Contents lists available at ScienceDirect

Hearing Research

Research Paper

The detection of 'virtual' objects using echoes by humans: Spectral cues☆

Daniel Rowan [a, *], Timos Papadopoulos [b], Lauren Archer [c], Amanda Goodhew [d], Hayley Cozens [e], Ricardo Guzman Lopez [f], David Edwards [g], Hannah Holmes [h], Robert Allen [a]

[a] Institute of Sound and Vibration Research, University of Southampton, Southampton, Hants, SO17 1BJ, UK
[b] Biodiversity Institute, Department of Zoology, and Machine Learning Research Group, Department of Engineering Science, University of Oxford, Oxford, UK
[c] Royal Berkshire NHS Foundation Trust, Reading, UK
[d] Guy's and St Thomas' NHS Foundation Trust, London, UK
[e] Plymouth Hospitals NHS Trust, Plymouth, UK
[f] Directorate General of Civil Aviation, Chile
[g] Yeovil District Hospital NHS Foundation Trust, Yeovil, UK
[h] InHealth, Southampton, UK

ARTICLE INFO

ABSTRACT

Some blind people use echoes to detect discrete, silent objects to support their spatial orientation/navigation, independence, safety and wellbeing. The acoustical features that people use for this are not well understood. Listening to changes in spectral shape due to the presence of an object could be important for object detection and avoidance, especially at short range, although it is currently not known whether it is possible with echolocation-related sounds. Bands of noise were convolved with recordings of binaural impulse responses of objects in an anechoic chamber to create 'virtual objects', which were analysed and played to sighted and blind listeners inexperienced in echolocation. The sounds were also manipulated to remove cues unrelated to spectral shape. Most listeners could accurately detect hard flat objects using changes in spectral shape. The useful spectral changes for object detection occurred above approximately 3 kHz, as with object localisation. However, energy in the sounds below 3 kHz was required to exploit changes in spectral shape for object detection, whereas energy below 3 kHz impaired object localisation. Further recordings showed that the spectral changes were diminished by room reverberation. While good high-frequency hearing is generally important for echolocation, the optimal echo-generating stimulus will probably depend on the task.

© 2017 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Some sighted, visually impaired and blind people are able to use auditory cues from echoes to determine various features of otherwise silent objects (Kolarik et al., 2014) and to avoid objects during locomotion (Kolarik et al., 2016). Furthermore, some blind people use echolocation in daily life to enhance their spatial navigation, especially in unfamiliar environments (Thaler, 2013). The ability to detect objects is fundamental to echolocation, including object localisation combined with head movement (Rowan et al., 2013). However, there is a lack of understanding of which acoustic features of echoes and the echo-producing sounds ('emissions') best support it. One challenge is that those features probably depend on a range of factors relating to the object, task and environment. If there is overlap in time between the emission and the echo (e.g. if the object is close), the cues available to detect the object may be quite different than those available when there is no overlap (e.g. if the object is much farther away) because of the acoustic interference between the emission, which contains no information about

the object, and the echo. For example, the difference in time of arrival at the ear of an echolocator between an emission from their mouth and an object 1 m away is around 5 ms, which is similar to the duration of transient sounds typically used by blind echolocators (Rojas et al., 2009, 2010; Schörnich et al., 2012). Interference between the emission and echo can therefore be expected when objects are at close range such as when the echolocator might need to take evasive action to avoid harm. We previously considered the effect of this interference on a binaural object localisation task (Rowan et al., 2013, 2015); one finding was that excluding energy in the emission below 2 kHz improved object localisation. This finding may not generalise to object detection. In this paper, we focus on elucidating the cues to object detection available when there is interference between the echo and the emission.

Studies of human echolocation typically distinguish between two types of auditory attributes that might be used for object detection: loudness and pitch (e.g. Cotzin and Dallenbach, 1950; Schenkman and Nilsson, 2011). By 'loudness', the studies presumably refer to the fact that if emission and echo arrive at the ear simultaneously, the overall level of the combined sound can be higher than with the emission alone, which might be heard as a change in loudness; we refer to this as the 'overall level' cue. However, there are two other level cues. Firstly, the level of the sound within a narrow frequency range, e.g. the bandwidth of an auditory filter, might be higher in the presence, compared to absence, of the object; we refer to this as the 'within-channel level' cue. The within-channel level cue can be substantially larger than the overall level cue. Secondly, if the within-channel level cue is not the same across auditory filters, the spectral shape of the sound must also change; that can be determined by comparing the changes in level across the outputs of auditory filters with different centre frequencies. We refer to this as the 'across-channel level' cue, also known in the literature as a spectral profile cue (Green, 1988) or an excitation pattern cue (Moore and Glasberg, 1983). A change in spectral shape could occur because of interference between the emission and the echo creating comb-filtering and because the echo may not have the same spectral shape as the emission due to imperfect reflection by the object. The distinction between these three types of level cue has not always been clearly made in the literature. For example, Schenkman and Nilsson (2011) distinguished only between overall level and pitch cues for detection, and they attempted to construct a stimulus with only 'pitch' information by removing only the overall level cue. This stimulus may still have contained within- and across-channel level cues. Judging from their Fig. 3, which plots the spectra of the stimuli in one-third-octave bands for the object absent and present, within- and between-level cues were available in the 'pitch only' stimulus at an object distance of 1 m, were apparent but subtle at 2 m and were not apparent at 3 m. Whether listeners can use within- and across-channel level cues with echolocation-related stimuli and tasks is unclear. These three level cues might be apparent when comparing the sounds over their entire durations ('static' level cues) or during

the sound with the object present ('dynamic' level cues). This is because of the latency in the arrival of the echo at the ears compared to the emission. Interaural level difference cues could arise, especially if the object is not straight ahead of the echolocator, and temporal cues could also contribute to object detection, such as via the perception of pitch (sometimes referred to as 'repetition pitch').

The first aim of this study was to characterise the level cues for object detection associated with several objects in otherwise anechoic conditions through the analysis of acoustic recordings. This was done over a range of distances so as to be relevant to a wide range of emission durations. The second aim was to acquire baseline data on object detection ability from a listening experiment using several objects and distances, and sighted listeners. The recordings were combined with synthetic emissions and played over earphones, known as the virtual auditory space technique, a method we used previously for object localisation (Rowan et al., 2013, 2015).

Of the static level cues that are potentially available when the echo and emission interfere, the across-channel level cue might be particularly useful because it is expected to be more robust to unpredictable changes in level and spectrum of the emission and to fluctuating background noise. Blind people might also be better able to exploit an across-channel cue than sighted people (Doucet et al., 2005). While it is well established that humans can detect changes in spectral shape using an across-channel cue (e.g. Green, 1988), few studies have investigated it for sounds with spectra above 4 kHz, where the prominent changes in spectral shape might occur with echolocation. One such study found that the detection of single peaks or notches was poorer at 8 kHz than at 1 kHz using noise bands of various bandwidths (Moore et al., 1989). The third aim of this paper was to determine whether humans can access the static across-channel cue in an object detection task. A second listening experiment was carried out, again using a virtual object in otherwise anechoic conditions, to determine if an across-channel cue can be used for object detection.

2. Analysis of recordings

2.1. Method of recording impulse responses

The geometrical arrangement used here was identical to that used previously (Rowan et al., 2013) except that the object was placed centrally. Binaural impulse responses (IRs) were measured from the electrical input to a small loudspeaker (KEF HTS3001) to the electrical outputs of the in-ear microphones of a human manikin (KEMAR, head only) in an anechoic chamber ($\geq$0.1 kHz) using a maximum-length sequence, with 16-bit resolution and 88.89-kHz sampling rate. The loudspeaker driver was positioned 0.25 m below and 0.05 m in front of KEMAR's interaural axis, which itself was 0.975 m above the chamber's grid floor; see Fig. 1. See Papadopoulos et al. (2011) for more details. Objects were a:

