HUMAN ECHOLOCATION USING CLICK TRAINS AND CONTINUOUS NOISE

Bo N. Schenkman[1], Mats E. Nilsson[2] and Nedelko Grbic[3]
[1]School of Management, Blekinge Institute of Technology,, SE-37179 Karlskrona, Sweden and Royal Institute of Technology, Stockholm, Sweden, [2]Gösta Ekman Laboratory, Stockholm University and Karolinska Institutet, SE-10691 Stockholm, Sweden, [3]School of Engineering, Blekinge Institute of Technology, Box 520, SE-37179 Karlskrona, Sweden.
Emails: [1]bsc@bth.se, [2]mnn@psychology.su.se, [3]ngr@bth.se

Abstract

Blind people may detect objects from the information in reflected sounds, echolocation. Detection as a function of the number of clicks compared to a continuous noise was tested by presenting clicks of 5 ms with rates from 1 to 64 clicks during a 500 ms period and a 500-ms continuous white noise. The sounds were recorded in an ordinary room through an artificial binaural head. The reflecting object was an aluminum disk, diameter 0.5 m, at distances of 1 and 1.5 m. These sounds were later presented to 3 blind and 16 sighted participants in a laboratory using a 2AFC methodology. The task was to detect which of the two sounds that contained a reflecting object. Feedback was provided. The blind participants had a higher detection than the sighted, but there were also differences among the blind. These results are put in relation to physical features such as the autocorrelation function and spectral variations.

Blind people use sounds to find out how the world around them is constituted. One source of information is the direct sounds emitted by sound sources in the environment. Another source is the echoes from reflecting objects in the environment, i.e. echolocation (sometimes the term “sonar” or “biosonar” is used for technical or animal systems). A number of factors affect the success of this ability for blind people. Some of them pertain to the structure and composition of the reflecting object, others to the characters of the room, and some to the abilities and experiences of the blind person. Another set of factors are the properties of the signal itself. The sound may be emitted by their own voice, or by using an external source, e.g. tapping with a long cane on the ground. The sound may also vary as to its duration in time, or the rate with which is repeated over a period of time.

The theoretical explanation for echolocation at close distances is usually by repetition pitch. According to a theory by Yost (1996), a number of parameters determine the strength of repetition pitch. One of these is the rate with which the stimuli are presented. We believe that an optimal number of clicks exist for which people may detect an object in front of them. In this study, we wanted (1) to determine this optimal rate of bursts for echolocation (2) to see the differences in detection for click trains and continuous noise, (3) to see the differences when the object was at 100 cm and at 150 cm, and finally, (4) to compare a selected group of high-performing blind persons with a group of sighted persons.

Method

Sound Recordings

Sound recordings were conducted in an ordinary lecture room using an artificial head placed with its ear entrances at the same height as the center of the reflecting object, 1.46 m above the floor. The object was an aluminum sheet, 1.5 mm thick and with a diameter of 0.5 m. Recordings were conducted at 100 and 150 cm distance between microphone (ear entrance) and the reflecting object. In addition, recordings were made with no obstacle in front of the artificial head. The emitted sound was either bursts of 5 ms each, varying in rates from 1 to 64 bursts per 500 ms or a 500 ms white noise. These sounds were generated by a loudspeaker (Genelec 1031A) placed 1 m straight behind the center of the artificial head. The obstacle was put on a microphone tripod in the room. The object's center position was at a height of 1.45 m in front of the artificial head at the two distances 100 and 150 cm. The loudspeaker was placed 1 m straight behind the center of the artificial head. The obstacle was put on a microphone tripod in the room. The object's center position was in front of the artificial head at the two distances 100 and 150 cm. The dimensions of the room were 9.4 x 7.2 x 3.8 m (l x w x h). The recording set up is shown in Figure 1. The ambient sound level in the room was 29 dB(A) and the reverberation time, T60, was 0.6 s. For further details on the principles of the recordings, see Schenkman and Nilsson (2010), although these recordings were made in different rooms than the one in the present study.

