Human echolocation: waveform analysis of tongue clicks


Some blind individuals have the ability to detect and classify objects in complex scenes by using echolocation based on 'tongue clicks'. In this paper, we present a waveform analysis of the tongue clicks collected from three blind individuals who use tongue-click based echolocation on a daily basis. It is found that the tongue clicks are wideband signals and that the spectrum of clicks varies within and between individuals. However, by using the wideband ambiguity function, we find that all of the clicks from three different individuals share some common characteristics.

Introduction: Since the invention of radar and sonar, studies of echolocation abilities of bats and dolphins have contributed to advancing technology [1]. Now the study on human echolocation has opened a new opportunity: the cognitive processes behind human echolocation can demonstrate the concept of a perception-action cycle that is highly desirable in future radar and sonar systems [10].

Human echolocation was first studied by Supa in 1944 [2]. In their experiment, blindfolded participants demonstrated the ability to sense objects via the noise made by scuffing their heels on the ground. The obstacle could be detected between 3m and 5m with best performance by blind participants. More research on human echolocation was conducted in the 1960s, e.g. [3, 4]. It was found that most participants choose either a long hissing sound or a punctuated tongue click. Participants showed accurate spatial localization and discrimination of objects with the same surface area but different shapes. These days more research has accumulated to suggest that echolocation may also give information about an object’s distance and azimuth, shape, material, size and motion [5, 6, 9]. It has been suggested that the spectral composition of echoes may provide a vital source of environmental information [5, 7, 8]. In the last 5 years research into human echolocation has increased with a particular focus on echolocation using tongue clicks. Those blind echolocation experts who echolocate on a daily basis choose tongue clicks instead of other waveforms, suggesting investigation on this specific signal would provide insight to human echolocation.

The first step in understanding echolocation is to analyze the signals being used, e.g. the tongue clicks. Various time-frequency analyses have been conducted on bats’ emitted signals in order to characterize and investigate them, however, the time-frequency characteristics of human generated signals for echolocation have scarcely been investigated.

In this letter, we provide an initial waveform analysis for echolocation to assist in understanding human cognition. We show how the tongue clicks can vary as well as presenting their common characteristics. The human tongue clicks used for this analysis were recorded from three blind experts who use echolocation on a daily basis. Compared with [10], in which an analysis was made based on a single tongue click, this study is a significant advance, since multiple clicks from three different participants were investigated.

Data set used in study: The click samples are from three blind participants who have been using click based echolocation since their childhood (see Table 1). All experimental procedures had been approved by the Durham University department of Psychology ethics committee and followed BPS code of practice and Declaration of Helsinki (1968). Information and consent forms were provided in an accessible format, and we obtained informed written consent from all participants.

The experiment was conducted in a sound-insulated and echo-acoustic dampered room (approx. 2.9m x 4.2m x 4.9m, lined with acoustic foam wedges that effectively absorb frequencies above 315 Hz). Participants were positioned in the centre of the room (40 cm away from the microphone) with their head stabilized. Recordings were made with DPA SMK-SC4060 (a protective grid removed) and TASCAM DR100MKII at 24bit and 96kHz. The noise floor of the recording system was measured to be -90 dB/Hz. Each recording contained a large number of clicks, that were extracted into individual click files using a fixed threshold. The threshold level was set manually for each file to a level of clicks, that were extracted into individual click files using a fixed measured to be -90 dB/Hz. Each recording contained a large number of MKII at 24bit and 96kHz. The noise floor of the recording system was SMK-SC4060 (with protective grid removed) and TASCAM DR100-microphone) with their head stabilized. Recordings were made with DPA dampened room (approx. 2.9m x 4.2m x 4.9m, lined with acoustic foam and we obtained informed written consent from all participants. Information and consent forms were provided in an accessible format, childhood (see Table 1). All experimental procedures had been approved.

Data set used in study: The click samples are from three blind participants who have been using click based echolocation since their childhood (see Table 1). All experimental procedures had been approved by the Durham University department of Psychology ethics committee and followed BPS code of practice and Declaration of Helsinki (1968). Information and consent forms were provided in an accessible format, and we obtained informed written consent from all participants.

