Acta Acustica united with Acustica* **89**, 339-345 (2003)

[14] Schlauch, R. S., "A cognitive influence on the loudness of tones that change continuously in level", *Journal of the Acoustical Society of America* **92**, 758 (1992)

[15] Teghtsoonian, R., Teghtsoonian, M., and Canévet, G., "Sweep-induced acceleration in loudness change and the "bias for rising intensities"", *Perception and Psychophysics* **67**, 699-711 (2005)

[16] Scharf, B., Loudness adaptation, in *Hearing Research and Theory*, J.V. Tobias and E.D. Schubert, Editors. 1983, Academic Press: Orlando.

[17] Neuhoff, J. G., "'Perception of changes in loudness': Reply", *Nature* **398**, 673-674 (1999)

[18] Neuhoff, J. G., "Perceptual bias for rising tones", *Nature* **395**, 123-124 (1998)

[19] Olsen, K. N., Stevens, C. J., and Tardieu, J., "Loudness change in response to dynamic acoustic intensity", *Journal of Experimental Psychology: Human Perception and Performance* **36**, 1631-1644 (2010)

[20] Neuhoff, J. G., "An adaptive bias in the perception of looming auditory motion", *Ecological Psychology* **13**, 87-110 (2001)

[21] Jenison, R. L., "On acoustic information for motion", *Ecological Psychology* **9**, 131-151 (1997)

[22] Bach, D. R., Neuhoff, J. G., Perrig, W., and Seifritz, E., "Looming sounds as warning signals: The function of motion cues", *International Journal of Psychophysiology* **74**, 28-33 (2009)

[23] Bach, D. R., Schachinger, H., Neuhoff, J. G., Esposito, F., Di Salle, F., Lehmann, C., et al., "Rising sound intensity: An intrinsic warning cue activating the amygdala", *Cerebral Cortex* **18**, 145-150 (2008)

[24] Hall, D. A. and Moore, D. R., "Auditory neuroscience: The salience of looming sounds", *Current Biology* **13**, R91-R93 (2003)

[25] Seifritz, E., Neuhoff, J. G., Bilecen, D., Scheffler, K., Mustovic, H., Schachinger, H., et al., "Neural processing of auditory looming in the human brain", *Current Biology* **12**, 2147-2151 (2002)

[26] Neuhoff, J. G., Planisek, R., and Seifritz, E., "Adaptive sex differences in auditory motion perception: looming sounds are special", *Journal of Experimental Psychology: Human Perception and Performance* **35**, 225-234 (2009)

[27] Rosenblum, L. D., Wuestefeld, A. P., and Saldana, H. M., "Auditory looming perception: Influences on anticipatory judgments", *Perception* **22**, 1467-1482 (1993)

[28] Schiff, W. and Oldak, R., "Accuracy of judging time to arrival: Effects of modality, trajectory, and gender", *Journal of Experimental Psychology: Human Perception and Performance* **16**, 303-316 (1990)

[29] Canévet, G., Scharf, B., Schlauch, R. S., Teghtsoonian, M., and Teghtsoonian, R., "Perception of changes in loudness", *Nature* **398**, 673 (1999)

[30] Olsen, K. N. and Stevens, C. J., "Perceptual overestimation of rising intensity: Is stimulus continuity necessary?", *Perception* **39**, 695-704 (2010)

[31] Susini, P., Meunier, S., Trapeau, R., and Chatron, J., "End level bias on direct loudness ratings of increasing sounds", *Journal of the Acoustical Society of America* **128**, EL163-EL168 (2010)

[32] Canévet, G., Teghtsoonian, M., and Teghtsoonian, R., "Assimilation and asymmetry of loudness change in magnitude-estimation measurements", *Acta Acustica united with Acustica* **89**, 530-539 (2003)

[33] Susini, P., McAdams, S., and Smith, B. K., "Global and continuous loudness estimation of time-varying levels", *Acta Acustica united with Acustica* **88**, 536-548 (2002)

[34] Susini, P., McAdams, S., and Smith, B. K., "Loudness asymmetries for tones with increasing and decreasing levels using continuous and global ratings", *Acta Acustica united with Acustica* **93**, 623-631 (2007)

[35] Jahnke, J. C., "Serial position effects in immediate serial recall", *Journal of Verbal Learning and Verbal Behavior* **2**, 284–287 (1963)

[36] Olsen, K. N. and Stevens, C. J., "Forward masking of dynamic acoustic intensity: Effects of intensity region and end-level", *Perception* **41**, 594-605 (2012)

[37] Ries, D. T., Schlauch, R. S., and DiGiovanni, J. J., "The role of temporal-masking patterns in the determination of subjective duration and loudness for ramped and damped sounds", *Journal of the Acoustical Society of America* **124**, 3772–3783 (2008)

[38] Oxenham, A. J., "Forward masking: Adaptation or integration?", *Journal of the Acoustical Society of America* **109**, 732-741 (2001)

[39] Kuwano, S. and Namba, S., "Continuous judgment of level-fluctuating sounds and the relationship between overall loudness and instantaneous loudness", *Psychological Research* **47**, 27-37 (1985)

[40] Dean, R. T., Bailes, F., and Schubert, E., "Acoustic intensity causes perceived changes in arousal levels in music: An experimental investigation", *PloS one* **6**, e18591 (2011)

[41] Olsen, K. N., Dean, R. T., and Stevens, C. J., "A continuous measure of musical engagement contributes to prediction of perceived arousal and valence", *Psychomusicology: Music, Mind, and Brain* **24**, 147-156 (2014)

[42] Geringer, J. M., "Continuous loudness judgments of dynamics in recorded music excerpts", *Journal of Research in Music Education* **43**, 22-35 (1995)

[43] Bailes, F. and Dean, R. T., "Comparative time series analysis of perceptual responses to electroacoustic music", *Music Perception* **29**, 359-375 (2012)

[44] Dean, R. T. and Bailes, F., "Time series analysis as a method to examine acoustical influences on real-time perception of music", *Empirical Musicology Review* **5**, 152-175 (2010)

[45] Olsen, K. N., Stevens, C. J., Dean, R. T., and Bailes, F., "Continuous loudness response to acoustic intensity dynamics in melodies: Effects of melodic contour, tempo, and tonality", *Acta Psychologica* **149**, 117-128 (2014)

[46] Meddis, R., "Auditory-nerve first-spike latency and auditory absolute threshold: a computer model", *Journal of the Acoustical Society of America* **119**, 406-417 (2006)

[47] Meddis, R., Hewitt, M. J., and Shackleton, T. M., "Implementation details of a computation model of the inner hair-cell auditory-nerve synapse", *Journal of the Acoustical Society of America* **87**, 1813 (1990)

[48] Smith, R. and Brachman, M., "Adaptation in auditory-nerve fibers: a revised model", *Biological cybernetics* **44**, 107-120 (1982)

[49] Sumner, C. J., Lopez-Poveda, E. A., O'Mard, L. P., and Meddis, R., "Adaptation in a revised inner-hair cell model", *Journal of the Acoustical Society of America* **113**, 893-901 (2003)

[50] Hancock, P. and Manser, M., "Time-to-contact: More than tau alone", *Ecological Psychology* **9**, 265-297 (1997)

[51] Manser, M. and Hancock, P., "Influence of approach angle on estimates of time-to-contact", *Ecological Psychology* **8**, 71-99 (1996)

LANGUAGE EXPERIENCE SHAPES PROCESSING OF PITCH RELEVANT INFORMATION IN THE HUMAN BRAINSTEM AND AUDITORY CORTEX: ELECTROPHYSIOLOGICAL EVIDENCE

Ananthanarayan Krishnan, PhD[1]; Jackson T. Gandour, PhD[1]

[1]Department of Speech Language Hearing Sciences, Purdue University, West Lafayette, IN 47907, USA

rkrish@purdue.edu

Pitch is a robust perceptual attribute that plays an important role in speech, language, and music. As such, it provides an analytic window to evaluate how neural activity relevant to pitch undergo transformation from early sensory to later cognitive stages of processing in a well coordinated hierarchical network that is subject to experience-dependent plasticity. We review recent evidence of language experience-dependent effects in pitch processing based on comparisons of native vs. nonnative speakers of a tonal language from electrophysiological recordings in the auditory brainstem and auditory cortex. We present evidence that shows enhanced representation of linguistically-relevant pitch dimensions or features at both the brainstem and cortical levels with a stimulus-dependent preferential activation of the right hemisphere in native speakers of a tone language. We argue that neural representation of pitch-relevant information in the brainstem and early sensory level processing in the auditory cortex is shaped by the perceptual salience of domain-specific features. While both stages of processing are shaped by language experience, neural representations are transformed and fundamentally different at each biological level of abstraction. The representation of pitch relevant information in the brainstem is more fine-grained spectrotemporally as it reflects sustained neural phase-locking to pitch relevant periodicities contained in the stimulus. In contrast, the cortical pitch relevant neural activity reflects primarily a series of transient temporal neural events synchronized to certain temporal attributes of the pitch contour. We argue that experience-dependent enhancement of pitch representation for Chinese listeners most likely reflects an interaction between higher-level cognitive processes and early sensory-level processing to improve representations of behaviorally-relevant features that contribute optimally to perception. It is our view that long-term experience shapes this adaptive process wherein the top-down connections provide selective gating of inputs to both cortical and subcortical structures to enhance neural responses to specific behaviorally-relevant attributes of the stimulus. A theoretical framework for a neural network is proposed involving coordination between local, feedforward, and feedback components that can account for experience-dependent enhancement of pitch representations at multiple levels of the auditory pathway. The ability to record brainstem and cortical pitch relevant responses concurrently may provide a new window to evaluate the online interplay between feedback, feedforward, and local intrinsic components in the hierarchical processing of pitch relevant information.

INTRODUCTION

Pitch is an essential perceptual attribute in the processing of language and music [1, 2]. Functional brain-imaging studies provide strong evidence for hierarchical processing of pitch [3], starting in subcortical structures [4] (Griffiths, Uppenkamp, Johnsrude, Josephs & Patterson (2001) and continuing up through Heschl's Gyrus on to the planum polare and planum temporale [5-7]. Therefore, pitch provides an excellent window for studying experience-dependent effects on both cortical and brainstem components of a well-coordinated, hierarchical processing network. It is our view that for a complete understanding of the neural organization of language it is necessary to treat these processes as a set of hierarchical computations that are applied to representations at different stages (subcortical and cortical) along the processing hierarchy. Such representations, particularly of linguistically relevant features or dimensions, in turn are shaped by experience within a specific domain.

Indeed, recent empirical data show that these neural representations of pitch, at both the brainstem and cortical level, are shaped by one's experience with language and music [8-19] While it is not known how language/music experience shapes subcortical and cortical stages of pitch processing, it is likely that the neural processes underlying such experience-dependent plasticity at each stage along the processing hierarchy are modulated by a coordinated interplay between ascending, descending and local neural pathways which involve both sensory and cognitive components [20]. That is, feedback from language–dependent cortical processes shape early sensory level processing at both the brainstem and cortical level. These enhanced sensory level outputs transform later functionally more salient cortical representations that drive processes mediating linguistic performance. This review is largely confined to crosslanguage (Mandarin vs English) electrophysiological studies evaluating processing of linguistically-relevant pitch contours in the brainstem and auditory cortex. Based on empirical evidence, we propose a theoretical framework that includes local, feedback, and feedforward components to account for experience-dependent enhanced representations of pitch at brainstem and cortical stages of processing.

WHY THE FOCUS ON TONE LANGUAGES?

Tone languages are especially useful for studying pitch because variations in pitch convey part of the meaning of a word [21,22]. It is well established that dynamic variations in voice fundamental frequency (F0) provide the dominant acoustic cue for tonal recognition [23-25]. Mandarin Chinese is a lexical tone language with four phonologically distinctive tones (see Figure 1) characterized by syllable-level fundamental frequency (f0) pitch contours. These pitch contours are commonly described as high-level (T1, e.g., *yi1* "clothing"), high-rising (T2, e.g., *yi2* "aunt"), falling-rising (T3, e.g., *yi3* "chair"), and high-falling (T4, *yi4* "easy"). Such languages are to be distinguished from those in which pitch variations are usually not contrastive at the syllable or word level (e.g., English). In nontone languages however, variations in pitch may be used to signal stress and intonation patterns at post-lexical levels of representation. Thus, tone languages not only give us a physiologic window to evaluate how neural representations of linguistically-relevant pitch attributes emerge along the early stages of sensory processing in the hierarchy, but they may also shed light on the nature of interaction between early sensory levels and later higher levels of cognitive processing in the human brain.

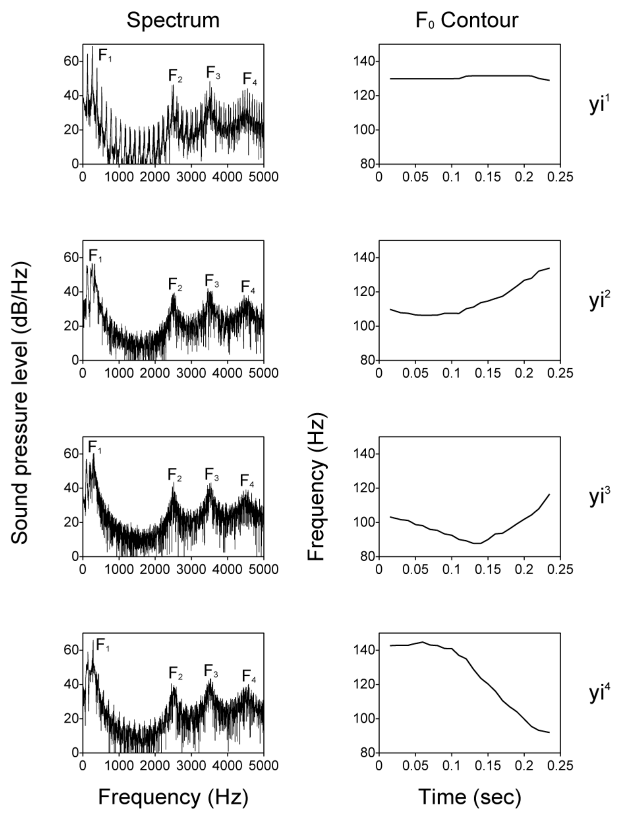


Figure 1. Spectra (left panel) and pitch contours of the four Mandarin lexical tones. Note that that were speech stimuli with invariant spectra across the four tones. The syllable is identified to the right of each panel.

LANGUAGE-EXPERIENCE SHAPES REPRESENTATION OF PITCH RELEVANT INFORMATION IN THE BRAINSTEM

Electrophysiological recordings are not only crucial for investigating questions about the hierarchy of pitch processing in the cerebral cortex, but also in subcortical structures [26] We present here empirical evidence suggesting that long-term language experience enhances the neural representation of linguistically relevant pitch in the human brainstem, well before evoked neural activity relevant to pitch is detected in the auditory cortex. Indeed, neural representation of pitch relevant attributes, as reflected in the scalp-recorded frequency following response (FFR), may emerge as early as 6-9 ms after stimulus onset [27]. In contrast, the pitch related neural activity in the auditory cortex emerges at about 140-170 ms post stimulus onset [28-32].

Characteristics of the brainstem frequency following response (FFR)

We use the FFR as a physiologic window into the early stages of subcortical processing relevant to pitch. The FFR reflects sustained phase-locked activity in a population of neural elements within the rostral brainstem, presumably the inferior colliculus [20, 33]. These responses can be easily recorded between scalp electrodes placed at high forehead and the seventh cervical vertebra (C7). The response is characterized by a waveform that follows the periodicities contained in both the envelope and temporal fine structure of complex sounds (see Figure 2).

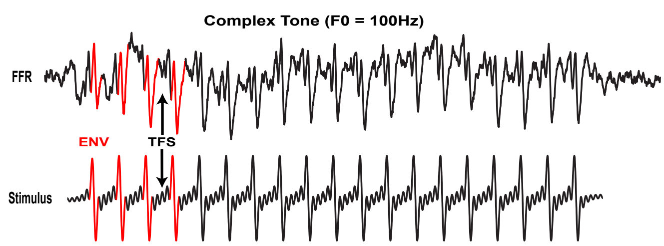


Figure 2. Frequency following responses (top trace) elicited by a complex sound (bottom trace). The response is characterized by neural phase-locking to the slower envelope periodicity (shown in red for a few cycles) and the faster temporal fine structure (TFS) of the complex sound (indicated by black arrow). Stimulus onset is shifted to the right to achieve temporal match between the stimulus and the response.

The temporal and spectral characteristics of complex sounds preserved in the FFR, can be extracted by frequency domain spectral and time domain (see figure 3) autocorrelation analysis (a measure of, correlation between the original signal and temporally delayed versions of the signal yielding high correlation for periodicities harmonically related to the fundamental frequency), respectively [34,35]. FFRs preserve spectrotemporal information relevant to the spectra (see Figure 4) [36] and pitch of steady-state [37] and dynamic speech [34,35] and nonspeech stimuli (see Figure 6) [38-41]. Importantly, the pitch-relevant information preserved in the FFR is strongly correlated with perceptual measures of pitch salience [38, 42]. Pitch salience is a measure of the strength of the perceived pitch. These findings suggest acoustic features relevant to pitch are preserved in the temporal pattern of phase-locked neural activity in the brainstem. Gockel, Carlyon, Mehta & Plack [43] observed that FFRs recorded using frequency shifted complex tones presented monaurally did preserve pitch relevant information, but that this information at the brainstem level was similar to that measured in an auditory nerve model. Also, they failed to observe any pitch relevant information in the FFRs to dichotically presented three tone harmonic stimuli. These findings led them to conclude that there was no additional pitch relevant processing at the level of the brainstem. Several arguments may be presented to counter this inference. First, if the temporal code for pitch available at the brainstem level also utilizes autocorrelation to determine the global distribution of interspike intervals from the temporal pattern of neural activity across a population of neurons, it would necessarily share certain fundamental attributes of the same temporal code operating at the level of the auditory nerve. Second, it is not clear that the dichotic stimuli used in their experiments would produce the same pitch as when all harmonics are presented to the same ear. Even if it does, its salience would be quite weak. It is possible that the FFR related neural activity is not sufficiently robust to preserve the less salient pitch for this stimulus. In our own experience, we have failed to measure FFR correlates of the less salient dichotic Huggins pitch. Finally, these inferences cannot convincingly account for the experience-dependent effects reflected in the FFR that are sensitive to specific attributes of dynamic pitch contours.

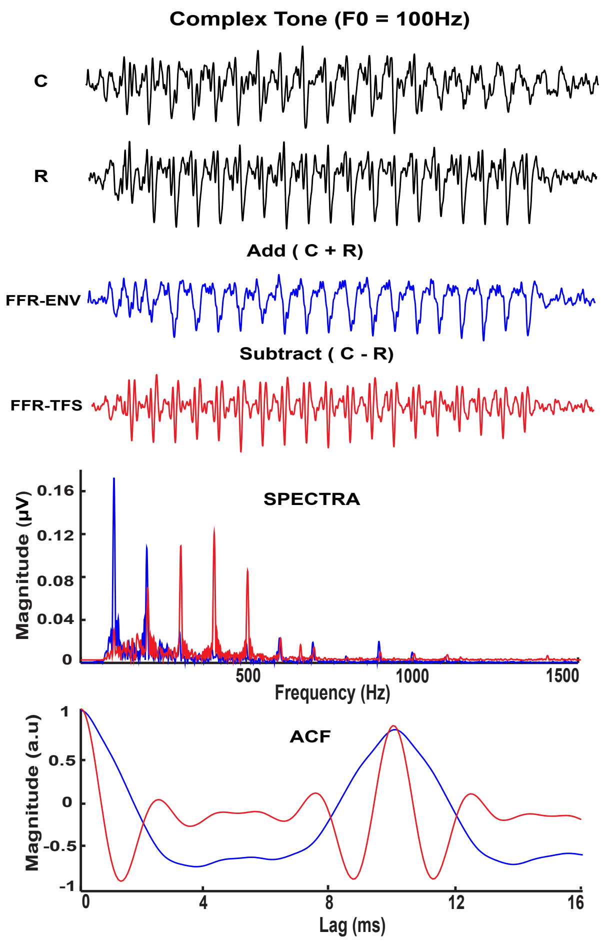


Figure 3. Frequency and time domain analysis of the FFR. The two top FFR waveforms (black) are responses to condensation (C) and rarefaction (R) onset stimulus polarity. The result of addition of these waveforms results in the FFR-ENV waveform (Blue), which is characterized by a prominent phase-locking to the envelope periodicity of the complex tone. In contrast, the subtracted (C-R) FFR shows phase locking to spectral components of the complex stimulus (Red). Frequency domain analysis (middle panel) shows that the envelope phase-locking, as expected, has a larger peak at F0 (blue) whereas phase locking to the spectral components of the stimulus are relatively smaller. The temporal analysis, (bottom panel) using autocorrelation, shows that for both responses there is a major peak at the fundamental periodicity of the complex tone.

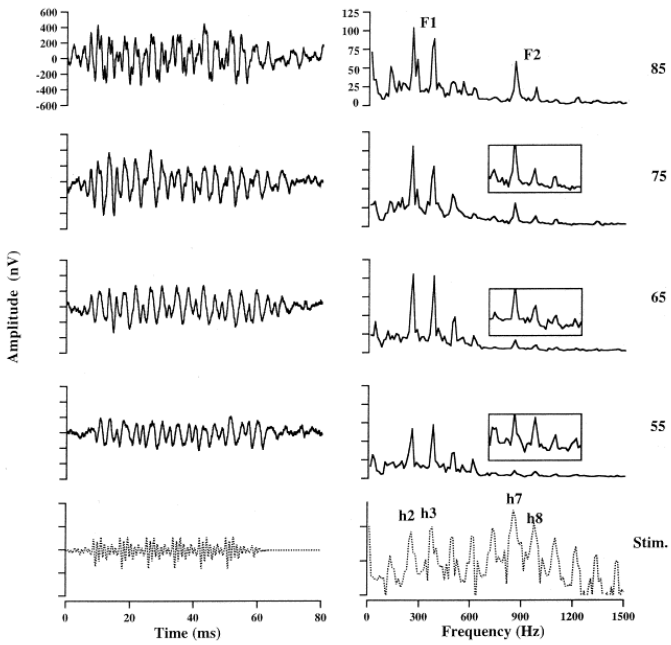


Figure 4. Grand averaged FFR waveforms and spectra are plotted as a function of stimulus level (dB nHL) for the English **vowel /u/**. The stimulus waveform and its spectrum (with F_1 and F_2 harmonics identified) are at the bottom of each panel. The amplified inset in the FFR spectral data clearly shows the F_2 harmonic peaks. Note the different amplitude scale for the stimulus spectrum. Note that the stimulus spectrum is without an amplitude scale and is used here to look for correspondence between the stimulus and the response spectrum. **Reprinted from Krishnan, 2003 with permission of Elsevier B.V.** A. Pitch tracking accuracy as a function of iteration steps. Note that pitch tracking accuracy increases more rapidly with increasing pitch salience for the Chinese listeners (red trace). B. Relationship between neural pitch strength and behavioral measure of pitch discrimination (F0 DL). Note that as F0 DL decreases with increasing pitch salience neural strength increases.

Evidence for experience-dependent plasticity in the human brainstem

The temporal pattern of the phase-locked neural activity generating the FFR preserves pitch-relevant information of lexical tones in both native (Mandarin) and nonnative (English) listeners. However, both pitch tracking accuracy (the strength of the correlation between the stimulus pitch contour, and the pitch contour extracted from the FFR) and periodicity strength (the strength of neural phase-locking to the fundamental periodicity of the stimulus, as reflected by the magnitude of the autocorrelation peak at a delay corresponding to the pitch period) is more robust in the former (Figure 5) because of long-term exposure to their native pitch contours [35].

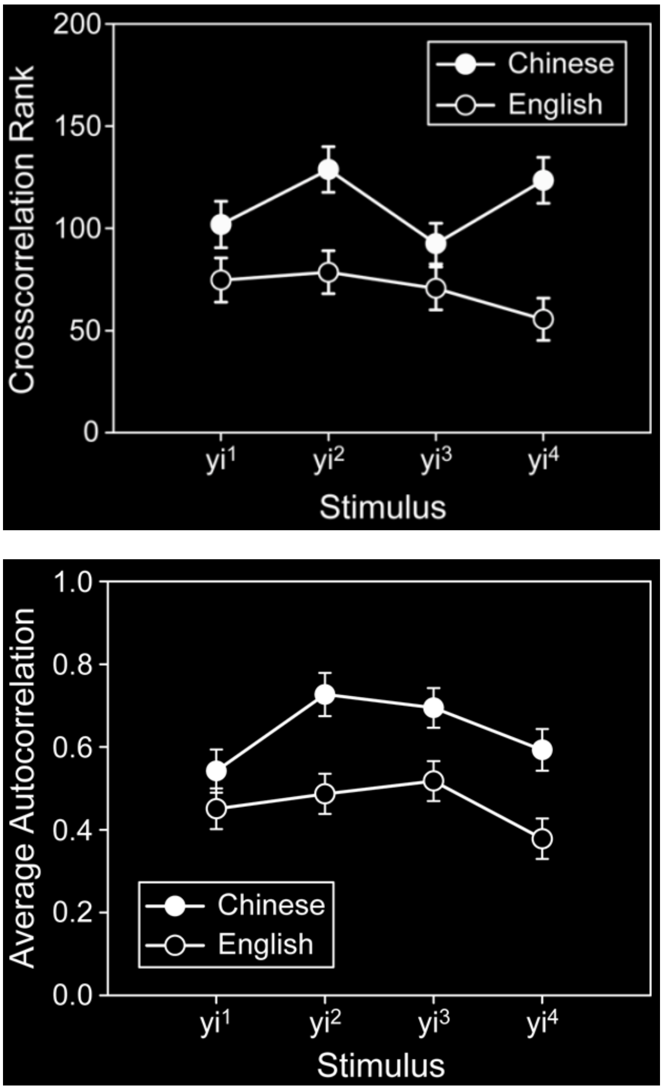
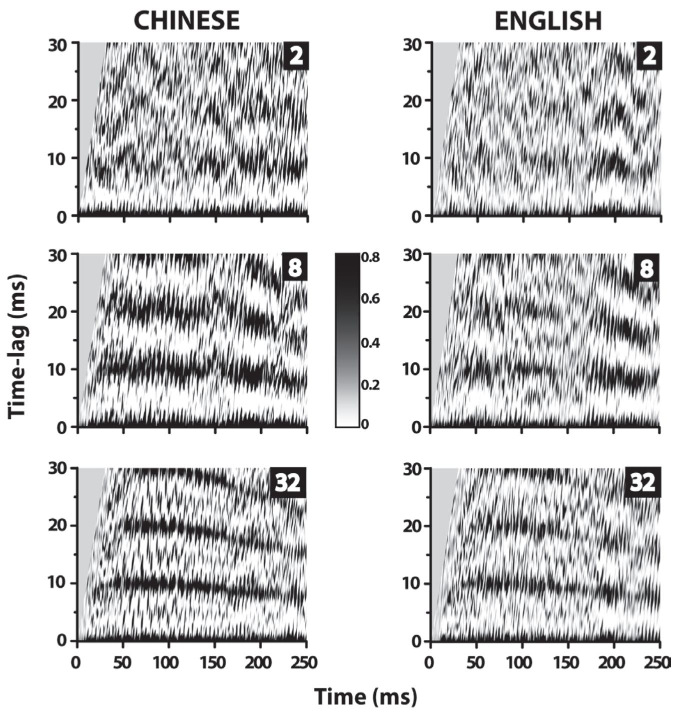


Figure 5. A & B. Comparison of pitch tracking accuracy (top panel) and pitch strength (bottom panel) for Chinese and English listeners for the four lexical tones. For both measure Chinese show better tracking accuracy and greater periodicity strength across the four stimuli compared to English listeners. Reprinted from Krishnan et al., 2005.

We therefore *conclude that early sensory representation of pitch-relevant information in the brainstem is sensitive to language experience*. This enhanced representation in Mandarin speakers is observed for native pitch contours whether they are presented in a speech or nonspeech context [35, 41]. Moreover, we find that Chinese exhibit more robust pitch representation precisely in those segments of Mandarin tones that contain rapid changes in pitch [40]. With respect to pitch features, brainstem representation of rising contours is found to be most important in discriminating Chinese from listeners of non-tone languages [44] Pitch features important to tone perception show increased resistance to degradation of the temporal regularity (decrease in temporal regularity reduces the perceived salience of pitch) of stimuli in Chinese listeners [45]. Iterated rippled noise (IRN) stimulus approximating the native lexical pitch contour of Tone 2 was used as a non-speech stimulus in this study.



Figures 6. Autocorrelograms derived from grand averaged FFR waveforms of Chinese (left) and English (right) groups in response to T2IRN at low (**n**=2), intermediate (**n**=8), and high (**n**=32) iteration steps. Little difference is seen between the Chinese and English group at low iteration steps (**n**=2; top row). By **n**=8 iterations, Chinese show clearer, tighter bands of temporal regularity (black) in FFR phase-locked activity at the fundamental period (1/f0) and its multiples as compared to the English group (middle row). At n=32, this superiority is even more pronounced (bottom row). Periodicity strength is indicated by the color gradient; darker shades indicate higher temporal regularity (i.e., more pronounced phase-locked activity). **Reprinted from Krishnan et al., 2010 with permission of Elsevier B.V.**

This IRN stimulus with a dynamic pitch contour was generated by applying a 32-step time-varying, delay-and-add iterative process to a broadband noise. These advantages in neural representation in the Mandarin listeners, however, diminish as the pitch acceleration within a pitch contour exceeds values beyond what occurs in natural speech [46]. Collectively, *these results lead us to conclude that experience-dependent neuroplasticity is primarily restricted to stimuli that are of functional relevance to the listener. Indeed, a strong correlation between neural and behavioral measures corroborates the view that neural representation of pitch-relevant information at a subcortical, sensory level of processing plays an important role in shaping tone perception [38].*

Language experience-dependent pitch processes in the brainstem are especially sensitive to the curvilinear shape of pitch contours that occur in natural speech. Enhanced representations do not occur for linear approximations of native rising or falling pitch contours [39, 47]. A nonnative, curvilinear pitch pattern similarly fails to elicit a language-dependent effect [39]. *Thus, experience-dependent effects in the brainstem are highly sensitive to specific fine-grained features of pitch patterns in one's native language.*

Finally, our comparisons between the language and music domains reveal overall enhancement in brainstem FFRs elicited by either musical or linguistic pitch patterns in musicians and tone language speakers alike [48,49]. Thus, long-term pitch experience seems to improve the brain's ability to represent pitch-relevant information regardless of the domain of expertise. However, subtle differences in these sensory representations suggest a domain-specific sensitivity to acoustic features that are part of the experience in each domain. Musicians, for example, show enhanced responses when pitch patterns intersect discrete notes along the musical scale; tone language speakers, on the other hand, during rapidly changing portions of tonal contours [48, 50]. Such cue weighting is consistent with the relative importance of these perceptual dimensions in their respective domains. *These findings collectively lead us to infer that both language and musical experience provide some mutual benefit to the neural representation of pitch-relevant information, but that specific features of the acoustic signal are highlighted in subcortical responses depending on their perceptual salience and function within a listener's domain of expertise.*

LANGUAGE-EXPERIENCE SHAPES REPRESENTATION OF PITCH RELEVANT INFORMATION IN THE AUDITORY CORTEX

Characteristics of the Pitch Onset Response (POR) using Magnetoencephalography (MEG)

At the cortical level, MEG has been used previously to study the sensitivity to pitch relevant periodicity, by investigating the N100m component. However, this component to a large extent simply represents sound onset and does not exclusively represent pitch [51-55]. To obtain pitch-specific response uncontaminated by the onset response, a novel stimulus paradigm was developed wherein two segments – an initial noise segment devoid of pitch evoked the onset responses only, followed temporally by an iterated rippled noise (IRN) segment, matched in intensity and overall spectral profile of the noise precursor, which elicited the pitch response components [28]. Interestingly, a transient pitch onset response (POR), with a latency of about 140-170 ms, was evoked from this noise-to-pitch transition. The reverse stimulus transition from pitch to noise failed to produce a POR. It has been proposed that the human POR, as measured by MEG, reflects synchronized cortical neural activity specific to pitch (28, 29; 56, 57]. POR latency and magnitude, for example, has been shown to depend on pitch salience. A more robust POR with shorter latency is observed for stimuli with stronger pitch salience as compared to those with weaker pitch salience. Source analyses [28, 52, 58] corroborated by human depth electrode recordings [31, 59] indicate that the POR is localized to the anterolateral portion of Heschl's gyrus, the putative site of pitch processing [5, 6, 60-62]. Given both its sensitivity to pitch and its salience, and consistency across a number of studies, the POR offers an excellent window for studying early sensory level representation of pitch specific information at the cortical level. Our preliminary POR data, extracted from scalp-recorded EEG, yielded *multiple* peaks (labeled as Pa (latency: 70-80 ms), Na (130-150 ms), Pb (200-215 ms), Nb (265-280 ms), and Pc (305-320 ms), where P and N represent peaks with positive and negative polarity, respectively) in addition to pitch onset response (POR). We therefore have chosen to designate this scalp-recorded neural activity as cortical pitch response (CPR).

Characteristics of the cortical pitch response (CPR) using electroencephalography (EEG)

Using a similar two-segment stimulus paradigm (see Figure 7), we demonstrated that the EEG-derived human cortical pitch response (CPR) elicited by IRN steady-state pitch stimuli increased in magnitude with increasing temporal regularity of the stimulus [32]. This change in CPR response amplitude with increasing stimulus temporal regularity was strongly correlated with behavioral measures of change in pitch salience (perceived strength of pitch). That the CPR is pitch specific is confirmed by its absence in response to a "no-pitch" IRN stimulus. Recently we examined the sensitivity of the multiple transient components of the CPR to specific temporal attributes associated with three (see Figures 8, 9 and 10), within-category variants of Mandarin Tone 2 (T2_150; T2_200; and T2_250) that varied in pitch acceleration (rate of change of fundamental frequency) and duration [63]. Our results showed that the latency of the pitch onset response (Na) was unaltered by changes in pitch acceleration (Figures 8 and 9). In contrast, response components Pa-Na, Na–Pb and Pb–Nb showed a systematic decrease in the interpeak latency, and a decrease in amplitude with increase in pitch acceleration (see figure 10 for amplitude changes with acceleration). These changes in interpeak latencies closely followed the time course of pitch change across the three stimuli. This is readily apparent if the stimulus pitch contours (in Figure 7) are compared to the 500-700 ms response window (in Figure 8). A strong correlation with pitch acceleration was observed for Na-Pb and Pb-Nb only - a putative index of pitch-relevant neural activity associated with the more rapidly-changing portions of the pitch contour. Pc–Nc marks unambiguously the stimulus offset. These data lead us to propose that in the early stages of cortical sensory processing of pitch, a series of neural markers flag different temporal attributes of a dynamic pitch contour: onset of temporal regularity (Na); changes in temporal regularity between onset and offset (Na–Pb, Pb–Nb); and offset

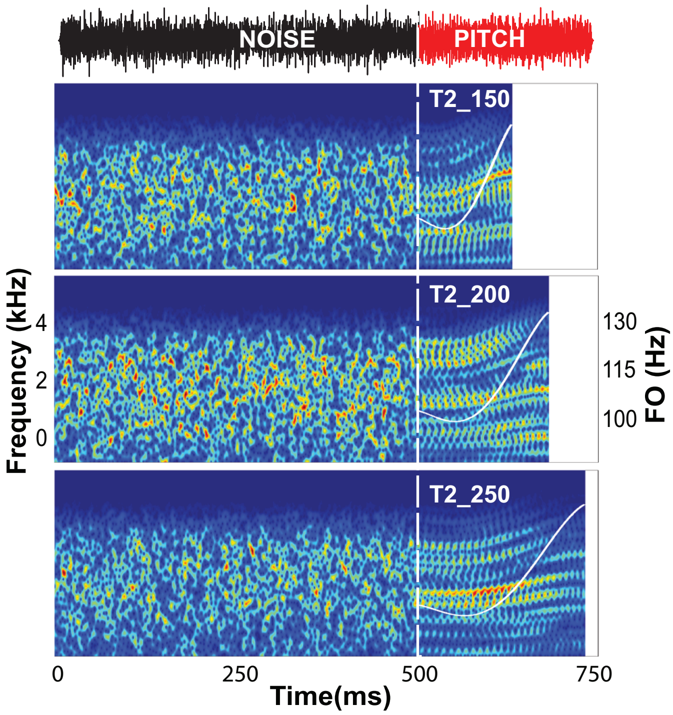


Figure 7. Noise-to-pitch Waveform (T2_250) and spectrograms of each of the three iterated rippled noise (IRN) stimulus conditions (T2_150; T2_200; T2_250) illustrate the experimental paradigm used to acquire brainstem and cortical responses concurrently. The vertical dashed line at 500 ms demarcates the transition from the initial noise segment to the final pitch segment. FFRs and CPRs were extracted from evoked responses beginning with the onset of the pitch. F0 contours (white) are superimposed on their respective pitch segment spectrograms. Within the pitch segment, the waveform (top) shows robust periodicity at a high IRN iteration step (n=32); the spectrograms show clear resolution of dynamic, rising spectral bands corresponding to the harmonics of the fundamental frequency. **Reprinted from Krishnan et al., 2014 with permission from Elsevier R.V.**

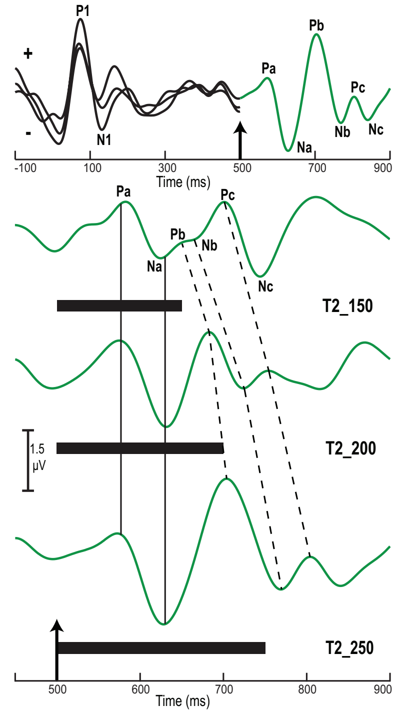


Figure 8. Grand averaged cortical evoked response components at the Fz electrode site per stimulus condition. The P1/N1 onset complex for the three stimuli (black) and the CPR component to T2_250 (green) are displayed in the top panel. The up arrow at 500 ms marks the boundary between the noise precursor segment and the pitch-eliciting stimulus. Na, Pb, and Nb are the most robust response components. CPR waveforms (green) elicited by the three stimuli (T2_150, T2_200, T2_250) are shown in the bottom panel. The up arrow at 500 ms marks the beginning of the pitch-eliciting stimulus. Solid black horizontal bars indicate the duration of each stimulus. Whereas Pa and Na do not change across stimuli (solid vertical lines), Pb, Nb, Pc, and Nc all show a systematic increase in peak latency (dashed vertical lines). Response amplitude for Na, Pb, and Nb increases from T2_150 to T2_250 in conjunction with decreasing pitch acceleration and increasing duration across stimuli. **Reprinted from Krishnan et al., 2014 with permission from Elsevier R.V.**

of temporal regularity (Pc–Nc). Furthermore, an asymmetry favoring the right hemispheric (RH) was observed only for the prototypical tonal variant (T2_250) suggesting the emergence of early stimulus-dependent hemispheric asymmetry (see Figure 13). The stimulus with the most gradual change in pitch acceleration, T2_250, appears to cross that threshold that leads to a RH asymmetry. Yet within the left hemisphere (LH), T2_250 was indistinguishable from T2_200. By assuming that the LH is preferentially engaged for mediating the categorical status of pitch, this finding suggests that T2_200 as well as T2_250 are better candidates as representations of Mandarin Tone 2. Thus, within this early cortical time window we begin to see the emergence of hemispheric preference, and how the complementary roles of the two hemispheres reflect influences from sensory and cognitive properties of the stimulus.

