

COVER PAGE

Document downloaded by @DAEL

Tue Nov 3 03:15:46 2026
For personal use

When automatic English translation is provided, only the original document is authentic.
The EAA cannot be held responsible of any translation error

Bibliographical reference

Audibility Thresholds of a Sound Reflection in a Classical Human Echolocation Experiment, David Pelegrín-García, Monika Rychtáriková and Glorieux, Acta Acustica **vol. 102** (Number 3), 2016, pp. 530-539

DOI
https://doi.org/10.3813/AAA.918970

Acta Acustica united with Acustica
Vol. 102 (2016) 530 – 539

DOI 10.3813/AAA.918970

Audibility Thresholds of a Sound Reflection in a Classical Human Echolocation Experiment

David Pelegrín-García[1)*], Monika Rychtáriková[1,2)], Christ Glorieux[1)]

[1)] Laboratory of Acoustics Div. Soft Matter and Biophysics, Dept. Physics and Astronomy, KU Leuven, Belgium. [david.pelegringarcia, monika.rychtarikova, christ.glorieux]@fys.kuleuven.be

[2)] STU Bratislava Faculty of Civil Engineering Dept. of Building Structures, Radlinskeho 11, Bratislava, 813 68, Slovak Republic

Summary

Given the renewed interest in human echolocation research, the results of a classical psychophysical experiment by Rice *et al.* [*J.Exp.Psychol.* **70**, pp. 246–251 (1965)] are revisited to derive the audibility thresholds of a single acoustic reflection in a human echolocation task, namely the detection of a reflecting disk in front of a blind person by using self-generated sounds. In the present work, the strength of the reflections caused by the disks in the original experiment is quantified by means of acoustic measurements of the airborne sound transmission from the mouth to the ears using a dummy head. A single-number metric, the Reflected-to-Direct Level Difference (RDLD), which uses the spectrum of a typical human echolocation signal and the inverted 40 phon equal-loudness level contour as weighting functions, is used for quantifying the reflection strength. The audibility thresholds for the sound reflections are re-interpreted in terms of RDLD, achieving typical values of −22 to −19 dB for reflection delays between 4 and 15 ms. These values are notably lower than single-reflection threshold values for externally played musical passages, found in the literature of subjective room acoustics, suggesting a superior performance of blind echolocators in actively identifying echoes with tailored signals.

PACS no. 43.55.Hy, 43.66.Dc

1. Introduction

Human echolocation is a technique used by some visually impaired or blind people to detect the spatial configuration of their environment by interpreting the way sounds, generated by the users themselves or externally, are reflected [1]. This technique is also used proficiently by some species in the animal world, e.g. bats or cetaceans [2]. In the present paper, we use the term human echolocation in the sense of an active perception modality that makes use of self-generated oral sounds. Oral clicks, such as the one shown in Figure 1, represent one of the most typical signals used in this active perception modality.

During the 1950s and 1960s, a number of quantitative studies were carried out within the field of psychology regarding human echolocation with self-generated sounds [3, 4, 5, 6, 7]. In particular, Rice *et al.* [5] investigated the ability of echolocators (i.e. people able to echolocate) to detect circular targets (disks) at different distances in a highly sound absorbing environment using self-generated oral sounds—subjects were free to use either oral clicks or sustained sounds—and reported the results in terms of disk diameter thresholds for which detection was possible on 50% of the cases. In that investigation, rigid disks with diameters ranging from 2.7 cm (non-detectable by any subject) to 38.6 cm (detected almost 100% of the times) were located at distances between 61 cm and 274 cm.

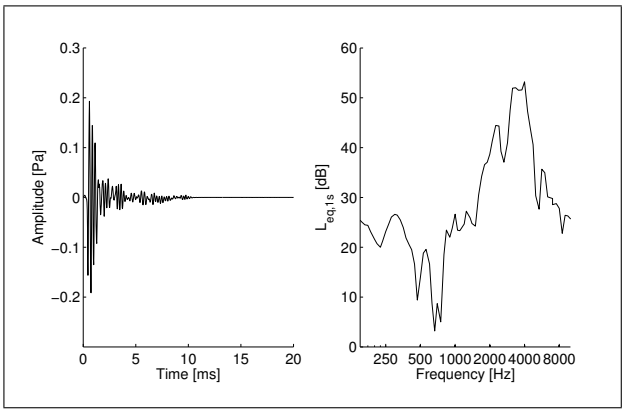


Figure 1. Waveform (left) and 1s-equivalent sound pressure level $L_{eq,1s}$ spectrum (right) of a typical oral click used in human echolocation, measured at the entrance of the left ear canal. The time origin was arbitrarily set to the start of the measurable perturbation. The spectrum of the oral click has a maximum between 3 and 4 kHz, coinciding with the resonance of the ear canal and the tuning of the mouth resonance to this frequency.

Received 29 May 2015,
accepted 01 February 2016.

[*] now with: KU Leuven, Dept. Electrical Engineering, STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Belgium).

© S. Hirzel Verlag · EAA

Document downloaded by @DAEL, 2026-11-03 03:15:48 - personal use only

This degree of detection performance, typical of blind experts, linked to an actual use of echolocation on a daily basis, may provide real-life advantages to blind users [8]. Although blind individuals often show superior echolocation abilities than normal sighted individuals [9, 10], some of the latter can acquire echolocation abilities close to those of blind echolocation experts through a sufficient training [11]. Besides training, age of blindness onset may also affect echolocation ability [12], given that some early blind individuals accommodate areas in the visual cortex for the processing of echolocation [13].

In this paper, we revisit the results of Rice *et al.* [5], re-interpreting the thresholds in acoustical terms. In the fields of animal bioacoustics and underwater acoustics, the metric of Target Strength (TS) is commonly used to characterise the backscattering of objects by comparing the incident intensity with the reflected intensity at a certain distance [14]. This metric, however, does not take into account the filtering of the head and the ears (the Head-Related Transfer Function) that any sound goes through before being sensed by a human listener. In order to include this effect—in addition to source directivity, distance attenuation and reflection effects—in section 2 we define the Reflected-to-Direct Level Difference (RDLD) metric. The RDLD is used in the posterior analysis for quantifying the subjectively perceived strength of the reflections generated by the obstacles in response to echolocation signals and in relation to the strength of the direct echolocation signal traveling through the air from the mouth to the ears. In this way, results from classical human echolocation experiments [5] can be expressed in a way that is familiar in the field of subjective room acoustics. In this field, there have been studies focusing on the audibility of reflections for external sounds like music signals [15], or in a variety of conditions with different signals, reflection delays and angular separation between direct sound and reflection (a summary of these studies is given in [16]). The thresholds obtained in these studies provide a valuable reference for the understanding of passive human echolocation [17], which is based on the detection of intensity or coloration variations in the ambient sound.