The experiment was conducted in a sound-insulated and echo-acoustic dampered room (approx. 2.9m x 4.2m x 4.9m, lined with acoustic foam wedges that effectively absorb frequencies above 315 Hz). Participants were positioned in the centre of the room (40 cm away from the microphone) with their head stabilized. Recordings were made with DPA SMK-SC4060 (a protective grid removed) and TASCAM DR100MKII at 24bit and 96kHz. The noise floor of the recording system was measured to be -90 dB/Hz. Each recording contained a large number of clicks, that were extracted into individual click files using a fixed threshold. The threshold level was set manually for each file to a level of clicks, that were extracted into individual click files using a fixed measured to be -90 dB/Hz. Each recording contained a large number of MKII at 24bit and 96kHz. The noise floor of the recording system was SMK-SC4060 (with protective grid removed) and TASCAM DR100-microphone) with their head stabilized. Recordings were made with DPA dampened room (approx. 2.9m x 4.2m x 4.9m, lined with acoustic foam and we obtained informed written consent from all participants. Information and consent forms were provided in an accessible format, childhood (see Table 1). All experimental procedures had been approved.

Data set used in study: The click samples are from three blind participants who have been using click based echolocation since their childhood (see Table 1). All experimental procedures had been approved by the Durham University department of Psychology ethics committee and followed BPS code of practice and Declaration of Helsinki (1968). Information and consent forms were provided in an accessible format, and we obtained informed written consent from all participants.

The experiment was conducted in a sound-insulated and echo-acoustic dampered room (approx. 2.9m x 4.2m x 4.9m, lined with acoustic foam wedges that effectively absorb frequencies above 315 Hz). Participants were positioned in the centre of the room (40 cm away from the microphone) with their head stabilized. Recordings were made with DPA SMK-SC4060 (a protective grid removed) and TASCAM DR100MKII at 24bit and 96kHz. The noise floor of the recording system was measured to be -90 dB/Hz. Each recording contained a large number of clicks, that were extracted into individual click files using a fixed threshold. The threshold level was set manually for each file to a level of clicks, that were extracted into individual click files using a fixed measured to be -90 dB/Hz. Each recording contained a large number of MKII at 24bit and 96kHz. The noise floor of the recording system was SMK-SC4060 (with protective grid removed) and TASCAM DR100-microphone) with their head stabilized. Recordings were made with DPA dampened room (approx. 2.9m x 4.2m x 4.9m, lined with acoustic foam and we obtained informed written consent from all participants. Information and consent forms were provided in an accessible format, childhood (see Table 1). All experimental procedures had been approved.

Data set used in study: The click samples are from three blind participants who have been using click based echolocation since their childhood (see Table 1). All experimental procedures had been approved by the Durham University department of Psychology ethics committee and followed BPS code of practice and Declaration of Helsinki (1968). Information and consent forms were provided in an accessible format, and we obtained informed written consent from all participants.

The experiment was conducted in a sound-insulated and echo-acoustic dampered room (approx. 2.9m x 4.2m x 4.9m, lined with acoustic foam wedges that effectively absorb frequencies above 315 Hz). Participants were positioned in the centre of the room (40 cm away from the microphone) with their head stabilized. Recordings were made with DPA SMK-SC4060 (a protective grid removed) and TASCAM DR100MKII at 24bit and 96kHz. The noise floor of the recording system was measured to be -90 dB/Hz. Each recording contained a large number of clicks, that were extracted into individual click files using a fixed threshold. The threshold level was set manually for each file to a level of clicks, that were extracted into individual click files using a fixed measured to be -90 dB/Hz. Each recording contained a large number of MKII at 24bit and 96kHz. The noise floor of the recording system was SMK-SC4060 (with protective grid removed) and TASCAM DR100-microphone) with their head stabilized. Recordings were made with DPA dampened room (approx. 2.9m x 4.2m x 4.9m, lined with acoustic foam and we obtained informed written consent from all participants. Information and consent forms were provided in an accessible format, childhood (see Table 1). All experimental procedures had been approved.

Data set used in study: The click samples are from three blind participants who have been using click based echolocation since their childhood (see Table 1). All experimental procedures had been approved by the Durham University department of Psychology ethics committee and followed BPS code of practice and Declaration of Helsinki (1968). Information and consent forms were provided in an accessible format, and we obtained informed written consent from all participants.