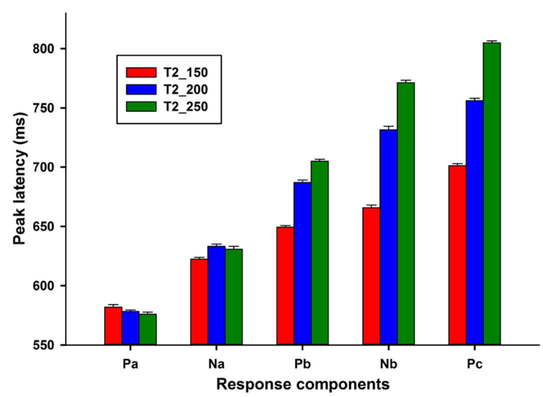


Figure 9. Mean peak latency of Fz response components as a function of stimulus. From stimulus onset (Pa) to stimulus offset (Pc), peak latencies show increasing differentiation across stimuli (T2_150; T2_200; T2_250) reflecting sensitivity to increases in pitch acceleration and duration. No changes in peak latency are evident for Pa in contrast to marked changes across stimuli for Pb, Nb, and Pc. Na (pitch onset) reveals that the peak latency of T2_150 is shorter than either T2_200 or T2_250. **Reprinted from Krishnan et al., 2014 with permission from Elsevier R.V.**

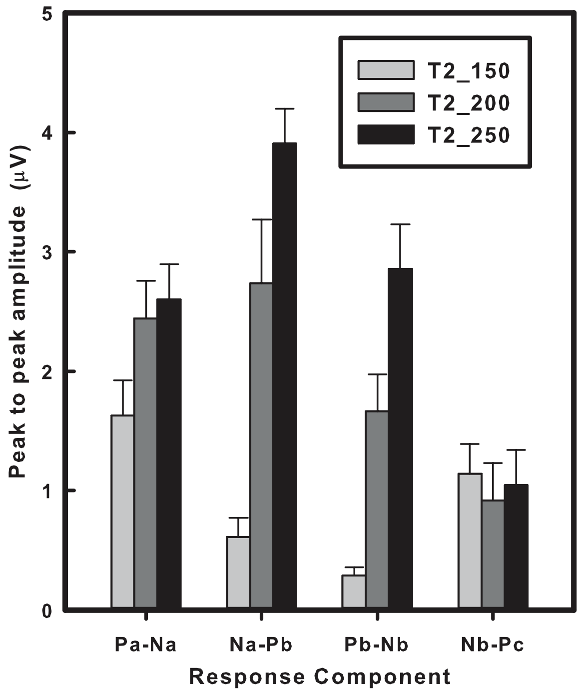


Figure 10. Mean peak-to-peak amplitude of Fz response components as a function of stimulus. Two pitch-relevant components (Pa-Na, Na–Pb, Pb–Nb) show that peak-to-peak amplitude increases steadily across the three stimuli (T2_150; T2_200; T2_250). In the case of Pa–Na, T2_250 is marginally higher in peak-to-peak amplitude than T2_150. Nb–Pc shows no differences in peak-to-peak amplitude between stimuli. Error bars.+/-1SE. **Reprinted from Krishnan et al., 2014 with permission from Elsevier R.V.**

Evidence for experience-dependent plasticity in the auditory cortex

Having described the basic characteristics of CPR components, we next examined how language experience (Mandarin vs. English) shapes the processing of different temporal attributes of pitch reflected in those CPR components using the same three variants of Tone 2 [64]. Results showed that the magnitude of CPR components (Na-Pb and Pb-Nb) and the correlation between these two components and pitch acceleration were stronger for the Chinese listeners (see Figures 11 and 12) compared to English listeners for stimuli that fell within the range of Tone 2 (acceleration rates presented in the pitch contours of T2_250 and T2_200 do occur in the language experience whereas the acceleration rate for the pitch contour T2_150 is well beyond rates in their language experience). Discriminant function analysis is used to determine which variables discriminate between two or more naturally occurring groups. In our application, it revealed that the Na-Pb component was more than twice as important as Pb-Nb in grouping listeners by language affiliation. This language-dependent effect suggests an experience-dependent increase in sensitivity to rapidly changing portions of pitch contours that occur in the native listeners' experience. Because enhanced sensitivity to time-varying dimensions (e.g., acceleration) is already present at the level of the brainstem [12, 27] it seems plausible that cortical pitch mechanisms may be reflecting, at least in part, this enhanced pitch input from the brainstem. We further note that experience-dependent enhancement of pitch was reflected primarily in the amplitude, instead of the latency, of CPR components. The more robust amplitude suggests greater temporal synchronization and improved synaptic efficiency of pitch-relevant neural activity among cortical neurons generating these CPR components. In contrast, absolute and interpeak latency may simply serve as discrete temporal event markers of neural activity indexing the temporal course of a pitch contour. That is, while amplitude may reflect the strength of the pitch relevant neural activity, the latencies may define the tine window in which this activity is occurring.

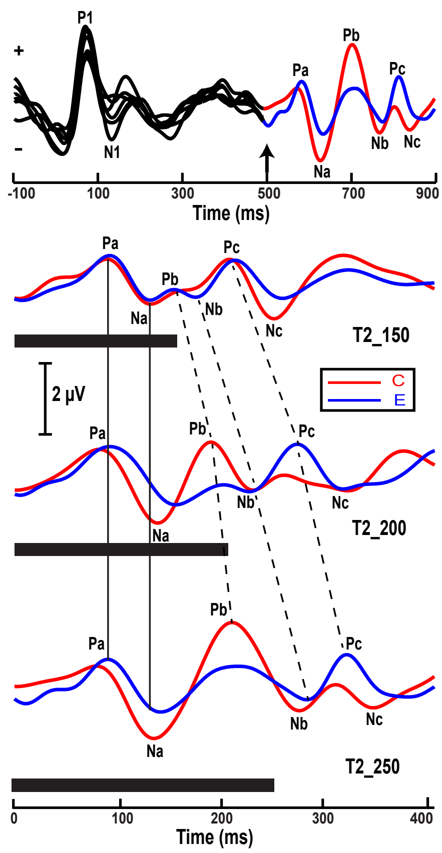


Figure 11. Grand averaged cortical evoked response components recorded at the Fz electrode site per stimulus condition. The P1/N1 onset complex for the three stimuli (black) and the CPR component to T2_250 (Chinese, red; English, blue) are displayed in the top panel. The up arrow at 500 ms marks the onset of the pitch-eliciting segment of the stimulus. Na, Pb, and Nb are the most robust response components. CPR waveforms elicited by the three stimuli (T2_150, T2_200, T2_250) are shown in the bottom panels. Solid black horizontal bars indicate the duration of each stimulus. Whereas Pa and Na do not change appreciably across stimuli (solid vertical lines), Pb, Nb, and Pc all show a systematic increase in peak latency (dashed vertical lines). Response amplitude for Na, Pb, and Nb increases from T2_150 to T2_250 in conjunction with decreasing pitch acceleration and increasing duration across stimuli. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.) **Reprinted from Krishnan et al., 2014 with permission from Elsevier, R.V.**

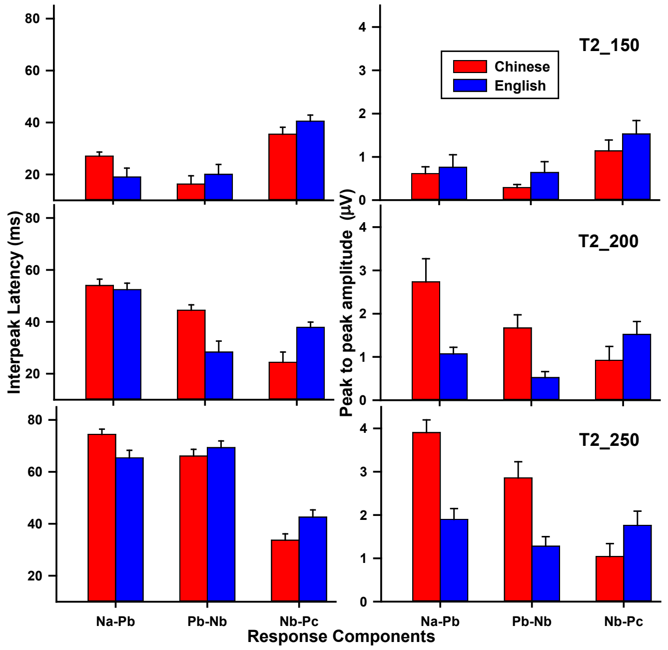


Figure 12. Mean interpeak latency (left panel) and peak-to-peak amplitude (right panel) of Na–Pb, Pb–Nb, and Nb–Pc components recorded at the Fz electrode site from T2_150 (top panel) to T2_250 (bottom panel) in both Chinese and English groups. Interpeak latencies increase across stimuli for Na–Pb and Pb–Nb in both groups. In the case of T2_200 (middle panel), Chinese interpeak latency is longer than English for Pb–Nb, but shorter than English for Nb–Pc. The Chinese group exhibits larger amplitude than the English group for Na–Pb and Pb–Nb in T2_200 and T2_250. Error bars = ±1 SE. **Reprinted from Krishnan et al., 2014 with permission from Elsevier, R.V.**

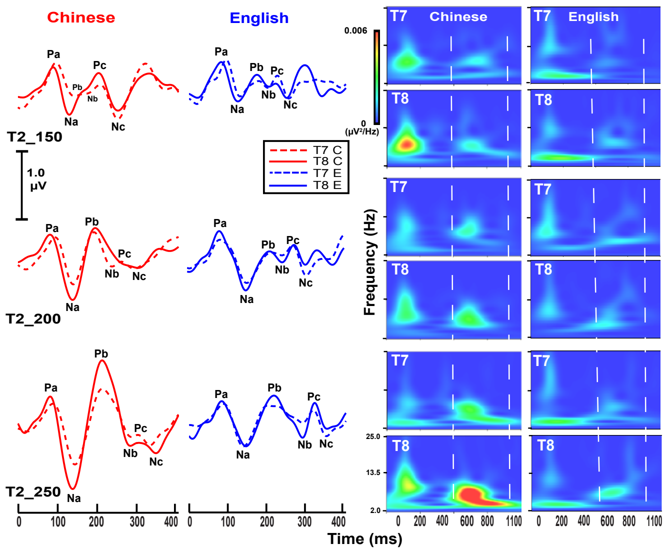


Figure 13. Grand average waveforms (two left panels) and their corresponding spectrograms (two right panels) of the CPR components for the two groups (Chinese in red and English in blue) recorded at electrode sites T7 (left temporal) and T8 (right temporal) for each of the three stimuli. Pitch-relevant CPR response components are generally greater in magnitude, and in particular for stimulus T2_250, for the Chinese group compared to the English group. The spectrograms in the right two panels show a large rightward asymmetry in the pitch-relevant neural activity (time window indicated by the dashed vertical lines) for the Chinese group only for stimulus T2_250. Note that while the waveform data time axis is referenced to the onset of the pitch segment (0-400 ms), the spectrogram is for the entire stimulus. The two vertical dashed lines in the spectrogram indicate the time window corresponding to CPR components. **Reprinted from Krishnan et al., 2014 with permission from Elsevier, R.V.**

In addition, a stronger stimulus-dependent, RH preference was observed for the Chinese group (see Figure 13). One view is that auditory processing occurs symmetrically in the core (primary auditory cortex), but asymmetrically in auditory-related areas beyond the core [65, 66]. We hypothesize that the language-dependent RH preference of the CPR response to T2_250 is due to an interaction with pitch-specific areas beyond the core that, in turn, are connected to higher order memory areas related to language. As such, it is an example of the essential interaction between sensory and cognitive components in pitch processing within the language domain in right auditory-related cortex. These findings are consistent with earlier ERP data showing emergence of experience-dependent hemispheric preference at early cortical levels of processing. For example, pure tones produced larger early cortical components (N19 mP30m) with a RH preference in musicians only; non-musicians show no hemispheric preference [67]. In response to musical stimuli, the N1 (change-N1, 100ms latency) component, related to pitch transition, showed greater amplitude at the right temporal electrode site in trained musicians [68]. Both these studies, using musically relevant stimuli point to experience-dependent enhancement of pitch processing in the right auditory cortex related to the music domain. We must also point out that our stimuli exhibit dynamic, curvilinear F0 trajectories that are representative of a Mandarin lexical tone. Interestingly, MEG recordings fail to observe any hemispheric differences with regard to either latency or amplitude of the pitch-relevant cortical component elicited by stimuli with flat pitch [29, 52, 69, 70]. This disparity in hemispheric asymmetry between dynamic and flat pitch patterns, seen in our Chinese listeners, further emphasizes both, the specialization of neural networks driven by behavioral relevance of sound and the importance of using ecologically-relevant stimuli to study pitch processing in the language domain.

Taken together, these findings suggest that long term language experience shapes early sensory level processing of pitch in the auditory cortex, and that the sensitivity of the CPR and hemispheric preference for processing pitch information may vary depending on the relative linguistic importance of specific temporal attributes of dynamic pitch. Our findings are consistent with earlier cross-language studies that have revealed experience-dependent neural plasticity at both subcortical and cortical levels of the brain [9, 12, 17, 19]. We believe that long-term experience shapes pitch mechanisms at early sensory levels of pitch processing in the auditory cortex. Like in the brainstem, they sharpen response properties of neural elements to enable optimal representation of linguistically relevant temporal attributes of native pitch contours.

NEURAL MECHANISM(S) FOR EARLY SENSORY LEVEL PITCH PROCESSING IN THE BRAINSTEM AND AUDITORY CORTEX

It is generally agreed that lateral Heschl's gyrus is the putative source for the pitch onset component (Na). Specific generator sources for the remaining pitch-relevant components (Pb, Nb, Pc, Nc) are yet to be determined. We speculate that these later components (Na-Pb, Pb-Nb) reflect neural activity from spatially distinct generators that represent later stages of sensory processing, relative to Na, along a pitch processing hierarchy. Whether pitch-relevant information extracted by these neural generators is based on a spectral and/or temporal code is unclear. At subcortical levels up to the midbrain, physiologic and computational modeling data support the possibility of either a purely temporal mechanism or a hybrid mechanism using both spectral and temporal information [2, 71-73]. There is evidence that neurons in primary auditory cortex exhibit temporal and spectral response properties that could enable these pitch-encoding schemes [74, 75], but it is not known whether they form a network with pitch-selective neurons to carry out this process. Unlike the subcortical auditory structures where periodicity and pitch are often represented by regular temporal patterns of action potentials that are phase-locked to the sound waveform, the most commonly observed code for periodicity and pitch within cortical neurons is a modulation of spike rates as a function of F0. It is possible that the wider temporal integration window at the cortical level may render the auditory cortical neurons too sluggish to provide phase-locked representations of periodicity within the pitch range [76]. Thus, it is not yet clear how cortical neurons transform the autocorrelation-like temporal analysis in the brainstem to a spike rate code to extract pitch-relevant information. One possibility is that the temporal code is transformed in to a response synchrony code where temporally coherent activity from the subcortical stages will produce greater spike rate, yielding larger response amplitude at the cortical level. Interestingly, Micheyl, Strater & Oxenham [77] show (based on analyses of statistical properties of the spike rates of virtual neural units, that have similar frequency tuning and spike rate characteristics as auditory cortical neural units) that sufficient statistical information is present in the population spike rate to account for small differences in frequency (pitch) and intensity (loudness).

NEURAL MECHANISMS MEDIATING EXPERIENCE-DEPENDENT PLASTICITY FOR PROCESSING PITCH RELEVANT INFORMATION IN THE BRAINSTEM AND AUDITORY CORTEX

Experience-dependent enhancement of pitch representation for Chinese listeners most likely reflects an interaction between higher-level cognitive processes and early sensory-level processing to improve representations of behaviorally-relevant features that contribute optimally to perception. It is our view that long-term experience shapes this adaptive process wherein the top-down connections provide selective gating of inputs to both cortical and subcortical structures to enhance neural responses to specific behaviorally-relevant attributes of the stimulus. The goal clearly is to achieve optimal correspondence between the sensory representations and the resulting percept at all levels of processing [78]. Evidence for this signal selectivity, mediated through top-down influence, comes from response properties of cortical neurons in animal models, that show a selective increase in responsiveness and shifts in best frequencies toward task-relevant, target stimuli [79-81]; and selective expansion of receptive fields for stimulus features that are being learned [82]. In the case of humans, the top-down influence mediated by the corticofugal system likely shapes the enhancement of brainstem pitch representation resulting from short-term auditory training [83, 84]; long-term linguistic experience [12, 27, 35]; and musical training [48-50, 85-87]. The reverse hierarchy theory (RHT) provides a representational hierarchy to describe the interaction between sensory input and top-down processes to guide plasticity in primary sensory areas [88, 89]. This theory suggests that neural circuitry mediating a certain percept can be modified starting at the highest representational level and progressing to lower levels in search of more refined high resolution information to optimize percept. The RHT has been invoked as a plausible explanation for topdown influences on cortical [90] and subcortical sensory processing [38, 91]. Consistent with this theory, it is possible that sensory-level representation of spectrotemporal features related to pitch in the brainstem is more precise than the more labile, spatiotemporally broader, pitch-relevant information in the auditory cortex [93-95]. Indeed, fine-grained, spectrotemporal details that are present in the sustained brainstem response are absent in transient, cortical pitch response components. We nevertheless observe a close correspondence between cortical and brainstem responses when manipulating the degree of pitch salience [32].

One proposed circuitry, mediating learning-induced plasticity; the colliculo-thalamo-cortico-collicular loop [96] may be invoked to mediate the RHT described above. This circuitry is comprised of bottom-up (colliculo-thalamic and thalamo-cortical) and top-down (corticofugal) projections that form a tonotopic loop. It incorporates several neuromodulatory inputs that form a core neural circuit mediating sound-specific plasticity associated with perceptual learning. Auditory stimuli and neuromodulatory inputs are believed to induce large-scale, frequency-specific plasticity in the loop. It is also possible that bottom-up as well as local top-down cortical inputs may jointly influence pitch processing as reflected in the CPR components. In the case of the former, enhanced representations from brainstem pitch mechanisms are functionally reorganized by top-down influence during the critical period of language acquisition. As a result, brainstem responses constitute an indirect reflection of inputs from the corticofugal system. Once this reorganization is complete, local mechanisms in the brainstem and auditory cortex would be sufficiently robust to extract linguistically-relevant pitch information optimally without an engaged, online corticofugal influence [97]. Indeed, the strong correlation between neural representations relevant to pitch salience at the brainstem and early cortical levels of processing suggests that sensory processing at the brainstem level may be driving early preattentive sensory processing relevant to pitch at the cortical level [32]. In the case of humans, top-down processes likely shape the reorganization of the sensory processing of pitch-relevant information in the brainstem and early sensory level processing in the auditory cortex to enhance pitch extraction in earlier stages of language development when adaptive plasticity presumably would be most vigorous [98, 99]. The slower time constants of corticofugal processing render it much too sluggish to effectively influence a dynamic pitch pattern over its entire duration [100]. Nonetheless, its adaptive properties would still be able to facilitate extraction of behaviorally-relevant information under degraded listening conditions and during training protocols.

CONCLUSIONS

While language experience shapes pitch processing at both subcortical and cortical levels, neural representations are transformed and fundamentally different at each biological level of abstraction. The representation of pitch relevant information in the brainstem is more fine-grained spectrotemporally as it reflects sustained neural phase-locking to pitch relevant periodicities contained in the dynamic stimulus. In contrast, the cortical representation is coarser. That is, the cortical pitch relevant neural activity reflects primarily a series of distinct transient temporal neural events marking only certain temporal attributes of the pitch contour. These differences notwithstanding, we believe that long-term language experience shapes adaptive, hierarchical pitch processing. Top-down connections provide selective gating of inputs to both cortical and subcortical structures to enhance neural representation of behaviorally-relevant attributes of the stimulus and instantiate local mechanisms that exhibit enhanced representation of behaviorally-relevant pitch attributes. The ability to record brainstem and cortical pitch relevant responses concurrently may also provide a new window to evaluate the online interplay between feedforward and feedback components in the processing of pitch-relevant information at the level of the brainstem and the auditory cortex, and how experience shapes pitch processing at cortical and subcortical levels. The selection of dynamic pitch contours representative of lexical tone will enable us to evaluate the influence of language experience on the latency and amplitude of these cortical pitch response components in subsequent experiments. The challenge is to develop experiments that systematically manipulate pitch attributes in order to optimally evaluate the relationship between representation of pitch relevant information at the brainstem and cortical levels. The results of these experiments would be essential to develop a framework to understand both the nature of interplay between levels of processing and, interactions between sensory and cognitive processes influencing pitch representation. Complementary studies using MEG will be crucial to determine the anatomical sources of these components in an effort to shed more light on specific cortical generators contributing to the hierarchical stages of pitch processing, and how experience may shape these processes.

ACKNOWLEDGEMENTS

Research supported by NIH 5R01DC008549-06 (A.K.). Thanks to Jingyi Zhu and Longjie Cheng for their assistance with statistical analysis (Department of Statistics); JillianWendel and Chandan Hunsur Suresh for their help with data acquisition and graphics, respectively. Reprint requests should be addressed to Ananthanarayan Krishnan, Department of Speech Language Hearing Sciences, Purdue University, West Lafayette, IN 47907-2038,USA, or via email: rkrish@purdue.edu.

REFERENCES

[1] Oxenham, A. J. (2012). Pitch perception. *Journal of Neuroscience*, **32**(39), 13335-13338. doi:10.1523/JNEUROSCI.3815-12.2012

[2] Plack, C. J., Oxenham, A. J., & Fay, R. R. (Eds.). (2005). *Pitch: neural coding and perception (Vol. 24)*. New York: Springer

[3] Kumar, S., Stephan, K.E., Warren, J.D., Friston, K.J., Griffiths, T.D. (2007). Hierarchical processing of auditory objects in humans. *PLoS Computational Biology* **3**: e100. doi:10.1371/journal.pcbi. 0030100.

[4] Griffiths, T. D., Uppenkamp, S., Johnsrude, I., Josephs, O., & Patterson, R. D. (2001). Encoding of the temporal regularity of sound in the human brainstem. *Nature Neuroscience*, 4(6), 633-637.

[5] Griffiths, T.D., Buchel, C., Frackowiak, R.S., & Patterson, R.D. (1998). Analysis of temporal structure in sound by the human brain. *Nature Neuroscience*, **1**(5), 422–427.

[6] Patterson, R.D., Uppenkamp, S., Johnsrude, I.S., & Griffiths, T.D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron*, **36**(4), 767–776.

[7] Gutschalk, A., Patterson, R. D., Scherg, M., Uppenkamp, S., & Rupp, A. (2007). The effect of temporal context on the sustained pitch response in human auditory cortex. *Cerebral Cortex*, **17**(3), 552–561, http://dx.doi.org/10.1093/cercor/bhj180

[8] Besson, M., Chobert, J., & Marie, C. (2011). Language and music in the musician brain. *Language and Linguistics Compass*, **5**(9), 617–634, http://dx.doi.org/10.1111/j.1749-818x.2011.00302

[9] Gandour, J. T., & Krishnan, A. (2014). Neural bases of lexical tone. In: *H. Winskel, & P. Padakannaya (Eds.), Handbook of South and Southeast Asian psycholinguistics* (pp. 339–349). Cambridge, UK: Cambridge University Press.

[10] Koelsch, S. (2012). *Brain & Music*. Chichester, UK: Wiley-Blackwell.

[11] Kraus, N., & Banai, K. (2007). Auditory-processing malleability: Focus on language and music. *Current Directions in Psychological Science*, **16**(2), 105–110.

[12] Krishnan, A., Gandour, J.T., & Bidelman, G.M. (2012). Experience-dependent plasticity in pitch encoding: From brainstem to auditory cortex. *Neuroreport*, **23**(8), 498–502.

[13] Meyer, M. (2008). Functions of the left and right posterior temporal lobes during segmental and suprasegmental speech perception. *Zeitshcrift fur Neuropsycholgie*, **19**(2), 101–115.

[14] Munte, T. F., Altenmuller, E., & Jancke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews: Neuroscience*, **3**(6), 473–478, http://dx.doi.org/10.1038/nrn843.

[15] Patel, A. D., & Iversen, J. R. (2007). The linguistic benefits of musical abilities. *Trends in Cognitive Sciences*, **11**(9), 369–372.

[16] Tervaniemi, M., Kruck, S., De Baene, W., Schroger, E., Alter, K., & Friederici, A. D. (2009). Top-down modulation of auditory processing: Effects of sound context, musical expertise and attentional focus. *The European Journal of Neuroscience*, **30**(8), 1636–1642, http://dx.doi.org/10.1111/j.1460-9568.2009.06955.x.

[17] Zatorre, R. J., & Baum, S. R. (2012). Musical melody and speech intonation: Singing a different tune. *PLoS Biology*, **10**(7), e1001372. doi:10.1371/journal.pbio.1001372

[18] Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, **6**(1), 37–46.

[19] Zatorre, R. J., & Gandour, J. T. (2008). Neural specializations for speech and pitch: moving beyond the dichotomies. Philosophical Transactions of the Royal Society of London. *Series B: Biological Sciences*, **363**(1493), 1087-1104. doi: J412P80575385013 [pii] 10.1098/rstb.2007.2161.

[20] Chandrasekaran, B., & Kraus, N. (2010). *The scalp-recorded brainstem response to speech: neural origins and plasticity. Psychophysiology*, **47**(2), 236-246. doi: PSYP928 [pii] 10.1111/j.1469-8986.2009.00928.x

[21] Maddieson, I. (1978). Universals of tone. In J. H. Greenberg (Ed.), *Universals of human language* (Vol. 2, pp. 335-365). Stanford, CA: Stanford University Press.

[22] Yip, M. (2003). Tone. New York: Cambridge University Press

[23] Abramson, A. S. (1962). *The vowels and tones of standard Thai: Acoustical measurements and experiments*. Bloomington: Indiana U. Research Center in Anthropology, Folklore, and Linguistics, Pub. 20.

[24] Gandour, JT. (1994). Phonetics of tone. In: Asher, R.; Simpson, J., editors. *The encyclopedia of language & linguistics*. New York: Pergamon Press; p. 3116-3123.

[25] Xu, Y. (2001). Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication*, **33**, 319-337.

[26] Griffiths, T. D., Warren, J. D., Scott, S. K., Nelken, I., & King, A. J. (2004). Cortical processing of complex sound: A way forward? *Trends in Neurosciences*, **27**(4), 181–185.

[27] Krishnan, A., & Gandour, J. T. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain and Language*, **110**(3), 135-148. doi: S0093-934X(09)00042-X [pii] 10.1016/j.bandl.2009.03.005

[28] Krumbholz, K., Patterson, R.D., Seither-Preisler, A., Lammertmann, C.,& Lutkenhoner, B. (2003). Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cerebral Cortex*, **13**(7), 765–772.

[29] Seither-Preisler, A., Patterson, R., Krumbholz, K., Seither, S., & Lutkenhoner, B. (2006). Evidence of pitch processing in the N100m component of the auditory evoked field. *Hearing Research*, **213**(1-2), 88–98.

[30] Griffiths, T.D., Kumar, S., Sedley,W., Nourski, K.V., Kawasaki, H.,Oya, H.,et al. (2010). Direct recordings of pitch responses from human auditory cortex. *Current Biology*, **20**(12),1128–1132.

[31] Schonwiesner, M., & Zatorre, R.J. (2008). Depth electrode recordings show double dissociation between pitch processing in lateral Heschl's gyrus and sound onset processing in medial Heschl's gyrus. *Experimental Brain Research*, **187**(1), 97–105.

[32] Krishnan, A., Bidelman, G. M., Smalt, C. J., Ananthakrishnan, S., & Gandour, J. T. (2012). Relationship between brainstem, cortical and behavioral measures relevant to pitch salience in humans. *Neuropsychologia*, **50**(12), 2849-2859. doi: 10.1016/j.neuropsychologia.2012.08.013

[33] Krishnan, A. (2007). Human frequency following response. In R.F. Burkard, M. Don & J. J. Eggermont (Eds.), *Auditory evoked potentials: Basic principles and clinical application* (pp. 313-335). Baltimore: Lippincott Williams & Wilkins.

[34] Krishnan, A., Xu, Y., Gandour, J.T., & Cariani, P. (2004). Human frequency-following response: representation of pitch contours in Chinese tones. *Hearing Research,* **189**:1–12.

[35] Krishnan, A., Xu, Y., Gandour, J. T., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, **25**, 161-168.

[36] Krishnan, A. (2002). Human frequency following response: Representation of steady-state vowels. *Hearing Research*, **166**,192-201.

[37] Krishnan, A., & Plack, C.J. (2011). Neural encoding in the human brainstem relevant to the pitch of complex tones. *Hearing Research*, **275**:110–119.

[38] Krishnan, A., Bidelman, G. M., & Gandour, J. T. (2010). Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. *Hearing Research*, **268**(1-2), 60-66. doi: 10.1016/j.heares.2010.04.016

[39] Krishnan A., Gandour, J.T., Bidelman, G.M., & Swaminathan, J. (2009). Experience-dependent neural representation of dynamic pitch in the brainstem. *Neuroreport*, **20**:408–413.

[40] Krishnan, A., Swaminathan, J., & Gandour, J. T. (2009). Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *Journal of Cognitive Neuroscience*, **21**(6), 1092-1105. doi: 10.1162/jocn.2009.21077

[41] Swaminathan, J., Krishnan, A., & Gandour, J.T.(2008). Pitch encoding in speech and nonspeech contexts in the human auditory brainstem. *Neuroreport*, **19**:1163–1167.

[42] Bidelman, G.M., & Krishnan, A. (2011). Brainstem correlates of behavioral and compositional preferences of musical harmony. *Neuroreport*, **22**:212–216.

[43] Gockel, H.E., Carlyon, R.P., Mehta, A., and Plack, C.J. (2011). The frequency following response (FFR) may reflect pitch-bearing information but is not a direct representation of pitch. *Journal of the American Association for Research in Otolaryngology*, **12**,767-782

[44] Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010b). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics*, **23**(1), 81-95. doi: 10.1016/j.jneuroling.2009.09.001

[45] Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010a). Brainstem pitch representation in native speakers of Mandarin is less susceptible to degradation of stimulus temporal regularity. *Brain Research*, **1313**, 124-133. doi10.1016/j.brainres.2009.11.061

[46] Krishnan, A., Gandour, J. T., Smalt, C. J., & Bidelman, G. M. (2010). Language-dependent pitch encoding advantage in the brainstem is not limited to acceleration rates that occur in natural speech. *Brain and Language*, **114**(3), 193-198. doi: 10.1016/j.bandl.2010.05.004

[47] Xu, Y., Krishnan, A., & Gandour, J.T. (2006). Specificity of experience-dependent pitch representation in the brainstem. *Neuroreport*, **17**:1601–1605.

[48] Bidelman, G.M., Gandour, J.T., & Krishnan, A. (2011a). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*. **23**:425–434.

[49] Bidelman, G.M., Gandour, J.T., & Krishnan, A. (2011b). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition*. **77**:1–10.

[50] Bidelman, G.M., Gandour, J.T., & Krishnan, A. (2011c). Musicians demonstrate experience-dependent brainstem enhancement of musical scale features within continuously gliding pitch. *Neuroscience Letters*. **503**:203–207.

[51] Alku, P., Sivonen, P., Palomaki, K., & Tiitinen, H. (2001). The periodic structure of vowel sounds is reflected in human electromagnetic brain responses. *Neuroscience Letters*, **298**(1), 25–28.

[52] Gutschalk, A., Patterson, R. D., Scherg, M., Uppenkamp, S., & Rupp, A. (2004). Temporal dynamics of pitch in human auditory cortex. *Neuroimage*, **22**, 755-766.

[53] Lutkenhoner, B., Seither-Preisler, A., & Seither, S. (2006). Piano tones evoke stronger magnetic fields than pure tones or noise, both in musicians and non-musicians. *Neuroimage*, **30**(3), 927–937.

[54] Soeta, Y., & Nakagawa, S. (2008). The effects of pitch and pitch strength on an auditory evoked N1m. *Neuroreport*, **19**(7), 783–787.

[55] Yrttiaho, S., Tiitinen, H., May, P.J., Leino, S., & Alku, P. (2008). Cortical sensitivity to periodicity of speech sounds. *Journal of the Acoustical Society of America*, **123**(4), 2191–2199.

[56] Chait, M., Poeppel, D., & Simon, J. Z. (2006). Neural response correlates of detection of monaurally and binaurally created pitches in humans. *Cerebral Cortex* (New York, NY: 1991), **16**(6), 835–848.

[57] Ritter, S., GunterDosch, H., Specht, H. J., & Rupp, A.(2005). Neuromagnetic responses reflect the temporal pitch change of regular interval sounds. *Neuroimage*, **27**(3), 533–543.

[58] Gutschalk, A., Patterson, R.D., Rupp, A., Uppenkamp, S., & Scherg, M. (2002). Sustained magnetic fields reveal separate sites for sound level and temporal regularity in human auditory cortex. *Neuroimage*, **15**(1), 207–216.

[59] Griffiths, T.D., Kumar, S., Sedley,W., Nourski, K.V., Kawasaki, H.,Oya, H.,et al. (2010). Direct recordings of pitch responses from human auditory cortex. *Current Biology*, **20**(12), 1128–1132.

[60] Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature*, **436**(7054), 1161–1165.

[61] Johnsrude, I.S., Penhune,V.B., & Zatorre, R.J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain*, **123**(Part1), 155–163.

[62] Penagos, H., Melcher, J.R., & Oxenham, A.J. (2004). A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *Journal of Neuroscience*, **24**(30), 6810–6815.

[63] Krishnan, A., Gandour, J. T., Ananthakrishnan, S., & Vijayaraghavan, V. (2014a). Cortical pitch response components index stimulus onset/offset and dynamic features of pitch contours. *Neuropsychologia*, 59C, 1-12.

[64] Krishnan, A., Gandour, J. T., Ananthakrishnan, S., & Vijayaraghavan, V. (2015). Language experience enhances early cortical pitch-dependent responses. *Journal of Neurolinguistics*, **33**, ,128-148.

[65] Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication*, **41**(1), 245–255.

[66] Poeppel, D., Idsardi, W.J., & van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. Philosophical Transactions of the Royal Society of London. *Series B: Biological Sciences*, **363**(1493).

[67] Schneider, P., Scherg, M., Dosch, H.G., Specht, H.J., Gutschalk, A. & Rupp, A. (2002) Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience*, **5**, 688–694.

[68] Itoh, K., Okumiya-Kanke, Y., Nakayama, Y., Kwee, I.L., & Nakada,T. (2012). Effects of musical training on the early auditory cortical representation of pitch transitions as indexed by change-N1. *European Journal of Neuroscience*, **36**(11), 3580–3592.

[69] Hari, R., Pelizzone, M., Makela, J.P., Hallstrom, J., Leinonen, L. & Lounasmaa, O.V. (1987). Neuromagnetic responses of the human auditory cortex to on-and offsets of noise bursts. *Audiology*, **26**(1), 31–43.

[70] Lutkenhoner, B., & Steinstrater, O. (1998). High-precision neuromagnetic study of the functional organization of the human auditory cortex. *Audiology and Neuro-Otology*, **3**, 191–213.

[71] Cariani, P.A., & Delgutte, B. (1996a). Neural correlates of the pitch of complex tones. I.Pitch and pitch salience. *Journal of Neurophysiology*, **76**(3), 1698–1716.

[72] Cariani, P.A., & Delgutte, B. (1996b). Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. *Journal of Neurophysiology*, **76**(3), 1717–1734.

[73] Cedolin, L., & Delgutte, B. (2005). Pitch of complex tones: Rate-place and interspike interval representations in the auditory nerve. *Journal of Neurophysiology*, **94**(1), 347–362.

[74] Lu, T., Liang, L., & Wang, X. (2001). Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nature Neuroscience*, **4**(11), 1131-1138. doi:10.1038/nn737

[75] Steinschneider, M., Reser, D. H., Fishman, Y. I., Schroeder, C. E., & Arezzo, J. C. (1998). Click train encoding in primary auditory cortex of the awake monkey: evidence for two mechanisms subserving pitch perception. *Journal of the Acoustical Society of America*, **104**(5), 2935-2955.

[76] Walker, K. M., Bizley, J. K., King, A. J., & Schnupp, J. W. (2011). Cortical encoding of pitch: recent results and open questions. *Hearing Research*, **271**(1-2), 74-87. doi: 10.1016/j.heares.2010.04.015

[77] Micheyl, C., Schrater, P.R., & Oxenham, A.J. (2013). Auditory frequency and intensity discrimination explained using a cortical population rate code. *PLoS Computational Biology*, **9**(11):e1003336. doi: 10, 1371/journal.pcbi.1003336.

[78] Gilbert, C. D., & Sigman, M. (2007). Brain states: top-down influences in sensory processing. *Neuron*, **54**(5), 677-696. doi: 10.1016/j.neuron.2007.05.019

[79] Fritz, J., Shamma, S., Elhilali, M., & Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience*, **6**(11), 1216-1223. doi: 10.1038/nn1141

[80] Lee, C. C., & Middlebrooks, J. C. (2011). Auditory cortex spatial sensitivity sharpens during task performance. *Nature Neuroscience*, **14**(1), 108-114. doi: 10.1038/nn.2713

[81] Weinberger, N. M. (2011). Reconceptualizing the primary auditory cortex: learning, memory and specific plasticity. In J. A. Winer & C. E. Schreiner (Eds.), *The auditory cortex* (pp. 465-491). New York: Springer

[82] Polley, D. B., Steinberg, E. E., & Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. *Journal of Neuroscience*, **26**(18), 4970-4982.

[83] Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., & Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioural Brain Research*, **156**(1), 95-103.

[84] Song, J. H., Skoe, E., Wong, P. C. M., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience*, **20**, 1892-1902.

[85] Bidelman, G. M., & Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *Journal of Neuroscience*, **29**(42), 13165-13171.

[86] Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the United States of America*, **104**(40), 15894-15898. doi: 0701498104 [pii] 10.1073/pnas.0701498104

[87] Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, **10**, 420-422.

[88] Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, **8**(10), 457-464. doi: 10.1016/j.tics.2004.08.01 1S1364-6613(04)00215-3 [pii]

[89] Nahum, M., Nelken, I., & Ahissar, M. (2008). Low-level information and high-level perception: the case of speech in noise. *PLoS Biology*, **6**(5), e126. doi: 07-PLBI-RA-2244 [pii]10.1371/journal .pbio.0060126

[90] Krizman, J., Skoe, E., Marian, V., & Kraus, N. (2014). Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling. *Brain and Language*, **128**(1), 34-40. doi: 10.1016/j.bandl.2013.11.006

[91] Banai, K., Abrams, D., & Kraus, N. (2007). Sensory-based learning disability: Insights from brainstem processing of speech sounds. *International Journal of Audiology*, **46**(9), 524-532. doi: 781872134 [pii] 10.1080/

[92] Chechik, G., Anderson, M. J., Bar-Yosef, O., Young, E. D., Tishby, N., & Nelken, I. (2006). Reduction of information redundancy in the ascending auditory pathway. *Neuron*, **51**(3), 359-368. doi: 10.1016/j.neuron.2006.06.030

[93] Warren, J. D., & Griffiths, T. D. (2003). Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *Journal of Neuroscience*, **23**(13), 5799-5804.

[94] Winer, J. A., Miller, L. M., Lee, C. C., & Schreiner, C. E. (2005). Auditory thalamocortical transformation: structure and function. *Trends in Neuroscience*, **28**(5), 255-263 doi: 10.1016/j.tins. 2005.03.009

[95] Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, **11**(10), 946-953.

[96] Xiong, Y., Zhang, Y., & Yan, J. (2009). The neurobiology of sound-specific auditory plasticity: A core neural circuit. *Neuroscience and Biobehavioral Reviews*, **33**(8), 1178-1184. doi: doi:10.1016/j.neubiorev.2008.10.006

[97] Bajo, V. M., Nodal, F. R., Moore, D. R., & King, A. J. (2010). The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nature Neuroscience*, **13**(2), 253-260. doi: 10.1038/nn.2466

[98] Keuroghlian, A. S., & Knudsen, E. I. (2007). Adaptive auditory plasticity in developing and adult animals. *Progress in Neurobiology*, **82**(3), 109-121. doi: S0301-0082(07)00073-1[pii] 10.1016/j.pneurobio.2007.03.005

[99] Kral, A., & Eggermont, J. J. (2007). What's to lose and what's to learn: development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Research Reviews*, **56**(1), 259-269. doi: S0165-0173(07)00187-7 [pii]10.1016/j. brainresrev.2007.07.021

[100] Dean, I., Robinson, B. L., Harper, N. S., & McAlpine, D. (2008). Rapid neural adaptation to sound level statistics. *Journal of Neuroscience*, **28**(25), 6430-6438. doi: 28/25/6430 [pii] 10.1523/JNEUROSCI.0470-08.2008

METHODS FOR ASSESSING THE QUALITY OF TRANSMITTED SPEECH AND OF SPEECH COMMUNICATION SERVICES

Friedemann Köster, Sebastian Möller, Jan-Niklas Antons, Sebastian Arndt, Dennis Guse, Benjamin Weiss
Quality and Usability Lab, Telekom Innovation Laboratories, Technische Universität Berlin, Berlin, Germany
{friedemann.koester, sebastian.moeller, jan-niklas.antons, sebastian.arndt, dennis.guse, benjamin.weiss}@tu-berlin.de

The quality of transmitted speech is the major indicator for telecommunication providers to classify their services. As a result, the assessment of quality is of high scientific and economic importance, and corresponding methods for assessing the quality of transmitted speech have been in the focus of multiple studies in the past. In this contribution, traditional methods are reviewed, their weaknesses are identified and improvements and extensions are proposed. The presented work covers subjective, diagnostic, instrumental, and physiological methods, as well as possibilities for evaluating long-term quality aspects.

