

Psychophysics of human echolocation

Sven Schörnich¹, Ludwig Wallmeier¹, Nikodemus Gessele¹, Andreas Nagy¹, Michael Schraner¹, Daniel Kish², Lutz Wiegrebe¹

¹Division of Neurobiology, Dept. Biology II, University of Munich, Martinsried, Germany

²World Access for the Blind, California, USA

Corresponding author

Lutz Wiegrebe

Division of Neurobiology, Dept. Biology II, Grosshaderner Str 2, 82152 Martinsried.

Phone: +49 89 2180 74314, email: lutzw@lmu.de

Abstract

The skills of some blind humans orienting in their environment through the auditory analysis of reflections from self-generated sounds, has received only little scientific attention to date. Here we present data from a series of formal psychophysical experiments with sighted subjects trained to evaluate features of a virtual echo-acoustic space, allowing for rigid and fine-grain control of the stimulus parameters. The data show how subjects shape both their vocalisations and auditory analysis of the echoes to serve specific echo-acoustic tasks. First, we show that humans can echo-acoustically discriminate target distances with a resolution of less than 1 m for reference distances above 3.4 m. For a reference distance of 1.7 m, corresponding to an echo delay of only 10 ms, distance JNDs were typically around 0.5 m. Second, we explore the interplay between the precedence effect and echolocation. We show that the strong perceptual asymmetry between lead and lag is weakened during echolocation. Finally, we show that through the auditory analysis of self-generated sounds, subjects discriminate room-size changes as small as 10 %. Moreover, the recruitment of virtual echo-acoustic space allows executing the psychophysical experiments in an fMRI scanner. The imaging data show that when subjects echo-acoustically classify the size of a virtual echo-acoustic space, activation patterns co-varying with perceived room size are seen in the subjects' auditory midbrain, thalamus, and auditory as well as parietal- and motor cortices. In summary the current data confirm the practical efficacy of human echolocation, and they provide a rigid psychophysical basis for addressing its neural foundations.

Keywords: Binaural hearing, echolocation, precedence, reverberation

1. Introduction

Echolocation, defined as imaging of the environment through the auditory analysis of acoustic reflections elicited by self-generated sounds, allows gaining information about one's surroundings even in complete darkness. Therefore, this ability is found in mammals whose habitat or way of life renders the use of vision difficult or impossible, like toothed whales or bats.

Several studies have shown that blind or blind-folded human subjects can echo-acoustically detect and discriminate objects of different shape or texture (Kellogg 1962, Rice and Feinstein 1965, Rice 1967, Schenkman and Nilsson 2010). Also object localization has been shown to be quite precise in blind human echolocation experts (Teng *et al.* 2011).

This paper summarizes results from three formal psychophysical studies addressing the efficacy of human echolocation in fully controlled virtual echo-acoustic space.

The principal outline of an echo-acoustic experiment is fully described in terms of a linear system. The outgoing sound is reflected by an ensonified object and perceived through the subjects' ears. The spatial

characteristics of the outgoing sounds, the way it is reflected by an object, and the path the reflection takes from the object to the subjects' ear drums can be described by acoustic impulse responses (IRs). To transfer such an experiment into virtual echo-acoustic space (VEAS), these IRs have to be known and applied in real time to the sounds generated by the subject.

VEAS was created by picking up the subjects' vocalisation from a headset microphone and feeding them back to ear phones. The latter block external sound quite effectively and thus, the subject would not perceive their own vocalisations as they would in the free sound field. Therefore, direct sound and echoes were presented separately via two paths: The first path was a direct, level-adjusted path from the microphone to the earphones. The level of the direct path was set such that the subjects' percept of their own voice in the anechoic chamber was most similar to the percept of their voice with the ear phones removed from the ear canals. The second (echo) path incorporated the IRs including the vocal IR (which describes the azimuth-depend spread of sound from the mouth), the acoustic IR from the ensonified object, and the head-related IR, which describes the azimuth dependent path from the object to the two ears.

2. Target ranging in VEAS

In contrast to vision, where distance information is relatively difficult to infer (Palmer 1999), the distance to a sound-reflecting surface can be echo-acoustically determined by estimating the time delay between emission and echo reception (Simmons 1973; Denzinger and Schnitzler 1998; Goerlitz *et al.* 2010). A formal quantification of human echo-acoustic sensitivity to target range is not available to date.