1. Small square medium-density fibre (MDF) board with dimensions 0.55 m × 0.55 m (0.01-m thick), as in Rowan et al. (2013, Experiment 1b)
2. Small square metal board: an aluminium sheet (0.002-m thick) was mounted on the square MDF board and had the same dimensions
3. Small circular MDF board with a diameter of 0.621 m (0.01-m thick) and thus approximately the same surface area as the small square boards
4. Large MDF board 1.6-m high and 1.22-m wide (0.01-m thick)
5. Human (author AG; 1.69-m tall, 0.43-m wide at the height of the loudspeaker).
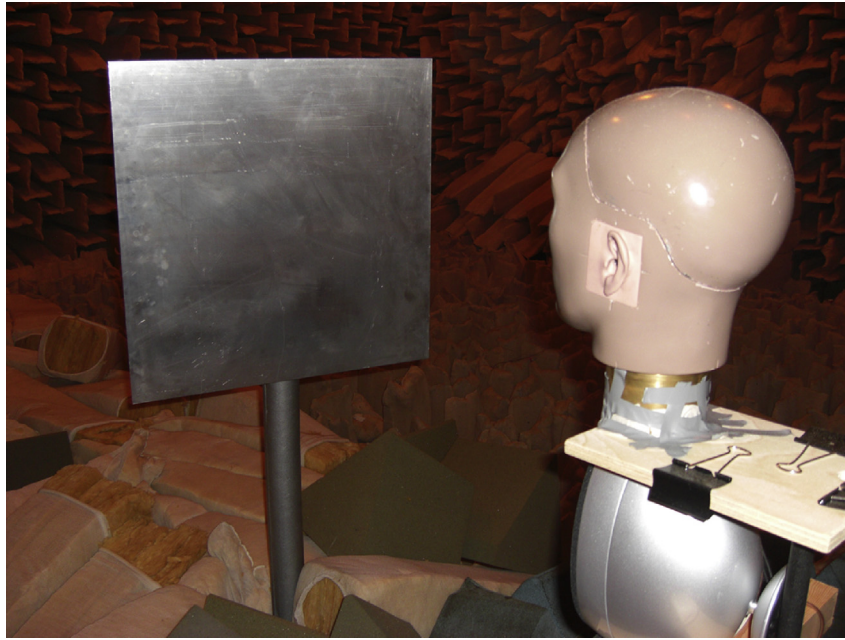
Fig. 1. The environment used for measuring impulse responses showing one of the objects.

For the recordings, the boards were placed vertically with their centre at the same height as, and on the midline of, KEMAR's interaural axis at distances from 1.5 m to 4 m in 0.5-m intervals plus 0.9 m, as in our previous studies. Fig. 1 illustrates the arrangement with the metal board.

The left panel of Fig. 2 shows three example pairs of binaural IRs: free field, metal board at 0.9 m and metal board at 2 m. The right panel of Fig. 2 illustrates the delay between the direct and reflection parts of the IR based on autocorrelation of metal board IRs (crosses) and that predicted from the physical arrangement (dotted line). The agreement between predicted and estimated delay for the boards was always within 2% and usually within 1%. Cross-correlation analyses of the output of a gammatone filter bank, as conducted by Rowan et al. (2015), suggested that no interaural time or interaural coherence cues were available (see also Section 6.3).

2.2. Static overall level cue

Fig. 3 plots the root-mean-square (rms) level of the *echo only* relative to the rms level of the emission only as a function of distance for all objects. The rms levels were estimated in MATLAB by taking the convolution of windowed versions of the measured IRs with a 50-s-long band of Gaussian noise (0.2−20 kHz). For the rms level of the echo only, the IRs with the object present were windowed between 4.7 ms and 29.7 ms to remove the part related to the emission. The rms level of the emission only was calculated based on the free-field IR from the right ear windowed between 0 and 4.7 ms. The rms level of the 'noise floor' was calculated from the free-field IR windowed between 4.7 ms and 29.7 ms. Different band-limited versions of Gaussian noise were used to generate the stimuli that were used in the listening experiments. The estimated measurement uncertainty of these rms levels based on comparisons of repeat measurements of the IRs (two standard deviations) is ± 0.3 dB. The square metal board is highlighted because it was used in the listening experiments. For distances below 2 m, the level did not drop as expected by the inverse-square law (−6 dB with a doubling of distance) due to near-field effects. The irregular, non-monotonic patterns are to be expected from natural, imperfect reflectors of finite size.
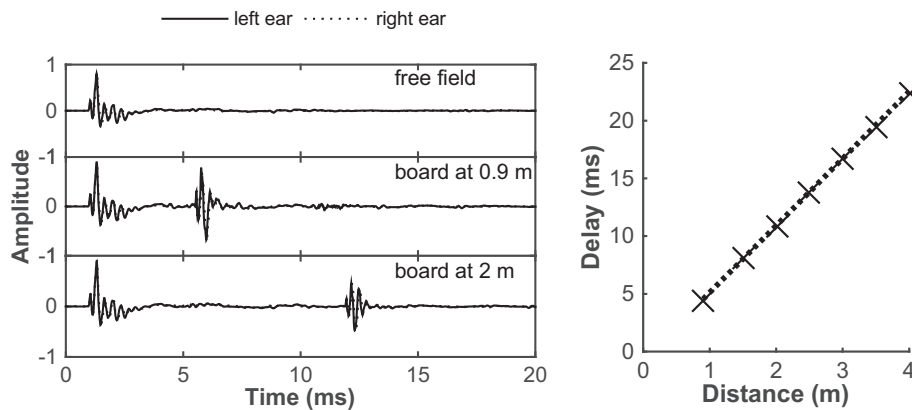

Fig. 2. Left panel: example binaural impulse responses without an object ('free field') and with the metal board at two distances. Right panel: the delay between the direct (related to the emission) and reflection (related to the echo) parts of impulse responses as a function of distance based on autocorrelation for metal board impulse responses (crosses) and predicted from the physical arrangement (dotted line). See Section 2.1 for a description of all objects used.
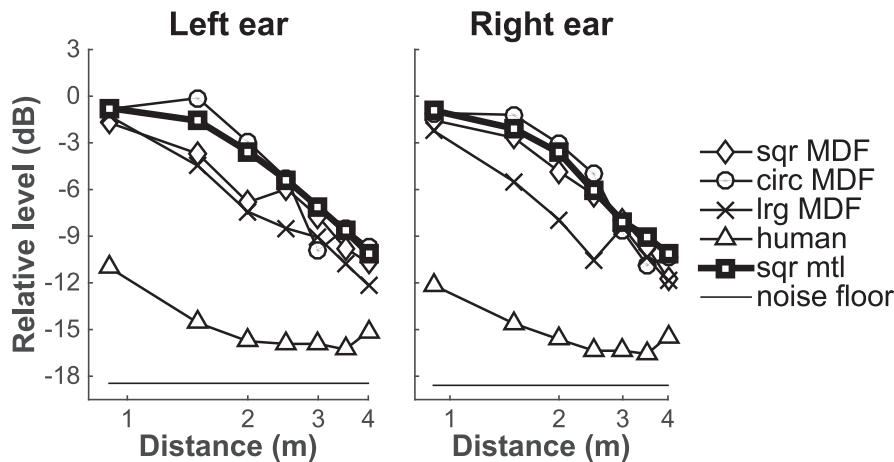
Fig. 3. Estimated overall rms level of the *echo only* relative to the rms level of the *emission only* as a function of distance for all objects, and the noise floor. See Section 2.2 for details.

Fig. 4 plots the difference in rms level between object present and absent (free-field) conditions, i.e. the static overall level cue, based on the complete IRs convolved with the same band of noise used previously. (Again, different band-limited versions were used in the listening experiments.) The estimated measurement uncertainty for each level difference is ±0.45 dB. Fig. 4 also plots data obtained by Schenkman and Nilsson (2010; Appendix 1) using an acoustic manikin in an anechoic chamber and a 500-ms-long noise emission; they do not give the values for each ear separately so the same values are included in both plots. Their emission source was placed in a similar position relative to the microphones as ours and their object was a circular metal disc with a diameter of 0.5 m. Their measurements were in direct response to the 500-ms-long noise emission rather than based on IRs like ours.