INTRODUCTION

The use of telephony for vocal human-to-human communication is deeply integrated into our daily life. The technological development made telephony even more useful, e.g., with the emergence of mobile phones or Voice-over-IP. However, creating and maintaining a reliably working telephony system has become an even more complex task; e.g., applying heterogeneous end-user devices and network infrastructure as well as transcarrier-interconnectivity.

Within a telephony service, a speech signal can be degraded while recording, coding, transmission, decoding, and reproduction. Here, the perspective of the end-users, i.e., customers, is of major interest: Understanding how end-users perceive and experience degradations allows for improving telephony service and reacting efficiently to occurring issues like network overload.

The *Quality of Experience* (QoE) "is the degree of delight or annoyance of the user of an application or service. It results from the fulfilment of his or her expectations with respect to the utility and / or enjoyment of the application or service in the light of the user's personality and current state." [1]. This definition of QoE and the development of methodologies to assess the subjectively perceived QoE is the objective of the European Network on quality of Experience in Multimedia Systems and Services (Qualinet COST IC1003). This is a collaboration of European QoE experts and some of the topics addressed in this article were part of this collaboration. Applied to telephony services, assessing and predicting QoE represents one of the major goals of current research. In order to understand QoE, experiments involving human participants are required. Typical methods for QoE assessment of transmitted speech are listening-only and conversational tests (with or without a task), where test subjects experience a stimulus and judge the quality. Such subjective studies enable to understand, if and how degradations are perceived under the presented condition and tasks of the study [2]. Quantitative feedback is often gathered using 5-point Absolute Category Rating (ACR) scale. The quantity evaluated from the scores is represented by the mean opinion score (MOS). A typical ACR listening-only scale resulting in a MOS (therefore also called MOS scale) can be seen in Figure 1 [3].

Quality of the speech:

excellent	good	fair	poor	bad
5	4	3	2	1

Figure 1: ACR Listening-quality scale according to [3].

For speech quality assessment usually short stimuli with a duration of up to a few seconds are used, which are sufficiently long to evaluate the quality related to codecs or network impairments. These types of subjective tests exhibit, however, certain inherent limitations:

- Stimuli must be carefully selected, so that all necessary conditions are covered while the amount of stimuli is still manageable.
- The effort of preparing and conducting the studies is significant.
- Only quantitative quality feedback is available.

The information collected during subjective studies forms the basis for the design of instrumental, so-called objective quality prediction models. A major complication is the fact that given subjective results are often only valid under the given conditions of the study in which they were collected (e.g. set of stimuli, choice of task etc.).

In the following, we present current work on the evaluation of speech quality that aims at overcoming the limitations outlined above. Such approaches include diagnostic methods, physiological measurements, and approaches to study temporal effects during the presentation of a single stimulus and over multiple distinct usage episodes. The ultimate goal is to include insights gathered by such recent approaches into traditional objective speech prediction models in order to extend their validity and therefore improve the accuracy of such models.

DIAGNOSING THE QUALITY OF TRANSMITTED SPEECH

Traditional subjective tests only provide little insight into the reason of quality impairments. More precisely, two different speech stimuli can both be rated with the same MOS value, while, for example, one is degraded by background noise and the other one due to clipping. Thus, the MOS score does not provide any diagnostic information regarding the cause of a quality impairment. Two important observations have been exploited in the design of approaches that allow for revealing the cause of a quality reduction:

- Naïve listeners identify perceptual dimensions related to impairments, such as degraded sound-color, noisiness or continuity
- Experts identify technical causes of the transmission channel, which result in impairments like sub-optimum speech-level, speechspectrum or noise-level

Two methodologies have been used to identify perceptual dimensions of transmitted speech that base on the first observation above: (1) scaling perceptual differences of pairwise presented stimuli, and then mapping the perceptual distance to a *multidimensional space* (MDS) [4]; or (2) rating all stimuli independently on a set of bipolar scales (*Semantic Differential*, SD [5]) and reducing the space of judgments with the help of a factor analysis (*Principal Component Analysis*, PCA). Applying both methodologies to narrowband (300-3400 Hz) and wideband (50-7000 Hz) transmitted speech stimuli allowed for identifying three perceptual dimensions of the quality judgments: coloration, noisiness, and discontinuity [6]. The results are illustrated in Figure 2. A fourth dimension representing loudness was added in a later study. In [7], it was shown that the identified perceptual dimensions can be quantified directly in a subjective test.

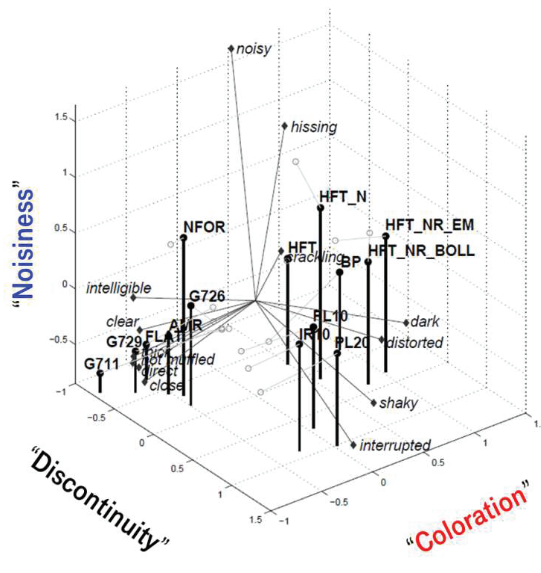


Figure 2: Results of the experiments for identifying the perceptual dimensions of transmitted speech [7].

The identification and estimation of perceptual dimensions is a current work item of the *International Telecommunication Union* (ITU-T) (working title P.AMD, *Assessment of Multiple Dimensions*). The instrumental estimation of perceptual dimensions will be discussed further in the next section.

The second observation mentioned above can be exploited as follows: First, expert-listeners identify the most dominant types of degradations ("Impairment type", Level 1, e.g. "Speech-level", "Speech-spectrum", etc., refer to Table 1) and rate them according to the three categories highly dominant, dominant, or less dominant. Afterwards, experts identify the detailed properties of each of the dominant degradations ("Degradation", Level 2, e.g. "Loud speech" or "Quiet speech" corresponding to "Speech-level", etc.) [8]. The experts can choose from the list of technical causes given in [9] where a total of 47 different impairments on Level 2, grouped into 9 categories on Level 1, are provided. Experts are asked to judge only those samples that received a MOS rating of 3.0 or lower in the subjective scores. The identification of technical causes is another work item of ITU-T with the working title P.TCA (*Technical Causes Analysis*).

Table 1: Extract of the P.TCA guidelines given in [9].

Impairment type (Level 1)	Degradation (Level 2)
Speech - level	*Loud speech*
	Quiet speech
	Loudness varies
	Speech level fluctuations
	Temporal speech clipping
	Choppy speech
	Self-clipping
	Speech cut-outs
Speech - spectrum	*Timbre varies*
	Muffled speech
	Sharp speech
	Coloured speech
Noise - level	*Line sounds dead*
	Loud noise
	Noise level fluctuations
	Temporal noise clipping
	Noise cut-outs

Obviously, there are links between the technical causes and the perceptual dimensions. In a first study, the results of a P.TCA annotation experiment have been analyzed with respect to the reliability of the annotations, as well as with respect to the relationships between technical causes, perceptual dimensions, and overall quality [10]. The results showed that there is a need for all - P.TCA cause analysis, P.AMD perceptual dimension analysis, and overall MOS scores - as these three metrics are only partly correlated and thus contain complementary information.

All of above mentioned approaches concentrate on a "passive" listening-only situation. Other phases appearing during the usage of a communication service, like speaking (impaired by e.g. side tone or echo) and interacting (impaired by e.g. delay) are not addressed. Concerning this, the following approach is currently under investigation [11]: The perceptual dimensions of the listening phase described above add up with the dimensions related to the speaking phase and with the dimensions related to the interacting phase. In initial studies, separate tests for the speaking and the interacting phases have been conducted, resulting in two dimensions for the speaking phase and in one dimension for the interacting phase [12]. These dimensions will have to be validated in subsequent studies, and until now, they have not proven to be orthogonal to each other. Subjective methods have to be developed for this target, which would allow for combining all phases of a conversation in one experimental paradigm.

INSTRUMENTAL ESTIMATION

Conducting subjective experiments to evaluate a telecommunication service is very time consuming, especially when large numbers of participants are required. Therefore, the need for instrumental (or so-called "objective") estimation models has grown over the past years. A variety of approaches are useful for estimating subjective ratings of the quality of transmitted speech. They can be divided into three groups based on the input information that they require:

1. Intrusive signal-based models; these models compare the input and output signal of a transmission channel and map the signal differences to a predicted rating, using a perceptual weighting
2. Non-intrusive signal-based models; these models rely only on the (degraded) output signal of a transmission channel and map signal characteristics to a predicted rating
3. Parametric models; these models use the information of a parametric description of the elements (e.g., Send Loudness Rating, Talker Echo Loudness Rating, or Roundtrip delay) of the transmission channel and map it to subjective ratings

A large amount of intrusive signal-based models have been developed in order to estimate the overall quality of the listening situation in a laboratory environment. The long-term standard of the ITU-T had been the so-called PESQ (*Perceptual Estimation of Speech Quality*) [13] model, which has been replaced by its successor POLQA (*Perceptual Objective Listening Quality Assessment*) [14]. These two models proved to be reliable for the estimation of the overall quality but do not address diagnostic features.

Estimators that represent the perceptual effects of certain system components (e.g. filters for coloration) were described in [15] in order to provide diagnostic information for the perceptual dimensions from [7]. These estimators have been further improved resulting in the DIAL (*Diagnostic Instrumental Assessment of Listening quality*) model [16]. This intrusive model combines the dimension estimators with a predictor for the overall quality in order to provide reliable instrumental assessment of both the overall quality and its corresponding perceptual dimensions. A block-diagram describing the DIAL model is depicted in Figure 3.

Furthermore, the DIAL model and the POLQA model were analyzed in terms of quality degradation indicators related to the perceptual dimensions [17]. More precisely, it was shown that indicators extracted from the two algorithms (e.g., *ERB* (Equivalent Rectangular Bandwidth), L_n (Noise Loudness), or *LTL* (Long-Term Loudness)) can be used to predict the subjective ratings of the perceptual dimensions. These results are intended to be used in the P.AMD project to develop a model to predict subjective ratings for the four perceptual dimensions coloration, noisiness, loudness and discontinuity.

The non-degraded input signal of a transmission channel is usually not available outside of the laboratory environment. Therefore, it is demanded by the industry being able to estimate the quality of transmitted speech on the basis of the degraded output signal alone. The ITU-T recommends its standard P.563 [18] for this purpose, which estimates an overall MOS of the transmitted speech quality. It provides reliable but not as highly correlated results as the intrusive models do. This is comprehensible, as the information carried by the input signal is not available. A model that estimates the perceptual dimension from [7] is currently under study so as to provide diagnostic information with a non-intrusive method. The approach is similar to that of the DIAL model, except that only the output signal is used. A non-intrusive estimator for each perceptual dimension is therefore required. While a first estimator for the perceptual dimension nosiness has already been part of a study [19] and showed good results, estimators for the other three dimensions coloration, discontinuity, and loudness are still under development.

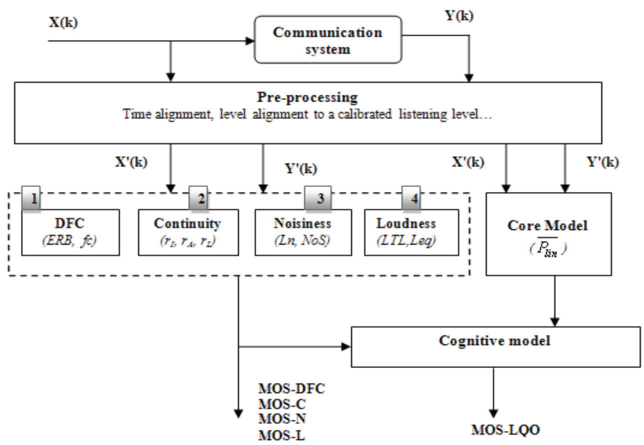


Figure 3: Overview of the DIAL model [17].

Before a transmission channel is installed, it is interesting to know during the planning phase what level of quality of the transmitted speech can be expected. For this purpose, so-called parametric models are used, which do not depend on the speech signals. These parametric models use a set of parameters that define each element of a transmission channel from the talker to the listener (e.g., loudness ratings, echo, and codecs). A prominent example of a parametric model is the E-Model, a network planning tool for the prediction of conversational and listening speech quality. The E-model is recommended by the ITU-T for narrowband and wideband network planning [20,21]. For the diagnostic information, the aforementioned three dimensions (discontinuity, noisiness, coloration) can also be estimated with a parametric approach similar to the E-model. Using parametric estimations of the perceptual dimensions, a dimension-based version of the E-model was developed, called the DNC (*Discontinuity, Nosiness, Coloration*) model [7].

The presented models all facilitate the assessment of the quality of transmitted speech and some of them also provide diagnostic information of the estimated quality. However, all of the presented models refer only to the listening situation, except for the E-Model, so that certain relevant properties of a communication channel are not covered. In [22] an intrusive model which combines listening, talking, and conversational features was developed for the estimation of the conversational speech quality. It estimates a quality value for each phase, and then maps the three values to asses an overall conversational quality. It does not provide any diagnostic information to its user, however.

PHYSIOLOGICAL QUALITY ASSESSMENT

Standard subjective tests lack information on the cognitive state of the test participant, and any physiological responses due to the presented stimuli. However, these tests currently build the basis for quality estimation algorithms. When physiological responses due to quality variations in the presented signal are better understood, they could enrich and improve current models significantly.

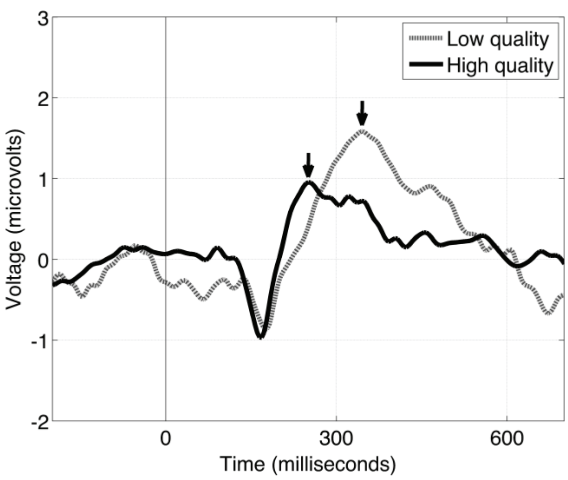


Figure 4: Exemplary grand average across all subjects of ERP plots for one stimulus in high quality and one stimulus in low quality at channel CPz. Reverberation was implemented as quality impairment of the low quality stimulus (reverberation time = 1500 ms). Arrows denote maximum amplitude (P300 peak for the low quality stimulus).

Brain activity measures are one methodology to better understand the cognitive processes underlying quality perception. Therefore, *electroencephalography* (EEG) has been introduced to the research area of speech QoE. EEG measures voltages on the participants' scalp. Using EEG, two basic analysing techniques can be considered – *eventrelated potentials* (ERPs) and *analysis of frequency bands*. ERPs are neural responses due to external events, such as presented audio signals. Measures of ERPs were used to confirm subjective results, as the amplitude of a positive deflection of the ERP after about 300ms of stimulus onset [23], P300, was shown to be gradually bigger for stronger degraded stimuli (see Figure 4). Furthermore, the P300 was shown in some cases to be even more sensitive than the behavioural answer, as in some trials subjects did not report a distortion on the behavioural level, but a similar EEG pattern as in cases where subjects did report a distortion was detected [24].

Frequency analysis is another possibility to analyse EEG signals. The obtained frequencies can be divided into several sub-bands, which provide information on the cognitive state of the listener. In several studies, it could be shown that participants tended to be more fatigued when listening to low-quality audio signals. This was detected when analysing the alpha and theta frequency power band, which are indicative for fatigue and drowsiness [25].

Another brain-based method is *functional near-infrared spectroscopy* (fNIRS), which is based on (de-)oxygenated blood flow and builds on neuroimaging techniques [26]. fNIRS was used in a study investigating the quality of synthetic speech. Significant correlations between subjective ratings and obtained fNIRS features were found [27]. So far, no study recording fNIRS using natural transmitted speech has been performed.

Similarly to brain-based measures, peripheral measures such as galvanic skin response, eye movement features, or heart-rate variability are of interest. To the authors' knowledge, only little research has been performed on peripheral physiology in the area of speech QoE, but would definitely be desirable to better understand the physiological processes underlying the speech quality perception and judgment.

LONG-TERM-EVALUATION

For longer usage periods of speech services, spanning for example over minutes or hours, or for repeated usage over weeks and months, subjective assessment methods can be used similarly to the ones for shorter samples. However, there are two major challenges with such longer time frames. Firstly, information on the variability in service quality is lost. Secondly, such retrospective judgements have proven to be far more difficult to estimate by instrumental means.

Continuous assessment methods have been used in (quasi-) laboratory conditions in order to address the first challenge mentioned above. Sliders were used to collect quality ratings continuously, applying, e.g., the *Continuous Evaluation of Time Varying Speech Quality* (CETVSQ) [28] similarly like in the experiment on noise and loudness perception [29,30]. Here, a MOS scale [3] is recommended to assist the usage of the sliders. At the end of each sequence, a retrospective overall quality rating is asked for on the same scale. An alternative is segmenting longer stimuli into smaller units and applying established assessment methods for short samples as discussed above. This was done, for example, to simulate conversational structures with alternating listening-only and talking phases in order to develop an estimation model for long-term speech quality [31,32].

Regarding the second challenge, a rich body of results indicating some kind of weighted average relationship between instantaneous or short-term ratings with retrospective, episodic judgments has been collected. The unweighted average tends to be too optimistic for estimating retrospective overall quality. Instead, extreme and longer degradations, as well as degradations with temporal proximity to the retrospective rating have to be weighted stronger. Refer to [33] for more information and estimation models.

Recent studies on repeated usage, such as regular calls every day, have provided controversial results. Applying weighting procedures as for episodic quality has failed so far at improving modelling performance [34]. A limitation that is inherent to all long-term methods is the required time effort. Only a limited number of conditions and stimuli can therefore be included. Hence, conditions must be very well selected.

CONCLUSIONS

In this contribution, traditional as well as new subjective and objective methods for assessing the quality of transmitted speech have been presented and discussed. Although speech quality is a well researched field, there still remains a vast number of open questions. This is partly also due to the fact that the technological infrastructure changes rapidly.

On the subjective side, special experiments can provide diagnostic information by extending them with the assessment of the perceptual dimensions noisiness, coloration, discontinuity, and loudness. Further research towards new assessment paradigms and the identification of additional conversational dimensions has to be conducted for being able to diagnose a whole conversational process that involves listening and talking.

Multiple intrusive and parametric approaches are available and provide reliable estimations regarding the objective estimation of the quality of transmitted speech. However, the need for nonintrusive models is still strong for monitoring purposes. This is also the case for models covering all conversational aspects. These models have to rely on new subjective paradigms, though.

Physiological measures are a valid test method for the assessment of speech based on neuronal states, and it could be a good complement to standard subjective tests in the future. This will particularly be of importance for high-quality media assessment in which the cognitive state of consumers is of interest.

Future work on long-term evaluation methods will provide insights into how quality evolves over longer episodes, as well as over multiple episodes. This will complement well-known short-term effects and thus enable telecommunication service providers to optimize their service for longer usage periods.

Studies combining long-term evaluation and physiological measures can build a foundation for better understanding the cognitive consequences of low-quality speech. This has already been successfully introduced in [25]. Finally, these consequences could be incorporated into objective metrics.

BIBLIOGRAPHY

[1] S. Möller, P. Le Callet, and A. Perkis, "Qualinet White Paper on Definitions of Quality of Experience," COST Action IC 1003, Lausanne, 1.1 edn., 2012.

[2] A. Raake and S. Egger, "Quality and Quality of Experience," in *Quality of Experience: Advanced Concepts, Applications, Methods*. Heidelberg: Springer, 2014, pp. 11 – 34.

[3] ITU-T Recommandation P.800, *Methods for Subjective Determination of Transmission Quality.* Geneva: International Telecommunication Union, 1996.

[4] I. Borg and P. Groenen, *Modern Multidimensional Scaling - Theory and Applications*, 2nd, Ed. New York, NY: Springer Series in Statistics, 2005.

[5] C. Osgood, The *Measurement of Meaning*. Urbana, IL: University of Illinois Press, 1957.

[6] M. Wältermann, A. Raake, and S. Möller, "Quality Dimensions of Narrowband and Wideband Speech Transmission," , 2010, pp. 1090 – 1103.

[7] M. Wältermann, *Dimension-based Quality Modeling of Transmitted Speech.* Berlin: Springer, 2012.

[8] ITU-T Temporary Document TD 686 (GEN/12), *Expert Listening for P.TCA*.: International Telecommunication Union; Rapporteur Q.16/12 (L. Malfait), 2011.

[9] ITU-T Temporary Document TD 650rev1 (GEN/12), *Requirement Specifications for P.TCA (Technical Cause Analysis)*.: International Telecommunication Union; Rapporteur Q.16/12 (L. Malfait), 2011.

[10] S. Möller, F. Köster, J. Skowronek, and F. Schiffner, "Analyzing Technical Causes and Perceptual Dimensions for Diagnosing the Quality of Transmitted Speech.," in *Proc. 4th International Workshop on Perceptual Quality of Systems (PQS 2013)*, Vienna, AT, 2013, pp. 30 – 35.

[11] F. Köster and S. Möller, "Towards a New Test Paradigm for the Subjective Quality Assessment of Conversational Speech," in DAGA, Meran, IT, 2013, pp. 440 – 443.

[12] F. Köster and S. Möller, "Analyzing Perceptual Dimensions of Conversational Speech Quality," in *Interspeech 2014,* Singapore, 2014, pp. 2041 – 2045.

[13] ITU-T Recommendation P.862.2, *Wideband Extension to Recommendation P.862 for the Assessment of Wideband Telephone Networks and Speech Codecs*. Geneva: International Telecommunication Union, 2007.

[14] ITU-T Recommendation P.863, *Perceptual Objective Listening Quality Assessment.* Geneva: International Telecommunication Union, 2011.

[15] K. Scholz, *Instrumentelle Qualitätsbeurteilung von Telefonbandsprache beruhend auf Qualitätsattributen*. Kiel: Shaker Verlag, 2008.

[16] N. *Côté, Integral and Diagnostic Intrusive Prediction of Speech Quality.* Berlin: Springer, 2011.

[17] S. Tiémounou, R.Le Bouquin Jeannès, and V. Barriac, "On the identification of relevant degradation indicators in super wiedeband listening quality assessment models," *Speech Communication,* vol. 55, no. 10, pp. 1047 – 1063, November – December 2013.

[18] ITU-T Recommendation P.563, *Single-ended method for objective speech quality assessment in narrow-band telephony applications.* Geneva: International Telecommunication Union, 2004.

[19] F. Köster, S. Möller, and G. Mittag, "Referenzfreie Schätzung der perzeptuellen Dimension Rauschhaftigkeit von übertragener Sprache," in DAGA, Oldenburg, 2014, pp. 501 - 502.

[20] ITU-T Reccomandation G.107, The E-model: a *Computational Model for Use in Transmission Planning.* Geneva: International Telecommunication Union, 2011.

[21] ITU-T Recommendation G.107.1, *Wideband E-model*. Geneva: International Telecommunication Union, 2011.
[22] M. Guéguin, R. Le Bouquin-Jeannès, V. Gautier-Turbin, G. Faucon, and V. Barriac, *On the Evaluation of the Conversational Speech Quality in Telecommunications*.: EURASIP J.Adv. Signal Process, 2008.
[23] C. Duncan, R. Barry, J. Connolly, C. Fischer, P. Michie, R. Näätänen, J. Polich, I. Reinvang and C. Petten, "Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400," in *Clinical Neurophysiol.*, vol. 120, 2009, pp. 1883 – 1903.
[24] J.-N. Antons, R. Schleicher, S. Arndt, S. Möller, A.K. Porbadnigk and G. Curio, "Analyzing Speech Quality Perception using Electro-Encephalography," in *Journal of Selected Topics in Signal Processing*. IEEE, 2012, pp. 721 – 731.
[25] J.-N. Antons, F. Köster, S. Arndt, S. Möller, and R. Schleicher, "Changes of vigilance caused by varying bit rate conditions," in *IEEE Int. Workshop on Quality of Multimedia Experience (QoMEX*). IEEE, 2013, pp. 148 – 151.
[26] A. Villringer and U. Dirnafl, "Coupling of brain activity and cerebral blood flow: basis of functional neuroimaging," in *Cerebrovasc. Brain Metab. Rev. 7*, 1995, pp. 240 – 276.
[27] R. Gupta, K. Laghari, S. Arndt, R. Schleicher, S. Möller, D. O'Shaughnessy and T.H. Falk, "Using fNIRS to Characterize Human Perception of TTS System Quality, Comprehension, Fluency: Preliminary Findings.," in *4th International Workshop on Perceptual Quality of Systems (PQS)*, 2013, pp. 73 – 78.
[28] ITU-T Recommendation P.880, *Continuous evaluation of time varying speech quality*. Geneva: International Telecommunication Union, 2004.
[29] S. Kuwano, S. Namba, and Y. Nakajima, "On the noisiness of steady state and intermittent noises," *Journal of Sound and Vibration,* vol. 72, no. 1, pp. 87 – 96, 1980.
[30] H. Fastl, "Evaluation and measurement of perceived average loudness," *5th Oldenburg Symposium on Psychological Acoustics,* 1991.
[31] B. Weiss, S. Möller, A. Raake, J. Berger, and R. Ullmann, "Modeling Call Quality for Time-Varying Transmission Characteristics Using Simulated Conversational Structures," *Acta Acustica united with Acustica*, vol. 95, no. 6, pp. 1140 – 1151, 2009.
[32] B. Lewcio, *Management of Speech and Video Telephony Quality in Heterogeneous Wireless Networks*, Dissertation ed. Berlin: Fakultät für Elektrotechnik und Informatik, Technische Universität Berlin, 2013.
[33] B. Weiss, D. Guse, S. Möller, A. Raake, A. Borowiak and U. Reiter, "Temporal Development of Quality of Experience," in *Quality of Experience*.: Springer International Publishing, 2014, pp. 135 – 147.
[34] D. Guse and S. Möller, "Macro-Temporal Development of QoE: Impact of Varying Performance on QoE over Multiple Interactions," in *DAGA 2013,* Meran, IT, pp. 452 – 455.

DETECTION OF SHARKS WITH THE GEMINI IMAGING SONAR

Miles J.G. Parsons[1], Iain M. Parnum[1], Kim Allen[2], Robert D. McCauley[1] and Christine Erbe[1]
[1]Centre for Marine Science and Technology, Curtin University, Perth, WA, Australia
[2]Intellipulse, Perth, Western Australia

Limiting environmental impacts of marine industrial operations and mitigating hazardous encounters between humans and marine fauna have become increasingly important as anthropogenic activity expands. To this end, significant effort has been made to develop sonar imaging of fauna and to increase detection and identification ranges. A Tritech Gemini imaging sonar was used to observe sharks of 1.4 to 2.7 m length, at ranges from 1 to 50 m, in various water depths ≤15 m. Within 5 m, shark shape, length and swimming action were readily discernible. However, as range increased, knowledge of movement patterns was required to discriminate a 'shark-like' object, before the shark became purely an acoustic target at greater ranges, where visual confirmation of the target was necessary for identification. Once the seafloor is ensonified by the acoustic beam, seafloor backscatter can dominate the image and mask shark detection. The results presented concur with other active acoustic detection studies that, for a given frequency and noise level, maximum detection and identification ranges are reliant on system source level, beam pattern, bathymetry, and target size and acoustic reflectivity.

INTRODUCTION

During typical fisheries or marine mammal surveys, traditional multi-beam sonar systems have been increasingly used for detecting, visualising and quantifying marine fauna in the waters below or to the side of a vessel [1-4]. In addition, anthropogenic marine activity over the past few decades has promoted the need for detection and identification of fauna within potentially hazardous areas, sometimes 24 hours a day. Partly as a result, multi-beam sonar systems are being modified and developed, producing forward-looking imaging sonars.

Studies on acoustic backscatter from teleosts (bony fishes) are abundant [e.g. 5-10], but when it comes to marine megafauna, while a significant amount of work has been conducted on marine mammals [11-14], there are few reports regarding the reflectance or acoustic imaging of elasmobranchs (sharks and rays). In 1970, Harden Jones [15] detected basking sharks (*Cetorhinus maximus*) at ranges of up to 180 m on a sector scanning sonar and Lieber et al. [16] furthered this work off Scotland, to begin looking at *C. maximus* ecology, using a Reson 7128 multi-beam sonar. However, these studies were conducted in deep waters, observing sharks of several metres in length. By contrast, in shallow waters, high-frequency acoustic cameras have been used in aquaria to image sharks and rays at ranges of less than 5 m [17, 18].

This study aimed to use a Tritech Gemini sonar, a low-power imaging sonar system, to investigate the detection and identification of sharks in shallow (<15 m) waters, similar to those of the beaches around the WA coastline, to gain a better understanding of the ranges at which a shark may be imaged with this system.

METHODS

The Gemini 720i 300M (Tritech, UK) system operates at 720 kHz, with 120° horizontal and 20° vertical beamwidths, and an elevation of -10°. Across the horizontal 120° the system comprises 256 dynamically focussed beams with effective azimuth-angular beam resolution of 0.5°. Along beam resolution is range setting dependent, but can be as high as 8 mm. In various depths ≤15 m, a pole-mounted Gemini was positioned 0.5-1.0 m below the water surface (Figure 1).

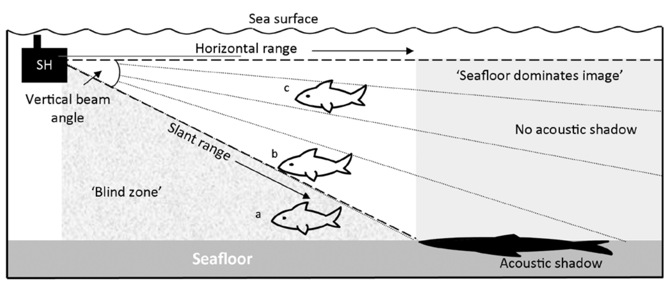


Figure 1. Schematic of Gemini imaging sonar beam and examples of sharks at similar range from the sonar head, but at varying depths that create different acoustic images, including a) outside the acoustic beam and in the 'blind zone', b) towards the lower region of the acoustic beam producing a weaker image of the shark, but a strong acoustic shadow and c) high enough in the water column to produce a strong image of the shark from backscatter in which case any shadow would be outside the designated range of the sonar. Horizontal and slant range are indicated and the "Blind zone" and region where "Seafloor dominates the image" are shown by the mottled and grey regions, respectively.

Theoretically, a target's range and position in the water column have considerable impact as to whether it is located within the sonar beam (Figure 1, conditions a, b and c). With decreasing horizontal range, it is increasingly likely that the target will be below the acoustic beam and therefore not ensonified or detected (Figure 1, condition a). Once within the beam the target will reflect acoustic backscatter to the receiver and targets at shorter slant range than the seafloor provide the highest signal-to-noise ratio (SNR). If the target is within the acoustic beam and at shorter slant range than the first contact of the acoustic beam with the seafloor, one expects not only some reflected backscatter, but also an acoustic shadow of the target (Figure 1, condition b). Those high in the acoustic beam create an acoustic shadow that is outside the range-setting of the sonar (Figure 1, condition c). As the return from the seafloor is high, in comparison with that of a mid-water target, once part of the acoustic beam insonifies the seafloor, the sonar image is dominated by backscatter from the seafloor, and any target at the same slant range will be difficult to discern from the reverberation. Therefore, the optimum system performance is when the target's range is less than that of the seafloor. Some systems are able to mitigate these issues in real-time or post-processing, by removing accumulated backscatter over a designated time to account for stationary objects. These "movement filters" are designed to remove background noise and highlight moving targets. However, the resulting images often lose resolution in the moving target and the process is non-trivial if either the system is moving (even minor movements relating to wave patterns or surge) or there is significant noise e.g. cavitation from waves, vessels or animal movement. In this study this removal of background noise was not used as the sonar head was not completely stationary.

The Gemini was deployed at the following locations around Australia:

- In waters off the Gold Coast, Queensland, carcasses of a recently captured 1.8 m bull shark (*Carcharhinus leucas*) and a 2.7 m great white shark (*Carcharodon carcharias*) were suspended at a depth of 3 m from plastic floats (one at the head and one at the tail), using detergent-covered monofilament fishing line. To remove air from body cavities and denticles that would have previously been filled by water or mucus before the shark was removed from the water, each shark was flushed with a deck hose, lowered, tail first into the water and bubbles allowed to escape. It was then brushed down and briefly dragged through the water. The sharks then drifted in the water column at two locations (7.5 and 15 m water depth) while being imaged using the Gemini system at ranges between 5 and 50 m. Imaging of the floats alone was also conducted at ranges of up to 25 m to ensure that these did not contribute to the sonar images of the shark carcasses. The seafloor in this area comprised a coastal sand substrate (based on visual classification).
- In Ocean Park Aquarium, Shark Bay, Western Australia, in a 3.5 m deep, 30 m diameter pond, images of 2.4, 2.0 and 1.7 m lemon sharks (*Negaprion brevirostris*), a 1.2 m nervous shark (*Carcharhinus cautus*) and a 1.5 m sandbar shark (*Carcharhinus plumbeus*) were collected as they swam past a near-stationary sonar system. The bottom of the pond comprised a concrete base, covered by a thin layer of fine sand.
- In the Eastern Gulf of Shark Bay, around channels that are approximately 7 m deep, the Shark Bay Ecosystem Research Project (SBERP) has run a shark tagging program for over 15 years. The Gemini system was pole-mounted on the starboard side of the Department of Parks and Wildlife *RV Sirenia II* and directed athwartships towards the SBERP vessel as a tagged shark was released from its port side. This occurred at a range of approximately 25 m.

In each case, system settings of range and gain were varied where possible, in an attempt to visually attain the optimum SNR for the intended targets. The Gemini system does not record the raw signal, but as a series of individual images, which can be reviewed as moving images, therefore analysis of the SNR for each situation was only through visual assessment of the colourbar.

All experiments were performed according to the Australian Code of Practice for the care and use of animals for scientific purposes.

RESULTS

All three locations provided images on the Gemini sonar and interesting information on the detection of sharks. In each case, an acoustic target could be detected at horizontal ranges up to the point where the acoustic beam encountered the seafloor (Figure 1). Therefore, the maximum detection range, in these tests, was a function of the beam pattern and water depth, rather than the system source level or acoustic reflectance of the shark. Unfortunately, given the logistical and time constraints, the maximum depth in which the study could be conducted was 7 and 15 m for live and deceased sharks, respectively, thus the maximum depth-independent detection range could not be tested. However, the study areas did reflect conditions of a significant number of shark encounters along the coast of Western Australia.

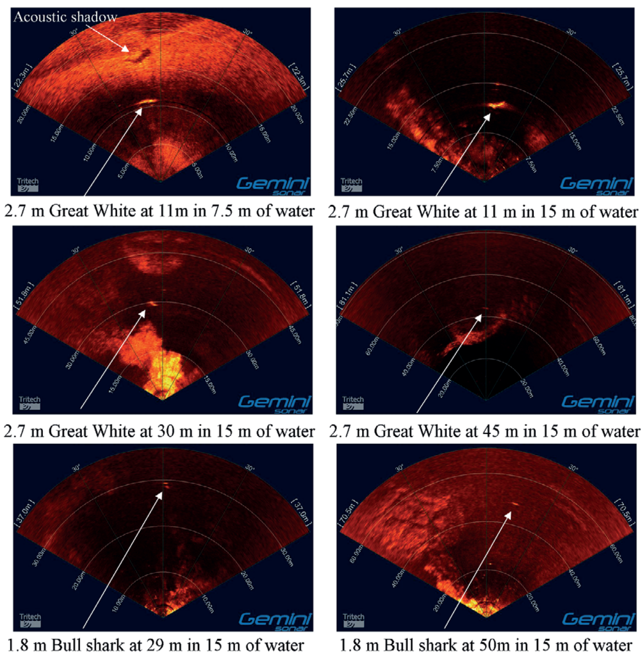


Figure 2. Screen shots from the Tritech Gemini imaging sonar system of a 1.8 m bull shark (*Carcharhinus leucas*) carcass and a 2.7 m great white shark (*Carcharodon carcharias*) carcass at various ranges, suspended at 3 m depth, in 7.5 and 15 m of water. Various gain settings were used to produce these images and the brighter responses represent a strong acoustic return.

The shark carcasses provided immobile targets of known location and therefore a more stable platform with which surveys could be conducted for a matter of minutes, rather than seconds. Both bull and great white carcasses were tested to a range of 50 m (Figure 2 bottom right images) and in both cases targets were discernible, though realistically only as a 'large blob'. In both water depths, even at ranges of 10-15 m (the closest range at which the Gemini was tested) the sharks were not always discernible as a shark-like object, but an acoustic target of length similar to that of the shark (Figure 2, top row). This was unlikely due to the orientation of the shark as each shark was imaged from all sides. The top left image in Figure 2 illustrates how the seafloor backscatter dominates the image, as well as the acoustic shadow of the shark against the seafloor.

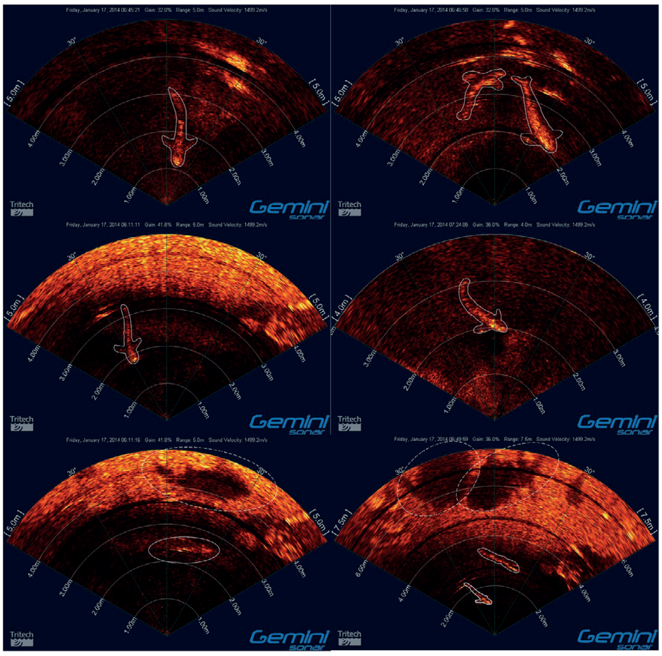


Figure 3. Screen shots from the Tritech Gemini imaging sonar system of a 1.7 and a 2.2 m lemon shark (*Negaprion brevirostris*) in 3.5 m depth at Ocean Park Aquarium, Shark Bay. Outlines of the backscatter from the sharks are shown by the continuous lines, and outlines of the acoustic shadows by the dashed lines.

The Ocean Park images (Figure 3) illustrated that at ranges <5 m it is possible to discern the shape of the shark. Once again, this was limited by the depth of the lagoon, thus in deeper water it may be possible to produce shark-like images at ranges in excess of 10 m. The swimming actions of the sharks, compared with those of fish, were clearly apparent through time and illustrated that at close range, the entire body of the shark could be imaged under certain conditions. Sharp turns and sudden movements of the sharks resulted in loss of image, but also often generated acoustically visible cavitation and vortices.

The release of a shark in the Eastern Gulf of Shark Bay highlighted that in approximately 7 m of water the live sharks could be discerned to at least 25 m (Figure 4). Though only observed briefly, no swimming pattern could be seen and, similar to the carcass, the shark merely an elongated acoustic target of >2 m length. In the case of the shark in Figure 4, an accompanying cobia (*Rachycentron canadum*) happened to be alongside the shark at the time of release. The difference between the elasmobranch and the swim-bladdered fish (cobia) was visible, though not at all times, in that the length of the shark target was mostly greater than that of the cobia. This release also highlighted the issue of the position in the water column as the shark (and its accompanying cobia) only remained in the beam for a few seconds before diving. Both animals presumably dived to the seafloor and into the "blind zone", beneath the acoustic beam, which would occur at ranges less than approximately 19 m.