Stimuli: The echo delay corresponding to the required range was applied by preceding the IR for the second path with so many zeros that, together with the digital IO delay of the hardware and convolution software, the delay corresponding to the required reflector range was generated. Next, the IR amplitude was scaled to match the range dependent geometric attenuation of a virtual echo. For each reference range, the geometric attenuation was globally set, i.e., compared to the reference range of 1.7 m, the IR amplitude was decreased by 6 dB and 12 dB for ranges of 3.4 and 6.8 m, respectively.

Procedure: To solve this task, the subjects had to produce tongue clicks and analyse the real-time generated virtual echoes from these clicks. In an adaptive two-alternative, forced-choice paradigm, subjects were trained to find the interval in which the reflective surface was further away from the subject; i.e., in which the delay between the emission and echo was longer than the reference delay. The latter was roved across trials by $\pm 5\%$. Each interval began with a 50 ms, 1 kHz tone pip. Directly after the tone pip, both the direct path and the echo path were activated for 5 s, such that when the subject produced a sound, the direct path would feed directly into its ears, while the echo path provided a real-time generated echo of the sound with the appropriate range, spectral content, and binaural characteristics. The end of the interval was signalled by another tone pip (50 ms, 2 kHz) which was presented directly after the direct and echo path had been switched off. Subjects were given audio feedback consisting of a 250 ms frequency chirp which was upward modulated for positive feedback and downward modulated for negative feedback

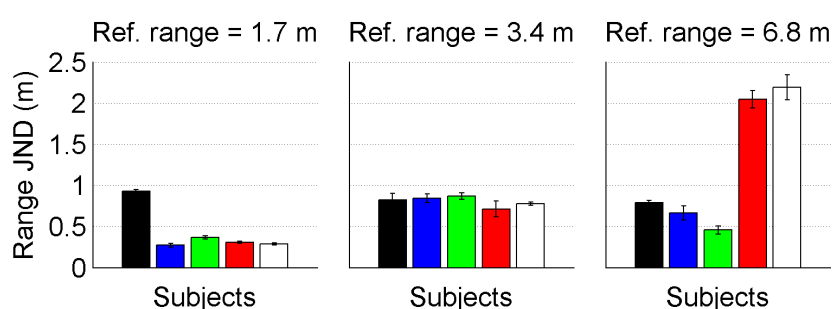


Fig. 1: echo-acoustic sensitivity to target range. Colour bars represent performances of individual subjects.

Results: At a reference range of 1.7 m (reference-echo delay = 10 ms, Fig. 1, top), most subjects could detect a change in target range of only 30-40 cm; only one subject performed significantly worse, with a

range JND of 93 cm. However, while for this subject the range JND stayed approximately constant, even when the reference range was increased to 3.4 or 6.8 m, two other subjects (red and white) showed increasing JNDs with increasing reference range.

To solve this task, subjects produced relatively short vocalisations (tongue clicks) with durations between 3 and 12 ms. The click SPLs, measured at the headset microphone, were more variable ranging from about 60 to 105 dB. Also the number of clicks produced to evaluate the reflection properties in a 5 s interval of the 2AFC task varied across the subjects with individual averages between only 6 and up to 23 clicks. Most subjects produced relatively high-frequency clicks with peak frequencies around 6 to 7 kHz and -15 dB bandwidths from 3 kHz to 10 or 15 kHz.

3. Echolocation vs. echo suppression in humans

The precedence effect predicts a conflict of echolocation and echo suppression: when localizing sound sources, the human auditory system suppresses spatial information of echoes, but just this information underlies effective echolocation. A common approach to investigate the precedence effect is the arrangement of two sound sources that present a direct sound (lead) and a delayed copy (lag), respectively. Several experiments on lag-discrimination suppression have quantified the deterioration of spatial information of the lag produced by the lead. Here, minimum audible angles (MAAs) were measured in VEAS. In the ‘listening’ version, the subjects had to discriminate between positions of a single sound source, the leading, or the lagging of two sources, respectively. In the ‘echolocation’ version, the sound sources were replaced by sound reflectors. Here, the subjects evaluated the echoes generated in real time from self-produced tongue clicks and thereby discriminated between positions of a single reflector, the leading, or the lagging of two reflectors, respectively.