At approximately 1 m, the overall static level cue was similar for our boards and Schenkman and Nilsson's disc. However, the change in overall level as distance increased varied between the boards. Assuming that humans can detect an overall level change of 0.5–1 dB for long-duration, broadband stimuli (Epstein and Marozeau, 2010), Fig. 4 suggests that this cue is viable for some hard objects for distances up to at least 4 m, which in turn suggests that Schenkman and Nilsson's conclusion regarding the limit of object detection to 2 m may not be generalizable. The overall level cue does not seem to be viable for the human reflector.
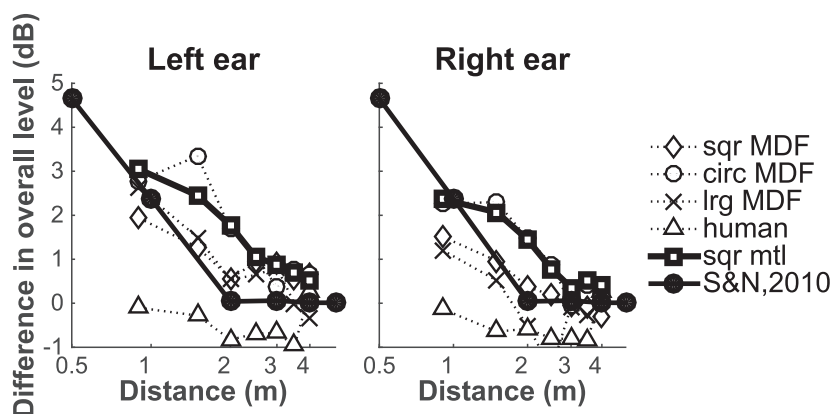
2.3. Static within- and across-channel level cues

Fig. 5 plots Fast Fourier Transforms of the IRs obtained with the free field (upper panel) and the metal board at 0.9 m (lower panel) for the right ear of KEMAR; broadly similar results were found for the other hard objects. The upper panel shows the full IR (emission only) and the windowed portion as in Fig. 3 (noise floor). The lower panel shows the full IR (emission and echo) and the windowed portion again as in Fig. 3 (echo only). In both panels, 0 dB represents the peak value for the full IR with the board present. The measurement uncertainty was estimated to be less than ±2 dB at most frequencies, although it was considerably greater at some frequencies, for example around the notch in the noise floor at 0.2–0.3 kHz. A notch around that frequency was apparent in some other IRs but not all. There was energy in the emission at KEMAR's ears across a wide range of frequencies and in the echo only above about 0.6 kHz (as predicted based on the size of the board) compared to the noise floor. When compared to the echo only, the full IR with the board present demonstrated the characteristic signs of comb filtering above 0.6 kHz, peaks and notches that occurred at regular frequency intervals (on a linear scale) corresponding to the reciprocal of the time delay between emission and echo.

To illustrate the within- and across-channel cues, Fig. 6 plots differences in auditory excitation patterns between the board present and absent, using a model of the auditory periphery (Chen

Fig. 4. Estimated difference in rms level between object present and object absent (free-field) conditions for our objects (open symbols; sqr = square; circ = circle; lrg = large; mtl = metal) and for the object used by Schenkman and Nilsson (2010; Appendix 1; 'S&N, 2010') as a function of distance. See Section 2.2 for details.
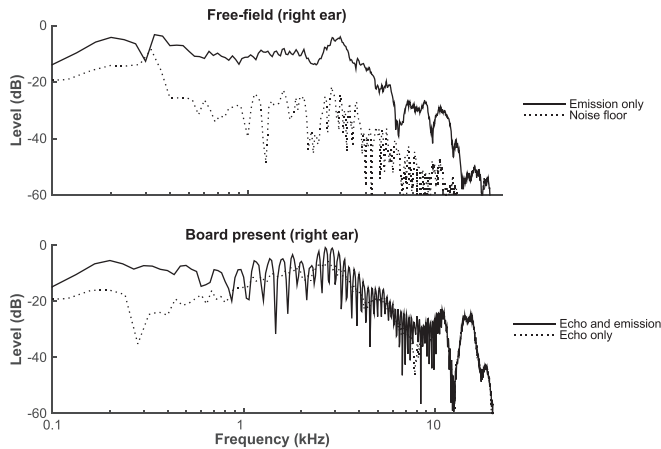
Fig. 5. Fast Fourier Transforms of impulses responses for the metal board at 0.9 m. Upper panel: with the echo removed and either the emission absent ('noise floor') or present ('emission present'). Lower panel: with the echo present and either the emission absent ('echo only') or present ('echo and emission'). See Section 2.3 for details.

et al., 2011). To obtain these, the full IRs were convolved with long-duration bands of Gaussian noise with a frequency range from 0.2 to 12 kHz for the small square MDF board and the metal board at all distances, the thickest lines showing the results at the distance of 4 m. The upper frequency limit of the noise was set to 12 kHz here for consistency with the listening experiments where the frequency response of the earphone was limited to 12 kHz. The peaks for frequencies up to approximately 0.3 kHz largely depended on which particular free-field IR recording was used. The excitation level differences above 0.3 kHz are more meaningful and associated

Fig. 6. Differences in auditory excitation ('ex') level between the board present and absent, using a model of the auditory periphery (Chen et al., 2011) for the MDF square (upper panels) and metal square (lower panels). The different lines indicate different distances, the thickest line being for the results at a distance of 4 m. See Section 2.3 for details. The vertical arrows indicate the main differences in excitation level.

with uncertainty of ±0.5–1.0 dB.

The comb-filtering effects were preserved in the excitation patterns for frequencies up to approximately 2 kHz, above which the auditory filters did not resolve the spectral ripples. The main differences in excitation level with the board present compared to absent occurred above approximately 2 kHz (see the vertical arrows in Fig. 6) and extended to at least 12 kHz; these are also apparent for distances up to at least 4 m. The differences between excitation patterns are often substantial and localised in frequency, providing clear potential within- and across-channel level cues; similar results were found with the other boards. The within- and across-channel level cues here are more distinct than for the metal disc reported by Schenkman and Nilsson (2011) using one-third-octave-band spectra (their Fig. 3). Differences in excitation level were apparent with the human reflector but were considerably smaller and restricted to distances of 0.9–1.5 m. Figs. 4 and 6 also indicate that the overall and within-channel level cues can be larger in one ear than the other, providing a potential interaural level cue, presumably arising from the board not being perfectly straight-ahead and orthogonal to both of KEMARs ears, or perhaps due to asymmetries with KEMAR itself.

3. General methods for listening experiments

The aim of Experiment 1 was to check object detection ability with our recordings under a range of conditions some of which allowing comparisons with Schenkman and Nilsson (2010). Experiment 2 focused on the static within- and across-channel level cues for a metal board at 4 m, specifically to determine whether an across-channel level cue can be used by sighted and blind listeners.

Approval of the ISVR Human Experimentation Safety and Ethics Committee was obtained before commencing these experiments. Sighted listeners were recruited from the university population and were otologically and ophthalmologically normal (excluding corrected short-sightedness). Some had prior experience of object localisation experiments. The blind listeners are described in Section 5. Unless indicated otherwise, all listeners responded to pure tones at 20 dB HL (as a screening level) for frequencies from 0.25 to 8 kHz at octave intervals and also at 12.5 kHz.

A three-interval, three-alternative forced choice format was used: two intervals contained a stimulus convolved with IRs with the board absent and one interval contained a stimulus convolved with IRs with the board present, in random order and each separated by a 400-ms silent gap. Listeners were required to determine which interval contained the 'odd one out' by selecting one of three buttons. The correct answer was displayed on a screen for 400 ms for sighted listeners and presented acoustically for blind listeners after a response was made.

Stereo sound files were generated and manipulated, and the procedure controlled and responses collected using custom-written MATLAB code; sound files were played out at 44.1 kHz and 16 bits (Creative, Extigy). Stimuli consisted of 400-ms-long bands of Gaussian noise convolved with bilateral IRs; each band of noise was generated independently prior to convolution and filtered using a 9th-order zero-phase filter. The resulting stimuli were played over Etymotic Research ER2 insert earphones to listeners seated in a quiet room. Stimuli were calibrated such that broadband stimuli with the board absent were presented at 65 dBA.

Data are presented as box plots, with circles indicating values for individual listeners ('outliers') if greater than 1.5 times the interquartile range away from the nearest quartile; the grey area represents the 99% range expected from guessing (a score outside of this range has a 99% confidence interval that excludes 50%). Statistical analysis was conducted on arcsine-transformed scores, using parametric methods when the data were at least approximately
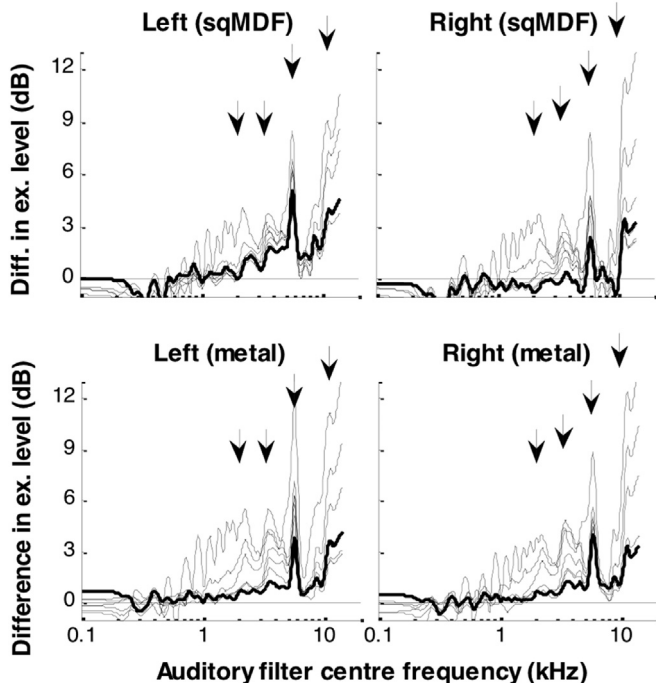
normally distributed. T-tests were two-tailed, paired-samples tests unless indicated otherwise.