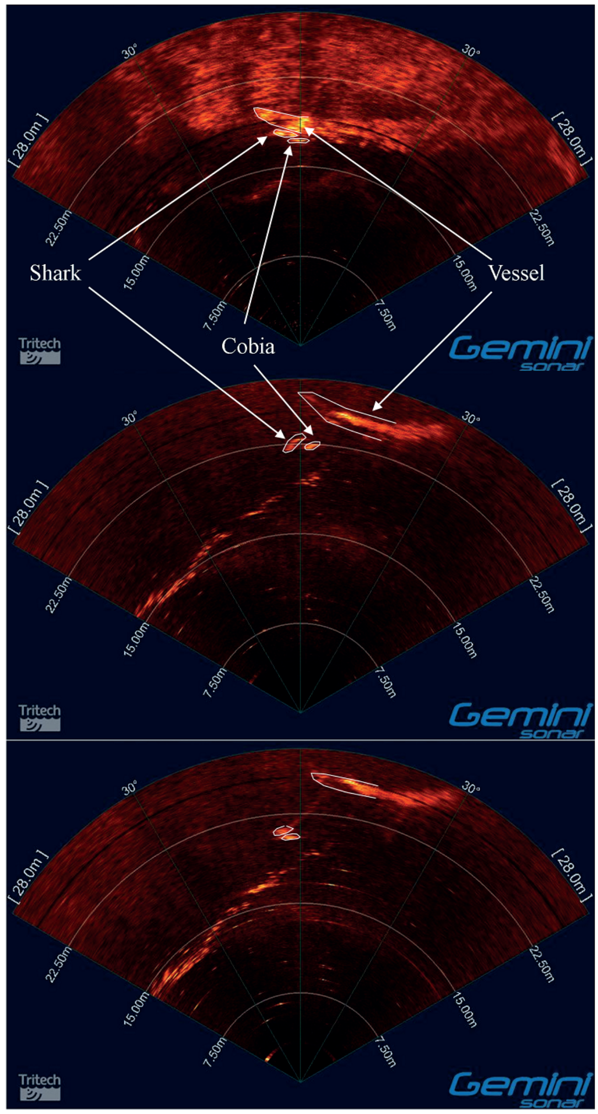


Figure 4. Screen shots from the Tritech Gemini imaging sonar system of a 2.2 m tiger shark (*Galeocerdo cuvier*) and its accompanying cobia (*Rachycentron canadum*) as the shark was released from a vessel at a range of 25 m.

DISCUSSION

This study has provided some estimates of likely ranges at which a shark could be detected with a 'low-power', wide-beam imaging sonar, such as the Gemini trialled in this study, under conditions similar to those of the WA coastline. Lieber et al. [16] used a Reson 7128 imaging sonar to detect 5-10 m length basking sharks (*Cetorhinus maximus*) at ranges of over 100 m and produced images where pectoral fins and the thunniform swimming action were visible in excess of 30 m from the transducer. These ranges are significantly longer than those observed here. However, there are significant differences between the two studies that are important to the development of a system for detecting sharks along WA beaches. Contributors to the differences in images with this study were, not only the size of the sharks (5-10 m, compared with the 1.8 m bull shark and 2.7 m great white in this study), but also the power of the system (220 dB for the Reson, compared with 160 dB for the Gemini used here), and importantly the water depth (>20 m compared with around 10 m in this study) and the resolution of the two sonar systems. In addition, the lagoon at Ocean Park comprises a rigid bottom that reflects significant backscatter and therefore noise on the sonar image. By contrast with performance, the advantage of the Gemini system in the possible detection of sharks off the coast is that it is a "low-power" system that can achieve the same results operating from 12 V battery power and thus could easily be deployed remotely.

This study has also conceptually highlighted several of the issues associated with sonar detection and identification of sharks in shallow water, not least of which is the need for an appropriate beam pattern. If positioned near the sea surface or seafloor the 20° vertical beamwidth of the Gemini system (with its offest of -10°) ensonifies the seafloor (or sea surface, respectively) at ranges of around three times the depth of water below the sonar head. At these ranges, discrimination of mid-water targets becomes problematic, particularly if the sonar head is mobile, as seafloor backscatter changes and persistent contributions to the image cannot be easily removed. To increase the detection range significantly would require a reduction in vertical beamwidth to increase the range at which the seafloor is ensonified. In contrast, too small a vertical beamwidth would only ensonify a small portion of the water column at close ranges (for example, a 1° beam ensonifies approximately 1 m of water column at a horizontal range of 50 m). Thus a compromise is required to detect targets at ranges more suitable for mitigating encounters between sharks and humans or excluding them from a hazardous area. Alternatively a vertical array of narrow beam systems would increase vertical coverage, though at ranges where these beams (or sidelobes) converge, issues of interference would increase. However, a vertical array in shallow waters (<15 m) may not provide suitable benefits at ranges of greater than 75 m due to the limited separation available.

The following is therefore suggested for testing as one possible, easy-to-deploy and integrate method of detecting sharks in shallow water. The sonar system would be positioned near to the seafloor, with enough altitude to prevent the bottom edge of the acoustic beam from contacting the seafloor. The sonar system would be dual frequency with different acoustic beam patterns for each frequency, designed to cover the short range and long range separately. The beamwidth either side of the sonar for both beam patterns would be as wide as possible (120-150°). The short range (e.g. 5-30 m, depending on water depth) would be covered by a high frequency beam, >700 kHz, with a larger vertical beamwidth (10-15°) that provides high resolution imagery over ranges of up to approximately three times the water depth and quickly includes a large percentage of the water column (Figure 5). To detect targets in the long range would require a lower frequency (400 kHz) beam of finer vertical beamwidth (3-5°) to ensure that neither seafloor or water surface are ensonified until at ranges of >10 times the water depth. The sonar would be set such that these two beams ping alternately each at 1 Hz ping rate. Thus any targets in the short range and the long range are being ensonified once each second. However, these beam patterns would leave a volume of water that is not ensonified by the long-range beam, and saturated by reverberation in the short-range beam (Figure 5), thus an ideal set requires investigation. A vertical array of sonars could limit such inadequately sampled volumes of water, but this cost at the cost of greater power requirements to operate them. Additionally, as the bathymetry of beaches can vary significantly it may be necessary to fine-tune these beam patterns to maximise detection ranges for a particular beach. Therefore a system would be required where beam patterns of the system can be scripted by the user. This process of beam steering is becoming more common in the use of multi-beam systems, though to the authors knowledge few systems exist with multiple frequency beams of differing beamwidth and with beam patterns and steering suitable to this application.

Lucifredi and Stein [19] designed a monostatic seafloor mounted system with an electronically steered vertical line array and 60 element receiver array to track gray whales (*Eschrichtius robustus*) in 70-100 m of water, off the California coast. This was a well-designed research tool, capable of vertical steering of a 6° vertical beam, providing good long-range detection with considerable contributions to automated detection. As very strong scatterers, gray whales, or pods of gray whales could be detected in excess 400 m, in these depths, however, interactions with surface, bottom and kelp were still present.

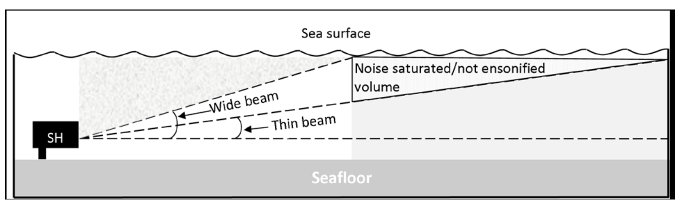


Figure 5. Schematic of the effective volumes of water ensonified by using two beam patterns of differing vertical beamwidth. The acoustic blind zone where no water is ensonified by either beam is shown to the left of the figure in mottled grey, while the volume of water only ensonified by the wider beam and at ranges where the response is likely to be dominated by the backscatter from the sea surface is shown in the hatched grey area to the right of the figure. Thus the area in white illustrates the volume of water where targets would be detected by both beams and the area in grey on the right hand side of the figure illustrates the volume of water in which targets are only likely to be detect by the thinner acoustic beam.

The study has also found that at large ranges, where the target is only covered by a few of the acoustic beams, while a shark may not be imaged, it can sometimes be discriminated from smaller targets, such as individual fish. Similar to teleosts, the reflectance by sharks is very stochastic, however, even at range when swimming motion is not discernible, and the shark appears simply as an acoustic target, the swimming may produce regular oscillations in target strength. The real-time monitoring of sharks at the Ocean Park Aquarium implied that this could be possible, but requires further testing for ranges useful to shark detection. One of the next steps in assessing the possible performance of sonar systems in detecting, identifying and tracking sharks involves identifying frequency and length dependent target strength (and its variation) and to verify computer models of the effect of varying vertical beamwidth.

ACKNOWLEDGEMENTS

The authors would like to thank Richard Rickett from Seatools, Western Australia, for supply of the Gemini sonar system and Mal Perry from CMST for fieldwork and equipment preparation. The Shark Bay Ecosystem Research Project and Department of Parks and Wildlife provided amazing logistical help, as always. Ocean Park Aquarium kindly provided access to the shark lagoon and logistical help, for which the authors are very grateful. This study was financially supported by the Western Australian Government through the Applied Research Program. The Western Australian Government does not endorse any information, product, process or outcome, arising from or in relation to this study

REFERENCES

[1] F.S.G. Gerlotto, M. Soria and P. Freon, "From two dimensions to three: the use of multibeam sonar for a new approach in fisheries acoustics", *Canadian Journal of Fisheries and Aquatic Sciences* **56**, 6-12 (1999)

[2] F.S.G. Gerlotto and E.K. Eriksen, "The application of multibeam sonar technology for quantitative estimates of fish density in shallow water acoustic surveys", *Aquatic Living Resources* **13**, 385-93 (2000)

[3] F.S.G. Gerlotto and J. Paramo, "The three-dimensional morphology and internal structure of clupeid schools as observed using vertical-scanning, multibeam sonar", *Aquatic Living Resouces* **16**, 113-122 (2003)

[4] M.J.G. Parsons, I.M. Parnum, and R.D. McCauley, "Visualizing Samsonfish (*Seriola hippos*) with a Reson 7125 Seabat multibeam sonar", *ICES Journal of Marine Science* **70** 665–674 (2013)

[5] J.E. Simmonds and D.M. MacLennan, Fisheries Acoustics, Theory and Practice, 2nd edn, Blackwell Science, Oxford. pp. 437 (2005)

[6] K.G. Foote, "Importance of the swimbladder in acoustic scattering by fish: a comparison of gadoid and mackerel target strengths", *Journal of the Acoustical Society of America* **67**, 2084-9 (1980)

[7] K. G. Foote and E. Ona, "Tilt angles of schooling penned saithe". *Journal du Conseil International pour l'Exploration de la Mer* **43**, 118-121 (1987)

[8] C.S. Clay, and J.K. Horne, "Acoustic models of fish: the Atlantic cod (*Gadus morhua*)", *Journal of the Acoustical Society of America* **96**, 1661-8 (1994)

[9] J.K. Horne, "Acoustic approaches to remote species identification: a review", *Fisheries Oceanography* **9**, 356-71 (2000)

[10] O. Nakken and K. Olsen, "Target-strength measurements of fish", *Rapports et Procès-Verbaux des Reunions du Conseil International pour l'Exploration de la Mer* **170**, 52-69 (1977)

[11] W.W.L Au, "Acoustic reflectivity of a dolphin", *Journal of the Acoustical Society of America* **99**, 3844-3848 (1996)

[12] J.L. Dunn, "Airborne measurements of the acoustic characteristics of a sperm whale", *Journal of the Acoustical Society of America* **46**, 1052-1054 (1969)

[13] A. Jochens, D. Biggs, K. Benoit-Bird, D. Engelhardt, J. Gordon, C. Hu, N. Jaquet, M. Johnson, R. Leben, B. Mate, P. Miller, J. Ortega-Ortiz, A. Thode, P. Tyack and B. Würsig, "Sperm Whale Seismic Study in the Gulf of Mexico (Contract No. 1435-01-02-CA-85186): US Department of the Interior, Minerals Management Service", 1 pp. (2008)

[14] R.H. Love, "Target strengths of humpback whales *Megaptera novaeangliae*", *Journal of the Acoustical Society of America* **54**, 1312-1315 (1973)

[15] F.R. Harden Jones, "Tail beat frequency, amplitude, and swimming speed of a shark tracked by sector scanning sonar", *ICES Journal of Marine Science* **35** 95-97, (1973)

[16] L.Lieber, B. Williamson, C.S. Jones, L.R. Noble, A. Brierley, P. Miller and B.E. Scott, "Introducing novel uses of multibeam sonar to study basking sharks in the light of marine renewable energy extraction", *Proceedings of the 2nd International Conference on Environmental Interactions of Marine Renewable Energy Technologies,* (EIMR2014), 28 April – 02 May, Stornoway, Isle of Lewis, Outer Hebrides, Scotland (2014)

[17] G.T. Kellison, J.R. Luo, S.R. Frias-Torres and J. Serafy, "Electronics in the prop-roots: application of multi-beam sonar and stereo video for fish community data collection in mangroves", *American Fisheries Society,* 137th Annual Meeting San Francisco, September 2-6, (2007)

[18] http://www.mickpeterson.org/Classes/Design/2011_12/Groups/Acou_cam/index.html

[19] I. Lucifredi and P.J. Stein, "Gray whale target strength measurements and the analysis of the backscattered response", *Journal of the Acoustical Society of America* **121** 1383 (2007)

A MATLAB TOOL FOR THE CHARACTERISATION OF RECORDED UNDERWATER SOUND (CHORUS)

Alexander N. Gavrilov[1], Miles J.G. Parsons[1]
[1]Centre for Marine Science and Technology, Curtin University, GPO Box U1987, Perth, WA 6845, Australia
a.gavrilov@cmst.curtin.edu.au

The advent of low-cost, high-quality underwater sound recording systems has greatly increased the acquisition of large (multi-GB) acoustic datasets that can span from hours to several months in length. The task of scrutinizing such datasets to detect points of interest can be laborious, thus the ability to view large portions of the dataset in a single screen, or apply a level of automation to find or select individual sounds is required. A toolbox that can be continually revised, the user-friendly "Characterisation Of Recorded Underwater Sound" (CHORUS) Matlab graphic user interface, was designed for processing such datasets, isolating signals, quantifying calibrated received levels and visually teasing out long and short-term variations in the noise spectrum. A function to automatically detect, count and measure particular signals (e.g. blue whale sounds) is integrated in the toolbox, with the ability to include categorised calls of other marine fauna in the future. Sunrise and sunset times can be displayed in long-term average spectrograms of sea noise to reveal diurnal cycles in the vocal activity of marine fauna. A number of example studies are discussed where the toolbox has been used for analysing biological, natural physical and anthropogenic sounds.

INTRODUCTION

The underwater "soundscape" comprises the combination of sounds of natural physical, biological and anthropogenic origin at a given location, and offers significant information on the local environment and its vocal inhabitants. Reports of soniferous behaviour by species, previously undocumented as vocal, are published consistently, highlighting how much can be learned from underwater passive acoustic observations, but also that there is still much more to discover by such means. In Australian waters, for example, cetaceans [1-4], fish [5-8], invertebrates [9-11], to name a few, all contribute to the soundscape. Along the coast, physical contributions to the soundscape are dominated by the sound of breaking waves or surf. Rain also generates consistent broadband noise during periods of heavy rainfalls. Anthropogenic activities, whether vessel use [12], offshore oil and gas exploration [13], construction and operation [14], or general horseplay [15] generate intense impulsive and tonal underwater sounds.

Natural and man-made underwater sounds can be impulsive or continuous signals with broad or narrow frequency bands and can contain frequency components from fractions of 1 Hz to several hundred kHz. They can occur individually, or so frequently that together they mask all other sounds [16]. The quantification of sound's acoustic characteristics, such as the sound source level, peak frequency and duration [8, 16, 17], can help understand the relationship between marine fauna and their acoustic environment, and the impacts anthropogenic sound has on that relationship. Periodic increases in sound pressure levels (SPLs) from biotic sounds can be indicative of multiple animals present, such as the choruses of vocal fish [18, 19] and whales [20]. These SPLs, or individually counted calls can be a proxy for the abundance of the calling species in the area [3, 21, 22]. Thus temporal patterns in the spectrum level of sound production can offer information on the temporal presence at the recording location and, if associated with influential factors in the local environment (temperature, salinity, anthropogenic noise), provide information on the responses of vocal fauna to such drivers of presence and absence [18, 19]. It is the long-term comparisons of these contributors to local ambient noise, or soundscape, that offers us this information and is becoming more readily accessible to those that study it.

Advances in data processing and storage have facilitated the acquisition of vast acoustic datasets through an increasing number of recording platforms, such as the Song Meter from Wildlife Acoustics, AURAL-M2 from Multi-Electronique and MARU form Cornell University. A comprehensive review of these and many others underwater sound recorders is given in [23] along with appropriate references. This has increased the amount of data to be analysed from individual minutes-long, MB-large audio recordings, to hundreds of GB, multi-month datasets. Processing of such long-term datasets requires a different approach to previous scrutiny of short, single recordings and automated data-mining software packages could provide the opportunity to do this. However, the advent of accessibility of underwater noise data as a sole or complementary data source also means that it is likely to be studied by researchers who may not be familiar with the physics behind underwater acoustics and acoustic signal processing. Thus, any processing package needs to be intuitive to those with little experience in acoustics.

Interrogation of long-term datasets requires the ability to view the entire set as a whole or in part, magnify any segment of interest to quantify signals (biotic or abiotic), across any given frequency range, and data-mine the entire dataset for said signals. It should be able to count signals, and display their occurrence on a timescale marine fauna respond to (daylight, sunset etc.), figuratively illustrating diel, lunar, seasonal or annual patterns. Received levels (RMS and peak acoustic pressure and sound exposure levels) need to be quantified to classify these signals and trends. Currently, many software packages are available that can aid in scrutiny of individual, or short-term audio sets to analyse small numbers of signals or extract and save sounds (some examples being Audition from Adobe, Ishmael from CIMRS Bioacoustics Lab of Oregon State University and Raven from Cornell Lab of Ornithology). Some other software packages, such as PAMGUARD (http://www.pamguard.org/), Song Scope (http://www.wildlifeacoustics.com/products/song-scope-overview) and Acousonde (http://www.acousonde.com), provide means for automatic detection of particular sounds, primarily whale and dolphin sounds, in long-term sea noise recordings. While these packages have a number of advantages, they do not provide a full set of tools to quickly review multi-month datasets with respect to long-term changes in the sea noise spectrum and its biological or anthropogenic components and, at the same time, to comprehensively analyse acoustic characteristics of noise sources of particular interest. The aim of this study was to design, and continually develop an adaptable software tool to reflect all of these needs, in this case, a Matlab toolbox to aid the Characterisation Of Recorded Underwater Sound, “CHORUS”.

TOOLBOX DESCRIPTION

The CHORUS toolbox is a library of Matlab routines with a Graphic User’s Interface (GUI) designed at the Centre for Marine Science and Technology (CMST), Curtin University to simplify analysis of large (up to several hundred GB) sets of sea noise data. It is easy-to-use, even for researchers who are not proficient in Matlab programming. CHORUS is written as an open source code with a comprehensible structure, which allows an advanced Matlab user to add new modules (GUI features and functions) in a tree-like manner. It works with large sets of individual audio recordings repeated with a certain interval. Long continuous recordings can also be analysed with the toolbox. CHORUS accepts various formats of audio data, including Wave Audio Files (WAV) and binary files with and without headers and/or footers. For uncommon binary data formats, a special function should be written to export audio data into CHORUS. The complete process of acoustic data analysis is divided into two steps: (1) data pre-processing and (2) review of preliminary results with a comprehensive analysis of individual sea noise recordings, which is detailed in the following sections.

Data pre-processing

Preliminary processing of sea noise data involves computationally expensive and hence time consuming operations, which does not require the user’s presence and expedites the post-processing data analysis. The pre-processing routines sort all recordings in a set in recording time order and calculate the Power Spectrum Density (PSD) averaged for each recording section of length chosen by the user. If the dataset consists of a large number of recordings and individual recording are not too long, then the PSD averaging window is recommended to be of the length of the whole recording, which is a default setting. The PSD is calculated with a 1-Hz resolution from 0 Hz to the Nyquist frequency which is automatically determined from the sampling frequency of recording.

When starting data pre-processing, the user has to provide some data and select settings as follows:

Calibration data are needed to correct the resulting PSD of sea noise for the frequency response of the recording system. It can be provided in the following forms:

- Calibration file name (with full path) of a data file with a recording of a signal from a white noise generator input in series with the hydrophone and recorded in the same data format and with the same or higher sampling frequency as the raw data files with sea noise recordings. If such a calibration file exists and this option is chosen, then the user is asked to input the calibration noise level in dB re 1 V2/Hz;
- ASCII or .mat file with the frequency response of recording system in dB re data units/Vin, where the data units can be either bits or Volts representing the signal amplitude in the data files and Vin is the voltage at the hydrophone output. The frequency response data need to be provided up to at least the Nyquist frequency. If the frequency response in the calibration data is gridded with a coarser resolution than 1 Hz, it will be interpolated into a 1-Hz grid.
- If no calibration data are available, then the PSD will be calculated in relative units.
- Hydrophone sensitivity in dB re 1 V/µPa. This value is needed only if the frequency response data are available.
- Full path to the folder with raw data files.
- Length of signal sections (time window) in seconds to calculate PSD. If it is different from the length of each individual recording, then the section length is recommended to be such that the whole recording consists of an integer number of sections.
- Length of each output file in days with resulting PSD. This setting is required to limit the size of the output files with PSDs, which depends on frequency band of recording and PSD window length.

The output files contain matrices of PSD and vectors of the time and frequency grids used for calculation. If the calibration data and hydrophone sensitivity are provided, then the PSD data of sea noise are saved in µPa2/Hz.

Data review and analysis

The pre-processed data can be quickly reviewed and analysed using the CHORUS GUI toolbox. For analysis, both raw data files and pre-processed data with PSDs are needed. Before the toolbox opens its main GUI window, it asks the user to input the time zone in order to display all results in local time and the coordinates of sea noise recorder, if the sunrise and sunset times are needed to be displayed in the CHORUS graphs.

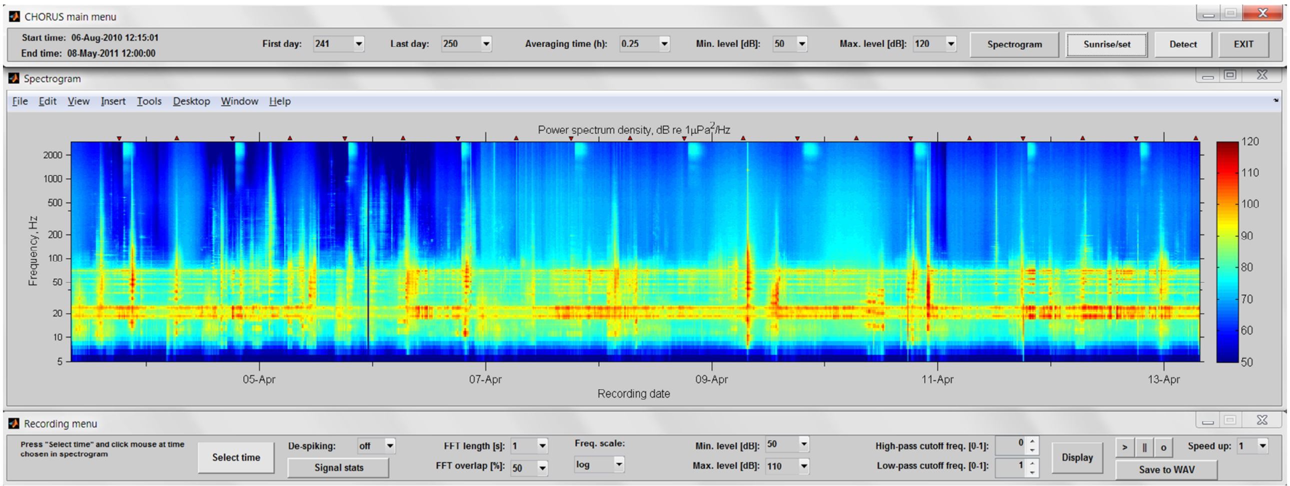


Figure 1: LTA spectrogram of sea noise recorded for 10 days compiled from PSDs of individual recordings. The red upward and downward triangles indicate the sunrise and sunset time respectively.

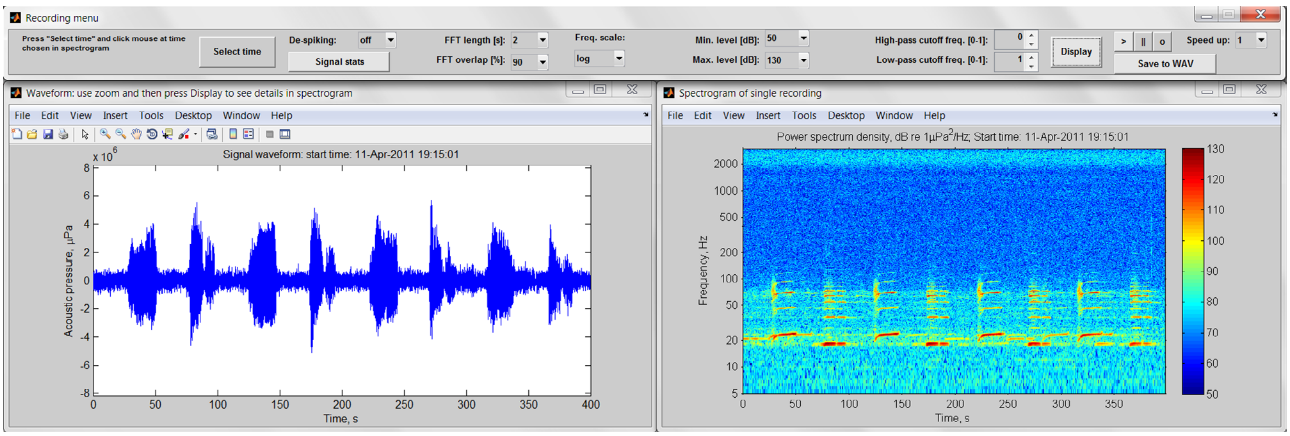


Figure 2: Recording menu (top panel) and waveform and spectrogram of selected recording (bottom left and right panels respectively), displaying blue whale calls.

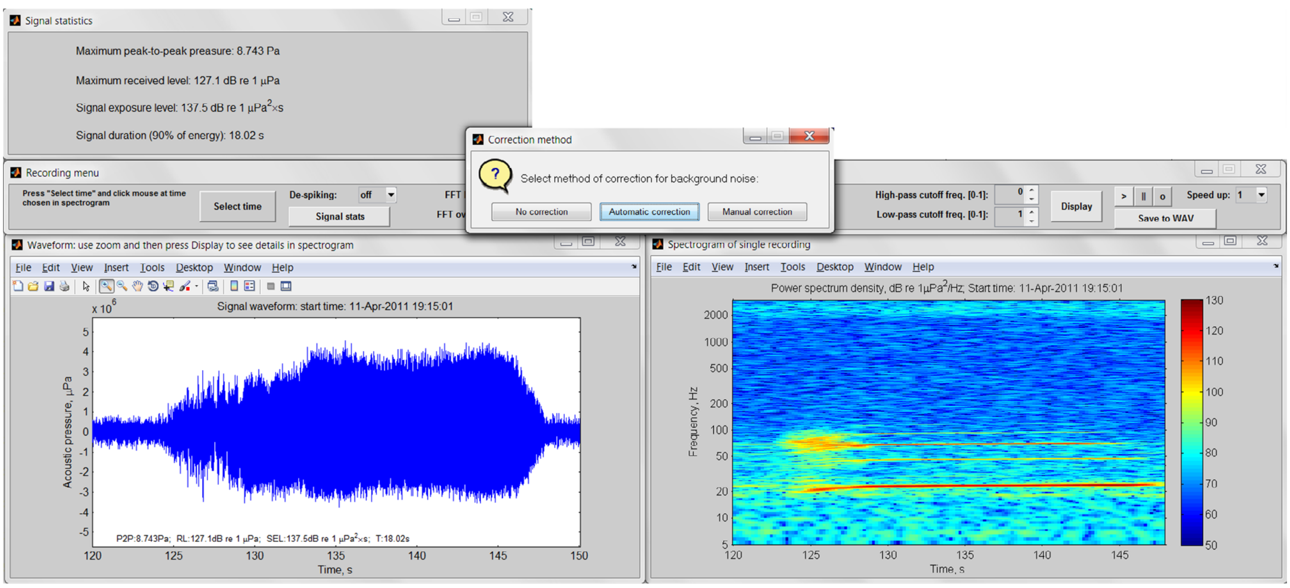


Figure 3: Waveform and spectrogram of a pygmy blue whale call localised in a sea noise recording (bottom panels), noise correction menu (middle panel) and signal statistics (top left panel)

The CHORUS GUI main menu appears at the top of the screen and displays the start and end dates and times of the whole dataset and user interface pop-up menus and push and radio buttons (Figure 1, top panel). The main menu controls parameters of a Long-Time Averaged (LTA) spectrogram to be displayed (middle panel in Figure 1). The LTA spectrogram is compiled from PSDs of individual recordings or signal sections chosen to calculate PSD data during pre-processing. It displays main spectral features of sea noise and their changes over a long time period of several days, which allows the user to quickly review the whole dataset of several months and identify the main contributors for a more detailed waveform and spectrum analysis. The user selects the first and last days of the time period from the dataset to be displayed in the LTA spectrogram. To display longer time periods, additional averaging of PSD may be needed by changing the averaging time in a pop-up menu. The averaging time can be chosen from one to four repetition periods of individual recordings (or length of section in a continuous recording chosen for PSD calculation). A logarithmic frequency scale is used to display the LTA spectrogram. The user can also adjust its colour-scale by selecting the minimum and maximum PSD levels in the corresponding pop-up menu. If coordinates of the recorder are entered at the program start, the sunrise and sunset time can be displayed by red triangle marks at the top of the LTA spectrogram. The times of new, full and quarter phases of the moon will also be displayed in the next version of the toolbox.

After pressing the “Spectrogram” button in the main menu, the LTA spectrogram is displayed with a Recording menu at the bottom of it (Figure 1, bottom panel). The Recording menu allows the user to select any time within the time period of the displayed LTA spectrogram and display the waveform and spectrogram of the recording (or recording section) made at that time. The waveform and spectrogram are displayed in two bottom panels, below the recording menu (Figure 2), after pressing the “Select time” button, clicking the mouse at the selected time in the LTA spectrogram and then pressing the “Display” button. The start time of recording is shown in the titles of the bottom panels.

The appearance of the waveform and spectrogram is controlled by setting the Recording menu, as shown in Figure 2. The signal can be low-pass, high-pass or band-pass filtered to highlight certain components of recorded noise. The spectrogram can be calculated using different FFT window length from 10 ms to 2 s and different overlap from 0% to 90%, depending on temporal and spectral characteristics of the signals to be satisfactorily displayed in the spectrogram. For example, long-lasting low-frequency tonal or quasi-tonal signals are displayed in a better way with a longer FFT window, while short impulsive signals need a short FFT window to be properly displayed in the spectrogram. The frequency scale of the spectrogram can be chosen to be either linear or logarithmic according to user’s preferences. If some part of the recorded signal needs to be displayed with more details, then the Matlab graphic zooming function can be used in the waveform window to select the time interval of interest. To update the waveform and spectrogram panels after any change in the Recording menu settings or zooming, the user has to press the Display button again.

The sound playback submenu of the Recording menu, which includes play, pause and stop buttons, allows the user to listen to the sound of the whole recording or its selected part. The progress of sound play-back in time is shown in the waveform panel by a moving vertical bar. As acoustic frequencies of many natural sounds in sea noise, e.g. blue whale calls, are beyond the typical range of human auditory perception, the playback submenu also has an option to speed up the sound playback by 2, 5 or 10 times in order to increase the sound frequency accordingly.

If raw acoustic data were recorded in a special binary format, the whole recording or its selected part can be saved in the .WAV format acceptable for conventional sound players in MS Windows and its applications, such as Power Point.

Characteristics of transient and continuous signal of interest can be calculated using the Signal statistics submenu of the Recording menu. If the signal of interest is transient, the user localises the signal in the recording waveform by using the zooming function and updating the waveform panel. Pressing the “Signal stats” button opens the background noise correction menu (Figure 3), where the user is asked to select one of three methods for correction. For a continuous signal, no correction is needed. For a transient signal, the correction can be made either manually or automatically. In the manual mode, the user localises the waveform section containing only the background noise by clicking the mouse at the start and end of the section. In the automatic mode, the program automatically selects the waveform section with minimum acoustic power assuming that it represents the background noise power. The calculated characteristics are the maximum peak-to-peak acoustic pressure in Pa, the maximum received level in dB re 1µPa, the signal exposure level in dB re 1µPa2×s, and the signal duration in seconds based on 90% of its energy (relevant only to transient signals). The measurement data are displayed in a separate window and added in the bottom of the waveform panel (for reporting purposes), as shown in Figure 3.

Sea noise recorded on the coastal shelf, especially in shallow tropical waters, is often dominated by intense impulsive noise from snapping shrimp. This high-amplitude noise can distort significantly the waveform of other underwater signals which are of interest for special analysis, e.g. calculation of signal parameters. The Recording menu contains a submenu which turns a de-spiking function on with parameters selected by the user. An autoregression method for localisation and interpolation of outliers (spikes) in a waveform [24] is employed for de-spiking, so that the waveform of underlying signals distorted by spikes is reconstructed. The default parameters shown in Figure 4 are optimum for filtering out spikes from snapping shrimp. Other settings can be chosen for filtering out parasitic impulsive signals of different origin. After applying de-spiking to the waveform, the signal waveform before and after de-spiking are shown in the waveform panel by red and blue lines respectively. All subsequent operations and calculations, including signal statistics, are performed with the de-spiked signal.

Signal detection

Detection and characterisation of signals of certain origin in sea noise are a common procedure. The CHORUS toolbox provides means to implement this procedure. The main purposes of this CHORUS function is: (1) to test performance of a signal detector of certain type and (2) to process long-term datasets in order to detect acoustic signals of certain origin and assess their major characteristics, such as the frequency content, duration and intensity. The detection menu is called from the CHORUS main menu. It opens a popup menu with a list of underwater sound sources to be selected for detection. The current version of CHORUS offers only two sound sources: pygmy and Antarctic blue whale vocalisations which are common for the southern, western and partly eastern parts of Australia's continental shelf and slope. Humpback whale calls and toothed whale echolocation clicks are planned to be added soon in the source list. Moreover, the architecture of the CHORUS toolbox makes it easy to add automatic detectors of other sound sources without any alteration of the basic routines and functions of the toolbox.

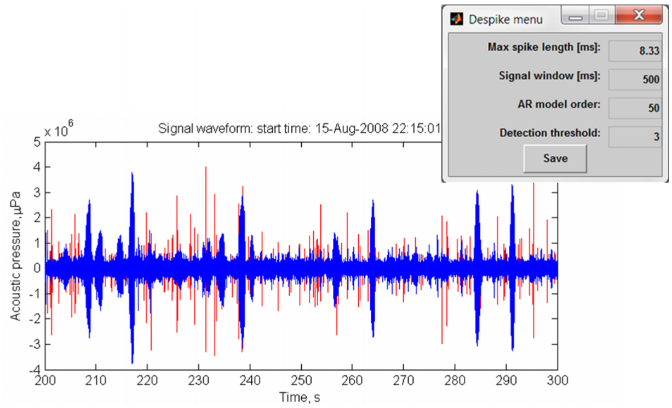


Figure 4: Waveform of sea noise with humpback whale calls before (red) and after (blue) de-spiking, which suppresses snapping shrimp noise. The top right panel shows settings of parameters in the de-spiking menu (see [24] for details of settings in the AR model).

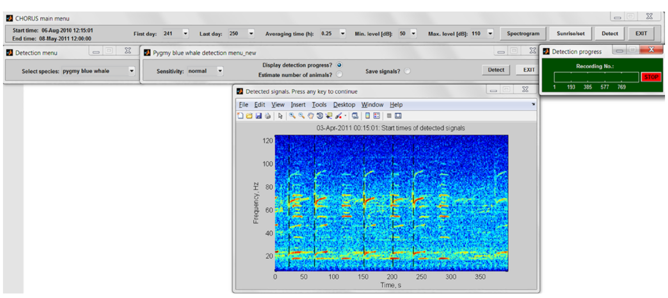


Figure 5: Detection menu with the window of detection settings for pygmy blue whale calls (unit 2 of whale song as described in [1]); sliding bar displaying progress of search over selected recording period (green panel) and spectrogram of an individual recording with the detection times indicated by the dashed vertical lines. The detector missed one of six targeted calls present in this particular recording (starting at about 370 s) and did not produce false detections.

Once the signal source is selected in the detection menu, a special Matlab routine with associated functions is called to implement automatic detection. The example given below is related to detection of calls from pygmy blue whales of the eastern Indian Ocean population that frequent waters off the western and southern Australian coast. Details of the detection method and its performance test are given in [1]. A similar approach can be implemented in the toolbox for other sources of underwater signals.

The detection routine opens a new window with settings governing the detection operation (Figure 5). Firstly, the sensitivity of detection has to be selected in order to control the proportion of missed detections and false alarms, as described in detail in [1]. The higher the sensitivity, the lower misdetection and higher false alarm rates will be. A trade-off between these two rates, when they are nearly equal, can be referred to as normal sensitivity. If the user wishes to control the detection results for verification purposes, then the spectrogram of individual recordings containing detection(s) is displayed with the detection times shown in the spectrogram (Figure 5, bottom panel). The detection operation is performed for the time period specified in the main menu by the first and last day. Once it is finished, the detection times are saved in a .MAT file along with other characteristics of detected signals, such as the signal frequency, RMS pressure and some others depending on signal type and functionality of the detector. Saving the waveform of detected signals can be useful for further verification, if the total number of expected detections is not too large. If the number of similar callers (or similar sources), making sound within the same recording period, can be estimated based on some criteria, then this number can also be saved. When processing large datasets spanning several months, it is useful to track the search progress over the observation time or recording number, using, for example, a separate window with a sliding bar, as shown in the green panel of Figure 5.

EXAMPLES OF RESULTS

Several tens of large datasets containing sea noise recordings from a few weeks to about a year long have been analysed for various purposes using the CHORUS toolbox. Employing CHORUS helps in preparing technical reports for industry and governmental agencies concerning with potential impacts of offshore industry operation on marine fauna. This includes measuring characteristics of natural sea noise and its components of physical and biological origin and assessing man-made noises resulting from offshore industrial and recreational activities. All panels with waveforms and spectrograms created by CHORUS can be easily copied and pasted into a MS Word document or a Power Point presentation.

CHORUS has also been extensively used for applied research in marine biology, underwater acoustics and oceanography. In particular, results of sea noise analysis with CHORUS have been used to study temporal and migratory patterns of various species of marine mammals and fish in several locations over the continental shelf and slope around Australia [2, 4, 6, 7, 8, 25-28]. As an example, the temporal and spatial patterns of pygmy blue whale migration off the western and southern coasts of Australia are explored using data from several autonomous sea noise recorders deployed on the shelf, including the passive acoustic observatories of the Australian Integrated Marine Observing System (IMOS, [29]). These patterns are derived in particular from the presence and frequency of occurrence of vocalisations by pygmy blue whales detected at different locations using CHORUS. Figure 6 shows the frequency of occurrence of pygmy blue whale calls detected at the IMOS passive acoustic observatory in the Perth Canyon. The plot reveals the temporal pattern of whale migration in this area of the continental slope. The whales appear in this area from late November - early December to mid January, when they migrate primarily to feeding grounds in the Southern Ocean, and then from early February to June, when they migrate to their breeding grounds in the tropical waters [30]. On the northern leg, the whales tend to spend more time in the canyon region, as it is believed to be a good feeding spot for them [31], so that the number of detections during the northern migration period is considerable larger. The automatic detector of pygmy blue whale calls in CHORUS is capable of measuring the call frequencies. This has allowed us to reveal a long-term decline in the frequency of their calls of about 0.12 Hz/year for the fundamental frequency [20].

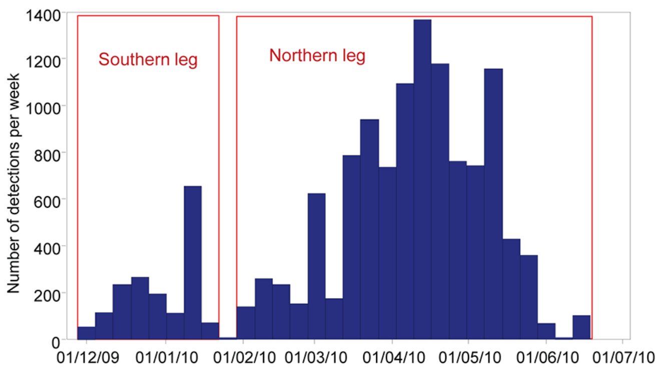


Figure 6: Histogram of the number of calls detected weekly at the IMOS acoustic observatory in the Perth Canyon from November 2009 to July 2010.

