UNIVERSITY OF
BIRMINGHAM

University of Birmingham
Research at Birmingham

Human echolocation

Antoniou, Michail; Cherniakov, Mikhail; Smith, Graeme; Zhang, Xinyu; Reich, Galen; Baker, Chris; Thaler, Lore; Kish, Daniel

DOI:
10.1049/el.2017.0454

License:
Other (please specify with Rights Statement)

Document Version
Peer reviewed version

Citation for published version (Harvard):
Antoniou, M, Cherniakov, M, Smith, G, Zhang, X, Reich, G, Baker, C, Thaler, L & Kish, D 2017, 'Human echolocation: waveform analysis of tongue clicks', *Electronics Letters*. https://doi.org/10.1049/el.2017.0454

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
This paper is a postprint of a paper submitted to and accepted for publication in Electronics Letters and is subject to Institution of Engineering and Technology Copyright. The copy of record is available at IET Digital Library.

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 11. Oct. 2024

Human echolocation: waveform analysis of tongue clicks

X. Zhang, G. Reich, M. Antoniou, M. Cherniakov, C. J. Baker,
L. Thaler, D. Kish, G. E. Smith

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 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 obstacles 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 for understanding the cognitive processes of human 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 dampened 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 (with protective grid removed) and TASCAM DR100-MKII 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 that identified the tongue clicks, but did not accept other sharp sounds that resulted from the participant moving. The output of this process were many individual files, each containing a single click with short periods of silence at the start and end.

Table 1: Description of participants in the experiment

Participant	Description
EE1	male, age 33; optic nerve atrophy; total blindness (since age 14); click based echolocation since age 15;
EE2	male, age 49; Retinoblastoma; total blindness (since age 1); click based echolocation as far as can remember;
EE3	male, age 31; Glaucoma; total blindness (since age 12); click based echolocation since age 12.

Tongue click signal properties: We first examined the duration of the clicks. To estimate a click duration, the envelop of the individual click files was extracted. This was achieved by first squaring all samples in the click file, then applying a zero-delay moving average window of length 20 samples and finally taking the square root of the filter's output. The click duration was determined based on a threshold set to a percentage of the envelop peak value. The threshold level was set as either 5% or 1% of the peak value. The start of the click was taken to be the last sample above the threshold when reducing the sample index from the peak's index to the start of the individual click file. Similarly, the end of the click was taken to be the last sample above the threshold when increasing the sample index from the peak's index to the end of the file. Once the start and end of the click was found, the duration was calculated by subtraction. The means and standard deviations of the durations, for the two threshold levels, are listed in Table 2, where SD is short for standard deviation. Based on the table, 3 ms was taken as a representative click duration. This was selected as a compromise between the 1% and 5% thresholds combined with the experimenters' empirical experience working with the tongue clicks. A duration of 3 ms would imply a range resolution of 51 cm, assuming a rectangular click envelope. As will be seen shortly, the actual resolution is better than this, implying the modulation of the click is significant.

Table 2: Duration and major comp. frequency of the clicks

Participant	Duration (ms) in 5% thresh.		Duration (ms) in 1% thresh.		Major comp. frequency(kHz)	
	Mean	SD	Mean	SD	Mean	SD
EE1	3.44	0.68	7.00	2.41	2.09	0.48
EE2	2.31	0.42	4.13	0.60	3.46	0.43
EE3	2.44	0.68	4.69	1.23	4.09	2.11

The spectrum of tongue clicks was estimated by periodogram with a Chebyshev window. Fig. 1 presents the power spectral density (PSD) of typical tongue clicks from each one of the participants. The clicks from all participants have a major frequency component and several minor components. The frequency of this main component varies both for individuals and between different participants. The mean and standard deviation of the frequency of this component are listed in Table 2. It can be seen that the major component frequency of clicks from EE2 and EE1 exhibited less variation than EE3. Besides, a number of minor components existed in the power spectrum. Although the power and frequency of these minor components vary for individuals, there is always a component located around 10kHz (referred to as the high frequency component). The power of this component in the clicks from EE1 and EE3 was usually higher than for EE2. The PSD shows that the human tongue clicks are wideband signals due to the large span of the multiple components, compared to the center frequency. However, it is not clear how these spectral features relate to the range resolution of the tongue clicks. If the total bandwidths of the clicks were used, the range resolution would be estimated as less than 1 cm. This estimate does not account for the noncontiguous nature of the spectrum, however. As will be seen in the following section, the wideband ambiguity function (WAF) analysis indicates the range resolution to be somewhat cruder than that implied by the total bandwidth.