Stimuli: Individual HRIRs and VIRs, measured at 10° azimuthal resolution were interpolated to 0.2° to construct the stimuli. In the ‘listening’ version, the stimuli were acoustic impulses at a repetition rate of 2.5 Hz, convolved with the required HRIRs. The lead-lag delay was 2 ms. In the ‘echolocation’ version, stimuli were generated by the listeners (typically tongue clicks), and, in addition to the direct path described above, they were convolved in real time with first the VIR and then the HRIR and played back with a delay of 10 ms (lead) or 12 ms (lag). Thus the lead-lag delay was fixed at 2 ms, the distance to the leading virtual reflector was 1.7 m, and the distance to the lagging reflector was 2.04 m. Note that in the echolocation version, the perception of the echoes from the leading and lagging reflections is always preceded by the percept of the outgoing sound.

Procedure: In an adaptive 2AFC paradigm with audio feedback, MAAs were measured following a three-down, one-up rule. The beginning and end of each 5 s interval was marked by 50 ms tone pips. Intervals were so long to allow listeners exploring the spatial layout of the reflectors in the ‘echolocation’ version with several tongue clicks. Listeners were extensively trained until they reached stable performance.

Results: MAA measurements for single objects presented in virtual space are shown in the top row of Fig. 2. Objects are either a sound source emitting impulses (Listening) or a surface reflecting the subjects’ tongue clicks (Echolocation).

The data show that sighted subjects can be trained to discriminate reflective surfaces by echolocation with an accuracy comparable to sound-source localization. Data from six subjects reveal a mean MAA of 3.4° in the listening version and 6.7° in the echolocation version, individual MAAs are shown in the first row of Fig. 2.

An appropriate method for comparing results across versions is the calculation of the discrimination deterioration factor (DDF), defined as ratio of lead/single or lag/single thresholds (Litovsky et al. 1997, Tollin and Henning, 1998). Individual DDFs for lead and lag discrimination are shown in the second and third row of Fig. 2, respectively.

In the listening version, the presence of a lagging source impaired lead-discrimination only slightly by a factor of 1.6 (second row, left), while a leading source impaired lag-discrimination considerably by a factor of 8.8 (third row, left). This asymmetry between lead- and lag-discrimination is consistent with previous studies on discrimination suppression (Litovsky et al., 1999).

In the echolocation version, however, this asymmetry was significantly weaker: lead and lag discrimination deteriorated by a factor of 4.8 and 6.2, respectively. To quantify the asymmetry between lead- and lag-discrimination as the defining measure for the influence of the precedence effect, the ratio of lag/lead thresholds was calculated and depicted in the bottom row of Fig. 2. The strength of precedence was significantly higher in the listening experiment.

These data indicate that the precedence effect is weakened in an echolocation context.

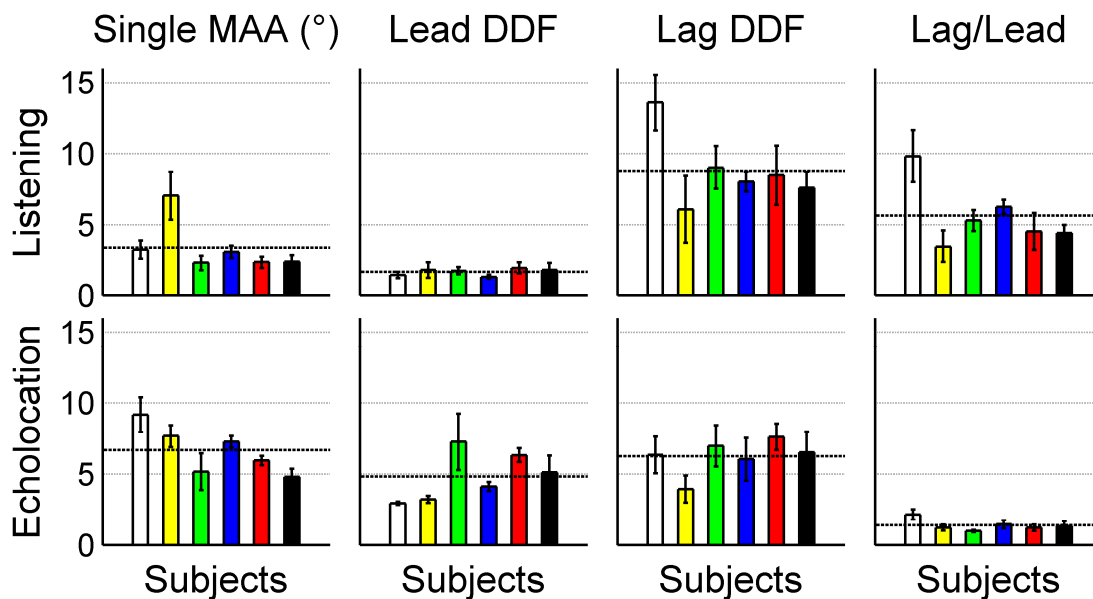


Figure 2: Precedence in a listening and an echolocation task; colour bars represent individual data.