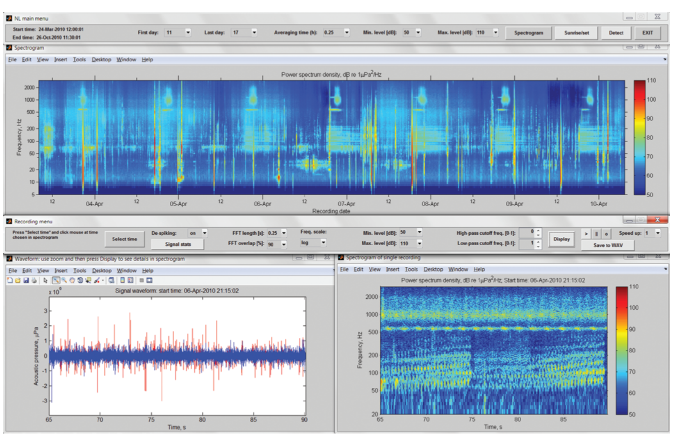


Figure 7. Stacked spectrogram of seven days of recording of fish choruses off the WA coast, produced using CHORUS (top panel) and waveform (bottom left panel) and spectrogram (bottom right panel) of a section of sea noise recording that contains sounds from two fish species.

Combining signal measurements of individual fish calls in CHORUS with single hydrophone localisation techniques [32, 33] has provided the source level of vocalisation by commercially important fish species [34]. Analysis of long-term datasets has been used to highlight the overlapping nature and periodicity of multiple fish choruses at a single location and infer movement of an aggregation over time [28]. Figure 7 illustrates how multiple fish choruses can occur with defined temporal patterns, at one location, over similar frequency bands. CHORUS helps quickly identify temporal patterns over a variety of timescales, and see how the callers might interact with each other. In this case, two fish choruses begin calling on the evening and call each evening for the following seven days (sound energy between roughly 500 and 2000 Hz), another chorus begins earlier in the day (sound energy between roughly 50 and 200 Hz), continuing through the night (Figure 7, main spectrogram panel). The bottom right panel shows how three of the choruses overlap closely in time, throughout the whole 15-second recording section and one fish producing the two upward sweeps beginning at around 65 and 83 s.

In acoustic modelling studies, specific characteristics of sound propagation over Australia's continental shelf have been investigated with the help of CHORUS [35].

CONCLUSIONS

The CHORUS toolbox has provided the basis to meet many needs of sea noise research and monitoring. Being under its fourth iteration it has also shown the ability to be revised simply to incorporate new aspects to acoustic ecology and quantification. This can be conducted on generic WAV files, generated by numerous recordings systems or the CMST-Defence Science and Technology developed sea noise loggers. An executable version of CHORUS will soon be available for academic purposes, particularly for processing IMOS data, for which the toolbox is most regularly used.

While the possibilities for development are numerous, the immediate future for CHORUS is anticipated to be an expansion of the detector functions to include other marine mammals and fish, and to export defined chorus levels of marine animals for long-term comparison with environmental and anthropogenic variables that may contribute to vocalisation levels.

ACKNOWLEDGEMENTS

The CHORUS toolbox has been driven by the use, requests and comments from several members of the CMST research group to meet the needs of those interrogating long-term datasets of sea noise. Special thanks to Dr Rob McCauley who formulated and developed the general concept of sea noise analysis. The contributions of the CMST researchers have, and will continue to, further the expansion and development of the toolbox.

REFERENCES

[1] A.N. Gavrilov and R.D. McCauley, "Acoustic detection and long-term monitoring of pygmy blue whales over the continental slope in southwest Australia", *J. Acoust. Soc. Am.* **134**(3), 2505-2513 (2013)

[2] C.P. Salgado Kent, A.N. Gavrilov, A. Recalde-Salas, C.L.K. Burton, R. D. McCauley and S. Marley, "Passive acoustic monitoring of baleen whales in Geographe Bay, Western Australia", *Proceedings of the Acoustical Society of Australia, Fremantle*, Western Australia, 21-23 November (2012)

[3] C.P. Salgado Kent, C. Jenner, M. Jenner, P. Bouchet and E. Restad, "Southern Hemisphere breeding stock 'D' humpback whale population estimates from North West Cape, Western Australia". *Jour. Cet. Res. Man* **12**, 29-38 (2012)

[4] C.P. Salgado Kent, R.D. McCauley, I.M. Parnum and A.N. Gavrilov, "Underwater noise sources in Fremantle inner harbour: dolphins, pile driving and traffic", *Proceedings of the Acoustical Society of Australia, Fremantle*, Western Australia, 21-23 November (2012).

[5] R.D. McCauley, "Fish choruses from the Kimberley, seasonal and lunar links as determined by long term sea noise monitoring", *Proceedings of the Acoustical Society of Australia, Fremantle*, Western Australia, 21-23 November (2012)

[6] M.J.G. Parsons, R.D. McCauley and F. Thomas, "Sound of fish calls off Cape Naturaliste, Western Australia", *Acoust. Aust.* **41**, 58-63 (2013)

[7] M.J.G. Parsons, R.D. McCauley and M.C. Mackie, "Characterisation of mulloway advertisement sounds", *Acoust. Aust.* **41**, 196-201 (2013)

[8] M.J.G. Parsons, P. Lewis, S.G. Longbottom, R.D. McCauley, and D.V. Fairclough, "Sound production by the West Australian dhufish (Glaucosoma hebraicum)", *J. Acoust. Soc. Am.* **134**, 2701-2709 (2013)

[9] M.J.G. Parsons, I.M. Parnum, and M. Legg, "Sounds of captive Western Rock Lobster", First International Conference on Acoustic Measurements, conference proceedings, Corfu, Greece, 23-27th June (2013)

[10] M. W. Legg, Non-Gaussian and non-homogeneous Poisson models of snapping shrimp noise. PhD Thesis, Curtin University (2010)

[11] N. Soars, "Characterization of sounds produced by temperate and tropical sea urchins from the east coast of Australia", *The Effects of noise on aquatic life*, (2013)

[12] J.A. Hildebrand, "Anthropogenic and natural sources of ambient noise in the ocean", *Mar. Ecol. Prog. Ser.* **395**, 5-20 (2009)

[13] D.H. Cato, M.J. Noad, R.A. Dunlop, R.D. McCauley, C.P. Salgado Kent, N.J. Gales, E. Kniest, J. Noad and D. Paton. "Behavioural response of Australian humpback whales to seismic surveys". *J. Acoust. Soc. Am.* **129**, 2396 (2011)

[14] C. Erbe, R. McCauley, C. McPherson and A. Gavrilov, "Underwater noise from offshore oil production vessels", *J. Acoust. Soc. Am.* **133**, EL465-EL470, (2013)

[15] C. Erbe, "Underwater noise of small personal watercraft (jet skis)" *J. Acoust. Soc. Am.* **133** (4), EL326-EL330 (2013)

[16] M.J.G. Parsons, R.D. McCauley, M. Mackie, and A.J. Duncan, "In situ source levels of mulloway (*Argyrosomus japonicus*) calls", *J. Acoust. Soc. Am.* **132**, 3559-68, (2012)

[17] M.J.G. Parsons, D. Holley and R.D. McCauley, "Source levels of Shark Bay dugong (Dugong dugon) calls" *J. Acoust. Soc. Am.* **134**, 2582-2588 (2013)

[18] M.J.G. Parsons "An investigation into active and passive acoustic methods to study fish aggregations", PhD thesis, Curtin University, 410 pp. (2010)

[19] A.T. Barrios, "Use of passive acoustic monitoring to resolve spatial and temporal patterns of spawning activity for red drum, *Sciaenops ocellatus*, in the Neuse River Estuary", North Carolina, Masters Thesis, North Carolina State University. 118 pp (2004)

[20] A.N. Gavrilov, R.D. McCauley and J. Gedamke, "Steady inter and intra-annual decrease in the vocalization frequency of Antarctic blue whales", *J. Acoust. Soc. Am.* **131**, 4476-4480 (2012)

[21] M. Sprague and J.J. Luczkovich, "Modeling fish aggregation sounds in very shallow water to estimate numbers of calling fish in aggregations", 161st Meeting of the Acoustical Society of America, Seattle, Washington, 23 - 27 May 2011, Vol. 1 (2011)

[22] J.J. Luczkovich, M.W. Sprague, S.E. Johnson and C. Pullinger, "Delimiting spawning areas of weakfish *Cynoscion regalis* (family sciaenidae) in Pamlico Sound, North Carolina using passive hydroacoustic surveys", *Bioacoustics* **10**, 143-60 (1999)

[23] R.S. Sousa-Lima, T.F. Norris, J.N. Oswald and D.P. Fernandes, "A review and inventory of fixed autonomous recorders for passive acoustic monitoring of marine mammals", *Aquatic Mammals* **39**(1), 23-53 (2013)

[24] S.V. Vaseghi and P.J.W. Rayner, "Detection and suppression of impulsive noise in speech communication systems", *IEE Proceedings* **137** (Pt. 1, No.1), 38-46 (1990)

[25] A. Recalde-Salas, C.P. Salgado-Kent, M.J.G. Parsons, S. Marley, A.G. Gavrilov and R.D. McCauley, "Social sounds of humpback whales off northwest Western Australia", (in review) *J. Acoust. Soc. Am.* (2014)

[26] A. Recalde-Salas, C.P. Salgado-Kent, M.J.G. Parsons, S. Marley and R.D. McCauley, "Non-song vocalisations of pygmy blue whales in Geographe Bay, Western Australia", *J. Acoust. Soc. Am.* **135**(5): EL213-EL218 (2014)

[27] A.N. Gavrilov, R.D. McCauley, C. Salgado-Kent, J. Tripovich and C. Barton, "Vocal characteristics of pygmy blue whales and their change over time", *J. Acoust. Soc. Am.* **130**, 3651-3660 (2011)

[28] M.J.G. Parsons, and R.D., McCauley, Port Hedland Sea noise logger program, Interim report, 16 pp. (2012)

[29] http://imos.org.au/

[30] R.D. McCauley and C.K. Jenner, "Migratory patterns and estimated population size of pygmy blue whales (*Balaenoptera musculus brevicauda*) traversing the Western Australian coast based on passive acoustics", IWC SC/62/SH26 9 pp. (2010)

[31] S.J. Rennie, R.D. McCauley and C.B. Pattiaratchi, "Thermal Structure above the Perth Canyon reveals Leeuwin Current, Undercurrent and Weather Influences and the Potential for Upwelling", *Mar. Fresh. Res.* **57**, 849-861 (2006)

[32] D.H. Cato, "Simple methods of estimating source levels and locations of marine animal sounds" *J. Acoust. Soc. Am.* **104**, 1667-78 (1998)

[33] M.J.G. Parsons, R.D. McCauley, M. Mackie, P.J. Siwabessy and A.J. Duncan, "A comparison of techniques for ranging close range mulloway (*Argyrosomus japonicus*) calls using a single hydrophone", *Acoustics. Aust.* **38**, 145-51 (2010)

[34] M.J.G. Parsons, S.M. Longbottom, R.D. McCauley and D.V. Fairclough, "In situ calls of Western Australian dhufish (*Glaucosoma hebraicum*)", *Acoustics. Aust.* **42**, 31-35 (2014)

[35] A.J. Duncan, A.N. Gavrilov, R.D. McCauley, and I.M. Parnum, "Characteristics of sound propagation in shallow water over an elastic seabed with a thin cap-rock layer", *J. Acoust. Soc. Am.* **134**, 207-215 (2012)

COMPARISON OF INFRASOUND MEASURED AT PEOPLE'S EARS WHEN WALKING TO THAT MEASURED NEAR WIND FARMS

Matthew Stead[1], Jon Cooper[1] and Tom Evans[1]
[1]Resonate Acoustics, 97 Carrington Street, Adelaide, SA 5000, Australia
Matthew.stead@resonateacoustics.com

Infrasound is observed in all environments at varying levels and is generated by a range of natural and anthropogenic sources. Some studies have suggested that modern wind turbines can generate a relatively low level of measurable noise at frequencies corresponding to the blade pass frequency of turbines. People walk at a variety of speeds, with typical walking frequencies similar to the blade passing frequency of modern commercial wind turbines. Measurements have been conducted of the levels of infrasound generated at the human ear when walking and compared to measured levels near wind farms. The measured level of infrasound generated at the ear at blade pass frequency when people walk can be considerably higher than the level near wind farms. In both cases, measured levels were significantly below the audibility threshold for very low frequency noise.

INTRODUCTION

Infrasound has been reported as a concern by some residents around proposed and operating wind farms and has been raised as an issue in a South Australian parliamentary inquiry [1]. Infrasound is observed in all environments, arising from a combination of natural and anthropogenic sources. A recent study conducted by the South Australian Environment Protection Authority (SA EPA) and Resonate Acoustics [2] has found infrasound levels to be no greater at houses near wind turbines than levels experienced in other urban and rural environments. Additionally, the contribution of wind turbines to the measured infrasound levels was found to be insignificant in comparison with the background level in the environment and with respect to human perception.

While the study [2] found that infrasound at houses located 1.5 kilometres from wind farms was at a level no more significant than at other locations, there were characteristics observed in the spectra at the one third octave band centre frequencies of 0.8, 1.6 and 2.5 Hz at the two locations near wind farms. These approximately correspond to the blade pass frequency and second and third harmonics of the nearby wind turbines. The highest measured $L_{eq,10min}$ levels in these one third octave bands were between 60 and 65 dB, with the highest typically occurring at 1.6 Hz. At the time the study [2] was not able to conclude whether this characteristic was a result of wind turbine operation or another source but acknowledged that it is possible that the wind turbines generated it. In light of other recent studies as discussed below, it is reasonable to conclude that these levels did result from operation of the wind turbine.

A 2012 study in Wisconsin [3] (the Shirley Wind Farm study) also measured infrasound in houses near wind turbines and found characteristics in the spectra at the blade pass frequency and related harmonics at a house located approximately 330 metres from the nearest turbine. These characteristics were not detected at two other houses located approximately 1 and 2 kilometres respectively from the nearest turbine. As for the SA EPA and Resonate Acoustics study [2], the level of the characteristics presented in the Shirley Wind Farm study [3] was well below the threshold of audibility, with a power spectrum level of approximately 60 dB measured at 1.4 Hz for a frequency resolution of 0.05 Hz.

Also of interest are recent measurement campaigns conducted by both the SA EPA [4] and the University of Adelaide [5] at Waterloo Wind Farm in response to complaints from some residents. Measurements at a number of sites were reported in the infrasonic range in both studies at sites between 1.3 and 3.5 kilometres from the nearest 3 MW wind turbines. In both reports, the blade pass signals measured in the one-third octave band 1.6 Hz were between 45 to 60 dB, including during periods of complaint. It therefore appears reasonable to use the levels measured during the SA EPA and Resonate Acoustics study [2] as a conservative basis for comparison with infrasound levels generated during walking.

Despite the measured levels of infrasound near wind farms presented in both [2] and [3] being significantly below the audibility threshold and only detectable through highly sensitive measurement equipment, the presence of infrasound that may be due to wind turbines has led to some community members expressing concern [6]. Therefore, there is interest in comparing these measured levels of infrasound to levels that people are exposed to on a daily basis.

The blade pass frequency of wind turbines is equal to the number of blades multiplied by the rotational speed of the turbine. With three blades and rotational speeds in the order of 15 to 18 rpm [7], modern commercial wind turbines of the type at the wind farms included in the SA EPA and Resonate Acoustics study [2] have a blade pass frequency of 0.8 to 0.9 Hz.

This paper investigates the hypothesis that, as people walk with a rate similar to the blade pass frequency of wind turbines, they may be exposed to infrasound with a similar characteristic and level to that measured near wind farms. This infrasound would result from the periodic change in pressure levels at the ear as people walk.

NATURE OF WALKING

When people walk there is a slight rise and fall of the head. The rate of the rise and fall (or walking pace) depends on the speed of walking, gate and stride length. There is also a side to side movement of the head at lower magnitude. There have been numerous studies of walker pace rate normally related to structural vibration assessments for building vibration response [8]. The rate of walking typically ranges from 60 paces per minute for slow walking to 120 paces per minute for fast walking [8, 9]. Using conventional engineering terminology this equates to 1 to 2 Hz. People hence commonly walk at a similar rate to the blade passing frequency and associated harmonics of modern commercial wind turbines.

Vibration testing was carried out with a test walker (walker A) at 1 Hz (60 paces per minute) and with the walker carrying a Brüel and Kjær Type 4100 head simulator. An accelerometer was attached to both the walker A (torso) and the head simulator during subsequent tests and the vertical acceleration was measured to confirm vibration levels. The measured walker vibration, measured vibration of the head simulator and a normalised sinusoidal 1 Hz signal are presented in Figure 1. The normalised sinusoidal signal has an amplitude of 0.18 g. The rms acceleration on the manikin head was 30% lower than on the test walker. Any measurements on the manikin head are hence conservative in terms of amplitude during walking.

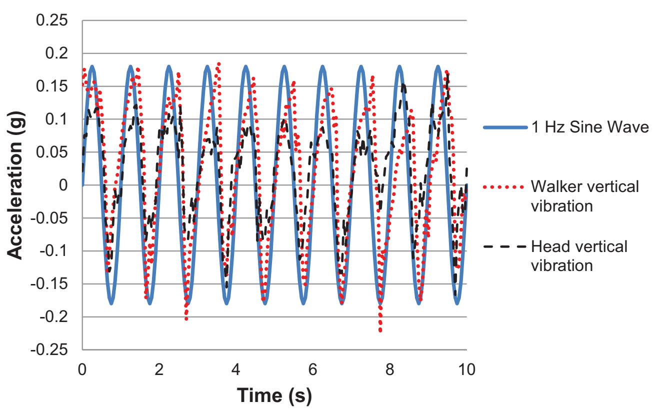


Figure 1 Measured test walker acceleration and 0.18 g amplitude normalised 1 Hz sinusoidal waveform.

To ensure that the results of the test walker are valid to the wider population a series of walking tests (an additional 6 walkers, walker B to G) was carried out where the walker was asked to walk at a comfortable pace. The measured walker vibration for a sample of each test walker is shown in Figure 2. Frequency analysis of the measured vertical vibration showed a range of walking speeds with a dominant frequency (around 100 paces per minute, with the frequency between 90 and 111 paces per minute). All test walking was inside a building, without coaching on walking style and with enough room such that their walking style was not impeded. The test walker A had similar walker vibration result (rms acceleration) to the walkers B to G and was +2% higher than the average rms acceleration at the walking speed.

The measured vibration levels for test walkers A to G show that people walk with acceleration levels similar to a sinusoidal wave (with a dominant primary frequency and harmonics) and with reasonable repeatability as demonstrated by the results for test walkers A to G. The measured levels for walker A also demonstrate that the vibration that the head simulator was subject to during the test is slightly lower than that of the walker A. This indicates that the levels of infrasound measured in the ear of the head simulator will provide a reasonable, if slightly conservative, representation of the infrasound levels at the ear of the walker.

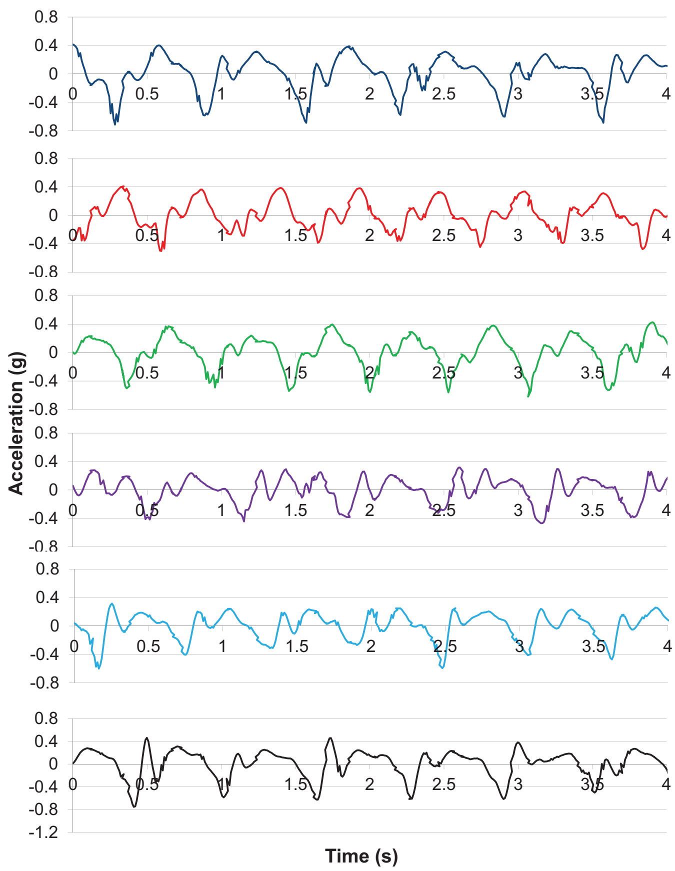


Figure 2 Measured walker acceleration for walker B (top of image) to G (bottom of image).

MEASUREMENT METHODOLOGY

The measurement of infrasound that walkers are exposed to requires a carefully considered experimental methodology. The key considerations for the measurement methodology were:

- Test walker
- Realistic head and ear interaction to represent the level of infrasound generated at the human ear
- Signal processing and low frequency system response

- Reference level to be measured within the space to identify any extraneous noise effects
- Isolation of extraneous vibration effects and other potential sources of error

It is noted that accurate measurement of environmental infrasound from anthropogenic sources requires the use of appropriate techniques to remove the effect of wind-induced noise at the microphone [2]. This is not considered to be a significant concern to this investigation as the human ear would also be subject to wind-induced infrasound when people walk. However, additional measurements were undertaken to determine whether wind-induced noise had affected the measurement results at the ear. In comparing wind farm infrasound to walker infrasound, another consideration is the indoor vs outdoor infrasound measurements. For the purpose of this comparison only the indoor levels were compared, as the variable of wind induced noise was able to be controlled. The level of infrasound measured while walking outdoors may be greater due to additional wind induced noise. As this study compares exposure infrasound levels (that humans might be exposed to) the indoor levels are taken for comparison purposes. The conclusions in this paper are taken on the basis of indoor infrasound levels during walking.

Test walker

The selected test walker (walker A) is experienced in walking at fixed frequencies with a natural gait primarily for the testing of structural vibration resulting from walker-generated vibration levels. A metronome was used to synchronise the walking frequency with a visual rather than audible beat, so that the frequency of all tests was consistent. The walker is approximately 90 kg and 180 cm tall. All walking was on a suspended timber structure floor with carpet covering. The test walker walked in a large circuit for over two minutes for each test condition to ensure sufficient data was obtained to obtain an accurate measured sound pressure level at very low frequencies.

Head and ear

Walker infrasound levels were measured with a Brüel and Kjær Type 4100 Sound Quality Head Simulator manikin designed for sound quality testing. Only the head was used for testing. A single microphone, positioned at the entrance to the ear canal on the manikin's head, was used to simulate the signal that includes the interference patterns caused by the head. The ears are moulded-silicone pinna simulator which sit around the microphones to provide directivity patterns similar to the human ear and are designed to have a frequency response to sounds coming from all directions which closely approximates the direction-dependent human response. The Brüel and Kjær Type 4100 head simulator is designed for measuring human exposure from a range of noise sources.

The head simulator was held in front of the test walker and carefully moved with a similar vertical displacement as the walker, as previously shown in Figure 1.

Signal processing and low frequency system response

Measurement of noise levels in the infrasonic range is complicated by factors that do not affect measurements in the normal audible range of sounds, in particular the use of equipment with an accurate measured response to a low enough frequency. The majority of sound level meters and microphones are generally designed to only measure noise levels accurately at the typical audio frequencies (20 Hz – 20 kHz), and are insufficient for the accurate determination of noise levels at frequencies below 20 Hz.

Measurements were carried out using the equipment listed in Table 1. All equipment held current calibration certificates from a National Association of Testing Authorities certified laboratory or were manufacturer calibrated in the case of the microphones.

Table 1. Measurement and analysis equipment

Analyser	Integrated microphone and preamplifier set	Frequency range [1]
Sinus Soundbook Quadro+ (S/N 06364)	Brüel & Kjær Type 4193-L-004 (S/N 2774943)[2]	0.2 Hz - 20 kHz
	Brüel & Kjær Type 4193-L-004 (S/N 2774944)[2]	0.2 Hz - 20 kHz

1. Frequency range determined based on the minimum of the analyser or microphone.

2. Fitted with a Brüel and Kjær Type UC-0211 low frequency noise adaptor.

Two matched Brüel & Kjær Type 4193-L-004 microphone and preamplifier sets have been used for simultaneous two-channel measurements with the Soundbook data acquisition system. One microphone was located in an ear of the Brüel and Kjær Type 4100 test head while the second microphone was located on a tripod pole within the space. The calibration chart for the data acquisition system shows negligible deviation of the instrument frequency response to frequencies as low as 0.1 Hz. The microphone calibration charts (dated 11 December 2012) also show 3 dB deviation of the frequency response to less than 28 mHz.

The equipment was setup to store linear (unweighted) 1/3rd octave band sound pressure levels from 0.2 Hz to 20 kHz over the duration of each test. The overall linear sound pressure level was also stored in 120 ms intervals during each test to allow the amplitude modulation corresponding to the walker frequency to be visualised.

Test location

All measurements described in this paper were undertaken indoors. By undertaking the measurements indoors, it was possible to minimise the influence of wind-generated turbulence on the measured infrasound levels. The stationary reference microphone was fitted with a 90 mm windshield and located approximately 2 m from two walls. For all measurement periods windows and doors were kept closed. The test walker microphone did not have a windshield as this represents the case of human exposure as discussed previously.

The approximate dimensions of the test room were 5 m (L) x 12 m (W) x 3 m (H). It is known that low frequency noise levels can vary within a room due to the modal response of rooms to noise at wavelengths on a similar scale to the physical dimensions of the room. However, studies have shown that there are negligible room effects within the infrasonic range of 20 Hz and below [2, 10].

Assessment of potential sources of error

Additional tests were carried out to identify potential sources of error in the measurements, namely:

- Extraneous vibration effects on the microphone
- Variation in measured sound pressure levels between the test head and reference microphone
- Noise generated by walking on the carpeted timber floor
- Wind-induced noise effects on the ear microphone when walking.

The results of the additional tests to assess these sources of error are discussed following the presentation of the measurement results.

MEASUREMENT RESULTS

Figure 3 presents the measured linear sound pressure levels over the infrasonic frequency range (0.2 Hz to 20 Hz) for both the microphone in the test head ear and the reference microphone within the room. It can be seen that there is a distinct characteristic peak at approximately 76 dB at the ear at the walker frequency of 1 Hz. The measured sound pressure levels are assumed to be as a result of change in static air pressure or as a result of dynamic pressure variation from side to side movement.

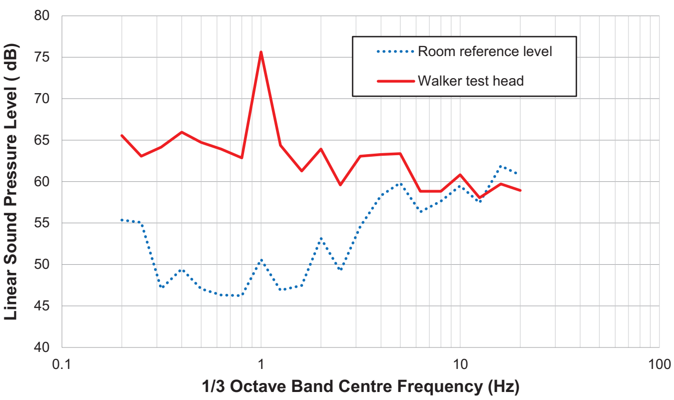


Figure 3 Walker and reference linear sound pressure levels over infrasonic range

The peak is also present in the room reference level but at a significantly lower level. This peak in room reference level is due to the repetitive noise generated by the walker on the carpeted timber floor. As the levels are significantly lower at the reference location than at the ear, despite the walking occurring very close to the reference measurement location, it demonstrates that airborne noise generated by walking on the floor did not significantly contribute to the measured levels at the ear.

It can also be seen that the levels at the ear are elevated in comparison to the room reference level across the frequency range from 0.2 Hz to approximately 5 Hz. This is most significant at the walker frequency of 1 Hz but, as it is elevated across the entire frequency range, it is believed to be the result of variations in pressure at these frequencies at a lower level than at the primary walking frequency.

Isolation of extraneous vibration effects

An obvious concern when measuring the infrasound levels walkers are exposed to is the potential influence of vibration on the microphone system. Condenser microphones have a metal diaphragm, which could be susceptible to vibration causing an elevated measured infrasound level generated by vibration rather than variation in static air pressure or as a result of dynamic pressure variation from side to side movement.

The Brüel & Kjær Type 4193 microphone has a published vibration sensitivity (<1000 Hz) of 65.5 dB equivalent SPL for 1 m/s^2 axial acceleration. The measured vibration levels with a walker at 1 Hz pace rate is nominally 0.18 g or 1.8 m/s2 amplitude in the vertical direction. This is the transverse direction for the in-ear microphone and is the less sensitive direction. The measured sound pressure level at the ear was 76 dB at the walker frequency, significantly above the microphone vibration generated level that would be expected for the vibration based on manufacturer data. Hence, vibration was considered unlikely to be the cause of measured infrasound levels at the walker frequency.

However, given the need to ensure that vibration was not the source of the induced infrasound, additional testing was carried out with the head rotated 90 degrees such that the microphone was orientated vertically. This corresponds to the most sensitive microphone direction relative to the direction of the vibration and, if vibration of the diaphragm was the cause of the measured infrasound levels, higher measured sound pressure levels would be expected.

Figure 4 presents the measured infrasound levels with the test head microphone rotated 90 degrees and in the standard orientation. Both measurements were undertaken with the test head stationary (no forwards motion) but moved up and down in the appropriate orientation. It can be seen that infrasound levels were no higher in the rotated orientation where higher vibration levels would be expected. This confirmed the measured infrasound levels were not generated through microphone vibration. It is noted that there was a change in the frequency of the motion between the two measurements, with the low vibration measurement having a greater range of motion at approximately 0.8 Hz whereas the frequency of the higher vibration measurement was more consistently 1 Hz. However, review of the measured vibration levels during both measurements indicated that there was only a minor change in the amplitude of the motion and that the two results can be compared.

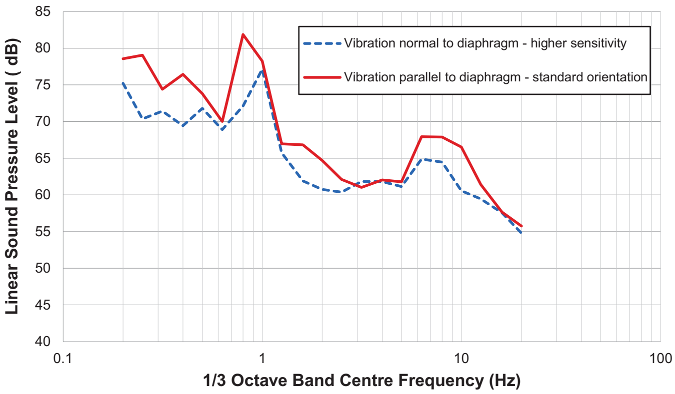


Figure 4 Walker infrasound levels with head in different orientations – rotated 90° (normal to microphone diaphragm, high vibration induced level expected) and standard (parallel to microphone diaphragm, low vibration induced level expected).

Response of test head compared to reference microphone

A stationary test was carried out to compare the measured infrasound level at the ear of the stationary test head to that measured at the reference microphone on the tripod. The measured infrasound levels were found to be within 1 dB of each other within each one-third octave band. This provided confidence that the reference microphone and test head observe the same infrasound levels and that the measured infrasound levels can be directly compared between microphones.

Wind-induced noise

It is necessary to check whether the measurement results were as a result of wind-induced noise (resulting from the walking motion) on the microphones, although this did appear unlikely given the distinct characteristic at the walker frequency.

The comparison of measurement results while walking shown in Figure 3, to measurement results with the head moved vertically with no forward motion (Figure 4) indicates the levels of the 1 Hz peaks are similar, when the vertical displacement is similar. The measurement of the same level of infrasound during both the stationary vibration tests as those in the walking tests supports the hypothesis that the primary source of measured infrasound is the change in pressure with height, rather than wind induced noise due to forwards movement through the air.

ANALYSIS

Pressure change at ear

The pressure change at the ear of the test head when walking can be approximately determined assuming a 1 Hz sine wave with amplitude of approximately 0.08g (Figure 1). This converts to a peak to peak displacement of 25 mm. The approximate rms sound pressure level at 1 Hz can then be determined for a standard air density (ρ = 1.2 kg/m^3) as:

$$p = \frac{\rho g h}{2\sqrt{2}} = \frac{1.2 \times 9.81 \times 0.025}{2\sqrt{2}} = 0.1Pa$$

$$SPL = 20\log\left(\frac{p}{p_{ref}}\right) = 20\log\left(\frac{0.1}{2 \times 10^{-5}}\right) = 74dB$$

This compares well to the measured SPL of 76 dB at 1 Hz as shown in Figure 3, particularly given the relatively simplistic approximation. In reality, the measured acceleration signal on the head does not represent a perfect sine wave and the typical peak amplitude may be slightly higher than the assumed 0.08g. However, it helps to confirm that the measured infrasound signal at 1 Hz is representative of the change in pressure at the ear and not due to extraneous sources. It should also be noted that there is side to side movement of the head with reduced magnitude compared to the vertical movement and some sound pressure fluctuation may be generated by this mechanism.

Comparison of walker infrasound levels to measured levels at wind farm locations

Indoor infrasound levels were measured at one house near the Bluff Wind Farm and another near the Clements Gap Wind Farm as part of the recent SA EPA and Resonate Acoustics study [2]. The houses were located 1.4 and 1.5 kilometres from the nearest wind turbine respectively. The measured indoor infrasound levels at the houses were reviewed to identify periods where potential characteristics at the blade pass frequency (0.8 Hz) and harmonics (1.6 Hz and 2.5 Hz) were identified.

Figure 5 compares two measurements of the test walker (walker A) to measured 10-minute averaged infrasound levels at the two houses. The 10-minute period at each house was selected such that it was representative of the highest sound pressure levels at the blade pass frequency and associated harmonics for which a potential characteristic could be identified. The corresponding hub height wind speed at the wind farm was 14 m/s for the Bluff Wind Farm and 12 m/s for Clements Gap Wind Farm. For reference, rated power of the wind turbine model installed at each wind farm occurs at a hub height wind speed of 14 m/s. The cause of the apparent lower frequency characteristic at 0.5 Hz during the second walker measurement was not obvious, but it is suspected that it may be due to a difference in the walkers step on the left and right foot during that measurement.

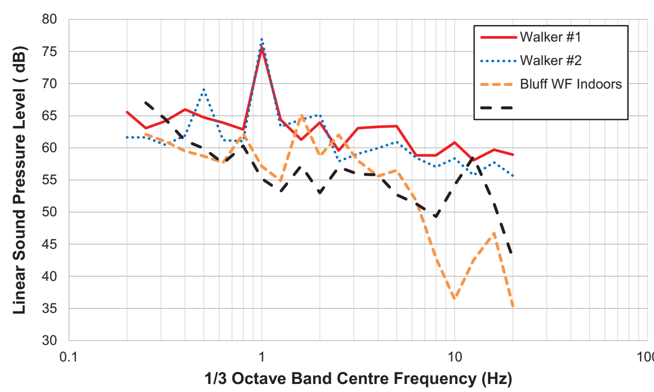


Figure 5 Comparison of measured walker infrasound levels to measured indoor infrasound levels at houses near wind farms

It is clear that, although the measured walker infrasound characteristic occurs at a slightly different frequency to the characteristics that may potentially arise from wind turbine operations, the measured sound pressure level is at least 10 dB higher for the relevant one-third octave band. The shape of the characteristic (or peak in the spectrum) is also significantly more pronounced for the test walker than for the wind farms, with a peak of approximately 12 dB above the adjoining one-third octave bands.

This comparison indicates that the human ear is regularly exposed to pressure variations at these very low frequencies (less than 2 Hz) that are significantly higher than that potentially resulting at houses from operation of a wind farm. It also demonstrates that the infrasound characteristic that may result from wind turbines at the blade passing frequency is not unique and is considerably less pronounced than that which walkers experience at similar frequencies.

The recent Shirley Wind Farm study [3] measured a power spectral level of approximately 60 dB (0.05 Hz bandwidth) at 1.4 Hz at a house approximately 330 metres from the nearest wind turbine. While difficult to directly compare the power spectrum level to 1/3rd octave band results, the level measured in the Shirley Wind Farm Study appears similar to the results measured in the SA EPA and Resonate Acoustics study [2] and measured sound pressure level at the ear when walking at a specific frequency would also be higher than that measured in the Shirley study. Note that 330 metres is significantly closer than the nearest residences to wind farms in Australia, with the nearest non-financially involved houses typically located about one kilometre away.

The SA EPA [4] and University of Adelaide [5] studies at Waterloo Wind Farm reported one-third octave band levels at 1.6 Hz of between 45 and 60 dB depending on location, including during times of complaints. The levels reported by these studies therefore appear lower than those measured in the SA EPA and Resonate Acoustics study [2] and significantly lower than those measured during walking and presented in Figure 5.

It is noted that the measured infrasound levels near wind farms in the various studies ([2], [3], [4] and [5]) include harmonics of the blade pass signal evident up to a maximum frequency of approximately 10 Hz. By contrast, the infrasonic signal generated at this test ear when walking does not exhibit as many obvious harmonics. The potential for multiple harmonics does however exist given the patterns in acceleration measured for the various test walkers which are included in Figure 2. In terms of absolute level, the blade pass signal at locations near wind farms is often measured to be highest below 2 Hz, and in this region the levels generated at the ear when walking have been measured to be higher than those measured at a typical residential distance from wind farms in Australia.

Comparison to the human hearing threshold

Møller and Pedersen [11] provide a summary of investigations into the human hearing threshold at infrasonic frequencies, presented here as Figure 6. The threshold is denoted a hearing threshold and it is noted that investigations indicate non-auditory perception occurs at levels approximately 20 to 25 dB above the hearing threshold [11].

While there is no information on the threshold at frequencies of 1 Hz and lower in Figure 6, the mean threshold at 1.5 to 2 Hz appears to be in the order of 110 to 130 dB. It is therefore clear that the measured walker infrasound levels are well below the mean hearing threshold, and the measured infrasound levels near wind farms even more so.

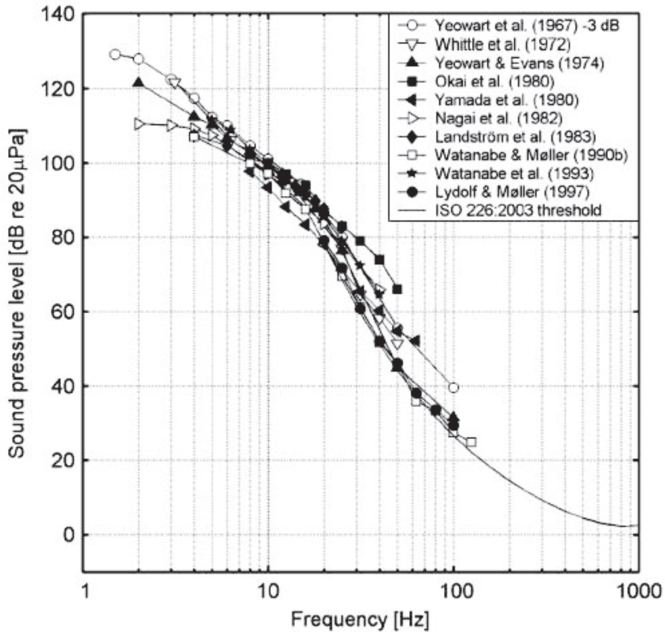


Figure 6 Summary of investigations into human hearing threshold covering frequency range at and below 20 Hz, from Møller and Pedersen [11]

CONCLUSIONS

A study has been undertaken into the infrasound levels generated at the human ear when walking. This investigation arose from findings as part of a recent study into infrasound levels measured at houses adjacent to wind farms [2] at approximately 1.5 km, where a characteristic was measured at frequencies corresponding to the wind turbine blade pass frequency and associated harmonics.

It has been found that the human ear is subject to sound pressure levels in the order of 75 dB at the one-third octave band centre frequency corresponding to the walking frequency. The characteristic that occurs at the ear during walking is similar in dominant frequency to that measured at the houses near wind farms (when walking at the same pace). However, it is significantly higher in level, with levels measured to be 10 dB higher than the highest levels measured near wind farms at 1.5 km away where residences may be located. This paper has not attempted to nor makes any conclusion on the human perception of infrasound.

ACKNOWLEDGEMENTS

The authors wish to thank the South Australian EPA for loan of the Soundbook and low frequency microphones. The authors also wish to thank Vipac Engineers & Scientists for loan of the head simulator.