Additional signal processing, the specific organisation of test sessions and any exceptions to these methods are described with each experiment.

4. Experiment 1: object type and distance

Experiment 1 measured object detection with four objects (small square MDF board; large MDF board; small square metal board; human) and three distances (0.9 m, 2 m and 4 m) with broadband noise (BBN). The human object was only tested at 0.9 m because the scores for all listeners in pilot testing at and beyond 2 m were indistinguishable from chance. Fifteen sighted listeners (3 male, 12 female) took part. They completed 26 trials for each of the 10 conditions in an order that was balanced across listeners and then revisited the conditions in reverse order, giving a total of 52 trials per condition. Before each condition, several familiarization trials were run with the metal board at 0.9 m, the responses from which were discarded. The BBN had a bandwidth of 0.1—12 kHz and listeners heard the full, binaural stimuli.

The results are shown in Fig. 7. All listeners scored better than chance in all conditions except for two listeners with the large MDF board at 4 m. Overall, scores were clearly higher for the boards than the human at 0.9 m and reduced with increasing distance; ceiling effects precluded sensible statistical confirmation of the effect of distance. T-tests confirmed that scores at 4 m for the large MDF board were lower than for both the small MDF board and the small metal board ($p < 0.001$), while the scores for the small boards were similar ($p > 0.1$).

This experiment demonstrates the potential for humans to detect some hard surfaces at distances beyond Schenkman and Nilsson's (2010, 2011) 2-m limit and probably beyond 4 m. This is consistent with Rice et al. (1965), who found that five blind listeners who could detect centrally placed physical surfaces to at least 2.7 m, and our acoustical analysis of the stimuli. The lower scores for our large MDF board compared to the smaller boards is consistent with the acoustical analysis although it is not clear if this represents a general effect of board size (e.g. perhaps because less reflections to the ears are generated from the edges as board size increases), the peculiarities of that board or peculiarities of the
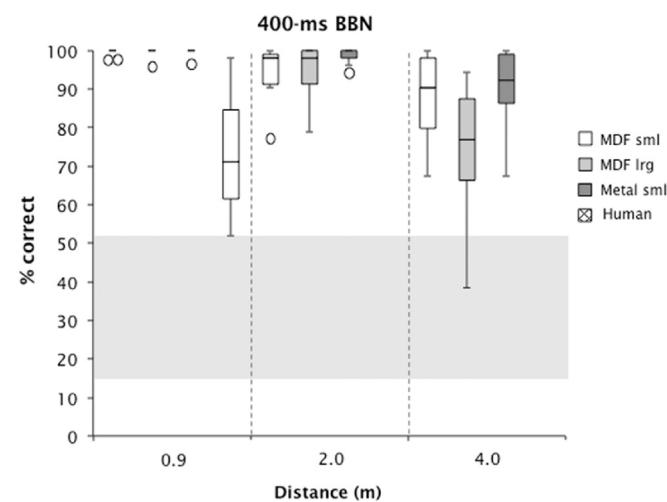
recordings for that board.

The object detection task used in Experiment 1 did not require listeners to recognise the presence of a board, but only to distinguish two stimuli. One could imagine real-world scenarios where one or other approach might be important. To check that sighted listeners can learn to recognise the presence of a board, a pilot study was conducted using a single-interval yes-no task. Listeners were randomly presented with one stimulus either representing the board absent or board present and had to report whether the board was present or absent; the 'correct' answer was then flashed on a screen. Using 20 new listeners, 90 trials were collected in blocks of 45 with four objects at 0.9 m and with the BBN as in Experiment 1. Signal detection theory was applied to the raw frequencies of stimulus-response outcomes to derive d' (a measure of sensitivity as opposed to bias) with Snodgrass and Corwin adjustment to avoid infinite values (MacMillan and Creelman, 2005); a d' of 0 indicates no ability to recognise the board and a d' of approximately 4.6 indicates perfect ability. The results are shown in Fig. 8. All listeners scored highly for the hard flat boards, with most achieving perfect recognition; scores were poorer with the human reflector. These findings mirror those from Experiment 1 and provide some confidence that the results of Experiment 1 are not peculiar to an 'odd-one-out' scenario.

5. Experiment 2: within- and across-channel level cues

5.1. Specific methods

The main aim of this experiment was to determine whether the across-channel level cue, arising from interference between the emission and echo, could be used to detect an object. We again used

Fig. 7. Results of Experiment 1, showing detection accuracy (%) for the three-alternative forced task (as also used in Experiment 2) as a function of object distance for four object types for a 400-ms broadband noise (BBN) emission. Data for the human were only collected at 0.9 m.
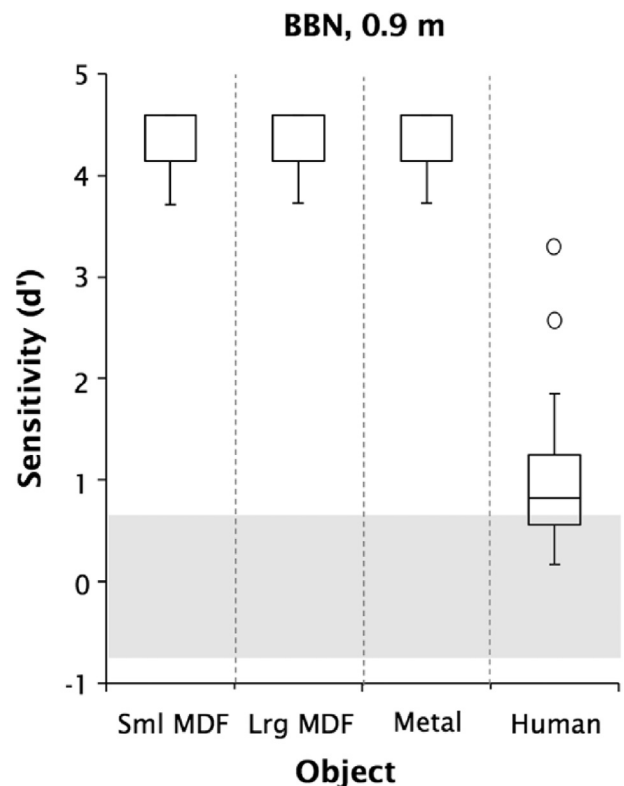

Fig. 8. Supplementary data for Experiment 1, showing 'sensitivity' (as opposed to bias), d', for a single-interval yes-no task at a distance of 0.9 m for various objects.

emissions with a duration of 400 ms to avoid floor effects that might have occurred for durations closer to those of the emissions used by expert echolocators. To remove binaural cues, the IRs from one ear (the right) were convolved with the bands of noise and presented to both ears, i.e. diotically. To remove dynamic cues, the first and last 12 ms of the stimuli were digitally removed and 1-ms-long cosine-squared onset and offset ramps were applied. To remove the overall level cue, the levels of the stimuli associated with board present and absent were equalized to within 0.1 dB. The stimuli may still have contained a temporal cue related to repetition pitch, and we will comment on that in the Discussion.

The five stimulus conditions were:

1. Broadband noise (BBN) from 0.3 to 12.5 kHz. A low-frequency edge of the filter of 0.3 kHz was used to avoid the potential spurious notches below 0.3 kHz in some IRs
2. Low-pass noise (LPN) from 0.3 to 3 kHz
3. High-pass noise (HPN) from 3 to 12.5 kHz
4. High-pass noise with a level rove, referred to as the 'HPN rove', which was included to disrupt the use of a within-channel level cue as described below
5. A low-frequency band of noise (0.3–2.8 kHz) together with a high-frequency band of noise (3.2–12.5 kHz) and the same magnitude of level rove (applied to the entire stimulus) as in HPN rove condition, referred to as the 'Frankenstein rove'. Importantly, *the low-pass band was always convolved with the IR from the board absent* condition and therefore provided no cue to the presence versus absence of the board. If the scores with this condition were higher than with the HPN rove condition, this would provide evidence for use of an across-channel cue.

Conditions were tested in blocks of 50 trials. The first four trials were for familiarization and were ignored. Two blocks were completed per condition per session, giving a total of 92 scored trials per condition per session. One set of blocks for every condition was completed before moving to the second set; one set of blocks typically took 30 min to complete. The first set of blocks was conducted in pseudorandom order. Testing started with the BBN, HPN and LPN conditions, found during piloting to be subjectively easier than the others, in random order. Testing then progressed to the HPN rove and Frankenstein rove in random order. The second set of blocks was conducted in completely random order.