4. Human sonar sensitivity to room size

Echolocation does not only allow to determine the distance and position of objects in space, but also to assess the dimensions of an enclosed space itself. From architectural acoustics it is known that the reverberant qualities of an enclosure can be quantified in terms of a binaural room impulse response (BRIR): The BRIR incorporates the spatial and temporal distribution of reflections which a sound undergoes on its way from a specific source to a human, binaural receiver. In the current experiment, we aimed to formally quantify the ability of sighted human subjects to detect changes in the size of an enclosed space by evaluating echoes of the subjects' own vocalisations.

Stimuli: We recorded the BRIR of a real room, (A chapel with a maximum width of 7.18 m, a maximum length of 17.15 m, and a maximum height of 5.54 m; Reverberation time ~ 1.8 s) and compressed the BRIR along the time axis to simulate decreases in room size (Blauert and Xiang 1993). As reference BRIRs the recorded BRIR was compressed with compression factors of 0.2, 0.5, and 0.7; a compression factor of 0.2 produced the smallest perceived room.

Procedure: Similar to the previous experiments, we recruited a 2AFC paradigm and VEAS to quantify the just-noticeable differences (JNDs) in acoustic room size. Room-size JNDs specify by which percentage each side of the virtual room must be increased such that the corresponding changes in the BRIR can be perceived via echolocation. In each interval, a virtual room was presented for 5 s, delimited by 50 ms tone pips. The subjects produced tongue clicks in each interval, and were to respond which of the two rooms was the smaller one after the sound interval. Again, listeners were given audio feedback.

Results: Figure 3 shows that listeners are quite sensitive for changes in the echo-acoustically perceived room size; their JNDs are on the order of 10 %. A graphic illustration of this echo-acoustic sensitivity is shown in the upper right panel of Fig. 3. For each of the reference room sizes, the psychophysical performance quantified here is sufficient to discriminate the grey-filled room from the transparent room surrounding it.

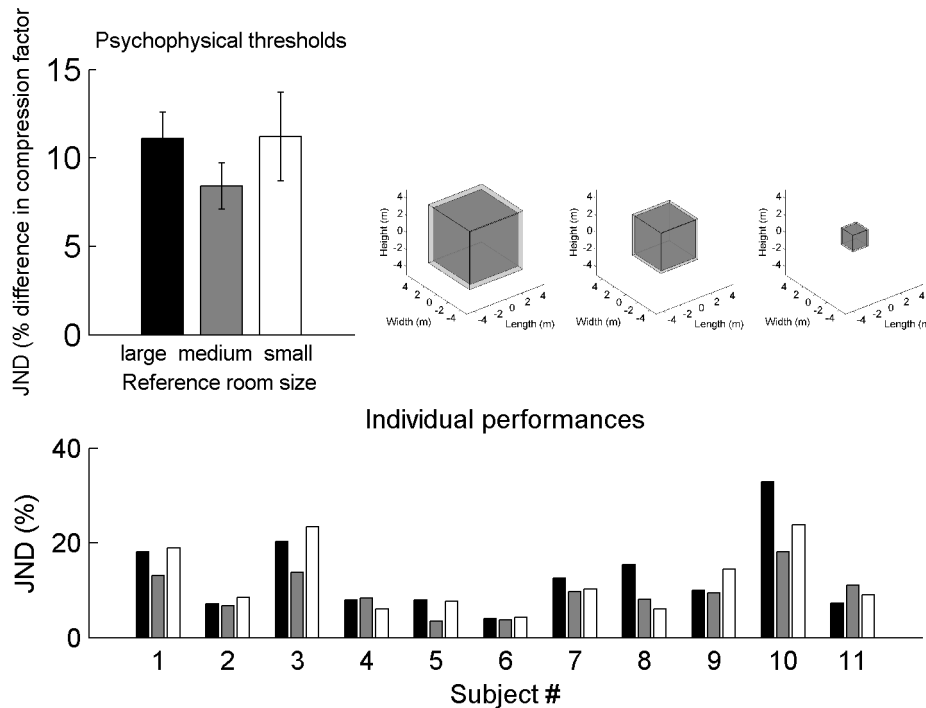


Figure 3: Echo-acoustic JNDs for room size. The upper left shows average JNDs as a function of the reference room size, these data are geometrically illustrated on the right. Individual data are shown in the bottom row.