Since variations in the properties of a target (sound absorption, size, shape and location) produce acoustical changes reflected in the RDLD, the re-interpretation of Rice's results in terms of RDLD may be helpful nowadays for establishing a reference human performance in the technique, given the renewed interest in psychoacoustic studies of human echolocation, e.g. [10, 18, 19]. For example, echo detection performance for objects made of different materials (and thus having different acoustic absorption) could be compared to echo detection performance for obstacles of different size, since both a decrease in object size or in acoustic reflection coefficient result in a lower reflection strength. Or more interestingly, the echo thresholds could also be determined by making use of electroacoustic systems that add, in a fully controlled way, artificial reflections to the sounds generated by an echolocator, without the presence of an actual obstacle; such systems have been used in recent studies of human echolocation [18, 19].

Section 2 introduces the concepts of Oral-Binaural Room Impulse Response, i.e. the function that characterises the airborne propagation of sound between the mouth and the ears and defines the metric for the quantification of reflection strength—the Reflected-to-Direct Level Difference. Acoustic measurements of RDLD in a number of cases are presented in Section 3. A study of the variability factors contributing to the uncertainty of the RDLD parameter is shown in Section 4. Section 5 re-interprets the thresholds obtained by Rice *et al.* [5] in terms of RDLD. The main conclusions of the work are presented in Section 6.

2. Relevant Concepts

We present the concept of Oral-Binaural Room Impulse Response (OBRIR) as a framework to discuss the airborne propagation of sound between the mouth and the ears, which reflects the input information available in the human echolocation technique with self-generated sounds, excluding body-conducted sound. Next, we define the Reflected-to-Direct Level Difference metric for quantifying the reflection strength, referred to the acoustic propagation paths (direct and reflected) present in the OBRIR.

2.1. Oral-Binaural Room Impulse Response (OBRIR)

The Oral-Binaural Room Impulse Response (OBRIR) [20] contains a pair of functions, one for each ear, that characterise the propagation of airborne sound between a point close to the mouth—a mouth reference point (MRP), 2.5 cm in front of the lips—and the left and right ears of the same person, at one position in a room and for one head orientation. The measurement point at the ears is usually the entrance of the blocked ear canal. The OBRIR contains both the direct sound and the reflections occurring at boundaries, as shown in the schematic example of Figure 2 for the reflection from a single obstacle.

In the depicted scenario, the OBRIR contains only two short-duration events: the direct sound and a single reflection. In simplified representations, the direct and the reflected sound are characterised by Dirac deltas. However, in real cases, the temporal evolution of the wave arrivals is affected by the frequency-dependent characteristics of radiation and propagation, including absorption and scattering effects from surface roughness and edges at obstacles and the head. High frequencies are radiated from the mouth mostly towards the frontal direction and suffer more attenuation than low frequencies when travelling around the head, due to shadowing.

The actual airborne signal received at the ears of the echolocator is the convolution of the echolocation signal generated at the MRP and the OBRIR. It should be noted that the oral sound not only propagates through the air, but also by body conduction, i.e., through body tissues

Document downloaded by @DAEL, 2026-11-03 03:15:48 - personal use only

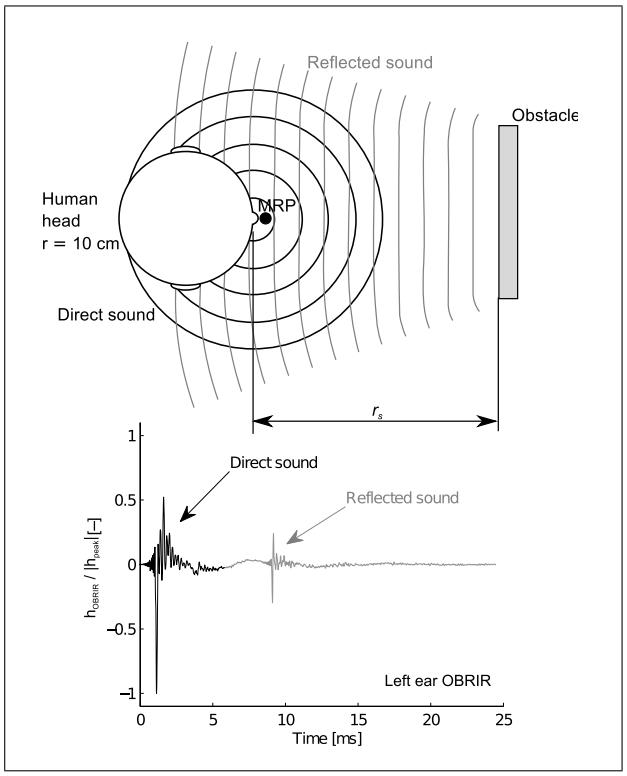


Figure 2. Example of an Oral-Binaural Room Impulse Response (OBRIR, bottom) h_{OBRIR} with amplitude $|h_{\mathrm{peak}}|$, showing the time characteristics of the airborne direct and reflected components of self-generated sounds in a scenario where a person is producing an oral sound in the presence of a single obstacle at a distance r_s (top). For clarity, only the OBRIR for the left ear is shown. Note that because the OBRIR is the representation of a system where both input (pressure signal the mouth reference point MRP) and output (pressure signal at the entrance of the ear canal) are in Pa, the OBRIR is dimensionless.

and bone, ultimately into the cochlea. The latter component of self-generated sound induces a loudness comparable to that of the airborne direct sound when producing stationary sounds [21, 22] (for transient sounds, measurements of this component have nevertheless not been determined). As a consequence, the use of long echolocation signals (e.g. stationary sounds) may result in an overlap and masking of reflections by the direct sound.

2.2. Reflected-to-Direct Level Difference (RDLD)

In order to quantify the strength of a single reflection (or of a group of reflections), it is convenient to introduce the Reflected-to-Direct Level Difference (RDLD) metric, which uses the level of the direct sound as a reference. Formally, the RDLD can be defined as the difference between the (logarithmic) energy levels of the reflected sound ($L_{me,\mathrm{refl}}$) and the airborne direct sound ($L_{me,\mathrm{dir}}$), which can be obtained either by windowing of the OBRIR itself or by providing a reference measurement of the direct sound in free-field [23]. For this, it is assumed that the sound pressure at the MRP is an ideal Dirac delta function and thus the sound at the ears is the OBRIR with units of Pa (its scaling, although arbitrary, is not relevant for the current analysis). Equivalently, the RDLD is derived from the sound pressure amplitudes due to the reflected ($p_{\mathrm{refl}}(f)$) and the direct sound ($p_{\mathrm{dir}}(f)$),

$$\begin{aligned}\mathrm{RDLD}(f) &= L_{me,\mathrm{refl}}(f) - L_{me,\mathrm{dir}}(f) \\ &= 10\log_{10}\left|p_{\mathrm{refl}}(f)/p_{\mathrm{dir}}(f)\right|^2.\end{aligned} \quad (1)$$

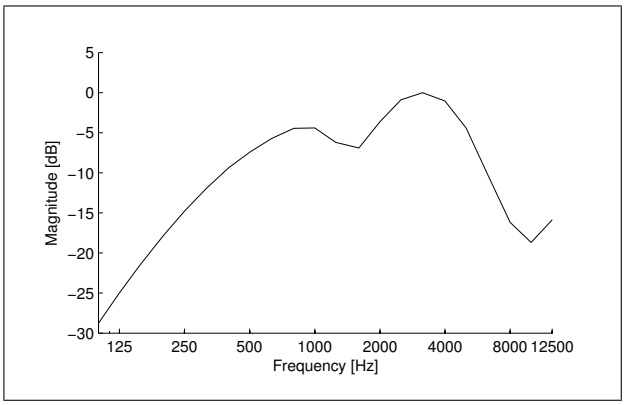


Figure 3. Inverted equal loudness curve of 40 phon according to ISO 226:2003 [24] used as an additional weighting function $T_W(f)$ to obtain the single-value RDLD.