The selection of the magnitude of level rove, object, object distance and ear were interrelated. On one hand, we wanted to use a flat, hard object such as an echolocator might need to avoid, and also for comparison with our previous data on object localisation using the small square MDF board. On the other hand, we wanted to prevent listeners from achieving high scores by using a within-channel level cue despite the level rove and to use a level rove of no more than ±15 dB (Lentz, 2005). Several options were considered in terms of the highest score that could be achieved using a within-channel level cue determined using a statistical model (Dai and Kidd, 2009) based on the output of a cochlear model (Chen et al., 2011). The process was described in detail in our previous paper (Rowan et al., 2015). The result was the selection of the IR from the right ear with the metal board at 4 m (see bottom panels of Fig. 6) and a level rove of ±14 dB using a rectangular distribution. (We argue in the Discussion that the main finding of this experiment will apply to shorter distances, for which the across-channel level cue may be more relevant.) That combination produced a maximum score expected from the use of the within-channel level of 63%. An individual listener must score 68% or higher to be statistically significantly higher than this with 99% confidence, given 92 trials per condition.

5.2. Experiment 2a: sighted listeners

Twelve new, sighted listeners (all postgraduate students; 3 male, 9 female aged between 22 and 30 years) participated in three test sessions; five wore lenses to correct their vision but were not classified as blind. The results are plotted in Fig. 9. The thin and thick horizontal lines show the 63% and 68% criteria described at the end of the previous section. The box plots and Shapiro-Wilk normality tests indicated that there were no marked deviations from normality ($p > 0.05$ in all cases but one, where $p = 0.04$). One-sample t-tests were conducted to compare the mean score for each stimulus condition against the expected probability from unbiased guessing using a Bonferroni-corrected criterion p value. The asterisks in Fig. 9 indicate statistically significant differences ($p \leq 0.003$; all others were $p \geq 0.02$). Correlations of the results across the sessions for each condition indicated that the findings were highly repeatable for the BBN ($r^2 \geq 83\%$; $p < 0.001$), HPN ($r^2 \geq 74\%$; $p < 0.001$) and Frankenstein rove ($r^2 \geq 84\%$; $p < 0.001$) conditions. Correlations for LPN were only statistically significant for Session 2 vs. Session 3 ($r^2 = 57\%$; $p = 0.005$; otherwise $r^2 \leq 26\%$); correlations for HPN rove were all non-significant ($r^2 \leq 29\%$).

Overall, listeners scored highest with the BBN. By the third session, scores were similar to those found in Experiment 1 in which all cues where available. Removing the information above 3 kHz in the LPN condition had a marked detrimental affect on scores, with only one or two listeners (the same ones across the sessions) appearing to score above chance on each session. This indicates that within- and across-channel level cues, and other monaural cues, between 0.3 and 3 kHz are weak; it is unclear whether any pitch cue would have been stronger in Experiment 1 where the bandwidth of the emission had a lower cut-off frequency of 0.1 kHz. Most listeners scored better than chance with the HPN and scored clearly better than with the LPN. Scores were higher overall for the BBN than for the HPN for every session ($p \leq 0.001$). The comparison between these three stimuli indicates that there is
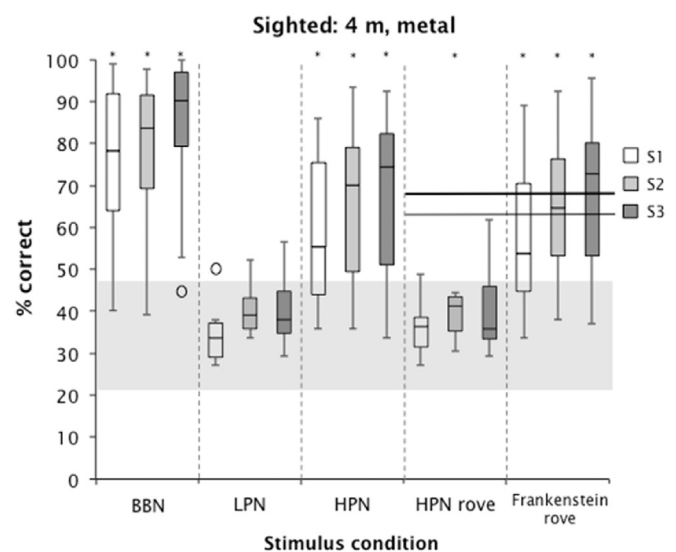

Fig. 9. Results of Experiment 2 with sighted listeners and the metal board only. Detection accuracy (%) is shown for each stimulus condition and test session (S1–S3). The thin horizontal line at 63% is the highest score expected from the use of within-channel level cues and the thick horizontal line at 68% is the score at or above which performance is statistically significantly better than that expected from the use of within-channel level cues. The asterisks indicate conditions where the sample mean was statistically significantly different from the expected percentage for guessing ($p \leq 0.003$; all others were $p \geq 0.02$). See Section 5.1 for an explanation of the stimuli.

a cue or cues available with the BBN that is not available with the HPN and LPN individually, possibly the across-channel level cue.

Scores for the HPN rove condition were indistinguishable from chance for most distances. This suggests that a within-channel level cue was being used in the HPN condition without rove. No listeners exceeded the criteria for ruling out the use of a within-channel level cue in the condition with the level rove. If an across-channel level cue was usable with the HPN, we would have expected scores in the HPN rove condition to be above chance; that they were not indicates that an across-channel level cue was not usable within the HPN.

In contrast, two aspects of the data shown in Fig. 9 indicate that an across-channel level cue was used in the Frankenstein rove condition. Recall that the Frankenstein rove condition was identical to the HPN rove condition with the exception that the stimuli contained a low-pass band of noise that was subject to an identical level rove. That low-pass band by itself provided no cue to the presence of the object since it was always processed as if the object were absent; rather, it provided a reference for changes in level in the HPN. Scores with the Frankenstein rove condition were clearly substantially better than for the HPN rove condition overall, even in Session 1. Also, most listeners had scores in the Frankenstein rove condition that exceeded both chance and the criterion for ruling out the use of a within-channel level cue by Session 3.

A repeated-measures analysis of variance comparing the three conditions with consistently above-chance average scores and good between session repeatability (i.e. BBN, HPN and Frankenstein rove) indicated a main effect of stimulus condition ($F_{2,22} = 26.6$; $p < 0.001$) and session ($F_{2,22} = 27.5$; $p < 0.001$) but no interaction ($F_{4,44} = 1.0$; $p > 0.1$). Post-hoc t-tests indicated that BBN gave a higher mean score ($p < 0.001$) than the other two conditions, which were not significantly different ($p > 0.1$), and that across all conditions the scores for each session were higher than for the preceding one ($p \leq 0.006$).

5.3. Experiment 2b: blind listeners

Twelve blind listeners were recruited via charities and societies for the blind and visually impaired within the Hampshire and Brighton area. Of the 12, one took part in a pilot study and six were excluded. Of the six excluded, three had residual vision meaning they could not be classified as blind; one had additional disabilities that led her to fatigue easily; two withdrew from the study before data collection. This left a sample of five; see Table 1. None of the listeners reported using specific self-vocalizations to navigate using echolocation. The listener identifier is consistent with our previous paper (Rowan et al., 2013); listeners B2, B4, B5 and B6 participated in both studies. As for Experiment 3 of Rowan et al. (2013), testing of blind listeners took place in their homes if they preferred. Ambient noise was monitored to be no higher than 30 dBA throughout. The audiometric criterion for inclusion was relaxed to 30 dB HL for three listeners at 8 kHz and 12.5 kHz bilaterally. Blind listeners participated in one session, which was identical in structure to Session 1 from Experiment 2a.

