Applied Acoustics 106 (2016) 77–86

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Human echolocation: Acoustic gaze for burst trains and continuous noise

Bo N. Schenkman [a,b,*], Mats E. Nilsson [c], Nedelko Grbic [d]

[a] CTT – Centre for Speech Technology, Department of Speech, Music and Hearing, Royal Institute of Technology, Stockholm, Sweden
[b] Psychological Sciences Research Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium
[c] Department of Psychology, Stockholm University, Stockholm, Sweden
[d] Department of Electrical and Information Technology, Lund University, Lund, Sweden

ARTICLE INFO

Article history:
Received 2 July 2015
Received in revised form 7 December 2015
Accepted 14 December 2015
Available online 12 January 2016

Keywords:
Blind
Echolocation
Bursts
Noise
Orientation

ABSTRACT

This study explored the ability of blind and sighted listeners to detect reflections, "echoes", of burst trains or continuous noise. Echo detection was compared by presenting 5 ms bursts, rates from 1 to 64 bursts, with a continuous white noise, all during 500 ms. Sounds were recorded in an ordinary room through an artificial binaural head, the loudspeaker 1 m behind it. The reflecting object was an aluminum disk, diameter 0.5 m, placed at 1 m. The sounds were presented to 12 blind and 26 sighted participants in a laboratory using a 2-Alternative-Forced-Choice methodology. The task was to detect which of two sounds contained an echo. In Experiment 2, 1.5 m distance sounds were presented to the blind only. At 1 m, detection for the blind increased up to 64 bursts/500 ms, but for the sighted up to 32 bursts. At 1.5 m, the peak performance for the blind was at 32 bursts. At the 1 m, but not at the 1.5 m distance, the blind performed best with continuous white noise. The overlap in time of signal and echo at 1 m for 64 bursts was 60%, but at 1.5 m 82%. Avoiding an overlap between emitted bursts and returning echoes seems important for echolocation, indicating that an acoustic gaze, analogous to in echolocating animals, may also exist in humans.

1. Introduction

This paper investigates how burst train compares to continuous noise in influencing echolocation for humans. Human echolocation pertains to the structure and composition of the reflecting object, to the characteristics of the room, to the abilities and experiences of the blind person, the properties of the signal itself and if the sound was emitted by the individual's own voice, or by using an external source, such as tapping with a long cane on the ground. The sound may be a short burst, or a longer sounding noise. This study evaluates how echolocation varies with the rate at which a sound is repeated over a specific period of time and how varying burst rates compare to continuous noise.

The success of echolocation by blind people depends on a number of factors. Research has shown that in some cases, the spatial acuity of blind people may approach that of bats [44]. The human capacity for echolocation has presumably evolved by other ways, and is dependent on other factors. Studies on bats have shown that they modulate the outgoing signals according to the environmental demands (e.g. [22]), many species for example change the signal when they approach a prey. We suppose that blind people act in a somewhat similar adapting way, when they are close to or far from an object, when there are many reflecting objects, or when there is much ambient noise. They may need to change and adapt the signals to their emitted rate, intensity or frequency content.

The avoidance of overlap of signals is prevalent among echolocating animals. Madsen and Surlykke [21] used the concept of acoustic gaze adjustments when echolocating animals update their acoustic sampling of the world. Both bats and toothed whales have to wait for their echoes to return before emitting the next sonar pulse (see also [5]). If they emit them too fast, range ambiguity will occur. This happens when they emit a new sonar pulse before previous generated echoes have arrived. Seibert et al. [39] for one bat species they studied, compared each pulse-pair to a visual saccade and regarded the sonar beam movements between pulses as acoustic gaze saccades. To the adjustments belonging to acoustic gaze, Madsen and Surlykke [21] lists hearing, call rates, levels, frequencies and beam width of the emitted sound pulses. We believe that similar processes to avoid overlap also take place for human

* Corresponding author at: CTT – Centre for Speech Technology, Department of Speech, Music and Hearing, Royal Institute of Technology, Stockholm, Sweden.

E-mail addresses: bosch@kth.se (B.N. Schenkman), mnn@psychology.su.se (M.E. Nilsson), Nedelko.grbic@eit.lth.se (N. Grbic).

http://dx.doi.org/10.1016/j.apacoust.2015.12.008
0003-682X/

echolocation when using clicks. However, other signals such as hisses, more akin to noise signals may depend on other processes for which avoidance of overlap is not critical.

Echolocation by blind people is a subset of auditory skills called spatial hearing or auditory space perception [4]. We will use 'echolocation' to describe when a sound is emitted and its reflection is perceived by a person. Ashmead et al. [4] discuss the ability to walk along a wall by using low frequency information. One may note that Stoffregen and Pittenger [41] credits human echolocation with a more central role for perception and action. There are auditory skills which are not echolocation, but may be useful for blind people, e.g. the ability to identify objects that obstruct sounds [32] and to perceive the distance to objects (for a review see [25]). A blind person might be able to use echolocation also in situations where the sound source is not close to his/her body. Most ambient sounds such as traffic sounds originate at a distance around or behind the person. Sounds may also be emitted from the long cane [35]. Room acoustics can also have an effect on how echolocation is used (see [26]).