The levels are calculated in frequency bands after spectral analysis of the direct and reflected sound. It should be noted that, since RDLD is aimed at characterising a single reflection from an object, unwanted reflections in echoic environments should be eliminated from the OBRIR by windowing only the reflection of interest.

It is convenient to derive a single-number RDLD indicator representative of the most important frequency bands in RDLD(f). We propose to calculate the single-number RDLD taking into account the frequency weighting of a typical oral click measured at the ears (see Figure 1), with a spectrum $L_{p,\mathrm{click}}(f)$, and the spectral sensitivity of the ear by means of the inverted equal loudness curve at 40 phon according to ISO 226:2003 [24], $T_W(f)$ (shown in Figure 3),

$$\begin{aligned}\mathrm{RDLD} &= L_{me,\mathrm{refl}} - L_{me,\mathrm{dir}} \\ &= 10\log_{10}\left(\frac{\sum_{k=1}^{N} 10^{\frac{L_{p,\mathrm{click},k}+\mathrm{RDLD}_k+T_{W,k}}{10}}}{\sum_{k=1}^{N} 10^{\frac{L_{p,\mathrm{click},k}+T_{W,k}}{10}}}\right).\end{aligned} \quad (2)$$

The subindex k is used to notate each of the N frequency bands. Note that RDLD depends on the spectrum of the click used, which acts as a weighting function emphasising different frequency regions of the OBRIR.

3. Acoustic measurements of RDLD caused by targets at different distances

We performed acoustic measurements with the aim of quantifying the RDLDs of a number of solid disks in a setting similar to that existing in the original experiments of Rice *et al* [5]. For this reason, OBRIRs of 4 disks, made of PVC with a thickness of 5 mm and diameters of 5.6 cm,

Document downloaded by @DAEL, 2026-11-03 03:15:48 - personal use only

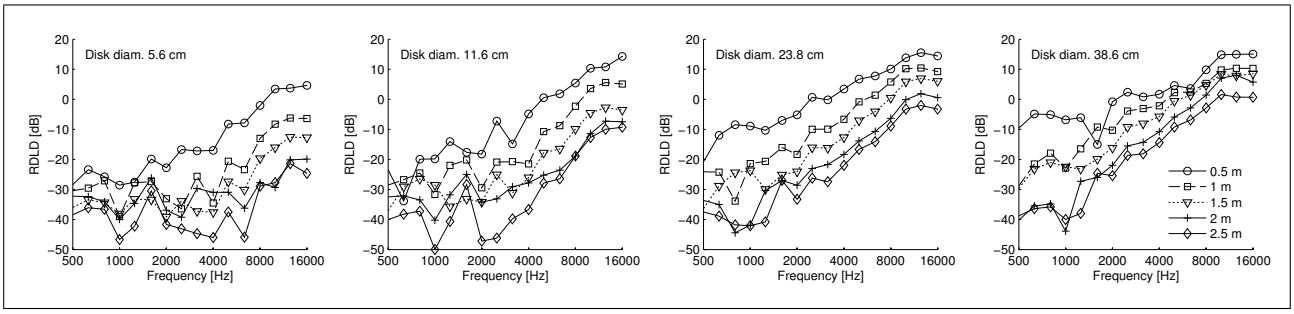


Figure 4. Frequency-dependent RDLD for the four measured disks at five distances (results from the custom dummy head).

11.6 cm, 23.8 cm, and 38.6 cm, at 5 distances (from 0.5 m to 2.5 m away from the mouth position in steps of 0.5 m) were measured in an anechoic room. A custom dummy head was used, with binaural microphones SP-TFB-2 (The Sound Professionals Inc., Hainesport, NJ) placed at the entrance of the blocked ear canals and a loudspeaker JBL Micro II (Harman, Stamford, CT) placed at the mouth position. The capsule of another microphone SP-TFB-2 was placed at the Mouth Reference Point (MRP), 2.5 cm in front of the loudspeaker centre. The axis of the disk was aligned with the mouth centre and the head always pointed towards the disk. Impulse response measurements were made with the ITA-Toolbox for MATLAB (developed at the Institute of Technical Acoustics, RWTH Aachen University, Germany), using an exponential sine sweep as a test signal. A sampling rate of 44.1 kHz was used in the measurements. Following the method of Cabrera *et al.* [20], the transfer function determined at the MRP microphone was used to compensate for the non-flat frequency response of the loudspeaker and the microphones in the measurements at the ears. In order to test the validity of the data obtained with the custom dummy head, the measurements were repeated with a head and torso simulator (HATS) from Brüel & Kjær type 4128 with an integrated loudspeaker at the mouth and microphones at the ears.

In all cases, direct and reflected sound did not overlap in time and they were separated by windowing each component separately with rectangular functions. The energy levels were calculated in 1/3rd octave bands and RDLD(f) was derived by means of Equation (1). The RDLD(f) values obtained for the four disks at five distances, as a function of frequency in the range 500 Hz–16 kHz, are shown in Figure 4. As expected, the general trend is that, the larger and the closer the reflector, the higher the reflection strength, and thus the RDLD. As happens with reflections at finite-size objects, low frequencies (where the wavelength is much larger than the object size) are reflected less efficiently than high frequencies.

By means of Equation (2), the overall RDLDs (click- and inverse 40 phon-weighted) were determined and are shown in Figure 5 as a function of target distance and reflection delay. The RDLD values for the smallest disk at 2 and 2.5 m, whose behavior differs from other data points, are likely affected by low available signal-to-noise ratios and other overlapping residual reflections from the anechoic room.