Experimental Sounds and Procedure

The experiment contained sounds recorded without the reflecting object (no-object-sounds) and sounds recorded in the presence of the reflecting object (object-sounds). The object-sounds were recorded with the object at 100 or 150 cm distance from the artificial head. In the experiment, pairs of object-sounds and no-object-sounds were presented and the task was to decide which sound that was the object-sound. The blind participants were tested at both distances, while the sighted only at the 100 cm distance. The reason was that we were certain that the 150 cm distance would be too difficult for the sighted group and would cause much frustration.

Fig. 1. Recording set up with the loudspeaker behind the artificial head and the reflecting object in front of the artificial head

A 2-alternative-forced-choice (2AFC) procedure was used. In each trial, a pair of sounds was presented sequentially in random order. The task was to decide which sound that contained reflections from the object. Each session contained 56 trials, all from one distance and one signal condition. The participant listened to the sound pairs through earphones and responded by pressing one of two keys on a computer keyboard. A recorded voice gave feedback on correct or incorrect response after each trial.

Participants

We tested 16 sighted and 3 blind participants. The sighted, 5 men and 11 women, were aged between 19 and 48 years (median = 24.5 years). The three blind participants, 2 men and 1 woman, were aged 58, 36 and 56 years. The hearing of the persons was tested by an audiometric test, using the Interacoustics Diagnostic Audiometer model, AD226. The measurement method used was Hughson Westlake which is a standard method for measuring pure-tone thresholds for detection (American National Standards Institute 1997).

Two of the blind persons had participated in previous experiments on echolocation conducted by the present authors, and they then had a much higher performance than most of the other blind participants. They may therefore be considered as expert echolocators. The third person was a mobile person, who seemed to rely much on his auditory abilities in his daily life, and therefore also could be assumed to be an expert in this area. All the blind participants had been blind since birth or early childhood.

The study was performed in accordance with the Swedish ethical standards as laid down in the 1964 Declaration of Helsinki and with the World Medical Association Helsinki Declaration as revised in October 2008 (http://www.wma.net/en/30publications/10policies/b3/index.html). The protocols were approved by the Regional Ethical Review Board in Lund, Sweden.

Signal Analysis

The frequency responses were calculated by a linear identification of the propagation channel between the loudspeaker and the receiving left ear of the artificial head. The used estimator is the H1 estimator given by the cross power spectral density of the transmitted signal and the received signal, divided by the auto power spectral density of the transmitted signal, i.e.

$$H1(f) = \frac{X(f) \cdot Y(f)^*}{|X(f)|^2}, \qquad (1)$$

where $X(f) = \int_{-\infty}^{\infty} x(t) e^{-i2\pi ft} dt$ is the Fourier transform of the transmitted signal (sent to the loudspeaker), and $Y(f)$ is the Fourier transform of the received signal (received at the left ear of the artificial head). The variables f, t denotes the real frequency (in Hz) and the time variable (in seconds), respectively, $(\cdot)^*$ denotes the complex conjugate operator. The autocorrelation function is calculated from the inverse Fourier transform of $H1(f)H1(f)^*$, i.e. the autocorrelation function is given by

$$\gamma_{hh} = \int_{-\infty}^{\infty} H1(f)H1(f)^{*} e^{i2\pi ft} df\ . \tag{2}$$