REFERENCES

[1] Australian Broadcasting Corporation, 2012, *Inquiry boss makes noise over wind farm infrasound*, http://www.abc.net.au/news/2012-08-22/inquiry-boss-makes-noise-over-wind-farm-infrasound/4215074, (accessed on 2 November 2014)

[2] T. Evans, J. Cooper and V. Lenchine, "Infrasound levels near windfarms and in other environments", *South Australian Environment Protection Authority*, Adelaide, Australia (2013)

[3] B. Walker, D. Hessler, G. Hessler, R. Rand and P. Schomer, "A Cooperative Measurement Survey and Analysis of Low Frequency and Infrasound at the Shirley Wind Farm in Brown County, Wisconsin", *Clean Wisconsin*, Madison, (2012).

[4] Environment Protection Authority, "Waterloo Wind Farm – Environmental noise study", *South Australian Environment Protection Authority*, Adelaide, Australia (2013)

[5] K. Hansen, B. Zajamsek and C. Hansen, "Noise Monitoring in the Vicinity of the Waterloo Wind Farm", *University of Adelaide*, Adelaide, Australia (2014)

[6] Wellington Times, 2013, *Low-level noise a fear,* http://www.wellingtontimes.com.au/story/1316642/low-level-noise-a-fear/, (accessed on 2 November 2014)

[7] Suzlon, "S88 – 2.1 MW Technical Overview" (Product Brochure), Suzlon, Pune.

[8] T.M. Murray, D.E. Allen, and E.E. Ungar, "Floor Vibration Due to Human Activity. AISC Design Guide, Series No. 11", *American Institute of Steel Construction*, Chicago, (1997).

[9] Wikipedia, "Military Step", http://en.wikipedia.org/wiki/Military_step, (accessed on 1 April 2013)

[10] DELTA, "Low Frequency Noise from Large Wind Turbines: Summary and Conclusions on Measurements and Methods", EFP-06 Project Report prepared for Danish Energy Authority, Hørsholm (2008).

[11] H. Møller and C.S. Pedersen, "Hearing at low and infrasonic frequencies", *Noise & Health*, **6**(23), 37-57 (2004)

Technical Note

Note: Technical notes are aimed at promoting discussion. The views expressed are not necessarily those of the editors or the Australian Acoustical Society.

REMOTE BEEHIVE MONITORING USING ACOUSTIC SIGNALS

Amro Qandour[1], Iftekhar Ahmad[1], Daryoush Habibi[1], and Mark Leppard[2]
[1]Centre for Communications and Electronics Research (CCER), Edith Cowan University, Joondalup,WA 6027, Australia
[2]Communications Engineer, Rio Tinto, Perth, WA 6000, Australia
a.qandour@ecu.edu.au

Recent developments in Wireless Sensor Networks (WSNs) have led to their use in remote data acquisition and automatic data analysis applications, which have proven to be an invaluable tool in a diverse range of fields including biosecurity. Further indications have been found that honeybee health can be monitored and determined through the use of acoustic analysis. In this paper, we present a system that has the ability to remotely detect the presence of pest infestation on a colony of honeybees by comparing the acoustic fingerprint of a hive to a fingerprint of known status. This will aid the goals of increasing surveillance programs by reducing the labour time and costs that are associated with managing and maintaining monitoring programs. Other benefits of the system proposed in this article include the ability to make available a collection of deterministic, standardised and nondiscriminatory statistical data for the purpose of research into determining the causes of colony collapse disorder.

INTRODUCTION

The honeybee (Apis-mellifera) [1] is unquestionably considered as the most important and significant contributor among the animal pollinators, playing an essential role in the prosperity of the world's ecosystems and indeed to life itself. It is estimated that the honeybee is responsible for the pollination of over 90% of global commercial pollination services, and approximately 35% of the world's food crops [2]. The honeybee is probably best known for its production and storage of honey; however the economic value of the pollinator is not attributed solely to the hive produce, but largely to the products derived as a direct result from honeybee pollination. This constitutes an estimated $2 billion in revenue per year for Australia and $198 billion worldwide. In recent times, there have been rapid increases in agricultural development and human population, both of which are heavily dependent on the success of the honeybee industry. This has led to greater than ever demands for honeybee pollination, placing mounting strain on managed honeybee colony populations worldwide [3].

Conversely to this trend of increasing global demand, bee colonies around the world are under an increasing number of threats from a range of sources. The rapid spread of exotic pests such as the Varroa-destructor, better known as the Varroa-mite, is undoubtedly the biggest mortal threat to honeybees [4]. The Varroa-mite has already proven to be extremely damaging to the international honeybee industry as it has advanced throughout the world, and alarmingly, since the first reports of the arrival of Varroa-destructor in New Zealand in early 2000, Australia is now the only country free of the pest.

The work reported in this Technical Note has been undertaken with the support received from the AAS education award in 2013.

	Freq. (Hz)	Signal Pattern	Sender	Possible Sig.
Tooting	300 ~ 500	Pulse sequence	Queen	Prevent hatching of further queens and trigger quacking
Quacking	300 ~ 350	Pulse sequence	Queen	Presence detection, viabilityof confined queens
Hissing	300 ~ 3600	Single pulse	Colony	Warning signal
Piping	100 ~ 2000	Single pulse	Scout	Triggers colony hissing, prepare for swarming
Recruit	200 ~ 350	Pulse sequence	Forager	Existence and quality of valuable food source

Figure 1. Acoustic signatures of honeybee colonies

Invasive pest surveillance

Currently there are surveillance programs operating on a state-by-state basis aimed at the early detection of the Varroa-mite and other foreign pests and threats arriving in Australia. This current beehive surveillance and monitoring is achieved through the use of bait hives, which are located around Australia's major harbours and ports. These hives are situated such that any foreign bee infested with unwanted diseases or pests arriving at a port will inhabit the bait hives before spreading further. It has been shown [5] that the early detection of pests is imperative if an infestation is to be contained and eliminated. Though there are treatments available, it is commonly agreed upon within the honeybee communities that the prevention of an outbreak is better than a cure. Through regular manual examinations of the bait hives and the bee colonies that settle in them, inspectors are able to ascertain if there is any potential threat and subsequently intervene with the necessary steps to eliminate the threat. To satisfy this requirement of increasing surveillance intensity with the current method of monitoring would necessitate large numbers of qualified inspectors to travel to every site individually. This would require large amounts of manpower, technical expertise, time and consequently money in order for their continued success. The lack of any better option is largely a result of the honeybee industry residing outside of the focus of modern technological developments.

Technological intervention

Information from relevant sources [2, 6] have indicated that honeybees change their acoustic behaviour as a result of being exposed to certain stressors. This being so, it is reasonable that a potential solution to the pest infestation problem could be to use wireless sensor devices that can automatically detect the presence of infected colonies based on the colony's measured acoustic signature. Such a device would operate as a remote surveillance system and would aid current and future surveillance programs by providing the inspectors with tools such as advanced warnings and detailed analysis of hive health. In close to real-time, the system could provide alerts as to the arrival of a new colony to a bait hive, as well as the current health status of the new colony. This information would ideally include as much detail as possible regarding the what, where and when such warning has occurred. Consequently, this would allow inspectors to utilise their time in the most efficient and informed way, could potentially save large amounts of money, and revolutionise national biosecurity monitoring programs.

BACKGROUND AND RELEVANT WORKS

Honeybees have been observed to produce a variety of different sounds [7, 8] as forms of communication within the colony. Most of the sounds produced have been characterised by a low fundamental frequency between 300 and 600Hz and their corresponding harmonics [8]. The sounds produced by honeybees are one of the primary forms of communication within the colony; however communication is also achieved through the use of chemical means [9]. As shown in Figure-1, there is a range of sounds of different acoustic frequencies used by bees for a variety of reasons. Interestingly to note, it is not only the range of frequencies that are produced that determine the meaning of the noise, but also the acoustic structure in terms of signal pattern. The accurate quantification of the characteristics of these signal patterns and frequency ranges will be the key in developing a system that can identify possible threats to the hive and colony.

Various studies [10, 11] have shown that the health, status and activity of a honeybee colony can be determined through the analysis of the acoustic characteristics of the hive. Through the analysis of these studies it becomes unquestionable that honeybee hive acoustics change to reflect the current status and circumstance of the hive. The idea of this project was to design and develop a system which could understand and recognise the different acoustic characteristics produced by a healthy colony and a colony infested with Varro-mites.

SYSTEM OVERVIEW

This section deals with the main components of the beehive monitoring system (system architecture) and acoustical analysis techniques executed using dedicated acoustical software.

System architecture

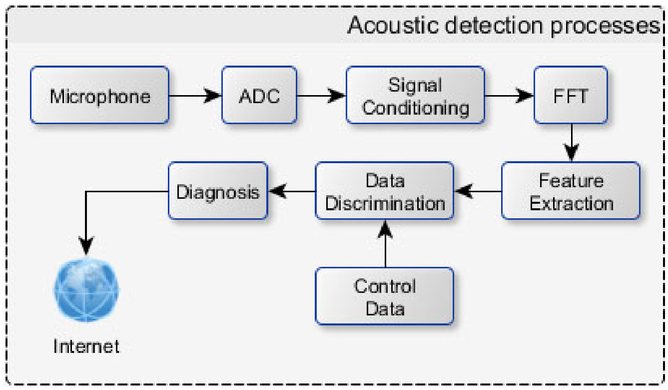


Figure 2. Acoustical detection process for honeybee colonies

The proposed honeybee monitoring system is designed to automatically acquire, process, and analyse audio data from remote honeybee hives to help alleviate the time and energy required for manual monitoring. To achieve this, the system must be as reliable and self-functional as possible. The key components of the honeybee monitoring system involve:

Sensor node

The Beagleboard [12] is an ideal platform for the honeybee system due to its miniaturised size (remote deployment) and intensive computational power (required for acoustical analysis tasks).

Sensors

The bee acoustics are acquired by using an electret microphone situated within the hive of the honeybee colony. The sound (honeybees) is picked up by the microphone and processed by the beagleboard using acoustical algorithms in order to discriminate between a healthy or infected hive.

Radio transmission

A low-power radio transceiver [13, 14] (Zigbee Link) is used to transmit the acoustical data from remote beehive sites to gateways. The data contains alarm messages in case an infection is detected, and diagnostics status to verify system operation (e.g. remaining power percentage).

Algorithm for acoustical analysis

The process depicted in Figure 2 is used to analyse the acoustic signatures of a honeybee colony. Training data (control) consist of sound samples of healthy and infected bee colonies. Acoustical comparison between training and live data (collected in real-time) will allow us to determine if the acoustic fingerprint of the hive correlate to the acoustic fingerprints of a hive infested with pests.

Acoustical analysis techniques

This section describes the techniques used to determine the acoustical properties of a beehive colony.

Acoustic features

The most commonly used acoustical features in acoustical analysis applications (e.g. honeybee monitoring) are:

- Peak Frequency (PF): PF can be defined as the frequency contains the highest (most) power for a given window of audio.
- Spectral Centroid (SC): SC is also known as the mean frequency – or gravity center – of the power spectrum of a frame.
- Bandwidth (B): The bandwidth of a signal is the range of frequencies present in a signal.
- Root Variance Frequency (RVF): The RVF feature component describes the convergence of the power spectrum for a given sample.

Data discrimination

One of the most fundamental goals of the honeybee monitoring system is to be able to accurately determine the status of hive health purely through the analysis of its acoustic fingerprint. The honeybee system is designed with binary categorical discriminant analysis functionality (rather than regression analysis). The classification tools used to perform this type of analysis are:

- Principle Component Analysis (PCA): PCA is an exploratory data analysis tool used for making predictive models [15]. Commonly implemented as a form of dimensionality reduction, it involves finding the eigenvalues and eigenvectors of the covariance matrix of the meansubtracted data set [8, 11, 16].
- Support Vector Machines (SVM): SVM are machine learning models built around algorithms designed to analyse data and recognise patterns. The application of SVMs to binary classification problems have been shown to perform exceptionally even for large dimensional vectors [15].
- Linear Discriminant Analysis (LDA): LDA is used to find the linear combination of a set of features which maximises the separability between the classes [17].

ACOUSTIC MODEL

This section describes the acoustical model used to analyse the audio signatures of beehive colonies. The aim is to be able to differentiate between a healthy or an infected beehive colony. The main processes involved are: a. control training, b. audio classification, and c. feature discrimination.

Training

The classifier training relies on input-signals from known data-sets (various infected/healthy beehive sound samples) in order to build accurate training data-sets [18, 19]. Four approaches were taken to perform the training of the classifiers. Firstly, the feature sets were passed through a PCA algorithm which narrowed the four features down to two. This reduced feature set is then put through an LDA algorithm and as well as an SVM algorithm. In these two scenarios, the LDA and SVM classifiers use the features chosen by PCA to establish their discriminant functions. They will be referred to herein as PCA_LDA and PCA_SVM.

Table 1. Selection from the feature list of the control data-set. Label "1" represents infected samples. Label "−1" represents healthy samples

F_{ID}	PF (Hz)	SC (Hz)	B (Hz)	RVF (Hz)	Label
1	1036	960	146	7211	1
2	998	971	168	7098	1
3	942	954	198	7239	1
4	1052	1025	125	4210	-1
5	1114	992	157	4767	-1
6	1028	1011	158	4709	-1

In the other two methods, the original feature sets are passed into an LDA algorithm and an SVM algorithm without first going through PCA. This means that the results of the last two methods are purely dependant on the features chosen by LDA and PCA. The idea of this analysis was to compare which of the four overall methods had the highest accuracy of classification, and thereby establish the best method of generating classifier functions for distinguishing between infected (labelled as 1) and healthy (labelled as -1) honeybee colonies as depicted by Table 1.

Classification

Both the LDA and SVM classification algorithms return prediction values for the given test data after applying their respective methods [20, 21]. The test data consist of a 10 second recording of either healthy or infected honeybee hive samples. The result of the prediction for both methods is an array of the same length as the number of frames observed. This means the classifying functions output either a '1' or a '1' for each frame depending on whether that frame matches the fingerprint of an infected hive or a healthy hive as seen in Table 1.

Implementation

A PCA algorithm was developed so that the feature set extracted from the system could be tested for suitability for use in a classifier system. To do this, a number of control experiments were conducted in order to establish what types of results could be expected from different inputs. The first experiment was conducted primarily to confirm that the set of acoustic features that had been extracted for use in the system were suitable enough to allow reasonably unique and independent fingerprinting functions to be generated for sets of predominantly similar audio data.

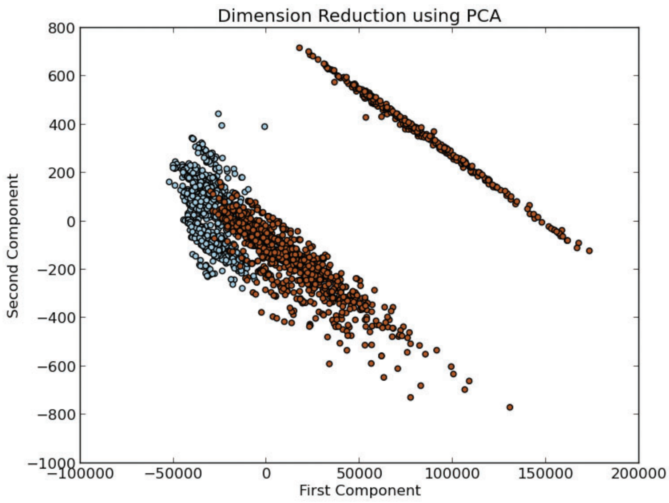


Figure 3. Separability of acoustic features using Principle Component Analysis (PCA) - healthy (blue) / infected (red) / first-component (RVF) / second-component (B)

As depicted by Figure 3, there is a clear separation (i.e. minimal overlapping between the coloured regions/points) between healthy and infected features in the acoustical domain of a beehive colony (mainly between RVF and B). This analysis confirms the suitability of PCA in determining the highest abnormalities in the acoustical spectrum components (i.e. healthy and infected feature separation). The next stage involves determining the linear relationship (classifier function) that best describes the highest degree of separability between healthy and unhealthy colonies using the principle components selected by the PCA. Once established (i.e. relationship is found), a set of test data is fed into the classifier function in order to make future predictions on the status of the bee colony involved.

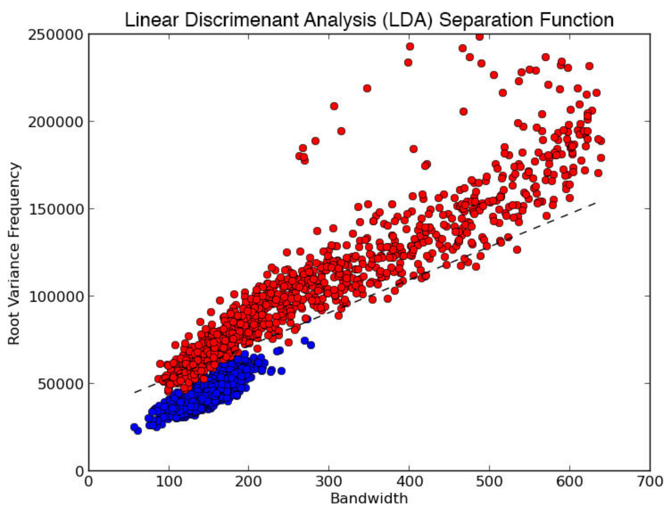


Figure 4. Classification analysis using Linear Discriminant Analysis (LDA) - healthy (blue) / infected (red)

Figure 4 shows the plot of the discriminant function generated by the LDA algorithm. The results indicate the best separation between healthy (blue) and infected (red) samples are derived from the bandwidth and root variance frequency components. The dashed line shows the threshold of classification based on the features chosen (i.e. B & RVF) to generate the corresponding plot.

In the next section, we will perform real-life experiments to test the validity of our acoustic model in detecting the presence of the Varro-mite pest in beehive colonies.

EXPERIMENTAL ANALYSIS

One of the most important factors to keep in consideration during the analysis of these results is that the training and test data was limited to a small number of low quality samples. This essentially means that the sounds had gone through a number of re-sampling processes by the time it was recorded onto our system for analysis, and as such, they had obviously undergone a significant degradation in quality. Additionally, with limited samples available, the choice was made to use the 5 healthy honeybee hive recordings obtained from [22] as both the training and test data for the healthy hive classification. Due to the fact that only one infested hive audio sample could be obtained, the training data was generated using 5 recordings of the same sample. In this section, we reveal the acoustical patterns associated with beehive colonies infected with the Varro-mite pest. We illustrate these patterns using classification techniques widely used in the acoustical analysis domain:

Support vector machine (SVM)

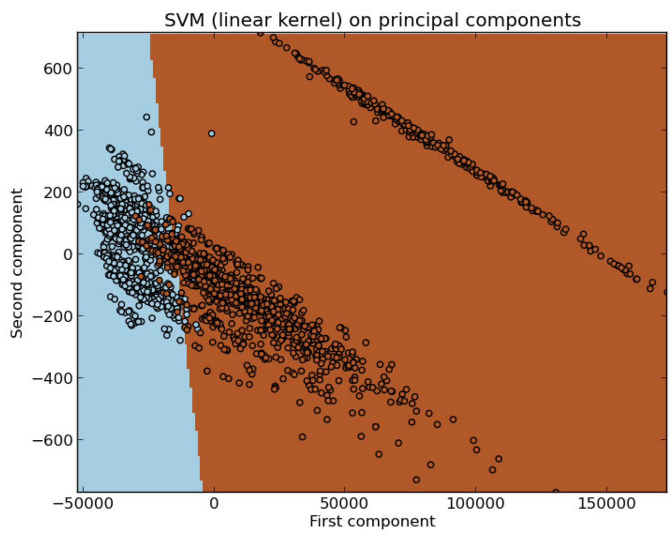


Figure 5. A depiction of healthy region (blue) vs infected region (red) using SVM - first-component (RVF) / second-component (B)

Figure-5 illustrates the regions where healthy (blue region) or infected (red region) components (PCA output data) reside within. This method of separation is generated using automatic scripts implemented on the target system (beehive node) and requires little human intervention in the final deployment.

Linear discriminant analysis (LDA)

Another similar approach which can be used to classify acoustic features is Linear Discriminant Analysis (LDA). This method is used to determine the highest separability function which can be generated from a given data set. The results in Figures 6,7 illustrate the use of separability functions in order to make future predictions on unknown data sets, and to classify the data as either healthy or infected.

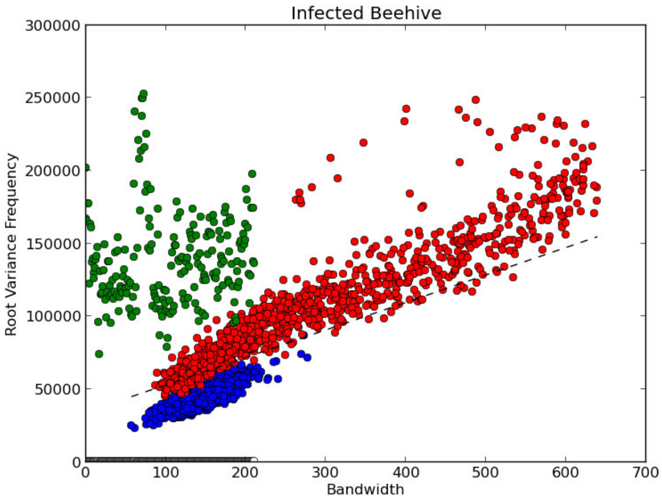


Figure 6. Varro-mite infected beehive acoustical response computed using LDA - healthy (blue) / infected (red) / test-data (green)

The blue and red features are used to differentiate between the healthy (blue) or infected (red) beehive acoustic characteristics. The green features represent a data set from an unknown beehive. By visual inspection, there is a clear indication that the data set (green) from Figure-6 resides mostly within the infected zone. Similarly in Figure 7, the green features statistically seem to indicate the beehive is healthy.

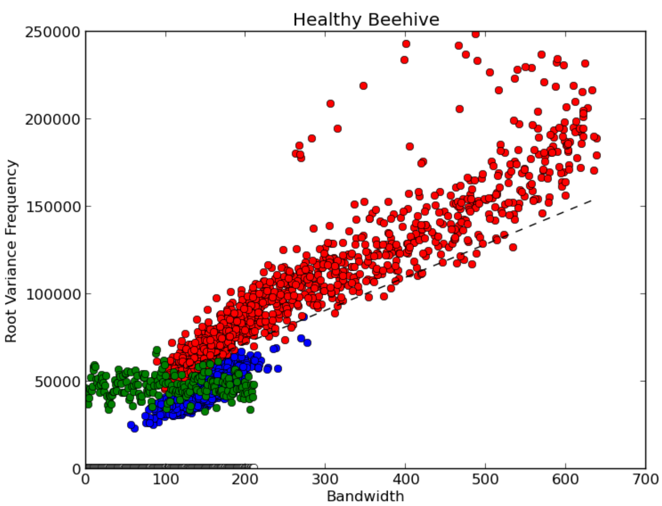


Figure 7. Healthy beehive acoustical response computed using LDA - healthy (blue) / infected (red) / test-data (green)

Discussion

Our beehive detection system demonstrated considerable accuracy with all four features used as discriminant functions. This however came at the cost of a higher computational time. With an improved quantity and quality of training data it would most likely be found that the discriminating capabilities of the system could be significantly improved. Also, it has been shown that both LDA and SVMs are both capable of generating predictions accuracy percentages which are better than "chance" based percentages. Depending on the outcome of these future tests, it may be decided that PCA is only needed to be incorporated as far as the development stages, to help identify which features should be extracted and used in the final design.

CONCLUSION

In this paper, we presented a system prototype for remote monitoring of beehives. This is rather a non-intrusive approach of dealing with pest infestation problem in the honeybee industry. The developed prototype is capable of capturing and analysing of acoustic samples collected from beehives. We showed how to extract features and train a classifier that can predict the infestation status of a beehive. However, further research is still required to complete the system. Several ways in which this research can be improved include:

1. Acquisition of more bee samples: One of the most important aspects for the next phase of the projects development will be to attain a more comprehensive set of control data to train the system with.
2. Expanded feature set and classifiers: There are a much greater range of features that could be used and tested for use in the honeybee system. It would therefore be a desirable process in the next stages of development that an expanded number of features be extracted from the audio samples, in order to create larger feature sets from which to compute the acoustic fingerprints.
3. Memory/Data management: Since the memory resources available to the system are limited, further methods of ensuring efficient memory and data management, as well as minimising any redundant processes should be implemented.

ACKNOWLEDGEMENT

The authors wish to acknowledge and thank the AAS committee for providing the necessary funds to realise the work presented in this technical note. Furthermore, we would like to extend our sincere gratitude to all peer-reviewers involved in the AAS for all their efforts and valuable feedback.

REFERENCES

[1] F. Ruttner et al., *Biogeography and taxonomy of honeybees*. Springer-Verlag, 1988. (document)

[2] Honeybee varroa vibration and ccd. [Online]. Available: http://modernsurvivalblog.com/natural-disasters/the-honeybee-varroa-vibration-and-ccd/ (document)

[3] A.-M. Klein, B. E. Vaissiere, J. H. Cane, I. Steffan-Dewenter, S. A. Cunningham, C. Kremen, and T. Tscharntke, "Importance of pollinators in changing landscapes for world crops," *Proceedings of the Royal Society B: Biological Sciences*, vol. 274, no. 1608, pp. 303–313, 2007. (document)

[4] D. Sammataro, U. Gerson, and G. Needham, "Parasitic mites of honey bees: life history, implications, and impact," *Annual review of entomology*, vol. 45, no. 1, pp. 519–548, 2000. (document)

[5] P. Boland, "A review of the national sentinel hive program," *Biosecurity Australia*, 2005. (document)

[6] T. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," *Magnetics in Japan, IEEE Translation Journal on*, vol. 2, no. 8, pp. 740–741, Aug 1987. (document)

[7] L. Fahrenholz, I. Lamprecht, and B. Schricker, "Calorimetric investigations of the different castes of honey bees, apis mellifera carnica," *Journal of Comparative Physiology B,* vol. 162, no. 2, pp. 119–130, 1992. (document)

[8] M. Hrncir, F. G. Barth, and J. Tautz, "32 vibratory and airborne-sound signals in bee communication (hymenoptera)," *Insect Sounds and Communication: Physiology, Behaviour, Ecology, and Evolution*, p. 421, 2005. (document)

[9] M. L. Winston, *The biology of the honey bee*. Harvard University Press, 1991. (document)

[10] D. Atauri Mezquida and J. Llorente Martínez, "Short communication. platform for bee-hives monitoring based on sound analysis. a perpetual warehouse for swarm apos; s daily activity," *Spanish Journal of Agricultural Research*, vol. 7, no. 4, pp. 824–828, 2009. (document)

[11] Y. Le Conte, M. Ellis, and W. Ritter, "Varroa mites and honey bee health: can varroa explain part of the colony losses?" *Apidologie*, vol. 41, no. 3, pp. 353–363, 2010. (document)

[12] G. Coley, "Beaglebone black system reference manual," 2013. [Online]. Available: http://elinux.org/Beagleboard:BeagleBoneBlack (document)

[13] Z. Alliance, "Ieee 802.15. 4, zigbee standard," *On http://www.zigbee.org*, 2009. (document)

[14] J. A. Gutierrez, E. H. Callaway, and R. L. Barrett, *Low-rate wireless personal area networks: enabling wireless sensors with IEEE* 802.15. 4. IEEE Standards Association, 2004. (document)

[15] V. N. Vapnik and V. Vapnik, *Statistical learning theory*. Wiley New York, 1998, vol. 2. (document)

[16] M. Shen, L. Cui, N. Ostiguy, and D. Cox-Foster, "Intricate transmission routes and interactions between picorna-like viruses (kashmir bee virus and sacbrood virus) with the honeybee host and the parasitic varroa mite," *Journal of General Virology*, vol. 86, no. 8, pp. 2281–2289, 2005. (document)

[17] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2005. (document)

[18] D. Albanese, R. Visintainer, S. Merler, S. Riccadonna, G. Jurman, and C. Furlanello, "mlpy: Machine learning python," *arXivpreprint arXiv:1202.6548*, 2012. (document)

[19] T. E. Oliphant, *A Guide to NumPy*. Trelgol Publishing USA, 2006, vol. 1. (document)

[20] B. Scholkopft and K.-R. Mullert, "Fisher discriminant analysis with kernels," *Neural networks for signal processing IX*, 1999. (document)

[21] I. Steinwart and A. Christmann, *Support vector machines*. Springer, 2008. (document)

[22] M. Ward, "Bee hive hums recorded to monitor insects' health," Technology correspondent, BBC News, http://www.bbc.com/news/technology-16114890. (document)

MODELLING THE INTERIOR SOUND FIELD OF A RAILWAY VEHICLE USING FINITE ELEMENT METHOD

WU Di[1] and GE Jianmin[1]
[1]Institute of Acoustics, Tongji University, Shanghai, China
wdi0225@163.com

The interior sound field of a railway vehicle was modelled using Finite Element method. The Finite Element model was composed of air cavity and seats. The boundary of the compartment was modelled as rigid surface and no fluid-structural coupling was considered. The simulated sound pressure level spectrum is in overall good agreement with the measurement result.

INTRODUCTION

At present, the statistical energy analysis and ray-tracing method are commonly used to predict the interior sound field of a railway vehicle [1-3]. These two methods could provide reasonably accurate predictions in middle and high frequencies, but will bring fairly great error in low frequency because of the diffraction effects. According to previous studies, low frequency noise is particularly dominant in modern high-speed trains [4]. In order to study the low frequency sound field with high accuracy, Finite Element (FE) method was used in this paper to model the compartment. The FE model was validated by experimental results.

FINITE ELEMENT MODELS

In this paper, two major elements were taken into consideration during the modelling process - the enclosed cavity of the compartment and the seats inside the compartment. In fact, coupling effect was created between the fluid and structure domain. However, accurately modelling the train's structure was almost unachievable with the existing processing power considering its complexity. In this case, greater error would have been produced if the effect of the structure was taken into account. According to V. Jayachandran's research [5], there is little difference between sound pressure distribution in rigid surface enclosure and in vibrating surface enclosure, though the difference is significant at the boundary. In this paper, rigid surface was used to model the boundary, and this simplification has proved to be reasonable according to the simulation results as shown in RESULTS AND DISCUSSIONS.

The geometrical model of the compartment was set up within CATIA, and its size and layout were determined based on the fifth compartment of CRH380AL, a Chinese electric high-speed train. The FE meshes were obtained using HyperMesh. The cavity and seats were meshed with tetrahedron elements using an average size of 0.17 m. The FE mesh of the compartment was shown in Figure 1. The software Actran was used for the FE simulation.

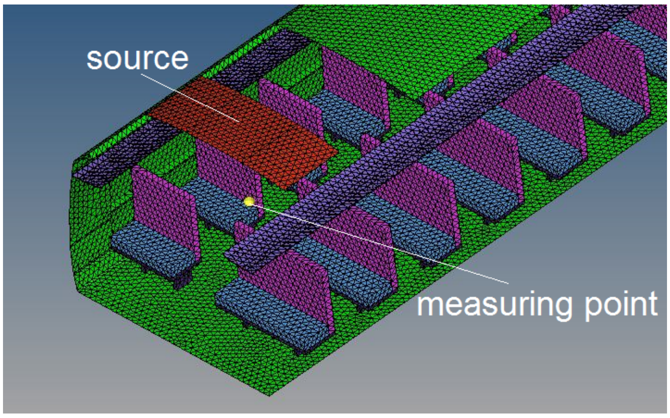


Figure 1. FE mesh of the compartment.

The most important source of interior noise of high-speed trains is aerodynamic noise generated by pantograph [4], and thus an area source was placed at the location of the pantograph. The simulation frequency is from 100 to 250 Hz, in which region the sound energy can be a source of human fatigue [4]. The interval of the simulation frequency is 1 Hz. The computing work was completed on a laptop with a 1.6 GHz dual-core processor and 4 GB memory, and the computing time was around 3 hours.

EXPERIMENT

To verify the FE model, a measurement was performed inside the CRH380AL's compartment. During the measurement, the train was stationary. A loudspeaker driven by white noise was set at the location of pantograph and it was 0.5 m above the roof. Data acquisition and processing were carried out by B&K PULSE platform with recording time of 15 s.

RESULTS AND DISCUSSIONS

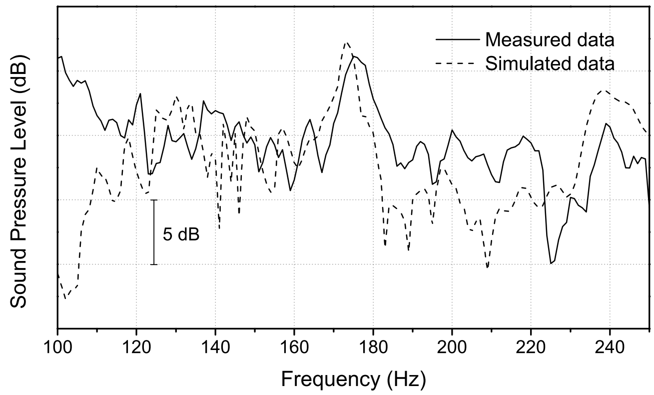


Figure 2. Comparison of SPL spectra obtained from measured (—) and FE simulated (- - -) data.

A measuring point was set inside the compartment, right below the pantograph and 1.2 m above the ground. As shown in Figure 2, the simulated spectrum basically agrees well with the measurement result. Compared to the measured curve, the simulated curve shifts about 2 Hz to lower frequencies. The reason is that the boundary is actually not rigid, using rigid surfaces in the model slightly affects the peak positions of the simulated result. Additionally, the error is relatively large at 100 Hz, the reason for which requires further analysis. The error is less than 5 dB in the frequency region from 110 to 250 Hz.

CONCLUSIONS

The FE method could predict low-frequency interior sound field of railway vehicle with high accuracy. The time cost of the computing work is acceptable.

ACKNOWLEDGEMENTS

The author is grateful to Yan Shuang for her help during the study.

REFERENCES

[1] F. Jens, T. Stefan, C.A. Can, F. Anders and K. Wolfgang, "Modelling the interior sound field of a railway vehicle using statistical energy analysis", *Applied Acoustics* **73**, 307-311 (2012)

[2] N.J. Shaw, *The prediction of railway vehicle internal noise using statistical energy analysis techniques*, Heriot–Watt University, UK, 1990

[3] B. Stegeman, *Development and validation of a vibroacoustic model of a metro rail car using Statistical Energy Analysis (SEA)*, Chalmers University, Sweden, 2002

[4] D. Thompson, *Railway Noise and Vibration*, Elsevier, 2009

[5] V. Jayachandran, S.M. Hirsch and J.Q. Sun, " On the numerical modelling of interior sound fields by the modal function expansion approach", *Journal of Sound and Vibration* **210**, 243-254 (1998)

Technical Note

Note: Technical notes are aimed at promoting discussion. The views expressed are not necessarily those of the editors or the Australian Acoustical Society. Contributions are not formally peer-reviewed.

CONSTRUCTIVE INTERFERENCE OF TONAL INFRASOUND FROM SYNCHRONISED WIND FARM TURBINES: EVIDENCE AND IMPLICATIONS

Andrew Bell
John Curtin School of Medical Research, Australian National University, Canberra ACT 0200
andrew.bell@anu.edu.au

SUMMARY

Noise from wind farms is contentious: people who live nearby complain of annoyance, and yet broadband measurements of infrasound seem to indicate the noise is generally not above audibility criteria. The paradox can be resolved by supposing that wind farms generate a strong tonal signal at the blade passing frequency, 0.8 Hz, and that this infrasound, with a wavelength of 400 m, can constructively interfere if two or more wind turbines operate in synchrony and the path lengths differ by a multiple of 400 m. Coherent infrasound at 0.8 Hz could propagate many kilometres, would tend to carry many harmonics due to the rapid changes within its waveform, and the high harmonics in the 20–30 Hz band have the potential to be heard by human ears. The existence of coherent infrasound from wind turbines has not been specifically recognised, but evidence of the phenomenon can be discerned in two anomalies contained in data from recent infrasound monitoring of wind farms in South Australia. This paper interprets the anomalies in terms of a model which suggests that wind farms produce enhanced sound pressure levels when the blades of multiple machines become mutually entrained and the sound from them becomes coherent. The inference is that acoustic measures, which assume wind turbine signals are stationary, may not be accurate indicators of peak noise levels.

BACKGROUND

There has been much debate about the impact of wind turbine noise on people living next to wind farms. In Australia, some of that debate has been conducted in the pages of this journal (e.g. [1]). While there have been many complaints of annoyance from neighbouring residents, the complaints have often been dismissed because acoustic measurements indicate the levels of infrasound should be inaudible since they do not rise above background noise levels [2]. The National Health and Medical Research Council recently commissioned a systematic review of the problem [3], and it issued for comment a Draft Information Paper summarising its findings [4]. The paper states that "Evidence suggests that levels of infrasound are no higher in environments near wind turbines than in a range of other environments. For example, a South Australian study [Evans et al. 2013] observed similar levels of infrasound at rural locations close to wind turbines, rural locations away from wind turbines, and at a number of urban locations" ([4], p.12).

This paper assesses the Evans et al. measurements, along with some related findings, and finds there are certain limitations which may put a question mark over such a conclusion. The synthesis here highlights the relevant measurements, describes possible limitations in interpretation, and recasts the findings in terms of coherent infrasound. It is pointed out that the waveform generated by blade–tower interaction (BTI) carries many harmonics that can reach into the audible range, and it is suggested that these high harmonics, when emitted by synchronised wind turbines, could be a major cause of the problem. A recent proposal [5] for minimising the infrasound problem by desynchronising the blades is endorsed.

INFRASOUND

As stated by the NHMRC, "infrasound is considered by some to be an important component of noise from wind farms" (p.12 of [4]). In this context the recent work of Thorne [6, 7] is particularly relevant, for he points out how constructive interference from synchronised wind turbines can lead to "heightened noise zones", a term developed by him in conjunction with Bakker and Rapley [8]. Thorne considers the synchronous rotation of wind turbines (p.42), although he appears to be more concerned with amplitude modulation of blade swish than with direct propagation of 1 Hz infrasound and its harmonics. His simulations were done with emission frequencies of 20, 48, and 66 Hz (p.51). The work of Doolan et al. (2012) [5] is also of interest, for they theoretically consider BTI, and derive a curve (shown here as Figure 1) for how the amplitude of the generated pressure varies over time. This theoretical waveform indicates that the signal has a large and sudden variation in pressure at the blade passing frequency (1 Hz). A spectrum of this waveform (their Fig. 5) shows the expected sequence of multiple harmonics.

Doolan and colleagues note that there is currently no methodology to accurately quantify BTI noise, but they say it could be important. They draw a diagram of how two BTI sources, which they describe as "temporally coherent", may constructively interfere (their Fig. 6), perhaps at a residence, and suggest that actively desynchronising the turbine blades may be a way to avoid the heightened pressure pulse.

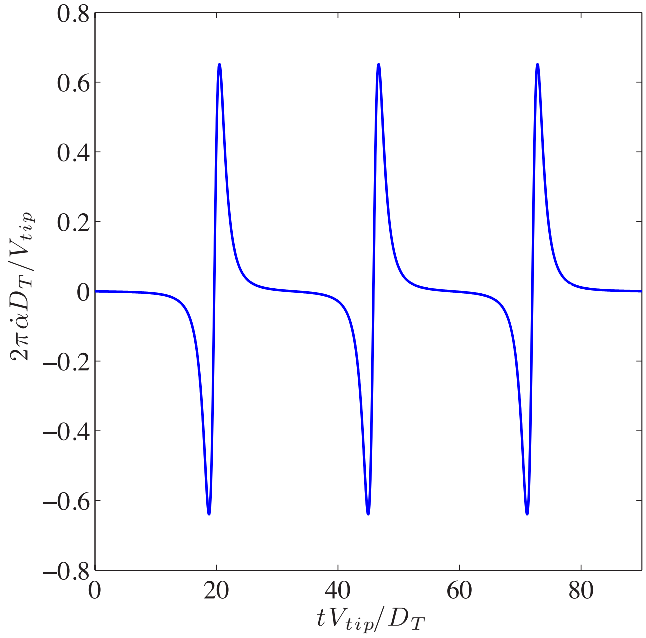


Figure 1. The acoustic waveform generated by a wind turbine as its blades pass the supporting tower (theoretical calculation of blade–turbine interaction from Doolan et al. (2012) [5], with permission). The frequency of the waveform is about 1 Hz and is tightly controlled, creating the possibility of the pressure pulse from one set of turbine blades to entrain another set. The steeply rising portion of the waveform makes it well suited for synchronisation with other such waveforms; it also produces many harmonics.