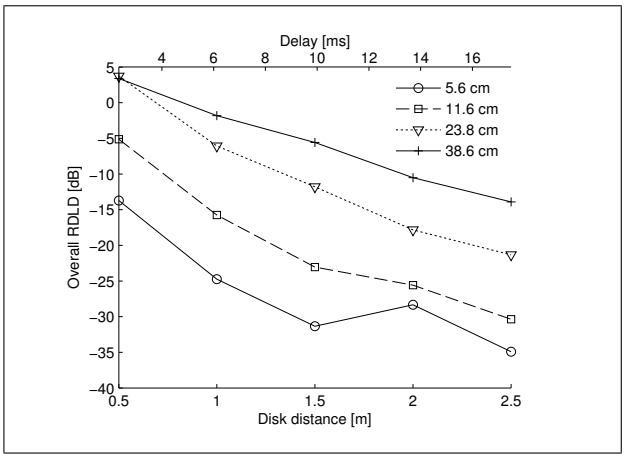


Figure 5. Overall RDLD versus distance and time delay, for disks of four different sizes (results from the custom dummy head)

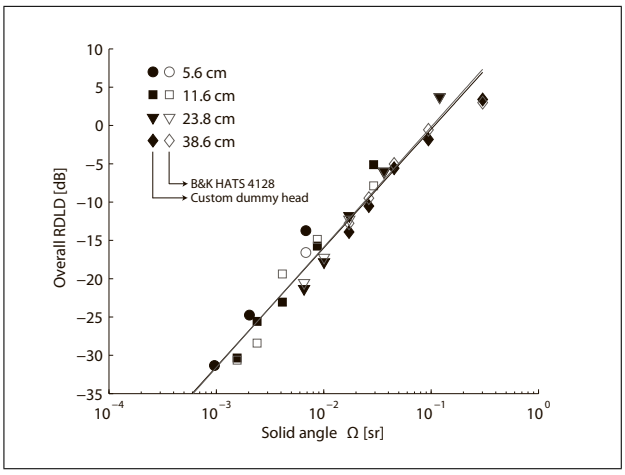


Figure 6. Overall RDLD versus solid angle Ω for a reflection from a solid disk in front of the mouth under free-field conditions. Dark filled symbols correspond to the custom dummy head data and light empty symbols correspond to measurements with the HATS

When expressing the RDLD values as a function of the solid angle Ω covered by the targets (see Figure 6) in an otherwise free-field, a linear model closely fits the data ($R^2 = 0.96$) in the studied range,

$$\mathrm{RDLD} = 15.5 \log_{10} \Omega + 15.0 \ (\mathrm{dB}). \qquad (3)$$

Document downloaded by @DAEL, 2026-11-03 03:15:48 - personal use only

The solid angle Ω is calculated as the area of a spherical cap defined by the intersection of a sphere of radius $(r_d^2 + r_s^2)^{1/2}$ centred at the mouth and a cone with apex at the mouth that passes through the circumference of the disk, and then this area is normalised to the area of a sphere with unit radius,

$$\Omega = 2\pi \left(1 - \frac{r_s}{\sqrt{r_d^2 + r_s^2}} \right) \approx \pi \left(\frac{r_d}{r_s} \right)^2, \qquad (4)$$

with r_d the disk radius and r_s the mouth-disk distance. The approximated term in Equation (4) corresponds to 4π times the ratio between the area of the disk and the area of a sphere with radius r_s. The value of the slope close to 15 determined by the particular choice of conditions leads to a dependence of RDLD with disk radius of approximately $30 \log_{10}(r_d)$ and a dependence with distance of $-30 \log_{10}(r_s)$ in far field ($kr_s \gg 1$).

The RDLD values obtained with the head and torso simulator are shown in Figure 6 together with the data obtained with the custom dummy head. Although individual data points show some random variation when using different equipment, the regression lines for both cases are nearly identical. For the rest of the analysis, we consider valid the RDLD measurements with the custom dummy head.

4. Variability of RDLD due to individual differences and the choice of spectral weighting function

In the previous section, the RDLD has been measured with two different equipment and the same trends have been observed in the data, although with slight differences in most conditions. In general, how large are the expected differences in RDLD among different equipment or among different individuals? And what are the factors potentially affecting the results? The following sections aim at answering these questions. Section 4.1 introduces a number of anatomical attributes that contribute to inter-subject variability in RDLD. Section 4.2 studies the variability in RDLD due to the spectral characteristics of the click signal used as a weighting signal. Finally, section 4.3 combines the different sources of variability to describe the overall spread in RDLD measurements.

4.1. Individual sources of variability

In Equation (1), the RDLD is defined by the frequency-dependent direct and reflected sound propagation paths between the mouth and the ears. These paths are affected by anatomical differences among subjects and among measurement equipment: on the one hand, the direct sound propagation is affected by the distance mouth-ears and the geometry and impedance of cheeks and ears (represented by the Oral-Binaural Transfer Function, OBTF, in 4.1.1). On the other hand, the reflected sound depends on the directivity of the source (i.e. the propagation from near-field to far-field, represented by the Vocal Transfer Function, VTF, in 4.1.2) and on the filtering of the head to incident sounds (i.e. the so-called Head Related Transfer Function, HRTF, in 4.1.3). Given that the direct sound component is also affected by body-conducted sound, impossible to measure with the used equipment, the following calculations are an indication of variability in the airborne component only. As a first approximation, we assume that individual variations in free-field sound propagation between the mouth and the ears (OBTF in 4.1.1) and near-field to far-field propagation (VTF in 4.1.2) are similar to those proposed in the recommendation ITU-T P.58 [25], and that variations in filtering of incident sound due to head and torso (HRTF in 4.1.3) are similar to those in Hammershøi and Møller [26]. Moreover, we assume that these three sources of variability (OBTF, VTF and HRTF) are independent and that the differences between the left and right ears are negligible (we consider that the geometry of the head and the directivity of the emission are symmetrical with respect to the sagittal plane).

The analysis that follows applies one-third octave band filtering as an approximation to the spectral resolution of the human auditory system.

4.1.1. Oral-Binaural Transfer Function (OBTF)

We define the modulus of the oral-binaural transfer function (OBTF, in dB scale) as the difference between the sound pressure level at an Ear Reference Point (ERP), located at the entrance of the ear canal, and a Mouth Reference Point (MRP) 2.5 cm in front of the lip center, in free-field conditions. The average OBTF $\overline{\Delta L_{\mathrm{OBTF}}}$ and its margin of variation are obtained from Tables 13 and 14 of reference [25]. We generated 1,000 random individual OBTF curves $\Delta L_{\mathrm{OBTF},i}$ uniformly distributed within the variability margins of the recommendation. These 1,000 curves and the average OBTF curve are plotted in Figure 7. The OBTF illustrates the fact that high frequencies are more attenuated than low frequencies, as a consequence of head shadowing and of the directivity of the mouth.

4.1.2. Vocal Transfer Function (VTF)

The Vocal Transfer Function (VTF), the frequency-domain version of the vocal impulse response used by Schörnich *et al.* [27], depends on the efficiency of radiation of sound from the near field to the far field and is defined as the ratio between the sound pressure at a point in far field (500 mm in front of the lips for the present calculations) and the sound pressure at the MRP, in free-field conditions. The average VTF $\overline{\Delta L_{\mathrm{VTF}}}$ and the recommended variability are given in Table 7d of reference [25]. The values of 1,000 individual VTFs $\Delta L_{\mathrm{VTF},i}$ generated within these variability margins, and their average value, are shown in Figure 8.