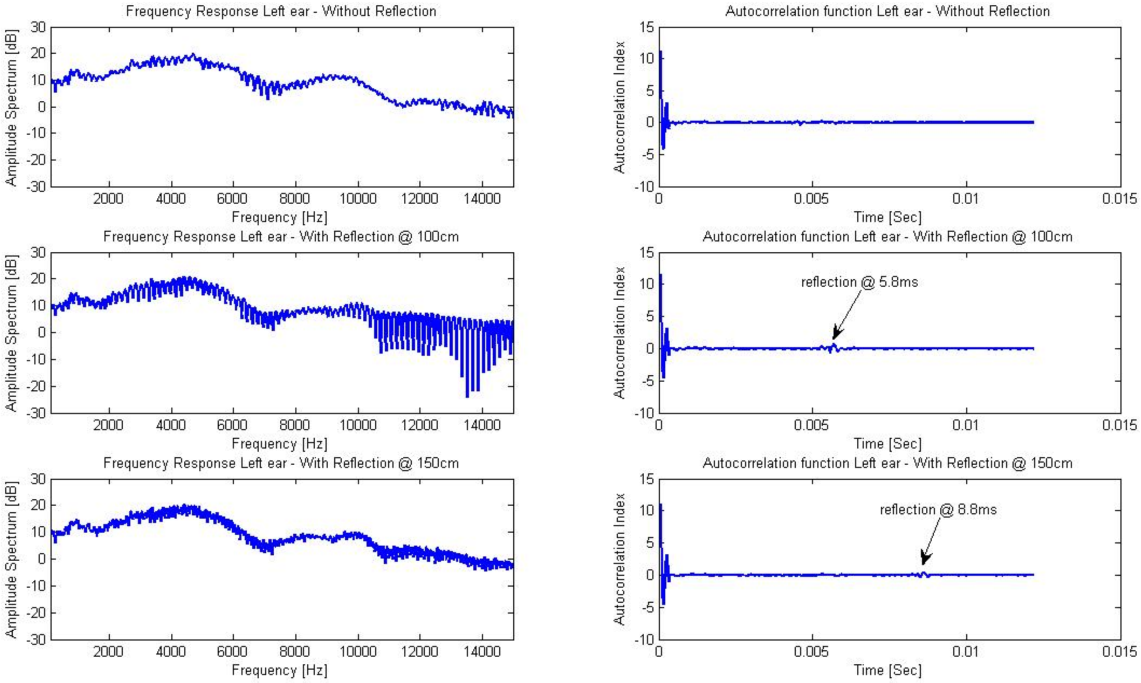

Fig. 2. Frequency response at left and autocorrelation function at right without reflecting object, and with reflecting object at 1 and 1.5 m

Left side of Figure 2 shows the resulting frequency response when a) there is no object present in the propagation part, b) when the object is placed 100 cm from the position of the artificial head and c) when the object it situated at 150 cm distance from the artificial head. It can be seen from the corresponding autocorrelation figures (to the right) that a reflection contribution occurs at 5.8 ms and 8.8 ms, respectively. For instance, when the object is situated 1 m from the artificial head, the propagation distance becomes 2 m (i.e. twice the distance, from propagation to the object and a return propagation) and the corresponding delay becomes $\tau = \frac{d}{c} = \frac{2}{342} = 5.8ms$, assuming that the speed of sound, c=342 m/s.

All the graphs, i.e. both frequency responses and ACF, are the same independently of the sound that was transmitted, since we were calculating the transmission of the room itself. That means, we were measuring the differences between the two cases, viz. if there was a reflecting object or not. The analysis that was made gives an answer primarily to the question for which excitation signals that a human person best can perceive the difference in the room when there is an object and when there is not an object. We show in Figure 2, by the use of the frequency response and the ACF, how this difference looks. Some excitation signals will mask this difference, while others will accentuate it.

Results and Discussion

The main results are shown in Figure 3. At the 100 cm distance, the sighted persons had their highest average detection at a rate of 32 bursts per 500 ms, mean percentage correct, p(c) = 0.73, SD= 0.14. The range of the sighted persons for this condition varied from p(c) = 0.46 to 0.78. The mean over all conditions for the sighted listeners was p(c) = 0.62, SD = 0.14.