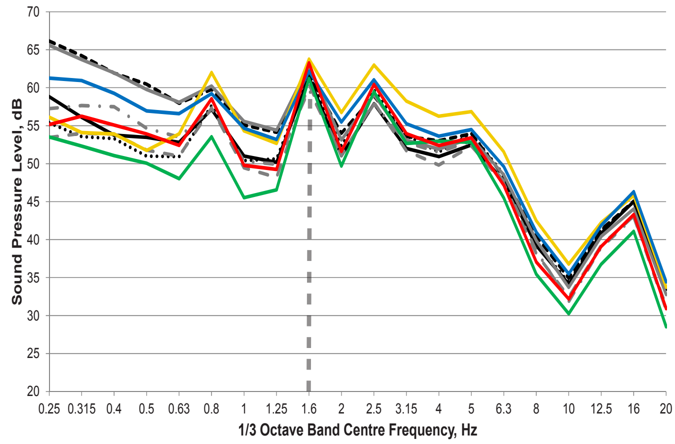


Figure 2. Absolute sound pressure levels in 1/3-octave bands from 0.25 Hz to 20 Hz measured by Evans et al. at Bluff Wind Farm, South Australia. Note the distinct peaks at 0.8 Hz, 1.6 Hz, and 2.5 Hz. The different traces are for different times, 4 with the wind farm operating (coloured lines), and 6 with it switched off (black and grey). Note the apparent anomaly that at 1.6 Hz (added vertical line) there is less than 3 dB difference between all measurements, *including between on and off*. Modified from [9], with permission.

This paper validates the idea that strong 1 Hz infrasound can arise by constructive interference of many wind turbines emitting signals such as in Figure 1. It also adds the idea that synchronisation of wind turbines may be promoted by the tendency of oscillators of closely matching frequencies to become physically entrained. The possibility that the blades of multiple wind turbines become locked together is a hypothesis that this technical note puts forward for further discussion.

Below, the implications for audibility of multiple entrained turbine blades are set out, and a mechanism is described whereby the BTI harmonics may be sensed by the ear. The paper emphasises that the control circuits used in South Australian wind turbines to regulate power output act to bring many turbines to almost the same rotational frequency, and hence are liable to produce entrainment and increase infrasound energy levels. Supporting evidence from recent South Australian monitoring data [9, 10] is discussed.

Signature of the blade pass frequency

Evans et al. (2013) [9] measured infrasound emitted from two wind farms, Bluff and Clements Gap, and the former showed particularly interesting results. The researchers found that at Bluff the infrasound contained clear peaks at the blade pass frequency of 0.8 Hz and at its harmonics of 1.6 Hz and 2.5 Hz, and the authors describe how these spectral peaks are a characteristic signature of a wind turbine's revolving blades. The peaks can be clearly seen in Figure 29 of their work, and are shown here in Figure 2. They are also visible in their Figures C3–C8, and in one figure for Clements Gap (the lower trace of Figure C9).

There are two important properties of these peaks which should be emphasised. First, these infrasonic signals have very long wavelengths – a frequency of 0.8 Hz, for example, has a wavelength of about 400 metres.

Second, these signals have a narrow bandwidth because the turbine blades in a wind farm are, for reasons of generation efficiency, regulated to maintain a constant rotational speed. The Suzlon S88 turbines installed at the wind farm measured by Evans and colleagues maintain a relatively constant rotational speed of 15–17 rpm, largely independent of wind speed variations (16 rpm = 3.75 sec/rev = 1.25 sec/blade for 3 blades = 0.8 Hz blade pass frequency). This factor deserves greater emphasis. For the Suzlon S88 the set speed at rated power is particularly precise, 15.79 rpm (p.36 of [11]), a speed which is electronically controlled and equivalent to a blade pass frequency of 0.7895 Hz. Therefore, if operating conditions permit, the rotation rate is fixed to a potential accuracy of 1 part in 1500 (a stability of 0.07%). Even if the windspeed fluctuates a little, the electronic controller will attempt to keep the rotational speed fixed at that optimum value. In other words, for windspeeds at and above the rated wind speed (12 m/s for the Suzlon turbine), the unit operates on a nearly *vertical portion* of the rotational speed–torque space, virtually independent of wind speed. The design of the Suzlon S88 allows a degree of slip between the blades and the generator, a factor which smooths out speed fluctuations but also has the side-effect of leaving the turbine (and its neighbours) open to mutual entrainment in which blade rotations become locked together.

Of course, there are a range of turbine types, each with its own operating characteristics set by a controller circuit. However, in South Australia the Suzlon S88 predominates [12]. The controller sets the operating characteristic, essentially a plot of torque against rotational speed, with the aim of optimising the power output as windspeed changes [13]. The controller and its algorithms are proprietary and are kept confidential by manufacturers. Nevertheless, once the wind speed reaches a point at which the turbine has reached its rated output, the general operating strategy is to actively adjust the torque to keep the rotational speed of the blades constant (Ch. 7 of [14]). Most turbines, including so-called 'variable speed' designs, have characteristics with vertical, or steep, portions along which the rotational speed is dynamically regulated for optimum power extraction (e.g., Fig. 4.9 of [13]; Fig. 7.12 of [14]; [15]).

A potential outcome of this tight, but slightly accommodating, regulation is that, for moderate windspeeds and above, all the blades in a wind farm might turn at almost exactly the same speed, making it highly likely, as explained in more detail below, that at least some will become synchronised. It is not uncommon to see the blades of a wind farm passing the tower at the same time (see Figure 3), and the likelihood that this occurs merely by chance are minute. Informal observation of a wind farm near Canberra, where Suzlon S88s are installed, showed that when two neighbouring 3-blade turbines were visually aligned so that one was in front of the other (producing a 6-pointed star), this configuration would sometimes stay perfectly fixed for more than a minute.

Figure 3. Monitoring wind turbine noise at an operating wind farm, Waterloo, South Australia. Notice that the blades of the four turbines have closely matched phase. Considered as a random event, the chance of 4 turbines having all phases equal to within 5° of each other are $(5/360)^3$ = about 1 in 2 million. Photo: EPA South Australia, with permission.

Such synchronisation implies that the blades will act as *coherent sources* of infrasound energy as they pass the towers. Whenever the generators come into synchrony, because of control circuits and physical entrainment, a wind farm could become a strong infrasound source. In this circumstance, it is inevitable that the waves will constructively interfere at various points (and destructively interfere at others). Doolan and colleagues discuss this theoretical possibility [5], and suggest a remedy in terms of adjusting each turbine so that the blades are never synchronous; however, at that time the data was not available to prove their point. This paper brings together findings which give the idea added plausibility. In fact, recent data from the same group [10] points directly to synchronisation at work, and this is evaluated in a later section.

Entrainment produces coherence

The key factor promoting a synchronous state is the universal phenomenon of mutual entrainment. It was first documented by Huygens when he noticed that two small pendulum clocks attached to the same wall became synchronous. Synchronisation of oscillators of closely matched frequency is now widely recognised in many branches of physics. It is therefore to be expected that physical coupling between turbine blades will occur aerodynamically, perhaps giving rise to the synchronicity seen in Figure 3. Interactions via electrical loading of the generators might also contribute to synchrony, and this idea has been explored for fixed-speed generators [16-18]. These studies, together with informal observations and the clear evidence of Figure 3, strengthen the possibilities outlined by Thorne [6, 7] and Doolan and colleagues [5].

Entrainment of a collection of oscillators is relatively simple to grasp intuitively (see the synchronisation of 32 metronomes at http://www.youtube.com/watch?v=kqFc4wriBvE), but in practice it is difficult to predict. It depends on multiple mutual interactions which are essentially chaotic [19]. If a number of wind turbine blades become synchronised and the sources become coherent, sound pressure levels will increase. For example, two wind turbines would have an intensity 6 dB louder than a single turbine, and four might have an intensity 12 dB louder. When one considers that a wind farm may have dozens of wind generators, the chances that some of them will become synchronised increases.

Unlike a set of metronomes, it is not suggested that all the turbines will become synchronous because their spacing may not support it, but the possibility is enhanced by the wavelength of infrasound (0.8 Hz = 400 m) being about the same as the spacing between wind generators, which are often placed in a line along a ridgetop (e.g., Fig. 3 of [20]). Allowance will need to be made for the propagation times between towers and for the distances from the towers to points of constructive interference. Such analysis is beyond the scope of this technical note. Although the analogy is imperfect, a simple comparison that conveys the concept is a swarm of cicadas. Each cicada emits its own regular chirping; however, because of mutual sensing and interaction, at some point the whole swarm can synchronise its sound and sing in chorus. At this point, the sound pressure becomes appreciably louder.

Another relevant factor is that a line of wind generators may be considered a line source rather than a point source. This means that, at right angles to the line, the intensity will fall off much more slowly with distance ($1/r$ rather than $1/r^2$); moreover, in directions along the line, the intensity might reach levels much greater than with a single generator and propagation could extend in a beam for tens of kilometres, especially since atmospheric attenuation of infrasound is small.

Coherence leads to nodes and antinodes

Constructive interference from coherent sources means that infrasound pressure at the blade passing frequency (and its harmonics) could be high, even at locations far away. It is also the case that coherent sources could interfere destructively, but given the generally equi-spaced arrangement of wind turbines in a farm (see for example Fig. 9 of [21]), the actual situation is likely to be an alternating pattern of constructive and destructive points (nodes and antinodes). Thorne [7] illustrates this for low frequency sound (his Figs 21–25), although his analysis does not extend to infrasonic frequencies where the effects will be even more marked. In terms of impact on local residents, the pattern of nodes and antinodes will depend on location and the wavelength of each harmonic, and will shift with wind, temperature, and atmospheric stability. Importantly, there will also be fluctuations in phase on the scale of tens of seconds due to the blades slipping in and out of synchrony – an entrainment effect which will lead to sudden phase jumps.

Broadband infrasound measurements can miss tonal components

If the coherent infrasound hypothesis is true, then the method of measuring the broadband sound pressure level as a single G-weighted measure encompassing energy over 0.15–315 Hz (as done by Evans and coworkers) may not be the most appropriate in terms of infrasound audibility. The authors measured the broadband sound as about 50–60 dB SPL and then compared it to natural sources over the same band. Their conclusion, endorsed by both the Systematic Review [3] and the Information Paper [4], is that there were "similar levels of [broadband] infrasound at rural locations close to wind turbines, rural locations away from wind turbines, and at a number of urban locations" (Section 6.1, paragraph 7). This creates a puzzle, for why is it that people living near wind farms complain of rumbling sounds, often with 1 Hz periodicity, while people in the city do not, and why are the annoyance levels so variable over time? [22]. Adding to the puzzle are some of the authors' one-third octave measurements below 20 Hz: on some occasions there was no difference in recorded level at 1.6 Hz between when the turbines were operating and when they were switched off (see Fig. 2).

There is a possible explanation for both these puzzles, and it is that the G-weighted measures over 0.15–315 Hz are not capturing the audibility of fluctuating narrow-band infrasound that slips in and out of phase. Although G-weighting is a general measure intended to reflect the audibility of low frequency sound, it may not accurately represent the audibility of sounds which incorporate many harmonics that reach into the audible range (20 Hz and above). There is work to show that the perception of tones and broadband noise is similar [23], but this is for single tones only, not sets of harmonics. In the latter case, if all harmonics are in phase, the *peak pressure* will be greater, and ref. 23 suggests that infrasound may be detected via a peak detection mechanism, so that the peak pressure is more important than the RMS pressure. In addition, since the loudness range of infrasound is very compressed, a small rise in pressure can lead to a substantial increase in loudness, and this applies both to constructive interference of multiple coherent infrasound and to the in-phase superposition of its harmonics. Moreover, ref. 23 describes how infrasound may be easily sensed by modulation of low frequency sounds, and so any harmonic content above 20 Hz could promote turbine audibility. More research is needed in this area. A related issue is whether the 1 Hz periodicity people hear derives directly from the infrasound or from modulation of the constant aerodynamic noise emitted by the moving blades (see Fig. 6 of [5] and [24]).

An additional factor that may make low-frequency sound more audible than a single measurement indicates is the use of long averaging times (several minutes). This means that instants when the infrasound is at its loudest (when the sources lock together) are averaged along with the quietest times (when the sources are out of phase). There is therefore the possibility that the measured signal is not stationary, as standard analyses assume, but a fluctuating sequence of in-phase and out-of-phase conditions. The question of the enhanced audibility of such a signal will be considered later, but at this point it is informative to look closer at the evidence that coherent infrasound exists.

DIRECT EVIDENCE FOR INFRASOUND COHERENCE

Two intriguing anomalies have recently been reported. The first, by Evans et al. (2013), is contained in a major report on infrasound measurements from windfarms and other locations in South Australia [9]; the second, by Zajamsek and colleagues [10], was presented to a 2013 conference on wind turbine noise and again concerns infrasound measurements at a South Australian wind farm. It is suggested that both anomalies can be explained by the presence of coherent infrasound.

No difference between 'on' and 'off'

The anomaly found by Evans and colleagues is reproduced here as Figure 2 and is indeed curious. In order to determine whether the Bluff Wind Farm was generating infrasound, it was arranged for the wind turbines to be switched off. As a result, the report found that "At Location 8 near the Bluff Wind Farm (Figure 29), the [0.8 Hz, 1.6 Hz, and 2.5 Hz] peaks were detected *at a similar level* during both operational and shutdown periods" (emphasis added). As shown here in Figure 2, the spectrum at 1.6 Hz shows less than a 3 dB variation between all measurements, which is strange given that the peaks are distinctive signatures of wind turbines. The interpretation suggested by the synthesis here is that the convergence of all the 1.6 Hz measurements must have been due to another wind farm (probably North Brown Hill, 8 km away) and that L8 happened to be at an infrasound antinode at the time.

Evans et al. also conclude that "there is a possibility that the peaks in the spectrum during the shutdown resulted from operation of North Brown Hill Wind Farm", a source which was "very faintly audible" during the shutdown (p.37). However, because the sources were below background levels and therefore not a problem, they did not attempt to resolve the paradox of why infrasound levels from a source 8 km away could be higher than from a source 1.5 km away. However, the coherent infrasound model makes it possible to appreciate how, if the antinodes are aligned correctly, this might happen. The situation has as much to do with the phases of the sources as with the distances.

Similarly, Figures C3–C10 of Evans et al. show the same signature peaks at 0.8, 1.6, and 2.5 Hz at two locations, Bluff and Clements Gap, and more comparisons are presented between when the wind farms were operating and when they were shutdown (pp. 61–65). Again, there was little difference for Bluff, although there was for Clements Gap where there is no adjacent wind farm. Once more, the Bluff anomaly was not considered important because recorded infrasound levels were below background (in a one-third octave band), especially for higher wind speeds. Nevertheless, it is significant that in Figure C3, which includes measurements at Bluff for low wind speeds (0 to 3 m/s, when wind turbines do not rotate – p.58), the distinctive peaks at 0.8, 1.6, and 2.5 Hz are still present, again indicating that infrasound from Brown Hill 8 km away (where the wind must have been above 3 m/s) was still contributing. These measurements point to the presence of coherent infrasound that can propagate considerable distances. The question it raises is whether the elevated infrasound levels could have led to troublesome audible sensations at higher frequencies, and this is now addressed.

High harmonics of the blade pass frequency

Another anomaly that can be interpreted as evidence for the presence of highly coherent infrasound comes from the acoustic monitoring of Zajamsek and colleagues [10]. These workers made continuous sound recordings in homes near the Waterloo wind farm in South Australia and matched the level and spectral content of the sound with occasions when the residents noted its degree of annoyance. Of particular interest, they found using narrow-band (0.1 Hz) analysis that in one 'slightly annoyed' case (their Fig. 11) multiple harmonics of the blade pass frequency were present. The harmonics occurred not only at 0.8, 1.6, 2.4, 3.2, 4.0, 4.8, and 5.6 Hz (corresponding to the 1st–7th harmonics, which have been seen in other work), but also at frequencies corresponding to harmonics in the 20–30 Hz band at 29, 33, 34, 35, 36, 37, and 38 times the fundamental.

Such high harmonics are unusual, although they have been seen in musical acoustics (clarinets, for example, see http://newt.phys.unsw.edu.au/music/clarinet/E3.html). Of more relevance, a recent detailed report on the same wind farm [25] records that, when using a spectral resolution of 0.0017 Hz, "each peak [up to 69 Hz] is an exact multiple of the blade-pass frequency" (p. 71). A theoretical explanation for such high harmonics comes from noting that at least the 14th harmonic is present (at about the –40 dB level re the peak) in Fig. 5 of [5], and that, if the plot were extended logarithmically to –60 dB, further harmonics would become visible. Although perhaps 1000 times smaller than the largest BTI harmonic, it is possible that such high harmonics are selectively amplified by room modes in the same way as a Helmholtz resonator can pick out high partials from a musical sound.

Consider that if the 38th harmonic is evident in a 0.1 Hz analysis, then the associated fundamental must have a bandwidth of 0.1/38 or 0.003 Hz. More explicitly, if the 38th harmonic appears at 30.2 ± 0.1 Hz, which it does in Figure 11 of [10], then the fundamental can be accurately specified as 0.795 ± 0.003 Hz. Curiously, this value, which here derives from Vestas V90 turbines, comes close to the specified 0.790 Hz at rated power for the Suzlon S88 controller (set out above). The similarity in BTI frequency suggests that both machines use a similar, tightly specified rotational frequency, and it is this condition which promotes turbine entrainment.

The conclusion is that the high harmonics are sensitive indicators of a narrow-band synchronised condition, and it is a short step to the hypothesis that these high harmonics may be the cause, or at least contribute significantly to, audible annoyance. It is notable that the levels measured in the 20–30 Hz band are within 10–20 dB of audibility (Fig. 8 of [10]; Fig. 3 of [26]), so it requires only a relatively small increase in sound pressure levels in this band for the sound to become audible. In comparison, the levels at 0.8 or 1.6 Hz are about 50 dB below audibility, so these infrasonic frequencies, in themselves, may be less troublesome than the harmonics they give rise to. However, since Zajamsek and colleagues observe that "the 23.3, 28, and 28.8 Hz frequencies are well below the perception threshold" of ISO:226–2003, could there be any other additional factor – over and above constructive interference – responsible for making the sounds audible?

Short-term amplitude fluctuations

The answer could lie in the particular signal processing used. As Zajamsek et al. note, the recorded signal may include "important characteristics that can be averaged out during normal statistical processing methods" (p.2 of [10]). Indeed, it is suggested here that short-term levels, perhaps over 10–15 seconds, might be higher at certain epochs within the 2-minute averaging window. If the period of in-phase synchronisation occurred for say just 10 seconds during a 10 minute sampling period, and if 5 turbines became coherent, then the peak recorded level would be about 7 dB above the average level. An additional 7 dB would place the identified harmonics close to the average threshold of audibility, especially if the ear uses a peak pressure detection scheme [23] and sensitively hears modulation of higher frequency sounds.

It is worth noting that high harmonics also appear at other occasions when the same resident reported 'high annoyance' (Figure 7) and when another resident in another household also did so. However, one anomaly is that the fundamental and its high harmonics do not always appear (Figures 6, 7, 11) every time a resident reports annoyance. It is possible the lack of correspondence might relate to timing uncertainties: intermittent coherence means the analysis windows may not always neatly correspond with the 'annoyance' windows.

A technique to successfully detect intermittently coherent infrasound might involve a refinement of a method used by Doolan and colleagues [10, 22, 26]. These workers analysed wind turbine noise and correlated it with resident annoyance, finding some weak trends. Circumstantially, their work points to acoustic energy below 10 Hz as a possible cause of objectionable thumping at a repetition rate of about 1 Hz. Significantly, when short-term (0.125 ms) fluctuations were investigated using a peak-detection algorithm, variations in unweighted SPLs of up to 10 dB were found. Such variations are consistent with the presence of coherent infrasound that is fluctuating in and out of phase (and may have been larger if the measurements had been made in the 20–30 Hz band instead of the selected 10–1000 Hz band).

Other wind farm studies have also reported short-term fluctuations, and an analysis by Cummings [27] has highlighted a pronounced low-frequency variability. In a separate English case study, a particular 60-second recording shows a clear 0.8 Hz infrasonic component in the waveform until two-thirds of the way through, at which point the infrasound suddenly fades away (reproduced as Fig. 8 of [7]). At the Macarthur wind farm in Victoria [21], Evans observed a sudden increase in infrasound levels which generally affected scores of 10-minute monitoring windows before suddenly disappearing (p. 33 and Figs 17, 18 of [21]). In this case, because the elevated levels (as high as 75 dB-G over 10 minutes) were not associated with changes in wind speed or direction, he believed they were due to 'extraneous' sources and hence they were systematically removed. The concern raised by the coherent infrasound hypothesis is that these instances of extraneous sources, and multiple other similar occasions (e.g., pp. 2, 30, 35, 40, 43, 45, 50, 51, 52, 58 of [21]), may represent an entrained condition of the wind turbines. It is therefore difficult to decide whether the Macarthur report's conclusion – that wind farms contribute little to infrasound and low frequency noise – is robust. This is especially the case when, curiously, not even the fundamental or harmonics of the blade pass frequency appear in any of the displayed spectral plots.

The major limitation of the analysis methods used so far is that they assume the signal is stationary – an almost universal assumption which tends to obscure infrasound that becomes coherent only intermittently. For example, some Canadian work [28] found that average sound levels increased by about 5–10 dB in the 20–30 Hz band after a wind farm was installed in Ontario; however the data was plotted as long-term averages, not peak readings. There could have been many short periods when sound levels were above the threshold of hearing. In some subsequent Australian monitoring work by Doolan and colleagues [26], the frequency resolution was limited to 2 Hz, which misses narrow harmonics; however, in a later undertaking [22] the frequency resolution was higher (0.1 Hz) but an averaging time of 2 minutes was used as standard. Perhaps because of this, narrow band analysis did not reveal harmonics beyond the 7th, and the harmonics were not always seen. Broad correlations with annoyance were found, but there were inconsistencies as well.

It therefore seems crucial that the selected bandwidth settings and signal processing routines are designed to detect nonstationary infrasonic signals. The analysis needs to be narrow band (so that all the BTI harmonics can be detected) and use little averaging, and this combination will involve some compromises in time and frequency resolution as well as with signal-to-noise ratio.

CONCLUSION

The long-standing puzzle has been that standard acoustic measures do not correlate well with reported wind farm annoyance [26] and that the average detected levels appear to be below the threshold of audibility [9, 22]. To resolve this paradox, the concept developed here has been that wind turbine infrasound can be narrow band, have multiple sources, and occur intermittently as the sources drift in (and out of) phase. When two or more sources become entrained, interference at the fundamental of the BTI frequency (0.8 Hz) will give a pattern of nodes and antinodes, and there will also be a different (but related) pattern of nodes and antinodes for each harmonic. If a particular harmonic in the 20–30 Hz band happens to have two or more sources in phase at the measuring point [5, 7], the increased sound pressure level produced has the potential to be audible, and the proposal here is that the intermittency of the in-phase and out-of-phase conditions might underlie wind turbine annoyance. Whenever the blades become synchronised (perhaps for many tens of seconds) the intensity of the fundamental and some of its harmonics could, at nodes, be at least 6 dB larger, but the levels will revert to baseline when the sources fall out of synchrony (Fig. 8 of [7]).

A lingering puzzle is why some people complain of effects from wind farms which persist for hours, not effects which come and go. Such long-lasting symptoms such as headaches and pressure in the ears might be the outcome of pressure effects within the middle ear [29], a possibility only more research can decide.

Taken together, the evaluations made here provide indications that intermittent coherence could be the physical basis for the annoyance of wind farm noise. One key factor is the precise frequency setting of the wind turbine control circuit, and the other is the universal tendency for coupled oscillators to synchronise. The work builds on the 'heightened noise zone' idea of Thorne and colleagues and gives it added weight by regarding the BTI signal as the primary contributor to the problem – this signal carries most of the acoustic energy and has the power to force entrainment of neighbouring blades that are turning at close to the same frequency. The impulse and its many harmonics generated by BTI could also set off resonant room modes in the 20–30 Hz band. The electronic controllers built into wind turbines are identified as exacerbating the noise problem by increasing the chances of entrainment.

If the human ear can hear wind turbine noise, while the measurements say the noise is inaudible, then the remedy is better measurement techniques. The Information Paper's position – that because instruments do not measure anything above ambient background levels then the sound cannot be a problem – may be giving precedence to measurements that could be inappropriate or incorrectly interpreted. If people living close to wind farms report troublesome noise then the preferred response is to find the source of the problem, not question the validity of the reports.

This paper suggests that the apparent paradox between what is heard and what is measured might be resolved by recognising the tonal nature of infrasound at the blade passing frequency (0.8 Hz) and of its harmonics, which may extend to 20–30 Hz. More measurements need to be made at these frequencies and at distances up to tens of kilometres. Measurements at multiple points are needed so that a map of infrasound nodes and antinodes under various conditions can be established. Theoretical studies of the interference patterns caused by regular arrangements of wind turbines would also provide insight.

A relatively simple solution to the problem of wind farm noise may be the one proposed by Doolan and colleagues:

ensure the blades never operate in synchrony. Synchronisation is inherent in the Suzlon turbines with their fixed speed–variable pitch design, although it is not the case with more modern variable speed designs [12]. Even with modern machines, however, the tendency for any set of physical oscillators to synchronise needs to be recognised.

Hopefully this paper will prompt a re-evaluation of wind farm infrasound and how it might lead to audible disturbances. Some of the statements in the NHMRC Draft Information Paper appear to be questionable and may be better reframed. For example, the NHMRC's statement that infrasound from wind turbines does not differ from other natural sources seems inconsistent with the observation that such infrasound has strong tonal components. Its reported effects, including annoyance and disturbed sleep, might be studied by simulated wind turbine noise generated by a speaker, as the draft report recommends (p.20), although reproducing sound with a flat frequency response down to 1 Hz and with accurate phase would no doubt be challenging.

Similarly, there is mounting evidence that wind farm noise can be heard at much more than the 500–1500 m specified by the NHMRC, particularly at night. The review recommends more research into "wind turbine signature" and for field measurements "ranging from 500 m to 3 km and beyond" (p.20). From the issues raised here of the behaviour of coherent infrasound, distances of 8 km and more may be more appropriate. Although measuring down to 0.8 Hz is possible, an easier approach technically may be to use narrow band analysis to detect high harmonics in the 20–30 Hz band, measurements which might relate directly to reported annoyance.

ACKNOWLEDGEMENTS

The author thanks the reviewer for a wide range of challenging questions which have considerably improved the text.

REFERENCES

[1] (various authors), "Special issue on windfarm noise", *Acoustics Australia*, **40**(1), 1-96 (2012)

[2] Doolan, C., "A review of wind turbine noise perception, annoyance and low frequency emission", *Wind Engineering*, **37**, 97-104 (2013)

[3] Merlin, T., et al., *Systematic Review of the Human Health Effects of Wind Farms*. Canberra: National Health and Medical Research Council (2013)

[4] NHMRC, *NHMRC Draft Information Paper: Evidence on wind farms and human health*. Consultation draft, February 2014. Canberra: NHMRC (2014)

[5] Doolan, C., D.J. Moreau, and L.A. Brooks, "Wind turbine noise mechanisms and some concepts for its control", *Acoustics Australia*, **40**(1), 7-13 (2012)

[6] Thorne, R., "The problems with 'noise numbers' for wind farm noise assessment", *Bulletin of Science, Technology & Society*, **31**, 262-290 (2011)

[7] Thorne, R., Wind Farm Noise and Human Perception. *Enoggera: Noise Measurement Services* (2013)

[8] Bakker, H.H.C. and B.I. Rapley, *Sound, Noise, Flicker and the Human Perception of Wind Farm Activity*. Palmerston North: Atkinson & Rapley Consulting (2010)

[9] Evans, T., J. Cooper, and V. Lenchine, *Infrasound Levels Near Windfarms and in Other Environments*. Adelaide: Environment Protection Authority of South Australia (2013)

[10] Zajamsek, B., et al., Simultaneous indoor low-frequency noise, annoyance and direction of arrival monitoring, in *Proceedings, Fifth International Meeting on Wind Turbine Noise*, Denver CO, 28-30 August 2013

[11] Harper-Somers-O'Sullivan, *Environmental Assessment for Black Springs Wind Farm*. Sydney: Wind Corporation Australia (2006)

[12] AEMO, *Wind Turbine Plant Capabilities Report*. Melbourne: Australian Energy Market Operator (2013).

[13] Bianchi, F.D., H. de Battista, and R.J. Mantz, *Wind Turbine Control Systems: Principles, Modelling and Design*. New York: Springer (2007)

[14] Hansen, M.O.L., *Aerodynamics of Wind Turbines*. London: Earthscan (2008)

[15] Leithead, W.E. and B. Connor, "Control of variable speed wind turbines: design task", *International Journal of Control,* **73**, 1189-1212 (2000)

[16] McSwiggan, D., et al., "A study of tower shadow effect on fixed-speed wind turbines", Universities Power Engineering Conference, Padova (2008)

[17] Cidras, J., "Synchronization of asynchronous wind turbines", *IEEE Transactions on Power Systems,* **17**, 1162-1169 (2002)

[18] Katayama, N., et al., "Theoretical study on synchronization phenomena of wind turbines in a wind farm", *Electrical Engineering in Japan*, **155**, 9-18 (2006)

[19] Pikovsky, A., M. Rosenblum, and J. Kurths, S*ynchronization: A Universal Concept in Nonlinear Sciences*. Cambridge: CUP (2001)

[20] EPA South Australia, *Waterloo Wind Farm Environmental Noise Study*. Adelaide (2013)

[21] Evans, T., *Macarthur Wind Farm Infrasound and Low Frequency Noise: Operational monitoring results*. Adelaide: Resonate Acoustics (2013)

[22] Zajamsek, B., D.J. Moreau, and C. Doolan, "Characterising noise and annoyance in homes near a wind farm", *Acoustics Australia*, **42**(1), 14-19 (2014)

[23] Moller, H. and C.S. Pedersen, "Hearing at low and infrasonic frequencies", *Noise and Health*, **6**(23), 37-57 (2004)

[24] Cooper, J., T. Evans, and D. Petersen, "Tonality assessment at a residence near a wind farm", in *Proceedings, 5th International Conference on Wind Turbine Noise*, Denver CO, August 2013

[25] Hansen, K., B. Zajamsek, and C. Hansen, *Noise Monitoring in the Vicinity of the Waterloo Wind Farm*. Independent report to M. Morris (2014). Available at http://waubrafoundation.org.au/wp-content/uploads/2014/08/Hansen-Zajamsek-Hansen-Noise-Monitoring-at-Waterloo1.pdf

[26] Doolan, C.J. and D.J. Moreau, "An on-demand simultaneous annoyance and indoor noise recording technique", *Acoustics Australia*, **41**(2), 141-145 (2013)

[27] Cummings, J., "The variability factor in wind turbine noise", *Fifth International Conference on Wind Turbine Noise*, Denver, August 2013

[28] Howe, B., N. McCabe, and S. Ferguson, "Infrasonic measurements, pre- and post-commissioning, Ontario wind farm", in *Proceedings, Fifteenth International Meeting on Low Frequency Noise and Vibration and its Control*, Stratford-upon-Avon, U.K., May 2012

[29] Bell, A., "Annoyance from wind turbines: role of the middle ear muscles [letter]", *Acoustics Australia*, **42**(1), 57 (2014)

NEWS

ACOUSTICIAN TURNS 100

Internationally renowned acoustician Leo Beranek turned 100 on 15 September 2014. Leo Beranek is still active in participation in acoustics. Among the celebrations have been a special session "Centennial Tribute to Leo Beranek's Contributions in Acoustics" at the Indianapolis meeting of the Acoustical Society of America. As well, a whole issue of Acoustics Today has been dedicated to Leo Beranek.

See some of the many congratulatory messages from the international acoustics community: http://www.icacommission.org/leo.html

Standards Australia - New vibration standard

Committee AV-010 Vibration And Shock Human Effects is pleased that AS ISO 2631.2:2014 Mechanical vibration and shock—Evaluation of human exposure to whole-body vibration-Vibration in buildings (1 Hz to 80 Hz) has been released on 27 November 2014. This is a direct text adoption of ISO 2631-2 of 2003 to replace the outdated AS 2670.2—1990, Evaluation of human exposure to whole-body vibration, Part 2: Continuous and shock-induced vibration in buildings (1 to 80 Hz). The objective of this Standard is to provide a method for the measurement and evaluation of whole-body vibration in buildings and to encourage uniform methods for collection of data on human response to building vibration.

It is important to note that this Standard does not provide guidance on the human response to vibration in buildings and the effects of that response in regard to human health and safety. For those aspects, usually workplace related, then AS 2670.1—2001, Evaluation of human exposure to whole-body vibration, Part 1: General requirements, should be referred to. Also his part of ISO 2631 does not provide guidance on the likelihood of structural damage, which is discussed in ISO 4866.

What this new standard does concern is human exposure to whole-body vibration and shock in buildings with respect to the comfort and annoyance of the occupants. It specifies a method for measurement and evaluation, comprising the determination of the measurement direction and measurement location. It defines the frequency weighting Wm which is applicable in the frequency range 1 Hz to 80 Hz where the posture of an occupant does not need to be defined.

Standards Australia - Revisions

There are two important standards currently under revision by Standards Australia committees. The public review period on AS/NZS 2107 led to over 40 comments on the draft and the committee is now working through the comments. The review of AS 2021 is progressing through the committee process.

AAS NEWS

Council News

At the AGM, 42 members attended in person and motions to accept the 2013-2014 Financial accounts and to reappoint the auditor W L Brown & Assocs were carried.

Other motions to amend the By-laws were also passed, ie:

- Item 6.1 Amend By-Law 36 "To replace the term 'dishonourable practice' with more precise and legally definable terms" as outlined in Appendix B of the Notice of Meeting document. This was carried with one(1) abstention.
- Item 6.2 Amend By-Law 37 "To provide clearer direction for any AAS Investigation Committee appointed by Council in exercising its responsibilities." was also carried subject to further legal review.

Further to previous announcements in 2014, development of the AAS Research Grant Assistance Program Offer is now complete, with outstanding legal and procedural issues recently cleared. At its most recent meeting in November 2014, the AAS Federal Council has set an application closing date of February 28, 2015. Another call for entries will be held towards the end of 2015.

Applicants are encouraged to visit the 'Awards and Grants' section of the beta AAS website (https://wp.acoustics.asn.au/awards/) for further details and the application form.

Federal Council regrets the inconvenience caused to some members during the changeover to our new website. We are working hard with the web designer to resolve the outstanding issues and look forward to our website being functional in 2015.

Finally, after nearly 20 years of involvement with AAS, many of those with Federal Council, Neil Gross stepped down as NSW Federal Council representative in November. Thank you Neil for all of your valuable input and insight (not to mention sense of humour) to Federal Council matters over the years. Neil has been replaced by NSW representative Matthew Harrison.

Gerald Riley Award Winner

Congratulations to presenting author Sebastian Oberst and co-authors Enrique Nava-Baro, Joseph C.S. Lai and Theodore A. Evans who were awarded the Victorian Division's Gerald Riley Award of $2500 for the best AAS member paper at Internoise 2014. The paper is entitled '*An innovative signal processing technique for the extraction of ants' walking signals*'.

AAS President Norm Broner and Sebastian Oberst at the presentation of the award.

Fellows of AAS

At the recent AGM, Norm Broner (VIC) and John Macpherson (WA) were elevated to Fellows of the Society.

Norm Broner (left) and John Macpherson (right).

Norm Broner

Citation: Dr. Norman Broner, past-President of the Australian Acoustical Society and current Chairman of the Victoria Division has had a distinguished career in the field of Acoustics spanning over thirty-five years. One of his particular interests has been in researching the human impact of and annoyance due to Infrasound. This developed whilst studying at Monash University and resulted in Norm choosing to investigate this further. He completed his Master of Engineering Science at Monash University and then continued his studies at the University of London where he was awarded his Doctorate (PhD) which focussed on his research in Psychoacoustics.

Having completed his studies, Norm returned to Melbourne where he took up a position with VIPAC Engineers and Scientists Ltd. as an Acoustic Consultant culminating in his role as Operations Manager of the Melbourne Office. At Vipac, he consulted in the fields of environmental, industrial, architectural, transport and defence noise and vibration over a twenty seven year period.

Norm continued his acoustics career with Sinclair Knight Merz (SKM) (and Jacobs) for a further nine years as Practice Leader, Acoustics, Noise and Vibration.

Of recent times Norm has established his own acoustic consultancy, Broner Consulting Pty. Ltd. consulting all areas of acoustics with an emphasis on high level acoustic work and peer review assessments.

During the course of his career, Norman has published many papers and written articles for Acoustics Australia dealing with noise and vibration issues. He has been a member of various International Acoustic Committees including the Technical Committee on Architectural Acoustics (TCAA) of the Acoustical Society of America and has been on the Wind Farms and Humans Reference Group of the NHMRC and international acoustics advisor to the Health Canada's Wind Turbine Noise and Health Study Expert Committee.

Norm began his association with the AAS in the early 1970's joining the Committee in his student days until he moved to London.

As a long standing Member and Council Member of the Society, Norm has served as Victoria Division Chairman and President of the Australian Acoustical Society on multiple occasions. He is enthusiastically welcomed by all State Divisions in their hosting Annual Conferences for his style of encouraging exhibitor participation ensuring their financial successes. Recently he has was also President of the very successful Internoise 2014 INCE Congress run in Melbourne.

Without doubt, Dr. Norman Broner is well worthy of receiving his elevation to a Fellow of the Australian Acoustical Society. His contribution to acoustics in Australia is invaluable and his involvement in the Society is legendary. He is indeed a role model for all Members.

John Macpherson

Citation: Like many acoustical consultants, John Macpherson started his career in the mid-1970s as a musician. After graduating from the University of Western Australia with a Bachelor of Engineering, he moved to New South Wales to work for the acoustical consultancy Louis Challis and Associates Pty Ltd in 1977. Ten years later, following his Master of Science in Acoustics at the University of New South Wales, John moved back to Western Australia to work at the then named Department of Occupational Health Safety and Welfare. Here he led various research and compliance projects in occupational noise, and authored a variety of informative publications for DOHSWA and Worksafe Australia, many of which are still current and in circulation. He is also credited with developing the occupational Noise Exposure Index method with Colin Tickell in the mid-1990s.

In 1994 he moved into an environmental officer role at the then named Department of Environmental Protection (now referred to as the Department of Environment Regulation), and has led the Noise Regulation Branch as Principal Environmental Noise Officer since 2001. In this role over the last 20 years, he has led the ongoing reform and update of Western Australia's environmental noise policy and its enforcement; an achievement all the more impressive given the relative change in the Western Australian economy over that time and relatively strong provisions compared to other Australian States for protecting the amenity of noise sensitive spaces.

Since he joined the AAS in 1987, he has provided outstanding contributions over the years, including roles as Federal Councillor in the early 1990s and WA State Division Secretary (1997-2003), and assisting with organisation of all national AAS conferences in WA since at least 1990. He has been involved in various Standards Australia technical committees including AV-003 (Effects of Noise) and EV-010 (Community Noise), and a member of the International Commission on Biological Effects of Noise (ICBEN) Team 9 - Noise Policy and Economics.

This year, following a career of 40 years, John retired from the Department of Environment Regulation. A musician to the end, his thoroughly enjoyable and memorable performance at his retirement function underlined not only the long history of success in his career, but also his contagious passion for acoustics. John has been appointed to the Grade of Fellow for his years of outstanding service to the Australian Acoustical Society, and for improving the science and practice of acoustics in not only Western Australia, but nationally and beyond.

AAS Education Grants

Congratulations to the successful AAS education grant winners for 2014:

- $10,000 awarded to Catherine Madill, Faculty of Health Sciences, University of Sydney "Towards development of a freely available and accurate analysis of vocal attack time on vowel onsets: Comparison of 3 measures of voice onset with glottal pulse identification in PRAAT."
- $5,000 awarded to Jamie McWilliam, Centre for Marine Science & Technology, Curtin University "Investigating the ecological importance of underwater sound in marine coastal ecosystems."
- $5,000 awarded to Lily Panton, School of Engineering and ICT, University of Tasmania "Investigating Auditorium Acoustics from Perspective of Chamber Orchestra Musicians."

Council thanks Emeritus Professor Colin Hansen and Professor Joseph Lai for their deliberations in selecting the successful applicants.

Report on InterNoise2014 by Norm Broner

By all accounts, Internoise 2014 was a very successful conference. 1024 abstracts were received with the final tally of 792 papers and 79 posters. Of these, 98 papers were reviewed and there were 158 oral sessions held in 15 parallel streams. 37 countries were represented with about one fifth from Australia. 11 Young Professional Grants were awarded by the Melbourne Convention Bureau while INCE awarded 15. All up, there were 1106 registrants including 80 accompanying persons.