4.1.3. Head-Related Transfer Function (HRTF)

The sound reflected from an obstacle is filtered by the head and torso. In the present case, the average frontal-incidence Head-Related Transfer Function (HRTF) $\overline{\Delta L_{\mathrm{HRTF}}}$ is

Document downloaded by @DAEL, 2026-11-03 03:15:48 - personal use only

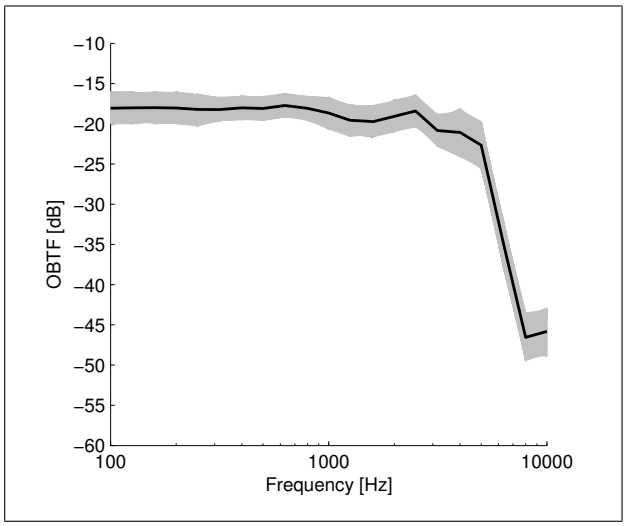


Figure 7. Average Oral-Binaural Transfer Function (OBTF, in black) and 1,000 individual OBTFs (in gray) generated within the margins recommended in ITU-T P.58 [25].

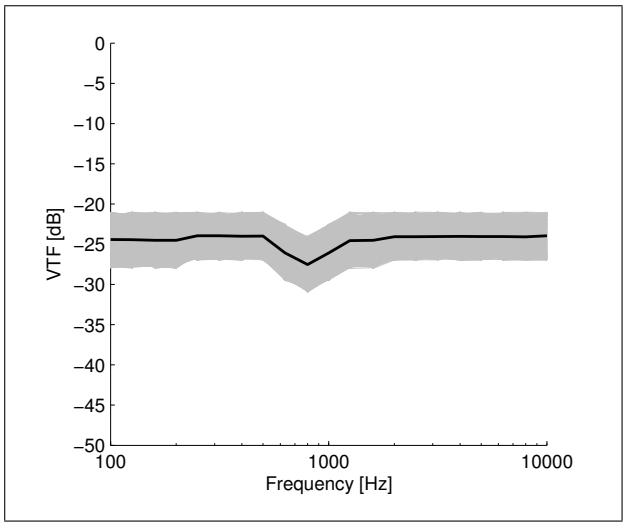


Figure 8. Average Vocal Transfer Function (VTF, in black) and 1,000 individual VTFs (in gray) for frontal direction, generated within the margins recommended in ITU-T P.58 [25]

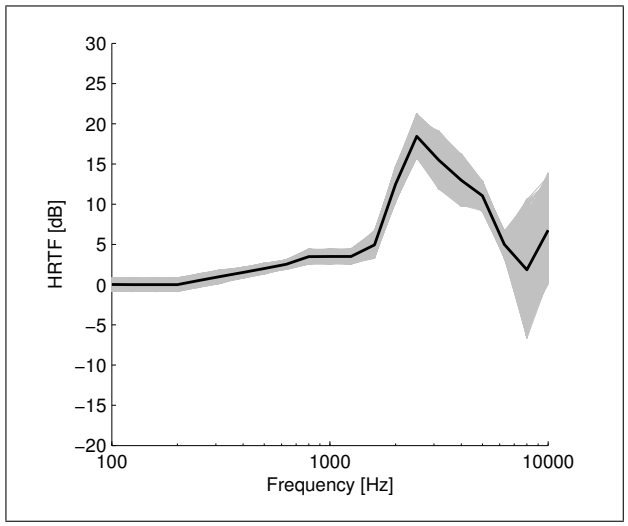


Figure 9. Frontal-incidence average HRTF (in black) and 1,000 individual HRTFs (in gray) randomly generated

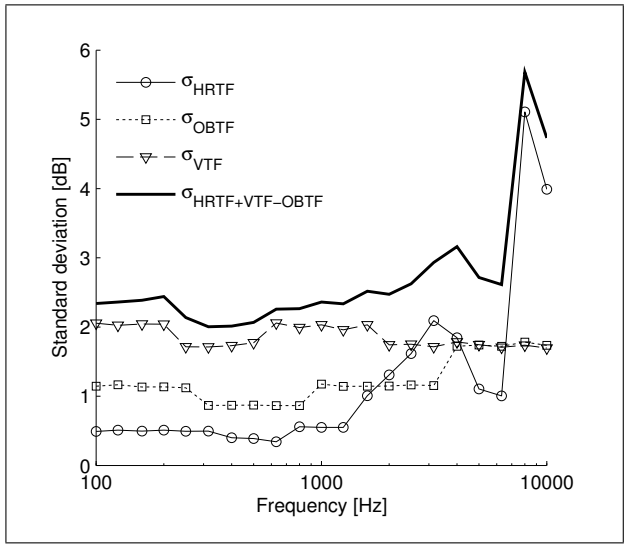


Figure 10. Standard deviation of the HRTF (σ_{HRTF}), the VTF (σ_{VTF}) and the OBTF (σ_{OBTF}) and their weighted sum ($\sigma_{\mathrm{HRTF+VTF-OBTF}}$), as a function of frequency, across 1,000 individual realizations.

given in Table 2 of reference [25] and the margins of variation between human subjects are taken from the measurements of Hammershøi and Møller [26] at the entrance of the blocked ear canals. With these specifications, 1,000 individual HRTFs $\Delta L_{\mathrm{HRTF},i}$ were randomly generated within the allowed variation margins following a uniform distribution (uncorrelated among frequency bands) and are shown in Figure 9 together with the average HRTF.

The inter-subject standard deviations of the OBTF (σ_{OBTF}), the VTF (σ_{VTF}) and the HRTF (σ_{HRTF}) are shown in Figure 10. These show different trends: the HRTF has very low standard deviation at low frequencies and increases to several dB above 6 kHz. The VTF has an approximately frequency-independent standard deviation of 2 dB and the OBTF has a standard deviation between 1 dB at low frequencies and 2 dB at high frequencies.