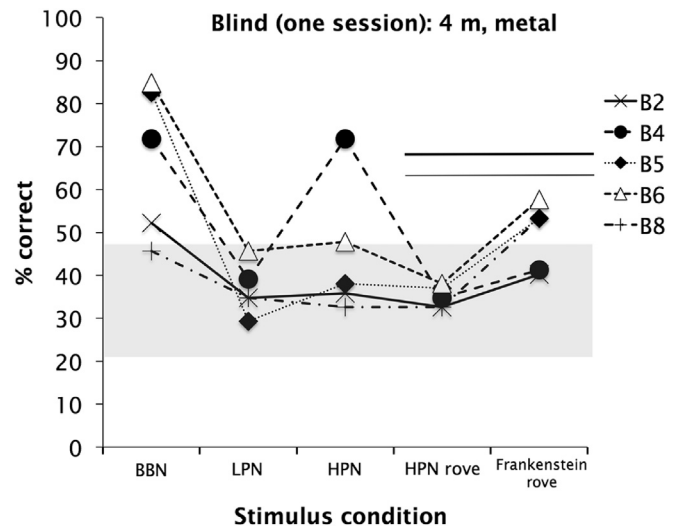The results are shown in Fig. 10. The general pattern is similar to

Fig. 10. Results of Experiment 2 with blind listeners. As in Fig. 9, except that each line shows the data for one blind listener and was one session only.

that for sighted listeners, in Session 1. Most blind listeners performed better than chance with the BBN. Scores were universally indistinguishable from chance in the LPN and HPN rove conditions; some listeners had above-chance scores for the HPN rove and Frankenstein rove conditions. None of the listeners exceed the criterion score to rule out use of a within-channel level cue in the HPN rove and Frankenstein rove condition. However, the scores for three listeners (B5, B6 and B8) were indistinguishable from chance with the HPN rove but above chance for the Frankenstein rove condition suggesting that they could make use of an across-channel level cue.

6. Discussion

6.1. General

There is a rich set of cues available for object detection, including various level cues when the emission and echo overlap in time such as would occur at close range even with the transient emissions typically used by blind echolocators. Those level cues may be important for object avoidance behaviour. They might also occur at farther distances if the emission duration was longer, such as with speech. We found that object detection was possible for inexperienced sighted listeners with an emission having a long duration relative to the duration of emissions typically used by expert echolocators, for three hard flat virtual surfaces placed centrally in an anechoic environment for distances up to at least 4 m, consistent with some previous research (Rice et al., 1965; Rowan et al., 2013). In contrast, Schenkman and Nilsson (2010) concluded that none of their sighted listeners and only two of their ten blind listeners scored above chance with a single hard flat surface at 4 m using similar emissions; most listeners were unable

Table 1
Details of the five blind listeners who took part in Experiment 2b. The listeners' identifiers are consistent with our previous paper (Rowan et al., 2013).

Listener	Sex	Age at testing (yrs)	Age at onset of blindness	Aetiology	Residual sight
B2	Female	46	13 years	Secondary glaucoma	None
B4	Male	44	42 years	Retinitis pigmentosa	Some light sensitivity
B5	Female	48	26 years	Trauma (road traffic accident)	None
B6	Female	64	11 months	Retinoblastoma	None
B8	Female	67	<1 year	Retinitis pigmentosa and secondary glaucoma	None

to detect the surface beyond 2 m. This can be explained, at least in part, by the weaker levels cues associated with Schenkman and Nilsson's hard surface compared to ours of similar size (see Fig. 4), presumably due to differences in reflection properties of the surfaces. Our study also differed from that of Schenkman and Nilsson in other ways, such as using a three-alternative rather than two-alternative forced choice task, which might contribute to the differences in findings. Other studies provide data related to a distance limit for object detection, but those are difficult to compare to our study due to differences in the object, emission, environment, and task (e.g. Cotzin and Dallenbach, 1950; Kolarik et al., 2016).

Rice et al. (1965) found that increasing object size improved object detection for metal discs ranging from approximately 0.03 m–0.40 m in diameter. We did not find an effect of object size for our two MDF boards, although they were both at least 0.5-m wide, the largest being 1.22-m wide, and thus larger than the largest used by Rice et al. Increasing the size of a flat surface may not lead to stronger echoes once a certain size is reached. From the perspective of a stationary listener centred on the object, echoes mostly arise from specular (mirror-like) reflections from around the centre of the surface and from diffraction from the edges of the object. Once the size of the surface is increased beyond a certain amount, the area of the surface that is effective in generating specular reflections that the listener can receive does not increase. As the surface increases in size, the lengths of the edges increase but the edges also get farther from the listener; the overall effect on the reflections from the edges that reach the listener is not obvious. The effect of object size requires further acoustical and psychoacoustical investigation, with greater control of potentially confounding variables, such as orientation, material and mounting (and hence rigidity), than in our current study.

Our experiments with object detection demonstrate considerable inter-individual variation in scores within both sighted and blind populations and that the scores of inexperienced sighted listeners can improve with practice across several hundred trials. The time course of the learning on this task is currently unknown. As discussed in our previous paper on object localisation (Rowan et al., 2015), the substantial inter-individual variation in scores and the short-term learning effects make it difficult to ascertain whether there are meaningful differences between populations. For that reason, and because of our small sample sizes, our results do not warrant any conclusions to be made about differences between blind and sighted people; our results simply confirm that the trends across stimuli we observed in a sample of sighted people were also observed in a sample of blind people. In any case, there are likely to be important sub-populations within the blind population: not all blind people are 'expert' echolocators.

6.2. Use of across-channel cue

Experiment 2 focused on whether listeners could use an across-channel level cue for object detection. Impulse responses for an object distance of 4 m were used. However, cues that would be specific to that distance were not useful to listeners in this experiment. Cues at the onset and offset of the emission were removed and listeners' scores were indistinguishable from chance with the LPN condition suggesting that the could not exploit any 'repetition pitch' cue (having a fundamental frequency of approximately 45 Hz for the 4-m IRs). We therefore argue that our findings on the question of whether the across-channel cue can be used for object detection generalises to other distances.

Our finding that object detection with LPN was poor is similar to our previous finding for object localisation (Rowan et al., 2013), presumably because the echo contains relatively little useful energy below 3 kHz for the object sizes we have used. An important,

qualitative difference between object detection and localisation tasks is how the LPN influences performance with the HPN when the two are combined in the BBN. For localisation, scores were higher for HPN than for BBN; for detection, scores were lower for HPN than for BBN. The information in the LPN interferes with that in the HPN for object localisation performance with the BBN, whereas it supports object detection performance. The interference with object localisation is presumably due to an obligatory use of unhelpful low-frequency (interaural time difference) information during binaural processing (Rowan et al., 2013). One possible explanation for the supportive effect of low-frequency information with object detection was that listeners used a static across-channel level cue (i.e. profile analysis); that explanation can be considered by comparing object detection performance across the other stimuli.

The above-chance detection performance for HPN without rove but not for HPN with rove indicates that a within-channel level cue was available for the former but not the latter. It also indicates that a within-channel level cue was not available for the Frankenstein-with-rove condition, since it was constructed using the same HPN and same magnitude of level rove. The only difference between the HPN-with-rove and Frankenstein-with-rove conditions was that the latter included a LPN that was always convolved with IRs for the object absent. The Frankenstein-with-rove condition therefore combined two stimuli associated with chance performance individually and that produced above chance performance together. This indicates that the unhelpful LPN provides a reference for the exploitation of a static across-channel level cue when combined with the HPN. This finding is important because that level cue may be more robust than the other level cues to unpredictable variations in the emission, echo or background noise, as may occur in real-world scenarios. It is currently unclear if the static across-channel level cue is used in more ecologically relevant tasks and with stimuli that have narrower spectra and shorter durations such as the vocalisations used by some blind people to echolocate. Also, in this study we did not consider temporal or dynamic cues, or object locations off to one side that produce binaural cues, which might be similarly or even more robust to uncertainty in the level and spectrum of the emission or background noise.

Why were the scores for the Frankenstein condition worse than for the BBN? One possible explanation is that the Frankenstein condition included level rove but the BBN did not. The level rove might lead to an increase in stimulus uncertainty, a change in listening strategy, a change in the internal representation of the spectra of the stimuli or a disruption of a within-channel level cue which might otherwise be used in combination with an across-channel level cue (Lentz, 2005). Alternatively, perhaps there was other information for object detection in the low-frequency portion of the BBN that was not available in the Frankenstein condition. This is unlikely to be a repetition pitch cue or a within-channel level cue since most listeners' scores were indistinguishable from chance in the LPN condition (without level rove).

While low-frequency information has generally been associated with relatively poor detection and localisation performance in the stimulus conditions featured in our studies, this does not mean that low-frequency information is not useful in general. For example, our stimuli were limited to frequencies above 0.3 kHz and the data with LPN in Experiment 2 were obtained only for an object distance of 4 m. Had we used stimuli with a lower cut-off frequency and shorter distances with LPN, scores might have been better. Ashmead and Wall (2002) have suggested that information below 0.1 kHz can be useful for detecting walls.