The location of the congress was at the Melbourne Convention and Exhibition Centre. All papers were given in rooms on Level 2. There was a very large exhibition held in the bay across the corridor from the convention foyer where 53 exhibitors from all around the world displayed their technologies. Ortech Industries and Embelton were Gold sponsors, Martini Industries was a Silver sponsors while CSR Bradford and Pyrotek were Bronze sponsors. Lunch and morning and afternoon coffees were provided in the exhibition area and allowed all attendees to interact with each other and with the exhibitors.

2629 cups of coffee and 566 cups of tea were drunk on both Monday and Tuesday. A very successful banquet was held at which an Aussies bush band provided entertainment. At the pre-dinner cocktails, attendees had the opportunity to see Aussie animals including koalas, a goanna, wallabies and dingos.

The congress started with a Welcoming ceremony in the Plenary on Sunday and was concluded after a very hectic three days of papers with the Closing Ceremony where we handed over to Internoise 2015 in San Francisco, who sponsored the Closing Reception.

Internoise 2014 was a very successful conference which gave Australians a good chance to be exposed to high calibre international papers on home soil. This would not have been achieved without the outstanding efforts of the organising committee. Thanks in particular to Charles Don, Technical Chair and John Davy, Co-Chair, Terry McMinn for the paper management and proceedings, Marion Burgess as advisor, Geoff Barnes for the social program, Dianne Williams our treasurer and our secretariat Lizz Dowsett. And of course thanks to all those volunteers who helped during the event.

The full proceedings of the conference are available from the AAS website under Conferences. Photos from the conference are available from the Internoise website www.internoise2014.org

The Internoise 2014 committee: from left Norm Broner, Geoff Barnes, Marion Burgess, Charles Don, John Davy, Terry McMinn, Dianne Williams and Lizz Dowsett.

Charles Don with the Internoise mascot 'Little Norm'.

Reports from recipients of the NSW AAS travel awards to attend InterNoise 2014

I received an Australian Acoustical Society NSW Division Travel Award to attend InterNoise2014, which was held in Melbourne, 16-19 November 2014. Firstly, I'd like to thank the AAS NSW Division for the award. For her continued support I'd also like to thank Tracy Gowen who is communicating and managing these awards for as long as I can remember which is throughout my entire career as PhD student and Postdoc.

I found the InterNoise2014 to be a very well organised conference with a lot of interesting topics in my field of numerical acoustics and hence this has been a very good conference for me. I presented three papers on the numerical analysis of structure-borne sound, sound radiation and vorticity induced vibrations. I also co-chaired one of the numerical sessions. Due to the large number of delegates at the conference and several sessions entirely dedicated to numerical approaches of sound and vibration, I had a very good audience for all of my presentations, which resulted in good discussions during question time and after the talks.

I was able to meet virtually all of my national and international collaborators, who are from Canberra, Melbourne, Perth and various places in Europe. This was unusual because I often tend to meet my national collaborators at the annual AAS Acoustics conferences only while I meet my international collaborators at international conferences such as NOVEM, Forum Acusticum, ICSV, ICA or InterNoise.

Lastly, let's not forget the entertaining evenings and thanks to some excellent Australian craft beer and many national and international friends who share the same passion and taste these turned out the be very successful, too!

Herwig Peters

The acoustic environment at InterNoise 2014 was very noisy! But it isn't always about low sound levels. What it is about is sounds that are appropriate to that place. On this occasion, the sounds orchestrated from wing tip vortices, vibrating hydrofoil to owl flight couldn't be more fitting and formed my experience of this acoustic environment. Why? I've always been interested in the machination of the Universe. That feeling of air lifting your hand up whenever you have it at an angle out of the car window has always fascinated me. Therefore, it was natural for me to study aerodynamics and structural dynamics at university. So seeing fellow acousticians working on state-of-the-art research on some of my favourite subjects was both exciting and inspiring – it makes me want to get back into conducting experiments in the wind tunnel again!

Apart from sitting in the audience seat I also presented at the conference to raise awareness on the benefits of using trees in noise control and with the hope to encourage its implementation in practice. I didn't really know what to expect from the audience because the topic of my talk have been the subject of much debate for many years, and tend to be on the negative side. To my surprise it was actually well received by the audience. I've also expanded my horizon on the application of trees from attending a plenary talk on soundscaping, more trees mean more foliage and biodiversity, which interestingly provide sounds people want or prefer. It's been an incredible journey which I never envisioned when I first started this project. To top it off, I had the opportunity to meet and exchange ideas with Dr Van Renterghem from Ghent University whose papers I have read many times and referenced in my research.

Jeffrey Peng

Many days of the last years I was standing in either an anechoic chamber or in an environmental chamber with 30 degreesC and about 80% relative humidity and some of the time did not know the value of non-stop observing and recording ant vibrations. Well, now I know why and the awards I received compensate a bit for this time and sweat. As a recipient of the AAS NSW Travel Award and the Gerald Riley best paper award I would like to express my gratitude for funding to participate at the conference and acknowledgement of my research.

In general, the Internoise 2014 offered high quality papers and interesting presentations. The presentations were usually followed by fruitful academic discussions, which certainly will give me new ideas, provide impetus and momentum for my own research for the coming year.

Congratulations on the very well planned successful conference in Melbourne. I was very pleased with the location as such: the convention centre in Melbourne offered the right number and size of rooms, the illumination was perfect to read e.g. the book of abstracts and not to be too disturbed while presenting. The provision of lunch in the exhibition zone was great and very beneficial for continuing useful discussions.

Sebastian Oberst

With the assistance provided by the Australian Acoustical Society NSW Division, I attended the InterNoise 2014 conference titled "Improving the World through Noise Control", held in Melbourne, 16-19th November 2014. My contribution to this conference was to present my paper titled "Prediction of the radiated sound power from a fluid-loaded finite cylinder using the surface contribution method". The concept of this paper was to predict the exterior radiated sound from a fluid-loaded finite cylinder. My paper includes comparison between the radiated sound power obtained from both surface contribution and the acoustic intensity methods. There were lots of technical audiences from different areas attending my presentation. I answered some great questions after finishing my presentation. It was my pleasure that they were interested in my topic.

Overall, presentations throughout the conference highlighted the need for better dealing with the noise and vibration problems all over the world, including traffic noise, construction noise and vibration, underwater noise and wind turbine noise.

As a student member, attending this conference provided an important learning experience and networking opportunity. I was able to know more members of the Australian Acoustical Society. I had many discussions on why it was such a successful acoustic conference in Australia, numerical simulation in vibro-acoustics and current noise and vibration control industry. These discussions, presentations and forums have provided much motivation and inspiration for my future PhD study in UNSW. I would like to thank AAS NSW Division for awarding me this travel grant and congratulate the conference organisers and presenters on an outstanding conference.

Daipei Liu

This conference, themed 'Improving the World through Noise Control', was by far one of the best international conferences I have attended. The conference was held at the MCEC (Melbourne Convention Exhibition Centre) from 16th to 19th of November. The 3-day conference involved 15 parallel sessions covering all fields of noise control. This conference opened my eyes to the variety of possibilities, opportunities and problems in the field of noise control. The speakers come from every continent of the world and I proudly found that Australia contributed most of the presented papers followed by China and Japan. I was impressed by the high quality of organisation and catering service.

The issue of wind turbine noise pollution and underwater noise control interested me at most. I presented my work on 'Sound radiation from the nested cylindrical shells' on the last day of the conference and received some invaluable comments from researchers based in DSTO, Melbourne. In my opinion, this is a great opportunity for me to contact other researchers and scholars who are working in a similar research field using different methods. For me, participating in this international conference helps to network and exchange ideas with other colleagues.

Along this journey, I have also made some friends within Australia and other countries of the world. I felt completely immersed in a group of researchers from Germany, UK and China. In such a short span of time, I was amazed that I have made many friendships of which I have treasured and hold dearly. Overall, InterNoise 2014 was a very unforgettable conference and I had a great time.

Nick Wu

Professional conferences such as InterNoise 2014 are a good platform for networking and to gain experience to present the research work. Most of the PhD students who are early in their career couldn't afford the whole expenses to attend these conferences. So it is great to get support from the Australian Acoustical Society. The grant can assist students to apply the knowledge gained from the conference in their PhD career.

I received the Australian Acoustical Society NSW division travel award to attend the InterNoise 2014 in Melbourne. It was a great opportunity to present a paper at this conference since it has a specialized session in the area in which I am doing my PhD research. I was asked several excellent questions after my presentation which I will definitely consider while researching. I had also received some intelligent insights from some of the leading professionals in the industry. By sharing the experiences with other attendees who are working on different noise problems, I feel that I need to expand my connections and not limit myself to people who have the same interests.

I really appreciate the financial support from the AAS NSW division to attend InterNoise 2014. It was a great motivation for me to participate in this high-qualified conference.

Samaneh Fard

The InterNoise 2014 conference was held from November 16-19 at the Melbourne convention centre. The conference started off after the wettest November day in Melbourne for the past 4 years. Luckily this weather didn't stick around, and by the Monday morning the weather had cleared and you were able to enjoy the close proximity of the venue to the Yarra River and restraints. The technical sessions of the conference started and ended with a Plenary session on sound-sketching and sound-scapes respectively. The Plenary sessions, along with the Keynote presentations at the conference, were certainly a highlight.

Personally, I was introduced to the concept of sound-scapes and how both environmental and man-made sounds are a part of our world, which add real value. The analogy for the response a person might have to a sound-scape could be similar to the attachment or wonder you might have to a landscape vista. Along with the great insights into sound-scapes during his presentation, Alan Lex introduced the work of the Museum of Endangered Sounds, and is certainly worth a look (listen).

The depth of technical parallel technical sessions was extensive at InterNoise with the opportunity to attend presentations which spanned from acoustics education to underwater acoustic radiation through to structural vibration. It is this extensive and high class research presented during these sessions that was the most valuable benefit of attending this year's InterNoise conference and what will bring me back again.

Gareth Forbes

I would like to acknowledge AAS (Australian Acoustical Society) for giving me the travel award for research students which enables me to attend this exciting and fruitful academic event. This congress offers me a great platform for sharing the latest research results and exchanging ideas with peers which is extremely important to the progress in my PhD research. The topic of my PhD research is analysis of disc brake squeal propensity which is highly relevant to the theme of this congress so that I was given a great opportunity to meet scientists and engineers all over the world who are doing exactly the brake squeal projects. I benefitted a great deal from communicating with them because they discussed my presented research with a view to the applicability to the realistic brake products, which triggered my fresh ideas on how to upgrade my research from being purely academic to application-oriented. Meanwhile, I attended several keynote speeches and sessions, especially those held by Mrs. Marion Burgess, from which I gained knowledge on the cutting-edge theories and technologies of improving passive and active modelling and understanding acoustical and vibrational performance of diverse application fields, such as automotive, architectures, structures, etc. I would like again, to take this opportunity to express my sincere gratitude to AAS for the student travel award.

Zhi Zhang

Educators in Acoustics group

A small group met during the Internoise/AAS conference in Melbourne – the busy program precluded greater participation. One of the intentions of the group is to consolidate and cooperate with course material. Anyone interested is welcome to join the group - contact Carl Howard at carl.howard@adelaide.edu.au

Technical Report for review

The I-INCE Technical Study Group (TSG) 9 has recently completed the draft of its report on **"Supplemental Metrics for Day/Night Average Sound Level and Day/Night/Evening Average Sound Level"**. The AAS, as a member of I-INCE, has the opportunity to provide a consolidated vote and comment on this draft report by 15th January 2015.

If you would like a copy of the report to review then please contact GeneralSecretary@acoustics.asn.au or aassec@iinet.net.au.

For those wishing to comment on the draft report, please forward your comments to Norm Broner (norm@broner.consulting) by 7th January 2015 at the latest.

REPORTS FROM THE AAS DIVISIONS

QLD Division

In August the Queensland Division members hosted a very informative and interesting talk titled "Acoustics and IEQ" from Dr Kenneth Roy, Principal Research Scientist at Armstrong World Industries. Focusing mainly on the acoustic design quality of LEED Certified office buildings in the USA, Dr Roy presented the results of research done into the subjective quality evaluation of open plan offices by the end users. The end result was that the "function drives form" approach to the acoustic design of buildings needs to be adopted to increase the acoustic quality within office tenancies.

On the 17th September Beau Weyers and Craig Beyers from the Calibre Technology Acoustics and Vibration Calibration Centre presented a talk titled "Acoustic Calibration 101 – Why? How? And who's New in the Zoo?" An overview of the history and basis for acoustic metrology was provided, along with an introduction to the requirements of the current international standards. The differences between the current international calibration standards and the historic standards previously adopted were highlighted. At the end of the presentation everyone who attended had an improved understanding of acoustic metrology and how it can impact upon acoustic measurement datasets.

On the 15th October the Queensland Division AGM was held at The Greek Club. The Divisional Committee would like to welcome Roger Hawkins to the Committee and thank outgoing Committee member Matthew Fishburn for his contribution over the last two years. After the AGM Steve Layfield and Andrew Richie from Ortech Industries presented a talk titled "Straw and R_w 74: High Performance Sound Attenuation in Lightweight Panelised Building Systems". A number of product samples and case studies were presented to demonstrate to the Queensland Members the practical application of the Ortech Durra Panel system. The ease of installation of the system and the acoustic performances that can be achieved from the system were of interest to our Members.

The Queensland Division will be concluding 2014 with a Christmas Party to be held at Buzz at Tennyson on the 3rd December. Next year promises to be another busy year for the Queensland Division, with several technical site visits and presentations already in the planning stage.

WA Division

On September 17, the WA Division held their State Seminar and AGM at the Rothschild's Room at Perth Zoo. Over 30 members enjoyed listening to a range of presenters covering a range of topics including environmental noise, underwater acoustics, and active noise control. The day began with a talk by Michael Ward from *Perth Airport titled Perth Airport's new ANEF and Aircraft Noise Information Portal*. Next there was a presentation by Luke Zoontjens from SLR Consulting regarding the *Trends in the acoustic design of Healthcare and Medical Research Facilities, and the new US and AAAC guidelines* which was quite relevant given the significant number of health projects under construction in WA at present.

Prior to morning tea Mahbub Alam Shiekh from Vipac Engineers and Scientists presented a talk titled *Assessing Indoor Aural Comfort for Environmental Noise – A Psychoacoustic Evaluation Approach.* The first of the talks on the day regarding underwater acoustics was presented by Marta Galindo Romero (CMST) titled *Predicting underwater peak pressure levels of propagating airgun array signals*. George Watts from Herring Storer then presented a very practical talk titled *Residential Air Conditioner Noise* which highlighted the difficulty in complying with the Environmental Protection (Noise) Regulations 1997 in the context of today's small residential lots. There was also a presentation by Julio Pieraldi from Embelton regarding *Fit for purpose isolated gym floors*. One of the most interesting presentations of the day was by Alec Duncan at the Centre for Marine Science and Technology. Alec discussed the *Underwater noise signals and the search for missing Malaysian Airlines Flight MH370*. Dr Jie Pan from UWA presented a talk regarding *Active noise controlled window with natural ventilation*, and there were also a few presentations by students who also presented at Internoise in Melbourne in November.

The Annual General Meeting was held just prior to lunch, with the required quorum easily achieved. Committee members and positions were ratified, with three member's renominating after standing down following two years of service.

On December 4, the 2014 Christmas Party was held for the WA member's at the Windsor in South Perth. The Christmas lunch was well attended, and was a great opportunity for member's to relax and mingle after another interesting year in acoustics.

NSW Division

The NSW Division has been fairly active since August, with three technical talks, all of which were videoed and are/were available on the AAS website. To access the videos you will need to login as a member on the AAS home page. The available videos are listed with the NSW Divisional Notices (http://www.acoustics.asn.au/joomla/divisional-notices.html). Note that some videos are time limited and may no longer be available to view.

In August we had a Panel Discussion entitled 'What is Offensive Noise?' We had three experts on 'offensive noise' present in this topic, followed by lengthy discussion on the topic. Gordon Downey, Noise Policy Section of the NSW EPA, presented on the existing regulatory framework with regard to offensive noise and current EPA guidance on how to determine offensive noise. Renzo Tonin presented a discussion of the application of the legislation, with reference to the Meriden case. Roslyn McCulloch, Special Counsel with Pikes & Verekers Lawyers, presented the legal interpretation of 'offensive noise', in particular in relation to the Meriden case.

October saw a presentation by David Hanson, Principal Technical Advisor – Noise and Vibration at Transport for NSW Freight and Regional Development Division (FRD) addressing noise from freight rail. The presentation focused on the FRD programs targeting wheel squeal and locomotive noise, including discussions of at-source noise control, modelling of freight rail noise, and results from recent monitoring campaigns. Dave also touched on aspects of railway engineering, including rolling stock design and track maintenance, and how a working knowledge of these areas is crucial for effective mitigation.

At the end of October we had a talk by Dr Stephen Conaty on 'The health effects of environmental noise (other than hearing loss)…' Stephen is a public health physician who trained in public health in both Australia and the UK and is currently responsible for the revision of the enHealth document "The health effects of environmental noise – other than hearing loss". Stephen presented a summary of the process of the enHealth review and the evidence emerging from systematic reviews completed to date.

This tech talk provided useful background to the Noise and Health Seminar, held on Friday 21 November 2014 following on from InterNoise. A broad range of topics were discussed at the workshop, with presentations from a policy perspective by Dr Stephen Conaty and Peter Marczan (Manager Noise Policy, NSW EPA). This was followed by presentations on noise exposure from aircraft and road traffic noise by Dr Rob Bullen (Wilkinson Murray) and Professor Lex Brown (Griffith University) respectively. After lunch Dr Irene Van Kamp (National Institute for Public Health and the Environment (RIVM), Amsterdam, the Netherlands) and Dr Anna Hansell (Small Area Health Statistics Unit, Imperial College, London) presented on some of the outcomes from studying the health effects of noise, which in some cases was quite alarming. The final session was a Panel Discussion aiming to look at future directions for noise and health research in Australia. From the discussion it was clear that there is a great need for further work to be done in this area, in particular in relation to Australia, to better inform the policy makers and subsequently decision makers.

SA Division

The South Australian division held its 38th Annual General Meeting on Wednesday 12 November. The AGM was followed first by a technical presentation by Kieran Doherty and Matthew Tripodi, and then by a dinner at Café Brunelli. Kieran and Matthew, who are students in the School of Mechanical Engineering at the University of Adelaide, presented findings from their honours-level undergraduate project. The project, which was undertaken by Kieran and Matthew along with fellow team member Francesco Larizza, was entitled "The adaptive-passive exhaust silencer" and was partially sponsored by the SA division of the Australian Acoustical Society. The project involved the development of an adaptive-passive quarter-wavelength tube (AQWT) with a novel tuning method to attenuate sound generated at the fundamental firing frequency of an engine. The developed tuning method implemented a control system which used a sliding-Goertzel Algorithm to calculate the phase angle of the transfer function of the sound pressure between the top and bottom of the tube, which was to be maintained at -90°, and also involved changing the air temperature within the AQWT. The method was seen to be effective over a range of exhaust gas temperatures and engine speeds. The project was awarded 'Best Acoustics - Vibration Project' at the University of Adelaide's Mech Expo 2014,

In other news, the South Australian division provided travel grants to assist five students to attend and present their papers at Internoise 2014. Each travel grant included sponsorship of the registration fees as well as up to $500 towards flights and accommodation. The successful grant recipients, all from the School of Mechanical Engineering at the University of Adelaide, were:

- undergraduate student Orddom Leav, for the paper "A novel semi-active quasi-zero stiffness vibration isolation system using a constant-force magnetic spring and an electromagnetic linear motor";
- undergraduate students Matthew Tripodi, Francesco Larizza and Kieran Doherty, for the paper "Adaptive quarter wavelength tube tuned by varying air temperature"; and
- postgraduate student Ric Porteous, for the papers "The flow-induced noise of square finite wall-mounted cylinders in different boundary layers" and "Three-dimensional beamforming of aeroacoustic sources".

BOOK REVIEW

Sonic Wonderland: A Scientific Odyssey of Sound
Trevor Cox
ISBN 9781847922106 Published by Norton, 2014

Trevor Cox is a Professor of Acoustic Engineering at the University of Salford. As well as his professional academic career he is a public face for acoustics, having presented radio documentaries, videos and participation in outreach activities promoting science using sound. With this background and reputation for communication it is not surprising that this book is an easy read yet still includes the scientific basis for the examples he discusses.

The book has nine chapters; the first has the intriguing title "The most reverberant place in the world". This sets the style for each chapter of the book. Cox tells the story of the investigation of the space and as reviewer I am not going to reveal the location! He describes the measurement techniques, the findings and within all of this the explanation of why it meets the description of the chapter title. The titles of the other chapters are similarly catching "Ringing Rocks", "Barking Fish", "Echoes of the Past", "Going Round the Bend", "Singing Sands", "Quietest Places in the World", "Placing Sound" and the concluding chapter is "Future Wonders".

Cox travels the world to investigate each of these "sonic" examples and Australia gets a number of mentions including whip birds and squeaky sand. He treats each in a thoroughly scientific way yet in the style of a gripping yarn so that, even if you know the outcome or the reason, there is still the desire to read through to the end.

This book is highly recommended to a range of readers. It would make an excellent gift for a high school student, to a person with a general interest in the world around them and to a person who deals with sound and considers themselves an acoustician. I'm sure that even the latter would find something new and interesting in this book.

Marion Burgess

Marion Burgess is a research officer in the Acoustics and Vibration Unit of UNSW, Canberra. She is involved with presenting courses, research projects and consulting.

WORKPLACE HEALTH & SAFETY NEWS

Report On Reversing Alarms

An English version of the informative report on worker safety using 3 different types of reversing alarms by Canadian researchers (Laroche et al) has now been published by IRSST. They found a considerably more uniform sound field behind heavy vehicles for the broadband alarm than for the tonal and multi-tone alarms. Abrupt variations in sound pressure levels of 15 to 20 dB over short distances (less than 1 m) were noted with the tonal alarm. Thus, auditory perception of the broadband alarm in the proximity of a heavy vehicle should also prove more uniform. Despite laboratory findings showing some advantages of the tonal alarm, the broadband alarm may ultimately prove superior for detection and perceived urgency in the actual work environment given its more uniform sound propagation behind vehicles.. Also, a 2 to 4 dB advantage in loudness of the broadband alarm over the tonal alarm noted in laboratory conditions should hold true in the field. Front/back sound localization is also best with the broadband alarm compared to the tonal and multi-tone alarms, a finding critical for worker safety given that front/back confusions can lead workers to move or to focus their attention towards the wrong direction. Overall, no contraindication to the use of broadband reverse alarms was identified during objective and subjective measurements. To ensure their optimal use, however, and to clarify regulations, the researchers give the following recommendations:

a. Revise the SAE J994 standard to specifically include broadband alarms;
b. Use alarms that can self-adjust based on ambient noise levels and determine the optimal characteristics of adjustment algorithms to ensure adequate audibility in a variety of background noise scenarios;
c. Determine optimal mounting positions on heavy vehicles that will provide the most uniform sound propagation;
d. Use earplugs instead of earmuffs for better sound localization;
e. Familiarize workers with the broadband alarm before it is used in work environments, in the hope of increasing perceived urgency for this warning signal in the field.

http://www.irsst.qc.ca/media/documents/PubIRSST/R-833.pdf

Report On Costs Of Occupational Injuries

The IRSST has also recently published a report on Costs of Occupational injuries and diseases in Quebec that includes a monetary estimate for 'human costs'. This shows that 'exposure to noise' is the condition with the highest average cost per case at more than $154,000.

http://www.irsst.qc.ca/media/documents/PubIRSST/R-843.pdf

Webinar For Small Businesses

The the AIOH, in conjunction with Safe Work Australia, has produced a Webinar "Practical challenges of workplace health and safety for small business – with a focus on occupational noise". Join Kate Cole and Kristy Thornton from the Australian Institute of Occupational Hygienists (AIOH) for some practical tips to help with the safety of your small business. They use examples from readily available guidance to help you identify the hazards, eliminate these where possible, or control the hazards to help manage the risk. This presentation is aimed at small business owners; however anyone with an interest in how to manage occupational noise will find this useful.

http://www.safeworkaustralia.gov.au/sites/swa/australian-strategy/vss/pages/practical-challenges.

NOHARMS Study

Safe Work Australia has commenced a project the 'National occupational hazard and risk management surveillance (NOHARMS)' This is a study of workers' exposures to noise, vibration and dusts and risk management practices in mixed crop farming. http://www.safeworkaustralia.gov.au/sites/SWA/australian-strategy/priority-industries/Documents/national-agriculture-activity-plan.pdf.

Buy Quiet Program

NIOSH in the USA has a new web resource on "Buy Quiet". Its easy-to-use materials highlight the benefits of a Buy Quiet program, explain how to establish a program in a workplace, and provide additional resources for finding quieter tools and machinery.

www.cdc.gov/niosh/topics/buyquiet/

Exposure Database

NIOSH has also put online a database of occupational noise exposure measurements taken during its Health Hazard Evaluation Program between 1996 and 2012.

http://www.cdc.gov/niosh/data/datasets/RD-1005-2014-0/

Work Hazards Database

The US Center for Construction Research & Training has created Construction Solutions, a database of work hazards and practical control measures to reduce or eliminate hazards. This includes sections on noise and vibration. www.cpwrconstructionsolutions.org/hazard/. It also has the ROI (Return on Investment) Calculator to help evaluate the financial impact of new equipment, materials and work processes introduced to improve safety and health.

http://www.safecalc.org/

ICBEN Proceedings

The 11th International Congress on the Biological Effects of Noise was held in Nara, Japan in June 2014. The Proceedings are now available online. The three keynote powerpoints are also available – Wolfgang Babisch's "Burden of Disease and Environmental Noise", Lex Brown's "Urban road traffic noise" and Staffan Hygge's "Classroom Noise and its Effects on Learning".

www.icben.org/Post_Congress_2014.html

UKHRV 2014 - Human Vibration

The 49th UK Conference on Human Response to Vibration was held 9-11 September 2014 in Buxton UK (www.hsl.gov.uk/health-and-safety-conferences/ukhrv2014/programme)

Of particular interest is the paper from Paul Pitts of the HS Laboratory "Development of example hand-arm data for HSE Guidance". If anyone is interested in reading this paper, please contact Pam Gunn (pam.gunn@commerce.wa.gov.au). Another interesting paper from Svantek on their improved HAV assessment method is available from: http://svantek.co.uk/wp-content/uploads/2014/10/UKHRV2014_06_JKuczyński.pdf

Hearing Protector Accreditation

A New Zealand laboratory, Imtest - SAI Global (NZ) has recently been accredited to test the attenuation of hearing protectors to AS/NZS1270. This adds to the one in USA (Michael & Associates) and one in UK (SATRA).

http://www.saiglobal.com/product-certification/Imtest-Testing/

Noise and Vibration Partnership Group

UK industries in the energy, extraction, manufacturing and construction sectors along with the HSE have formed a Noise and Vibration Partnership Group. This industry led group will work together, long term, to increase awareness of the risks associated with noise and vibration in the workplace and promote effective management and control. They have two new posters:

http://www.hse.gov.uk/noise/nv-partnership-group.htm and

http://www.hse.gov.uk/construction/healthrisks/physical-ill-health-risks/noise.htm

Hearing Protection Tool

The UK HSE has a new on-line tool showing the detrimental effect removing hearing protectors has on attenuation of noise achieved.

www.hse.gov.uk/noise/hearingprotection/

Tinnitis Research

Researchers at the UWA Auditory Laboratory (Helmy Mulders and colleagues) are studying the mechanisms in the brain that produce tinnitus after noise trauma and the possibility of using drugs such as furosemide as a treatment.

http://www.actiononhearingloss.org.uk/community/blogs/our-guest-blog/one-step-closer-to-a-cure-for-tinnitus.aspx

They have recently published two scientific papers that can be found at:

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4126040/ and

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4023991/

WA WHS Bill Tabled

On 23 October 2014, the Western Australian Minister for Commerce tabled in Parliament the Work Health and Safety Bill 2014 (WHS Bill) and announced a public comment period closing on 30 January 2015 for this draft legislation. The WHS Bill is a WA version of the model Work Health and Safety Bill developed by Safe Work Australia for implementation in Australian states and territories. The WHS Bill contains the core provisions of the model WHS Bill with some modifications to suit the WA working environment. For more details on the changes to the Model WHS Bill and public comment see the WA Department of Commerce website.

Compiled by Pam Gunn

New Products

01dB Smart Building Acoustics Solution.

The verification of building acoustic quality requires technicians to take numerous measurements from multiple locations in the building. 01dB offers a simple, effective and versatile solution to help acoustic engineers improve their productivity:
Comprising a module integrated into a sound-level meter (FUSION or DUO), a software tool (dBInside) and a suite of noise sources and shock generators, this new solution presents some innovative characteristics. Unique on the market, the patented measurement recognition function reduces the time between each measurement. The smart organization of data inside the sound-level meter allows final reports to be produced more quickly. The dBInside software allows normalized indexes to be obtained as soon as the data is imported, without user intervention. In one click, the acoustics engineer can obtain a report with all tests performed.
Video and information from www.acoemgroup.com

Future Conferences

ICSV 22 Florence, 12 to 16 July 2015

The 22nd International Congress on Sound and Vibration (ICSV22) will be held from 12 to 16 July 2015 in Florence, Italy. The program will include invited and contributed papers and the following distinguished plenary lectures: Otto von Estorff, Germany, Finite Element and Boundary Element Modeling; Kirill Horoshenkov, United Kingdom, Porous Material Characterization via Acoustical Methods; Dick Botteldooren, Belgium, Modelling Noise Effects Including Soundscapes and Quiet Areas; Semyung Wang, South Korea, Sound Focusing and Practical Applications; Lily M. Wang, USA, Room Acoustic Effects on Speech Comprehension; Wim Van Keulen, Netherlands, Traffic Noise; Roberto Pompoli, Italy, Opera House Acoustics.
ICSV22 participants will be able to take part not only in a congress with a first-rate scientific program but will also be able to experience the vibrant culture of Italy. The Congress venue will be the Firenze Fiera Congress & Exhibition Centre, an exclusive area in the centre of Florence. There will be an extensive exhibition of sound and vibration control technology, measurement instrumentation and equipment, and various social activities will be featured.
For more information including deadlines http://www.icsv22.org.

Wind Turbine Noise, Glasgow 20-23 April

The Series of international conferences on Wind Turbine Noise started in 2005 and takes place every two years.
After a successful 2013 conference held in Denver USA, we are pleased to announce the 2015 Wind Turbine Noise meeting. We welcome delegates from all over the world with a wide range of backgrounds – academics, industrial and consultants as well as objectors, all of which means there are lively debates at the end of each session.
The conference will be held Monday 20th April to Thursday 23rd April 2015 at Radisson Blu Hotel, Glasgow, Scotland.
Information: http://windturbinenoise.eu/

EuroNoise 2015, 31 Maastrict May-3 June

Euronoise 2015, the 10th European Congress and Exposition on Noise Control Engineering, will be held at the heart of Europe where the first treaties leading to the creation of the European Union were signed. Acousticians and noise experts from all over the Europe will gather for the event on noise control, and soundscape in Europe, organised by the European Acoustics Association.
The Belgian and Dutch acoustical societies, ABAV and NAG, warmly welcome you to Maastricht for Euronoise 2015
Information http://www.euronoise2015.eu/

INTER-NOISE 2015, San Francisco 9 -12 August.

The 2015 International Congress and Exposition on Noise Control Engineering, will be held in San Francisco, California, United States of America, 9 -12 August 2015. The Congress is sponsored by the International Institute of Noise Control Engineering (I-INCE), and is being organized by the Institute of Noise Control Engineering of the United States of America (INCE/USA) and in cooperation with the Korean Society of Noise and Vibration Engineering (KSNVE). In addition, joint meetings with the ASME Noise Control Acoustics Division (NCAD) and the TRB Noise & Vibration Committee ADC 40 summer meeting which will overlap the Congress Tuesday through Thursday.

Check the website for updates but currently abstract submission deadline is 28 March with final papers due 22 May. Papers seeking optional review will be due four weeks before the normal abstract deadline. Early bird registration closes 11 May 2015.

Information http://internoise2015.com/

Diary

2015

13-15 April, Dubrovnik
NOVEM 2015
http://novem2015.sciencesconf.org

20 - 23 April, Glasgow, Scotland
Wind Turbine Noise 2015
http://www.windturbinenoise.eu/

11 – 15 May, Metz, France
International Congress on Ultrasonics (ICU 2015)
http://2015-icu-metz.gatech.edu/

31 May - 3 June, Maastricht, Netherlands
Euronoise 2015
http://www.euronoise2015.eu/

12 - 16 July, Florence, Italy
22nd International Congress on Sound and Vibration (ICSV22)
www.icsv22.org

9-12 August, San Francisco, USA
Inter-Noise 2015
http://internoise2015.com

6-10 December Singapore,
Wespac 2015
www.wespac2015singapore.com

2016

10-14 July, Athens, Greece
23rd International Congress on Sound and Vibration (ICSV23)
http://iiav.org/index.php?va=congresses

21 - 24 August, Hamburg, Germany
INTER-NOISE 2016
http://www.internoise2016.org/

5-9 September, Buenos Aires, Argentina
22nd International Congress on Acoustics (ICA 2016)
http://www.ica2016.org.ar/

12-16 September, Terrigal, NSW, Australia
International Workshop on Rail Noise (IWRN)
http://iwrn12.acoustics.asn.au/

2017

25 - 20 June, Boston, USA
Acoustics 2017
Joint meeting of the Acoustical Society of America and the European Acoustics Association
http://www.acousticalsociety.org

Meeting dates can change so please ensure you check the conference website: http://www.icacommission.org/calendar.html


Specifier: RAL Architects

Sustaining Members

The following are Sustaining Members of the Australian Acoustical Society.
Full contact details are available from http://www.acoustics.asn.au/sql/sustaining.php

3M AUSTRALIA
www.3m.com

ACOUSTIC RESEARCH LABORATORIES
www.acousticresearch.com.au

ACOUSTICA
www.acoustica.com.au

ACRAN
www.acran.com.au

ACU-VIB ELECTRONICS
www.acu-vib.com.au

ADAMSSON ENGINEERING
www.adamsson.com.au

AERISON PTY LTD
www.aerison.com

ARMSTRONG World Industries
www.armstrongceilings.com.au

ASSOCIATION OF AUSTRALIAN ACOUSTICAL CONSULTANTS
www.aaac.org.au

BARRISOL AUSTRALIA
www.barrisol.com.au

BORAL PLASTERBOARD
www.boral.com.au/plasterboard

BRUEL & KJAER AUSTRALIA
www.bksv.com.au

C-MAC INDUSTRIES (AUST) PTY LTD
www.cmac.com.au

CSR BRADFORD INSULATION
www.bradfordinsulation.com.au

EMBELTON
www.vibrationisolation.com.au

HOWDEN AUSTRALIA
www.howden.com

IAC ACOUSTICS (AUSTRALIA)
www.iac-acoustics.com/au/

INSTYLE CONTRACT TEXTILES
www.instyle.com.au

JACOBS
www.jacobs.com

KNAUF AMF AUSTRALIA PTY. LTD.
www.amfceilings.com.au

MASON MERCER AUSTRALIA
www.masonmercer.com.au

OFFICE OF ENVIRONMENT AND HERITAGE NSW
www.environment.nsw.gov.au

PEACE ENGINEERING
www.peaceengineering.com

PYROTEK NOISE CONTROL
www.pyroteknc.com

REGUPOL (AUSTRALIA) PTY LTD
www.regupol.com.au

SOUND CONTROL
www.soundcontrol.com.au

SOUNDSCIENCE
www.soundscience.com.au

VIPAC ENGINEERS AND SCIENTISTS
www.vipac.com.au

ACOUSTICS AUSTRALIA INFORMATION

EDITORIAL ENQUIRIES

The Editor, Acoustics Australia
AcousticsAustralia@acoustics.asn.au
www.acoustics.asn.au

SUBMISSIONS

All Articles and Technical Notes
www.springer.com/engineering/journal/40857

News, New Products & Books for Review
acousticsaustralia@acoustics.asn.au

GENERAL BUSINESS

Advertising

Mrs Leigh Wallbank
PO Box 70, OYSTER BAY 2225
Tel (02) 9528 4362
Fax (02) 9589 0547
wallbank@zipworld.com.au

SUBSCRIPTION RATES

From 2015, Acoustics Australia will be published by Springer. For Subscription details please visit: http://www.springer.com/engineering/journal/40857

SUBSCRIPTION ENQUIRIES

www.springer.com/engineering/journal/40857

ADVERTISING ARTWORK

Cliff Lewis Printing
91-93 Parraweena Road
CARINGBAH NSW 222
Tel (02) 9525 6588 Fax (02) 9524 8712
email: matthew@clp.com.au

ADVERTISING RATES

Casual Rates - Single Issue

	Non-member	Sus Mem
Full Page	862.40	768.90
Half Page	559.90	503.80
1/3 Page	431.20	388.30
1/4 Page	359.70	323.40

COLUMN RATE $29.70
WEB ADS $467.50

All rates are AUD and include GST

All enquiries to:
Mrs Leigh Wallbank
Tel (02) 9528 4362 Fax (02) 9589 0547
wallbank@zipworld.com.au

AUSTRALIAN ACOUSTICAL SOCIETY ENQUIRIES

NATIONAL MATTERS

* Notification of change of address
* Payment of annual subscription
* Proceedings of annual conferences

Richard Booker - General Secretary
Australian Acoustical Society
PO Box 1843 Toowong DC QLD 4066
Tel: (07) 3122 2605
GeneralSecretary@acoustics.asn.au
www.acoustics.asn.au

AAS MEMBERSHIP

For 2014/15 Financial Year (Inc. GST):

Fellow and Member	$150
Graduates	$115 - $150
Associate and Subscriber	$115
Retired and Maternity	$45
Student	No Charge

DIVISIONAL MATTERS

Enquiries regarding membership and sustaining membership should be directed to the appropriate State Division Secretary

AAS - NSW Division
Angus Leslie
c/- AECOM
Level 21, 420 George Street
Sydney, NSW 2000
Tel: (02) 8934 0035
Fax: (02) 8934 0001
angus.leslie@gmail.com

AAS - QLD Division
PO Box 760
Spring Hill Qld 4004
Sec: Richard Devereux
Tel: (07) 3217 0055
Fax: (07) 3217 0066
rdevereux@acran.com.au

AAS - SA Division
The Secretary
Australian Acoustical Society
South Australia Division
Level 28, 91 King William Street,
Adelaide, SA 5000
Sec: Darren Jurevicius
Tel: (08) 7100 6400
Fax: (08) 7100 6499
darren.jurevicius@aecom.com

AAS - VIC Division
The Secretary
Australian Acoustical Society
Victoria Division
6 Milburn Grove
East St Kilda, VIC 3183
Sec: Peter Pirozek
Tel: (03) 8668 6356
ppirozek@globalskm.com

AAS–WA Division
Australian Acoustical Society
Western Australia Division
Unit 3/2 Hardy Street
South Perth WESTERN AUSTRALIA 6151
Sec: Kingsley Hearne
Tel: (08) 9474 5966
kingsley@gabriels.net.au

ACOUSTICS AUSTRALIA ADVERTISER INDEX - VOL. 42, No. 3

Noise Monitoring Station

SV 200 Noise Monitoring Station is a fully integrated solution for unattended noise monitoring applications featuring an internal web server for system configuration, live data streaming, data management and battery powered operation providing true flexibility for both short and long term measurements.

Svantek
SV 200

All In One Monitoring Station

APPLICATIONS

- Unattended noise monitoring with remote communication
- Airport nose monitoring
- Community noise monitoring
- Construction site noise monitoring
- Industrial noise
- Noise monitoring during concert and sport events
- Noise mapping
- Wind farm noise monitoring

FEATURES

- ‘All in one’ design for portable, mobile and permanent installation
- Real-time 1/1 or 1/3 octave analysis
- Audio events recording
- Rugged housing protecting the system against harsh environmental conditions (IP66)
- Integrated electrostatic actuator for full system verification
- Type 1 according to IEC 61672
- Integrated high speed 3G or Wi -Fi modem
- Automatic time synchronization
- Visit our website for full specs.

Call Today

Acu-Vib Electronics
www.acu-vib.com.au
Unit 14 /22 Hudson Ave CASTLE HILL 2154

(02) 9680-8133

WE CAN HELP YOU MAXIMISE OPERATIONS BY MINIMISING NOISE

Exceeding noise limits can bring your mining operation to a costly standstill. It is hard to gauge, unless you use Noise Sentinel from Brüel & Kjær. This alerts you before you reach critical levels identifying the source so you know what action to take. Now you can maximise equipment capacity safe in the knowledge that you are within limits: good for production, good for community relations and, most of all, good for growth. Let us handle your noise management, so you can focus on business.

www.bksv.com/mining
Sydney: +61 2 9889 8888 · Melbourne: +61 3 9508 4907 · auinfo@bksv.com.au

Brüel & Kjær
BEYOND MEASURE

BN 1396 – 11