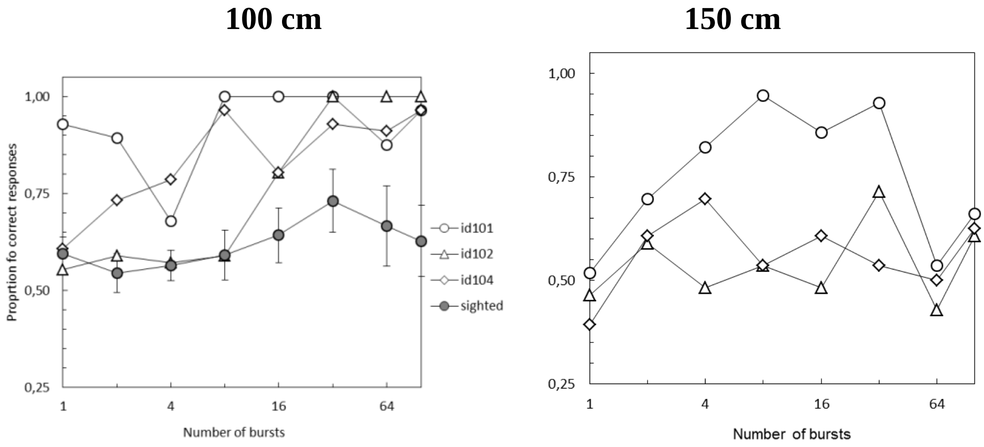

Fig. 3. Mean proportion of correct responses, p(c) for each rate of bursts and for the noise for the sighted persons (filled legends) with 95% confidence intervals and individually for the three blind persons (empty legends). The continuous noise condition is the rightmost level on the x-axis.

One may note that two sighted persons had 100 percent correct detections, both at the 64 bursts/500 ms condition. One of the blind persons had almost perfect performance at all rates higher than 8/500 ms. His mean p(c) = 0.92 over all the sound conditions. The second blind person had a rising performance with p(c) =1.0 at bursts greater than 32 burst/500 ms. Her mean p(c) = 0.76 over all sounds. The third blind had a mean p(c) = 0.83 at 100 cm. The

blind persons had almost 100 percent correct responses, for the continuous noise, p(c) = 0.98, while the sighted only had slightly more than 63% correct (range 0.30 to 0.93), SD= 0.17.

At the 150 cm distance, one blind person had the maximum performance at rates of 8 to 32 bursts/500 ms. His mean p(c) = 0.76. The second person had a rather irregular performance, and it is possible that the peak at 32 bursts was a result of guessing. Her mean p(c) = 0.54. The third blind person had a mean p(c) = 0.55. These three persons had a mean detection of only about p(c) = 0.63, when using the noise stimulus at the distance of 150 cm.

As to our four research questions, we found that (1) for the average sighted listener an optimal rate at 100 cm was 32 clicks/500 ms, whereas for the blind no single peak was seen. For the blind at 150 cm, a peak at 32 clicks/500 ms was suggested. (2) The blind participants used the continuous noise at 100 cm as proficiently as they used the best burst rates. (3) It was easier for the blind to detect the object at 100 cm than at 150 cm at all burst rates, also when using the noise. (4) The few blind persons we tested were better than the average for the sighted in their performance of echo detection.

The blind participants were highly proficient echolocators and are not typical of the blind Swedish population. The former persons are active, mobile and have a professional life, which is not the case for many blind people. It is probable that the general performance of the total Swedish blind population would be closer to the sighted group tested here than to these expert echolocators. One should note that most blind people become blind as adults, and some of the sensory compensations that probably takes place must do so in childhood, occurring at some critical period (Voss et al, 2010).

One common explanation for repetition pitch is by the autocorrelation function and since repetition pitch forms part of the explanation at close distances for human echolocation, this would also explain the present results (e.g. Schenkman and Nilsson, 2010). Another important factor is the time separation for the different bursts, so that they can be associated with the right originally emitted sound. In order to use the function, we believe that the human hearing has to be able to judge that a reflection belongs to a previously emitted sound. If the echo of a sound may not be resolved into which emitted sound that caused it, the perception may be ambiguous, and thus cause mistaken identities. We see this for the blind persons e.g. at the reduction of p(c) with 64 bursts at 150 cm distance. Still, even at the restricted laboratory conditions used in our experiment, where the blind persons could not move, or listen to the room or use their own head related function associated with their own ears, they still had an impressive performance.

Acknowledgments

This work was supported by a grant from the Swedish Council for Working Life and Social Research, FAS, (2008-0600). M.Sc. Maria Rådsten Ekman assisted in collecting the data.

References

American National Standards Institute (1997) Method for Manual Pure-tone Threshold Audiometry. ANSI S3.21-1978 (R1997) (New York: American National Standards Institute).

Schenkman, B.N. & 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.

Voss, P., Collignon, O., Lassonde, M. & Lepore, F. (2010). Adaptation to sensory loss. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1, 308-328

Yost, W. A. (1996). Pitch strength of iterated rippled noise. *Journal of the Acoustical Society of America*, 100, 3329-3335.