4.2. Spectral weighting

In order to determine the variability introduced by the choice of the weighting spectrum, a number of oral clicks were recorded on binaural microphones Sound Professionals SP-TFB-2P, with the capsule positioned at the entrance of the ear canal. The clicks were generated by one subject, intentionally shaping the mouth cavity in various ways (modifying the effective volume) and with different points of contact between the tongue and the palate at the moment of generating a vacuum release. The 1s-equivalent sound pressure level spectra $L_{eq,1s}$ of 26 selected clicks, averaged between the recordings at the two ears, are shown in Figure 11. In all cases, energy contained under 800 Hz was much lower than above 1 kHz. The main peak of sound pressure level occurred at frequencies between 1 kHz and

Document downloaded by @DAEL, 2026-11-03 03:15:48 - personal use only

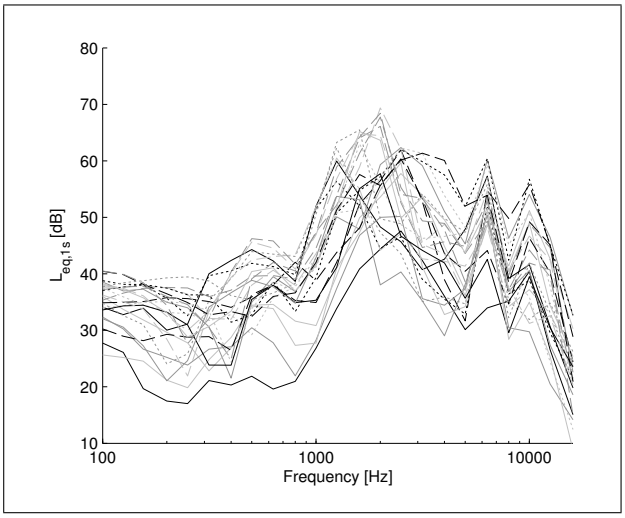


Figure 11. One-third octave band sound pressure level spectra of 26 oral click recordings.

3 kHz with important secondary peaks at around 6 kHz and 10 kHz.

4.3. Variability in RDLD

For calculating the variability of the single-value RDLD, due to the combination of different Head-Related Transfer Functions (HRTF), Vocal Transfer Functions (VTF), Oral-Binaural Transfer Functions (OBTF) and different click spectra, we assumed that the dummy head used for RDLD measurements had average HRTF ($\overline{\Delta L_{\mathrm{HRTF}}}$), VTF ($\overline{\Delta L_{\mathrm{VTF}}}$) and OBTF ($\overline{\Delta L_{\mathrm{OBTF}}}$) characteristics. In a first step, we constructed a set of 1,000 possible individual realisations by applying the randomly generated HRTF, VTF and OBTF of section 4.1 (taken as differences from the mean values) on the measured data. Thus, for the i-th individual realisation, the direct sound level $L_{me,\mathrm{dir},i}$ derived from the measured level $L_{me,\mathrm{dir,meas}}$ was:

$$L_{me,\mathrm{dir},i} = L_{me,\mathrm{dir,meas}} - \overline{\Delta L_{\mathrm{OBTF}}} + \Delta L_{\mathrm{OBTF},i}. \quad (5)$$

The reflected sound level $L_{me,\mathrm{refl},i}$ for the i-th realisation, corrected from the measured reflected sound level $L_{me,\mathrm{refl,meas}}$, was

$$\begin{aligned} L_{me,\mathrm{refl},i} = {} & L_{me,\mathrm{refl,meas}} - \overline{\Delta L_{\mathrm{VTF}}} \\ & + \Delta L_{\mathrm{VTF},i} - \overline{\Delta L_{\mathrm{HRTF}}} + \Delta L_{\mathrm{HRTF},i}. \end{aligned} \quad (6)$$

Thus, the frequency-dependent RDLD for the i-th realisation, corrected from the measured values of RDLD(f), resulted in

$$\begin{aligned} \mathrm{RDLD}_i(f) = {} & \mathrm{RDLD}(f) - \overline{\Delta L_{\mathrm{VTF}}} + \Delta L_{\mathrm{VTF},i} \\ & - \overline{\Delta L_{\mathrm{HRTF}}} + \Delta L_{\mathrm{HRTF},i} \\ & + \overline{\Delta L_{\mathrm{OBTF}}} - \Delta L_{\mathrm{OBTF},i}. \end{aligned} \quad (7)$$

The frequency-dependent standard deviation of this measure, due to the joint variations in HRTF, VTF and OBTF ($\sigma_{\mathrm{HRTF+VTF-OBTF}}$) is shown in Figure 10.

In a second step, the spectrum of each click was used as a weighting function (in addition to the inverted 40-phon equal loudness level contour $T_W(f)$) for each of the 1,000 computed frequency-dependent RDLDs (which contained the effect of potential individual variations). Hence, the single-value RDLD for the i-th individual realisation and the j-th click weighting spectrum was

$$\begin{aligned} \mathrm{RDLD}_{i,j} &= L_{me,\mathrm{refl},i,j} - L_{me,\mathrm{dir},i,j} \\ &= 10\log_{10}\left(\frac{\sum_{k=1}^{N} 10^{\frac{L_{p,i,k}+\mathrm{RDLD}_{i,k}+T_{W,k}}{10}}}{\sum_{k=1}^{N} 10^{\frac{L_{p,i,k}+T_{W,k}}{10}}}\right), \end{aligned} \quad (8)$$

where $\mathrm{RDLD}_{i,k}$ is the value of RDLD_i of Equation (7) in the k-th frequency band. Note that in Equation (8), the spectrum of the direct sound was forced to be the recorded one, whereas usually the recorded sound would implicitly contain the effect of OBTF on the transmission of sound from the mouth to the ears. In the current case, inter-subject variations of OBTF were passed to the reflected sound in the numerator of Equation (8) as part of RDLD.

For each combination of disk and distance, 26,000 RDLD values were obtained (as a combination of 1,000 individual realisations and 26 weighting spectra). The inter-subject variation σ_S was estimated by averaging $\mathrm{RDLD}_{i,j}$ values for all weighting spectra and calculating the standard deviation of the 1,000 resulting values. Results are shown at the left of Figure 12. The inter-subject variation is only residually affected by the distance to the obstacle or by its diameter. By averaging $\mathrm{RDLD}_{i,j}$ across subjects, 26 values of RDLD were obtained for each combination of disk and distance. The standard deviation of these values (called inter-click variation σ_C, due to differences in weighting spectra) is shown at the centre of Figure 12. Finally, the overall standard deviation σ_O derived from all values of RDLD_i, j is shown at the right of Figure 12. The average standard deviation of the RDLD across disks and distances is $\sigma_S = 1.5$ dB for inter-subject variation, $\sigma_C = 2.5$ dB for inter-click variation, and $\sigma_S = 2.8$ dB for combined inter-subject and inter-click variation. These quantities are an estimation of the expected spread or uncertainty linked to RDLD measurements of a given scenario when using different equipment or different weighting spectra.