The experiment was conducted in a sound-insulated and echo-acoustic dampered room (approx. 2.9m x 4.2m x 4.9m, lined with acoustic foam wedges that effectively absorb frequencies above 315 Hz). Participants were positioned in the centre of the room (40 cm away from the microphone) with their head stabilized. Recordings were made with DPA SMK-SC4060 (a protective grid removed) and TASCAM DR100MKII at 24bit and 96kHz. The noise floor of the recording system was measured to be -90 dB/Hz. Each recording contained a large number of clicks, that were extracted into individual click files using a fixed threshold. The threshold level was set manually for each file to a level of clicks, that were extracted into individual click files using a fixed measured to be -90 dB/Hz. Each recording contained a large number of MKII at 24bit and 96kHz. The noise floor of the recording system was SMK-SC4060 (with protective grid removed) and TASCAM DR100-microphone) with their head stabilized. Recordings were made with DPA dampened room (approx. 2.9m x 4.2m x 4.9m, lined with acoustic foam and we obtained informed written consent from all participants. Information and consent forms were provided in an accessible format, childhood (see Table 1). All experimental procedures had been approved.
Fig. 1. Power spectrum density of human tongue clicks

compression must be considered. The WAF is defined as

$$WAF(\tau, \alpha) = \sqrt{|\alpha|} \int_{-\infty}^{\infty} s(t) s^*(\alpha t - \tau) \, dt$$

where $s(t)$ is the complex sampled tongue click signal, $\tau$ is the range delay and $\alpha$ is the Doppler compression factor which, for a target moving with constant radial velocity $v$, is given by

$$\alpha = \frac{c - v}{c + v}$$

where $c \approx 340 \text{ m/s}$ is the speed of sound in air. For electromagnetics, the Doppler compression factor is typically neglectable, but in acoustics, a target with 20 m/s velocity will cause $\alpha = 0.89$ in the echo. The zero-velocity cut, i.e. range profile of the WAF is the autocorrelation function of the signal, while the zero-delay cut, i.e. Doppler profile of the WAF can be used to evaluate the velocity resolution of the click.

Fig. 2 shows typical normalized WAFs from EE1, EE3 and EE2, respectively, with the Doppler compression axis converted to velocity using (2). Fig. 3 shows the corresponding range and velocity profiles of the WAF in Fig. 2. It can be observed that the WAF of all three clicks have a clear central peak as required for range and velocity estimation. Additionally, high side lobes can be observed in the range profile of EE1. These side lobes are close to the main lobe and might compromise the ability to separate two close-in-range targets. We found that the time interval between these lobes was approximately equal to the inverse of gap width in the frequency domain (i.e. between the major frequency component and the high frequency component). This gap can be observed from Fig. 1.

We further examined the 3dB width of the peaks in the range and Doppler profiles as well as the range offset at high velocity for all the clicks. We choose to evaluate the range offset at $v = 20 \text{ m/s}$, because such speed is relatively high for the targets often met in the participants’ life. The results are listed in Table 3 where $\Delta r$ is the 3 dB range resolution and $\Delta v$ is the 3 dB velocity resolution and RO is range offset. It can be observed that the average $\Delta r$ is relatively small compared to the size of real world targets while the average $\Delta v$ is relatively large compared to the speed of expected targets. This suggests that using such waveform would allow the localization of targets that are separated by just a few centimeters. Conversely, it would not be possible to separate targets that are close in range but have velocities differing by less than 10 m/s. Another common characteristic between all clicks is that the waveforms are Doppler tolerant, meaning the estimates of range will only be slightly offset even when target is moving with relatively high speed (e.g. 20 m/s).

Conclusion: In this paper, human tongue click signals were evaluated to understand their utility for echolocation. The average duration of a click was found to be 3 ms. The spectrum of the clicks showed them to be wideband signals. By using wideband ambiguity function, we identified common characteristics among all clicks. First, the tongue clicks provided fine range resolution but crude velocity resolution. Second, the human tongue clicks are a Doppler tolerant signal, since the range offset at substantial target speed was small. Using such a waveform, the range of a real world moving target, can be expected to be estimated accurately even if the target is moving. We hope that this analysis of human echolocation tongue clicks provides insights that are useful for evaluating cognitive processes used by echolocating humans such that it may be exploited by future radar and sonar systems.

Acknowledgment: This work was supported by the British Council and the Department for Business, Innovation and Skills in the UK (award SC037733).

X. Zhang (School of Information and Electronics, Beijing Institute of technology), G. Reich, M. Antoniou and M. Cherniakov (Department of Electronic Electrical and Systems Engineering, University of Birmingham), C. J. Baker and G. E. Smith (Department of Electrical and Computer Engineering, Ohio State University), L. Thaler (Department of Psychology, Durham University), D. Kish (World Access for the Blind)

E-mail: zhangxinyu90111@gmail.com

References

2 M. Supa; M. Cotzin and K. M. Dallenbach: ‘Facial vision’: the perception of obstacles by the blind’, American Journal of Psychology, 1944, 57, p. 133-183