The duration of the emission we used in Experiment 2 was long relative to the duration of transient emissions typically used by blind echolocators. Both the detection of differences in spectral

shape using an across-channel level cue in profile analysis experiments (Green, 1988) and the detection of objects in echolocation-related experiments (Schenkman and Nilsson, 2010) weakens with reducing stimulus duration. However, the size of the across-channel level cue is substantially larger for much shorter object distances (e.g. 1 m) than the nominal distance (4 m) used in Experiment 2, as illustrated in Fig. 6. Furthermore, blind listeners might be better able to exploit an across-channel level cue than sighted listeners (Doucet et al., 2005). The combined effect of these factors is difficult to predict. Nevertheless, the goal of this paper was not to estimate real-world object detection accuracy using an across-channel level cue but to establish that it can be used at all with echolocation-related stimuli.

It is not clear which features of the spectrum above 3 kHz were important for object detection via an across-channel cue in Experiment 2 and it is difficult to compare the results for the stimuli used here (bands of noise with complex changes in spectral profile) to those of studies of profile analysis (usually multi-tone complexes with changes in relative level of one of the tones). A general finding from profile analysis research is that changes in spectral profile near the centre of the spectrum are more easily detected than changes at the edges of the spectrum (Green, 1988); listeners also seem to rely more on peaks than troughs (Lentz, 2006). We might therefore expect from Fig. 6 that the peak at approximately 6 kHz was more important than the peak near 10 kHz. Further research to clarify the frequency regions that are important for echolocation has clinical relevance since hearing loss usually affects such high frequencies earlier in life and more than lower frequencies. Furthermore, supra-threshold changes associated with sensorineural hearing loss in adults impair the use of an across-channel cue (Lentz, 2006). Some blind listeners who have used echolocation for much of their lives have told us that echolocation has become increasing difficult as they have aged into their 60s, reducing their confidence and independence. In one such case, the individual's hearing threshold levels were better than 30 dB HL up to 4 kHz and worsened to 70 dB HL from 8 to 12.5 kHz. He reported using a variety of emissions, all non-vocalisations, including taps of a ring on his finger against his cane. Of course, it is difficult to say from this whether the self-reported difficulties were due to reduced audibility, supra-threshold factors or other factors. Nevertheless, there is reason to investigate whether hearing loss, including loss that would not normally be considered clinically significant or would not even be detected by conventional clinical testing, might have a significant impact on object detection for blind echolocators.

6.3. Environment

Our experiments on both object detection and localisation to date used recordings obtained from an anechoic (at least above 0.1 kHz) environment. This presumably provides insight into real-world echolocation in open spaces or when the emission and echo are well separated in time from reflections from other objects. The results may have been different had the recordings been made in reverberant rooms. As indicated by Kolarik et al. (2014), reverberation might alter the spectral cues. To investigate this, we recorded additional binaural IRs from four conveniently located rooms to cover a range of reverberation times (RTs); see Fig. 11. Room 1, the anechoic chamber, was the same one as used for the recordings described in Section 2. Binaural IRs were measured with the interaural axis of KEMAR, this time including the torso, 1.2 m above the floor; a loudspeaker (Mackie HR824) driver was placed 1.0 m above the floor and just below and in front of KEMAR's chin. The IRs were measured using a pure tone sweep from 0.02 to 20 kHz. The responses were recorded with a sampling rate of

96 kHz and with 32-bit amplitude resolution. The recordings were convolved with an inverse filter to derive the IRs. The object was an aluminium plate that was 0.5 m in diameter and 1.5 mm thick, as described by Schenkman and Nilsson (2010), and it was positioned with its centre 1.2 m above the floor and directly in front of KEMAR's head at various distances.

Fig. 12 plots the difference in excitation level between the board present and absent, as in Fig. 6, for the left ear, all four rooms and four distances from 1 to 4 m. The analysis was the same as for Fig. 6, using the cochlear model, except that a temporal window was applied to the stimuli to remove the first 0.05 s and the last 10 s in order to focus only on the portion where the echo, emission and reverberation are all present and in a steady state. The difference in excitation level below 3 kHz is similar in all rooms except in Room 4, with the largest reverberation time, where the peak near 2 kHz is weaker. The difference in excitation level above 3 kHz is weaker for all the three reverberant rooms than for the anechoic chamber, such that there are no or only weak excitation level cues below 10 kHz in the reverberant rooms when the object was at 2 m and beyond. This confirms that the level cues can be detrimentally affected by room reverberation. In practice, the effect of room reverberation on object detection will be dependent on whether the listener is able to separate the echo from the reverberation. For example, author DR's and RGL's informal listening experience is that it is easier to detect the object in the reverberant rooms using the raw IRs than when convolved with noise of durations from 10 to 400 ms for object distances beyond 1 m. Schenkman and Nilsson (2010) found that object detection was better in a room with a reverberation time of 0.4 s, similar to our Room 2, than in an anechoic room for an emission duration of 500 ms and object distance up to 2 m, beyond which object detection was typically not possible. It has also been reported that performance on other echolocation tasks can improve in the presence of reflections from additional, 'background' objects (Schörnich et al., 2012; Wallmeier and Wiegrebe, 2014a). Our findings suggest that this is unlikely to be due to enhanced excitation level cues.

An alternative explanation is that room reverberation leads to an enhanced interaural coherence cue. Interaural coherence would typically be lower in a reverberant room than in an anechoic chamber (Aaronson and Hartmann, 2010) and an echo from an object could have the effect of increasing the interaural coherence by a more detectable amount in the reverberant room, heard as a change in the diffuseness of the auditory image. We analysed similar stimuli as used with Fig. 12 to determine the peak interaural correlation coefficient at the output of a bank of gammatone filters; here, the band of noise was modulated at 125 Hz with a half-wave rectified sinewave prior to convolution in order to clarify interaural coherence in the envelope (Rowan et al., 2015). An increase in interaural coherence in the waveform fine-structure between the board absent and present below 1.5 kHz was found for the reverberant rooms only, between 0.5 and 1.5 kHz. The magnitude of the increase varied with frequency between 0 and approximately 0.10. Above 1.5 kHz, substantial differences in interaural coherence in the waveform fine-structure were observed, although the human binaural system is insensitive to it (Bernstein and Trahiotis, 1992). An increase in interaural coherence in the waveform envelope was found for the reverberant rooms only, at most frequencies above 1.5 kHz and varied in magnitude with frequency between 0 and approximately 0.15. These increases in interaural coherence in the waveform fine-structure and envelope were from a wide range of baseline values with the object absent from as low as 0.40; typically, the higher the room reverberation time and the higher the frequency, the lower the interaural coherence with the object absent. While humans can detect a reduction in interaural coherence from a baseline of 1.00 of as little as 0.02 in the temporal fine

Fig. 11. The rooms used for recording the impulse responses. Top left: Room 1 (anechoic chamber, as in Fig. 1) with mid-frequency reverberation time (RT$_{0.5-1\ kHz}$) of 0.04 s. Top right: Room 2 with RT$_{0.5-1\ kHz}$ of 0.45 s. Bottom left: Room 3 with RT$_{0.5-1\ kHz}$ of 2.00 s. Bottom right: Room 4 (reverberation chamber) with RT$_{0.5-1\ kHz}$ of 7.63 s. Those reverberation times (RTs) are based on impulse responses measured in the rooms (BS EN ISO 3382–1:2009, 2009) analysed using the Schroeder integral (Schroeder, 1965).
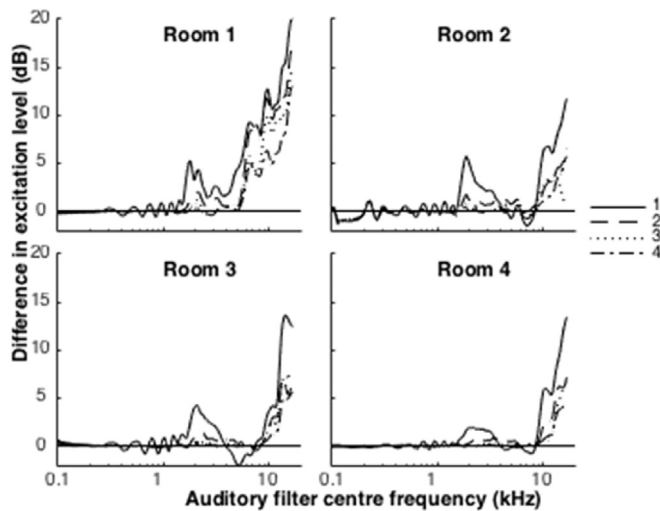
Fig. 12. Differences in auditory excitation level for the left between the board present and absent, using a model of the auditory periphery as Fig. 6, for an aluminium disk (similar to that reported by Schenkman and Nilsson, 2010) placed in the four rooms shown in Fig. 11. Room 1 is the anechoic chamber. The different lines indicate different distances. See Section 6.3 for details.

structure at 0.5 kHz (e.g. Gabriel and Colburn, 1981) and 0.05 in the envelope at 4 kHz (e.g. Bernstein and Trahiotis, 1992) it is unclear if it is a viable cue with our echolocation-related stimuli. Informal listening trials by authors DR and RGL suggest not. It is important for future research on the effect of reverberation on echolocation to connect any changes in listeners' behaviour to specific, quantifiable changes in the stimuli the listeners received.