Wideband ambiguity function of tongue clicks: The WAF was used since the tongue clicks are wideband signals and for acoustic signals, Doppler

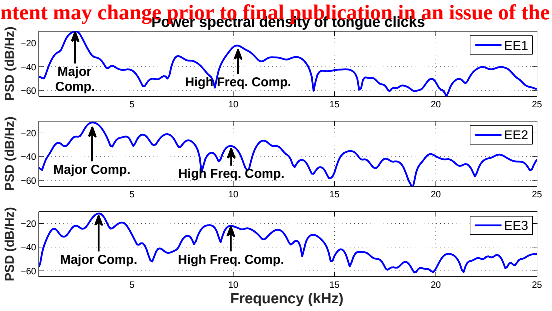

Fig. 1. Power spectrum density of human tongue clicks

compression must be considered. The WAF is defined as

$$WAF(\tau,\alpha) = \sqrt{|\alpha|} \int_{-\infty}^{\infty} s(t)\, s^*(\alpha t - \tau)\, dt \tag{1}$$

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

$$\alpha = \frac{c - v}{c + v} \tag{2}$$

where $c \approx 340$ 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.

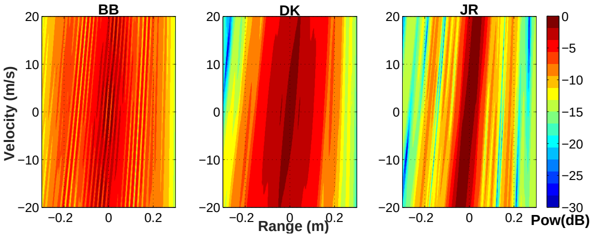

Fig. 2. Typical wideband ambiguity function of tongue clicks

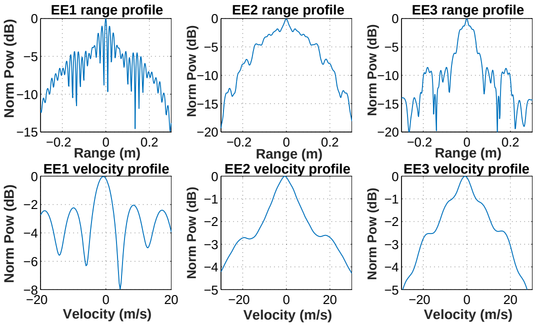

Fig. 3. Typical wideband ambiguity function profiles

Table 3: Characteristic parameters from WAF

Participant	Δr (cm)		Δv (m/s)		RO (cm)	
	Mean	SD	Mean	SD	Mean	SD
EE1	8.6	3.3	36.2	12.3	2.7	0.8
EE2	3.7	1.9	23.2	7.2	1.8	0.5
EE3	2.1	1.6	21.2	14.6	1.8	2.2

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 $v = 20$ 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 Δr is the 3 dB range resolution and Δv is the 3 dB velocity resolution and RO is range offset. It can be observed that the average Δr is relatively small compared to the size of real world targets while the average Δv 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. Such a characteristic implies that velocity information is probably not estimated by Doppler for echolocating humans. 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

1 Michele Vespe; Gareth Jones; Chris J. Baker: 'Lessons for radar', *IEEE Signal Processing Magazine*, 2009, **26**, p. 65-75
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
3 C. E. Rice: 'Human echo perception', *Science*, 1967, **155**, p. 656-664
4 C. E. Rice: 'Perceptual enhancement in the early blind', *Psychological Record*, 1969, **19**, p. 1-14
5 T. A. Stoffregen and J. B. Pittenger: 'Human echolocation as a basic form of perception and action', *Ecological Psychology*, 1995, **7**, p. 181-216
6 Thaler, L.; Goodale, M.A.: 'Echolocation in People: An Overview', *WIREs Cogn Sci*, 2016, doi: 10.1002/wcs.1408
7 B Hughes: 'Active artificial echolocation and the nonvisual perception of aperture passability', *Human movement science*, 2001, **20**, p. 371-400
8 Bo, N. Schenkman; Mats, E. Nilsson: 'Human echolocation: Pitch versus loudness information', *Perception*, 2011, **40**, p. 840-852
9 Kolarik, AJ; Cirstea, S; Pardhan, S; Moore, BC: 'A summary of research investigating echolocation abilities of blind and sighted humans', *Hear Res*, 2014, **310C**, p. 60-68
10 Graeme, E. Smith and Christopher, J. Baker: 'Human echolocation waveform analysis', *Radar Systems (Radar 2012), IET International Conference on*, 2012, p. 1-5