Another factor that may determine if echolocation is successful is the amount of information that the blind person has available during a limited time. In a previous study by Schenkman and Nilsson [36], it was found that people performed best in echolocating tasks with 500 ms white noise as compared to 50 ms and 5 ms burst sounds, where the worst performance was with the shortest sounds. However, the observers were given different time to listen to the sounds. If they missed the first burst of a 5 ms sound, e.g. for reasons of non-attention, the opportunity for attending to the quality change would be small. In contrast, for a 500 ms long signal they had half a second to attend to the sound, which made it less likely that it should be missed. This would mean that the short time duration of the bursts is not the essential factor, but rather the rate with which they are emitted. The potential information for the 500 ms signal is inherently larger than for a single 5 ms sound, which we in a previous article [36] called the 'information-surplus principle' implying that redundant information or information from many sources gives a more veridical perception.

Some blind echolocating persons are advocating methods based on clicking (e.g. [18]). On the other hand, some echolocating researchers have shown that blind people may use various methods (see e.g. [28]) including e.g. hisses. Physical analyses have been conducted on the acoustic properties of palatal signals, i.e. orally produced pulses or clicks [29] as well as hand and finger produced pulses [30]. The former were found to be better for echolocation, at least for the group of sighted persons who functioned as test persons. Gougoux et al. [16] showed that early-blind persons have a better pitch discrimination than either late-blind or sighted persons. That blind people are more sensitive to echoes has also found support in a study by Kolarik et al. [19], where totally blind listeners were better than sighted controls in distance-discrimination tasks.

Echolocation may be studied in different ways. In many of the earlier echolocation studies with people, sounds were orally emitted (e.g. [28,42]). This method is also used successfully today by many researchers, see e.g. Wallmeier et al. [49] and Kolarik et al. [20]. One may also construct a virtual room using a Kemar head (e.g. [33]). An alternative method is to use sounds produced mechanically or electronically and presented by a loudspeaker placed close to the mouth or ears (e.g. [36,37]).

The pitch phenomenon termed repetition pitch (see e.g. [7]), is commonly used as a theoretical explanation for echolocation at close distances. This pitch is created when a sound is repeated in a short time, and the signal with its repetition is perceived as a whole, a result of the interference patterns between emitted and returning waves. The signal with its repetition gives the original sound a coloration, a timbre, which is part of what is the basis for much of human echolocation. According to Yost [52], iterated ripple noise can function as the physical stimulus for the perception of repetition pitch, and 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 in any situation an optimal number of clicks exists for which people may detect an object in front of them. This optimal number may be due to psychoacoustical factors, as well as to room acoustical conditions, including the distance to a reflecting object. The focus of the present study is on how rates of 5 ms bursts of emitted signals affect the performance of echolocation, and how this performance compares to when white noise is used. In particular, we wanted (1) to determine an optimal rate of bursts for echolocation for the experimental situation chosen, (2) to see the differences in detection for click trains and continuous noise, (3) to see the differences when the object was at 1.0 m and 1.50 m, and finally, (4) to compare, at 1 m distance, a selected group of blind persons with a group of sighted persons.

2. Method

2.1. 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 recordings were made in a room with reverberations, since a previous study [36] had shown that for recorded signals such as those used in this study, and when the participants do not use their own vocalizations or emissions, then echolocation was better in an ordinary rooms than in an anechoic room. The object was an aluminum sheet, 1.5 mm thick and with a diameter of 0.5 m. Recordings were conducted at 1.0 and 1.5 m distance between the microphone (ear entrance) and the reflecting object. In addition, recordings were made with no obstacle in front of the artificial head.

The equipment used for the binaural recordings of the sounds consisted of an artificial manikin, a Head and Torso Simulator for binaural recordings (Brüel & Kjær type 4100), including two internal microphones (Brüel & Kjær type 4190) and pre-amplifiers (Brüel & Kjær type 2669). The microphone membranes were mounted at a position corresponding to the entrance of the ear canal of human listener.

Using the sounds recorded from the artificial head in the ensuing behavioral tests was presumably less efficient for the performance of blind people, than when they use their own ears. However, when comparing different stimulus conditions the differences are probably similar.

The emitted sounds were either bursts of 5 ms each, varying in rates from 1 to 64 bursts per 500 ms or a 500 ms white noise. We write the number of bursts that the participants were presented as bursts/500 ms, since this is what was actually presented to them. A conventional description as bursts/s, would be less accurate and misleading. The bursts had identical wave forms. The rise and fall time of the continuous white noise was 10 ms.

In this study, the geometry of the sound source relative to the head was changed compared to previous studies by the present authors, where the loudspeaker had been placed on the chest, close to the mouth. In this study, the sounds were generated by a loudspeaker (Genelec 1031A) placed 1 m straight behind the center of the artificial head. The main reason was that we wanted to see how ambient sound sources such as traffic noise can be used for echolocation. The reflecting object was placed on a microphone tripod in the room. The object's center position was at a height of