5. Echo thresholds for self-generated sounds

As pointed out in the introduction, Rice *et al.* [5] reported the diameter detection thresholds for disks at different distances (i.e. the diameter of the disk which could be detected 50% of the times), when subjects performed an echolocation task using their preferred signal (either a click or a hissing sound). The thresholds increased monotonically with distance in most cases. Some subjects were able to detect targets with a diameter of 16 cm at a distance of 274 cm. By using Eqs. (3) and (4), the diameter thresholds, expressed in terms of the reflected-to-direct level difference (RDLD), are shown in Figure 13. These thresholds

Document downloaded by @DAEL, 2026-11-03 03:15:48 - personal use only

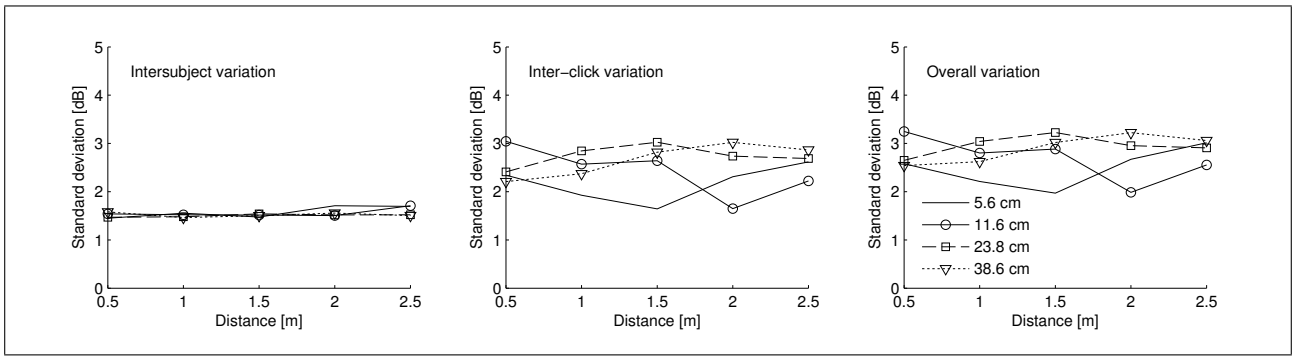


Figure 12. Inter-subject variation of RDLD values averaged across 26 weighting spectra (left), inter-click variation of RDLD values averaged across 1,000 simulated individual characteristics (centre) and overall variation of RDLD due to the combination of subject and weighting spectrum variability (right), as a function to the distance to the disks and disk diameter (in different symbols).

appear to be quite independent of distance, varying between -22 and -19 dB, though with inter-subject variations larger than intra-subject variations. As Rice observed, and according to Equation (3), invariance of RDLD with increasing mouth-disk distance is equivalent to keeping the solid angle covered by the disk fixed (shown as a horizontal line at RDLD = -20.5 dB in Figure 13, which corresponds to an angle of 4.63° from the observer, or a solid angle of $5.13 \cdot 10^{-3}$ sr). The inter-subject standard deviation of the RDLD thresholds in Figure 13, averaged across distances is 2.5 dB. This variability is higher than the inter-subject variability of the RDLD measurement as found in section 4.3 ($\sigma_S = 1.5$ dB). Additional sources of variability in the classical experiment can be attributed to the variability of the signal ($\sigma_C = 2.5$ dB) used but also to the experience of the subjects and to the hearing threshold of subjects [28]. Subjects who had a moderate hearing loss at high frequencies performed worse than their counterparts with normal hearing thresholds [28].

Given the substantial variability in RDLD when combining different spectra with individual Head-Related Transfer Functions (HRTF), Vocal Transfer Functions (VTF) and Oral-Binaural Transfer Functions (OBTF), subjects may have benefited from tuning their clicks to their own HRTF, VTF and OBTF characteristics so as to enhance the reflected energy reaching the ears and thus the likelihood of detecting objects through echolocation.

The thresholds in Figure 13 may be compared with other studies of audibility of a single reflection added to the direct sound. For example, audibility thresholds for a single reflection, averaged among different musical passages as sound signals [15], were between -15.5 dB and -13 dB for delays between 4 and 15 ms for normal sighted subjects. These thresholds could roughly be related to the passive human echolocation ability of normal sighted untrained individuals. The echolocation thresholds of Figure 13 also contain the effect of body-conducted sound, which effectively increases the strength of the direct sound. In this case, the effective RDLD thresholds would be about 3 dB lower (thus between -25 and -22 dB), assuming that the direct airborne sound and the body-conducted sound are of the same order of magnitude [22].

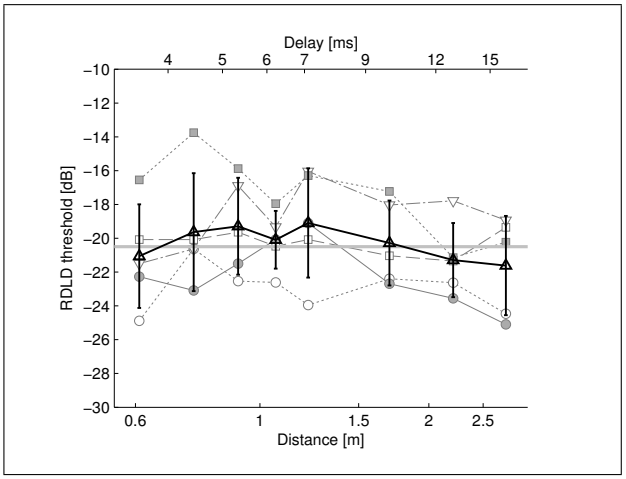


Figure 13. Audibility thresholds of a single sound reflection in the classical human echolocation experiment of Rice *et al.*[5], expressed in terms of the reflected-to-direct level difference, RDLD, as a function of the distance to the obstacle or the equivalent time delay of the reflection (average in black, error bars show ± 1 standard deviation, individual results in gray symbols). The average RDLD, corresponding to a constant solid angle covered by the target, is shown as a horizontal gray bold line.

The improvement of almost 10 dB in reflection thresholds for the case of active echolocation with respect to external music signals may be due to training by the subjects, who were functional echolocators. Moreover, the shorter time duration of echolocation clicks in comparison with the long time duration of music signals, contributes to limit the overlap between the direct and the reflected sound and may be the cause for lowering the detection thresholds in the echolocation task. Further research will address training effects and RDLD thresholds in more general conditions of targets (shape, size, material, location), reverberation and noise, looking at performance in groups of normal sighted people, early blind and late blind individuals. In this context, measuring the RDLD thresholds attainable by a group of early blind individuals who are echolocation experts could set a reference performance for evaluating the progress in training made by other groups. Thus, in successive training sessions, their RDLD thresholds should become closer to the results of the reference group.

Document downloaded by @DAEL, 2026-11-03 03:15:48 - personal use only

It is worth mentioning that, although the RDLD measure as defined in section 2.2 is independent of reverberation (provided that one can separate the reflection of interest by windowing of the oral-binaural room impulse response) or background noise, the RDLD detection thresholds may be affected by both reverberation and noise, and the intensity of the emission during the threshold determination (which has not been taken into account in the experiment of Rice *et al.*[5]) may be adjusted so that the reflected sound becomes audible above the background noise.