5. Discussion and Conclusions

The current paper summarizes data from a series of formal psychophysical experiments on the efficacy of human echolocation. Placing the experiments in virtual echo-acoustic space allows for unprecedented experimental control of stimulus parameters and detailed documentation of the sensory-motor interactions underlying echolocation in humans. The data show that sighted subjects can be successfully trained to echo-acoustically detect changes in the range of a reflector positioned in virtual echo-acoustic space. The subjects accomplish this task by vocally emitting short broad-band sounds and evaluating the echoes generated by the reflector or reflectors. The data show that range JNDs were typically below 1 m and, for a reference range of 1.7 m, they were typically below 0.5 m.

The fact that listeners recruited clicks with a high spectral centre of gravity and short durations argues for a temporal auditory analysis as opposed to a pitch-based spectral analysis which would benefit from long-lasting, low-frequency stimulation to produce spectrally resolved harmonics.

The second set of experiments sheds some light on the interplay between echo suppression (the precedence effect) and echolocation. Using a lag-discrimination-suppression paradigm, both in a 'listening' and an 'echolocation' version, the data show that sighted subjects can be trained to discriminate locations of single reflective surfaces almost as good as locations of external sound sources. Second, the data show that in the listening version, the presence of a leading sound source impaired lag-discrimination much more than vice versa. This strong asymmetry between lead- and lag-discrimination was not observed in the echolocation version. These data indicate that the precedence effect, which facilitates the localization of a leading sound source at the expense of the lagging source, is weakened in an echolocation context. While this decreases presumed sonar localization accuracy of the closer reflective

surface, it allows for a more balanced assessment of complex spatial layouts through echolocation.

The final experiments investigated human sonar sensitivity to the size of an enclosed space. Again, using a formal 2AFC procedure with virtual echo-acoustic space stimuli, the data show that sighted human subjects can discriminate the size of ensonified spaces with JNDs around 10 %. Subjects were very little affected by the sound level of the reverberations but evaluated the temporal decay of the reverberations to solve this task.

The current data were all gathered from highly trained but sighted listeners. Research has shown that blind human echolocation experts often perform significantly better even in such formalised tasks (e.g. Teng et al. 2011). These findings stress the need to extend the current techniques and paradigms to those subjects which rely on echo-acoustic information every day.

Acknowledgements

This work was supported by the 'Deutsche Forschungsgemeinschaft' (Wi 1518/9 to Lutz Wiegrebe) the 'Studienstiftung des Deutschen Volkes' (Stipend to Ludwig Wallmeier) and the Danish Research Foundation.

References

- Blauert, J. and Xiang, N. 1993. Binaural scale modelling for auralization and prediction of acoustics in auditoria. *J.Appl.Acoust.* 38, 267-290.
- Denzinger, A. and Schnitzler, H. U. 1998. Echo SPL, training experience, and experimental procedure influence the ranging performance in the big brown bat, *Eptesicus fuscus*. *J Comp Physiol [A]* 183, 213-224.
- Goerlitz, H. R., Geberl, C., and Wiegrebe, L. 2010. Sonar detection of jittering real targets in a free-flying bat. *J Acoust.Soc.Am.* 128, 1467-1475.
- Kellogg, W. N. 1962. Sonar system of the blind. *Science* 137, 399-404.
- Litovsky, R. Y. 1997. Developmental changes in the precedence effect: estimates of minimum audible angle. *J.Acoust.Soc.Am.* 102, 1739-1745.
- Litovsky, R. Y. *et al.* 1999. The precedence effect. *J.Acoust.Soc.Am.* 106, 1633-1654.
- Palmer, S. E. 1999. Cambridge, MA: MIT Press.
- Rice, C. E. 1967. Human echo perception. *Science* 155, 656-664.
- Rice, C. E. and Feinstein, S. H. 1965. Sonar System of the Blind: Size Discrimination. *Science* 148, 1107-1108.
- Schenkman, B. N. and Nilsson, M. E. 2010. Human echolocation: Blind and sighted persons' ability to detect sounds recorded in the presence of a reflecting object. *Perception* 39, 483-501.
- Teng, S., Puri, A., and Whitney, D. 2011. Ultrafine spatial acuity of blind expert human echolocators. *Exp.Brain Res.* 216, 483-488.
- Tollin, D. J. and Henning, G. B. 1998. Some aspects of the lateralization of echoed sound in man. I. The classical interaural-delay based precedence effect. *J Acoust.Soc.Am.* 104, 3030-3038.