6.4. Final comments

The technique that we have used to study echolocation creates virtual objects and enables careful control of the stimuli as well as detailed investigation of the auditory cues and processes involved in echolocation. It can be extended to make the tasks more ecologically valid. For example, listeners' own vocalisations and movement can be included in real time, which has been found to

influence echolocation performance (Wallmeier et al., 2013; Wallmeier and Wiegrebe, 2014b; Fiehler et al., 2015). The inclusion of motion can allow detailed investigation of the combination of object detection with head movement to locate objects, which we previously referred to as 'scanning' (Rowan et al., 2013). These and other developments to make laboratory echolocation more ecological relevant are important. Nevertheless, there remains a lack of detailed understanding of the specific acoustical cues and auditory processes involved in more basic echolocation scenarios, such as the object detection task used in the current study, which we hope our paper contributes to improving.

7. Conclusions

  (i) Changes in overall level, within-channel level and spectral shape (an across-channel level cue) are available for object detection when the emission duration is similar to, or longer than, the delay between the echo and emission arriving at the ears, as might be relevant to short-range object detection by blind echolocators using transient emissions such as mouth clicks. The within- and across-channel cues occur predominantly for frequencies above 3 kHz and for some hard flat objects sized 0.5 m^2 and greater.
 (ii) Inexperienced sighted listeners could detect hard flat objects 4 m away, and probably further, using 400-ms-long broadband emissions.
(iii) Changes in spectral shape could be used to detect objects. While the relevant level changes for that cue occur above 3 kHz, audible energy in the emission is also required *below* 3 kHz to act as a reference. In contrast, it was found previously that the addition of energy below 2 kHz impaired object localisation (Rowan et al., 2013). Hence, the optimal emission will probably depend on the task.
(iv) The object detection scores of inexperienced listeners improved over several hundred trials.

References

Aaronson, N.L., Hartmann, W.M., 2010. Interaural coherence for noise bands: waveforms and envelopes. J. Acoust. Soc. Am. 127 (3), 1367–1372.

Ashmead, D., Wall, R., 2002. Low frequency sound as a navigational tool for people with visual impairments. J. Low. Freq. Noise Vib. Act. Control 21, 199–205.

Bernstein, L.R., Trahiotis, C., 1992. Discrimination of interaural envelope correlation and its relation to binaural unmasking at high frequencies. J. Acoust. Soc. Am. 91 (1), 306–316.

BS EN ISO 3382-1:2009, 2009. Acoustics - Measurement of Room Acoustic Parameters. Part 1: Performance Spaces. Br. Stand., pp. 1–26.

Chen, Z., Hu, G., Glasberg, B.R., Moore, B.C., 2011. A new method of calculating auditory excitation patterns and loudness for steady sounds. Hear. Res. 282 (1–2), 204–215.

Cotzin, M., Dallenbach, K.M., 1950. "Facial vision:" the role of pitch and loudness in the perception of obstacles by the blind. Am. J. Psychol. 63, 485–515.

Dai, H., Kidd, G., 2009. Limiting unwanted cues via random rove applied to the yes-no and multiple-alternative forced choice paradigms. J. Acoust. Soc. Am. 126 (2), EL62–EL67.

Doucet, M.E., Guillemot, J.P., Lassonde, M., Gagné, J.P., Leclerc, C., Lepore, F., 2005. Blind subjects process auditory spectral cues more efficiently than sighted individuals. Exp. Brain Res. 160 (2), 194–202.

Epstein, M., Marozeau, J., 2010. Loudness and intensity coding. In: Plack, C.J. (Ed.), The Oxford Handbook of Auditory Science: Hearing. Oxford University Press, Oxford, pp. 45–69.

Fiehler, K., Schütz, I., Meller, T., Thaler, L., 2015. Neural correlates of human echolocation of path direction during walking. Multisens. Res. 28 (1–2), 195–226.

Gabriel, K.J., Colburn, H.S., 1981. Interaural correlation discrimination: i. bandwidth and level dependence. J. Acoust. Soc. Am. 69 (5), 1394–1401.

Green, D.M., 1988. Profile Analysis: Auditory Intensity Discrimination. Oxford University Press.

Kolarik, A.J., Cirstea, S., Pardhan, S., Moore, B.C., 2014. A summary of research investigating echolocation abilities of blind and sighted humans. Hear. Res. 310, 60–68.

Kolarik, A.J., Scarfe, A.C., Moore, B.C., Pardhan, S., 2016. An assessment of auditory-guided locomotion in an obstacle circumvention task. Exp. Brain Res. 234 (6), 1725–1735.

Lentz, J.J., 2005. Profile analysis: the effects of rove on sparse spectra. J. Acoust. Soc. Am. 118 (5), 2794–2797.

Lentz, J.J., 2006. Spectral-peak selection in spectral-shape discrimination by normal-hearing and hearing-impaired listeners. J. Acoust. Soc. Am. 120 (2), 945–956.

MacMillan, N.A., Creelman, C.D., 2005. Detection Theory: a User's Guide. Lawrence Erlbaum Associates, Inc., Mahwah, N.J.

Moore, B.C., Oldfield, S.R., Dooley, G.J., 1989. Detection and discrimination of spectral peaks and notches at 1 and 8 kHz. J. Acoust. Soc. Am. 85 (2), 820–836.

Moore, B.C., Glasberg, B.R., 1983. Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. J. Acoust. Soc. Am. 74 (3), 750–753.

Papadopoulos, Timos, Edwards, David S., Rowan, Daniel, Allen, Robert, 2011. Identification of auditory cues utilized in human echolocation—objective measurement results. Biomed. Signal Process. Control 6 (3), 280–290.

Rice, Charles E., Feinstein, Stephen H., Schusterman, Ronald J., 1965. Echo-detection ability of the blind: size and distance factors. J. Exp. Psychol. 70 (3), 246–251.

Rojas, Juan Antonio Martínez, Hermosilla, Jesús Alpuente, Montero, Rocío Sánchez, Espí, Pablo Luis López, 2010. Physical analysis of several organic signals for human echolocation: hand and finger produced pulses. Acta Acustica United Acustica 96 (6), 1069–1077.

Rojas, Juan Antonio Martínez, Hermosilla, Jesús Alpuente, Montero, Rocío Sánchez, Espí, Pablo Luis López, 2009. Physical analysis of several organic signals for human echolocation: oral vacuum pulses. Acta Acustica United Acustica 95 (2), 325–330.

Rowan, D., Papadopoulos, T., Edwards, D., Allen, R., 2015. Use of binaural and monaural cues to identify the lateral position of a virtual object using echoes. Hear. Res. 323, 32–39.

Rowan, D., Papadopoulos, T., Edwards, D., Holmes, H., Hollingdale, A., Evans, L., Allen, R., 2013. Identification of the lateral position of a virtual object based on echoes by humans. Hear. Res. 300, 56–65.

Schenkman, Bo N., Nilsson, Mats E., 2010. Human echolocation: blind and sighted persons' ability to detect sounds recorded in the presence of a reflecting object. Perception 39 (4), 483–501.

Schenkman, B.N., Nilsson, M.E., 2011. Human echolocation: pitch versus loudness information. Perception 40 (7), 840–852.

Schörnich, S., Nagy, A., Wiegrebe, L., 2012. Discovering your inner bat: echo-acoustic target ranging in humans. J. Assoc. Res. Otolaryngol. JARO 13 (5), 673–682.

Schroeder, M.R., 1965. New method of measuring reverberation time. J. Acoust. Soc. Am. 409–412.

Thaler, L., 2013. Echolocation may have real-life advantages for blind people: an analysis of survey data. Front. Physiol. 4, 1–8.

Wallmeier, L., Geßele, N., Wiegrebe, L., 2013. Echolocation versus echo suppression in humans. Proc. Biol. Sci. 280 (1769), 20131428.

Wallmeier, L., Wiegrebe, L., 2014a. Ranging in human sonar: effects of additional early reflections and exploratory head movements. PLoS One 9, 1–28.

Wallmeier, L., Wiegrebe, L., 2014. Self-motion facilitates echo-acoustic orientation in humans. R. Soc. open Sci. 1 (3), 140185.