6. Conclusions

The audibility thresholds of a single acoustic reflection in response to self-generated sounds used in human echolocation, calculated by Rice *et al.* [5] in terms of disk size, have been reanalysed in terms of the reflection strength of those disks, by using the overall Reflected-to-Direct Level Difference (RDLD) as the main metric, a quantity obtained by applying the weighting of a typical echolocation signal spectrum and the inverted 40-phon equal loudness level contour to the measured reflection strengths. It was found that the RDLD was closely related to the solid angle covered by the solid disks. Variability in this metric was investigated with randomly generated data (with parameters adjusted according to actual measurements and international standards) that characterised individual differences in free-field sound propagation from the mouth to the ears, directivity of the mouth and filtering for frontally incident sound.

Inter-subject standard deviation in RDLDs of a target at one distance, due to individual anatomical differences, were found to be in the order of 1.5 dB. The use of different echolocation signals as weighting spectra for a single-number RDLD introduced an additional standard deviation of 2.5 dB.

The audibility thresholds of a single sound reflection derived from experiments by Rice *et al.* [5], in terms of RDLD, were between -22 and -19 dB for reflection delays between 4 and 15 ms. These values are slightly lower than the thresholds of a single reflection on external sounds using different music passages as stimuli and indicate a benefit of active echolocation with transient signals in the detection of more subtle acoustic reflections.

Acknowledgements

DPG is grateful to FWO-V for supporting this research by financing his postdoctoral grant no. 1280413N. Jonas Brunskog, Jørgen Rasmussen and Mehran Sadri from the Acoustic Technology group at the Technical University of Denmark are acknowledged for kindly making their facilities available for our measurements.

References

[1] A. J. Kolarik, S. Cirstea, S. Pardhan, B. C. J. Moore: A summary of research investigating echolocation abilities of blind and sighted humans. Hearing Research **310C** (2014) 60–68.

[2] N. Fletcher: Animal Bioacoustics. – In: Springer Handbook of Acoustics. T. D. Rossing (ed.). Springer, New York, NY, 2007, Kap. 19, 785–804.

[3] M. Cotzin, K. Dallenbach: Facial vision: The role of pitch and loudness in the perception of obstacles by the blind. The American Journal of Psychology **63** (1950) 485–515.

[4] W. N. Kellogg: Sonar system of the blind. Science **137** (1962) 399–404.

[5] C. E. Rice, S. H. Feinstein, R. J. Schusterman: Echo-Detection Ability of the Blind: Size and Distance Factors. Journal of Experimental Psychology **70** (1965) 246–251.

[6] P. Worchel, J. Mauney, J. Andres: The perception of obstacles by the blind. Journal of Experimental Psychology **40** (1950) 746–751.

[7] I. Kohler: Orientation by aural cues. American Foundation for the Blind. Research Bulletin **4** (1964) 14–53.

[8] L. Thaler: Echolocation may have real-life advantages for blind people: an analysis of survey data. Frontiers in Physiology **4** (2013) 98.

[9] B. N. Schenkman, M. E. Nilsson: Human echolocation: Blind and sighted persons' ability to detect sounds recorded in the presence of a reflecting object. Perception **39** (2010) 483–501.

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

[11] S. Teng, D. Whitney: The Acuity of Echolocation: Spatial Resolution in Sighted Persons Compared to the Performance of an Expert Who Is Blind. Journal of Visual Impairment and Blindness **105** (2011) 20–32.

[12] S. Teng, A. Puri, D. Whitney: Ultrafine spatial acuity of blind expert human echolocators. Experimental Brain Research **216** (2012) 483–488.

[13] L. Thaler, S. R. Arnott, M. A. Goodale: Neural correlates of natural human echolocation in early and late blind echolocation experts. PloS One **6** (2011) e20162.

[14] H. Medwin, C. Clay: Sound scattered by a body. – In: Fundamentals of Acoustical Oceanography. Academic Press, San Diego, CA, 1997, 235–286.

[15] P. Schubert: Die Wahrnehmbarkeit von Rückwürfen bei Musik (On the perception of reflections with music). Z. Hochfrequenztechn. u. Elektroakust. **78** (1969) 230–245.

[16] H. Kuttruff: Subjective Room Acoustics. – In: Room Acoustics. Taylor & Francis and Spon Press, London and New York, 2009, 204–250.

[17] D. Ashmead, R.S. Wall, S.B. Eaton, K. A. Ebinger, M. M. Snook-Hill, D.A. Guth, X. Yang: Echolocation reconsidered: Using spatial variations in the ambient sound field to guide locomotion. Journal of Visual Impairment and Blindness **92** (1998) 615–632.

[18] L. Wallmeier, N. Gessele, L. Wiegrebe: Echolocation versus echo suppression in humans. Proceedings of the Royal Society of Biology **280** (2013) 20131428.

[19] D. Rowan, T. Papadopoulos, D. Edwards, R. Allen: Use of binaural and monaural cues to identify the lateral position of a virtual object using echoes. Hearing Research **323** (2015) 32–39.

[20] D. Cabrera, H. Sato, W. Martens, D. Lee: Binaural measurement and simulation of the room acoustical response from a person's mouth to their ears. Acoustics Australia **37** (2009) 98–103.

[21] C. Pörschmann: Influences of bone conduction and air conduction on the sound of one's own voice. Acta Acustica united with Acustica **86** (2000) 1038–1045.

Document downloaded by @DAEL, 2026-11-03 03:15:48 - personal use only

[22] S. Reinfeldt, P. Ostli, B. Håkansson, S. Stenfelt, P. Östli: Hearing one's own voice during phoneme vocalization—Transmission by air and bone conduction. The Journal of the Acoustical Society of America **128** (2010) 751–762.

[23] D. Pelegrín-García: Comment on "Increase in voice level and speaker comfort in lecture rooms" [J. Acoust. Soc.Am. 125, 2072–2082 (2009)]. Journal of the Acoustical Society of America **129** (2011) 1161–1164.

[24] International Standard Organisation: Acoustics – Normal equal-loudness-level contours (ISO 226:2003). 2003.

[25] International Telecommunication Union: Recommendation ITU-T P.58: Head and torso simulator for telephonometry. Tech. Rept. International Telecommunication Union, Geneva, 2011.

[26] D. Hammershøi, H. Møller: Sound transmission to and within the human ear canal. The Journal of the Acoustical Society of America **100** (1996) 408–427.

[27] S. Schörnich, A. Nagy, L. Wiegrebe: Discovering your inner bat: echo-acoustic target ranging in humans. Journal of the Association for Research in Otolaryngology **13** (2012) 673–682.

[28] C. E. Rice: Human echo perception. Science **155** (1967) 656–664.

Document downloaded by @DAEL, 2026-11-03 03:15:48 - personal use only